Volume Title: Studies in Income and Wealth

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

Volume Publisher: NBER

Volume ISBN: 0-870-14166-X

Volume URL: http://www.nber.org/books/unkn49-1

Publication Date: 1949

Chapter Title: International Comparison of Real National Income: A Note on Methods

Chapter Author: Hans Staehle

Chapter URL: http://www.nber.org/chapters/c5708

Chapter pages in book: (p. 221 - 272)

Part III

International Comparison of
Real National Incomes:
A Note on Methods

Hans Staehle

International Monetary Fund

1 Nature of the Problem

Although other methods, especially the comparison of volumes of
production, could be devised, it has become customary to meas-
ure real national income by the same method as real individual
income, i.e., by dividing a cost of living index into money incomes.
This approach, which focuses attention upon real income as a
measure of economic welfare, rather than of bulk, will be followed
here. Consequently, the international comparison of real national
incomes calls, first, for solving the problem of comparing the cost
of living to individuals in different countries. A modest attempt
will also be made to state, if not to answer, one or two neglected
points that arise in the comparison, both interlocal and intertem-
poral, of aggregate incomes of many individuals.

The generally accepted definition of a cost of living index is the
ratio of money incomes that will make a consumer (individual or
family) equally well off in two or more situations. The problem of
measurement then consists in finding a criterion for the equiva-
lence of money incomes in different situations.

The theory of this problem and its solution are complete for the
special case in which the situations to be compared differ only with
respect to prices, and where all commodities present in the market
can be purchased freely. All other factors, such as tastes, habits,
and wants, and the range of available commodities, are assumed
to be constant throughout. The criterion of equivalence is the
indifference of choice between different incomes, and the charac-
teristic feature of the solution consists in interpreting consumers'
behavior as the expression of preferences.

For example, suppose a man's income in situation I is just suffi-
cient to purchase the same quantity of each commodity he bought
in situation 0. Whether or not in the new situation he buys the
old collection, we know that he will be at least as well off in situ-
ation I as in situation 0. The ratio between the cost of the old
collection of goods valued at the new and at the old prices, respec-
tively, will thus be a maximum estimate (or an 'upper limit') of
the change in the cost of living.[1]

[1] The official cost of living indexes published in most countries, being computed
by the method described (i.e., on fixed 'base-year' weights), actually overstate
the rise and understate the fall in the 'true' cost of living.

Without going further into this body of theory,[2] let us note that its application is definitely limited to situations where we can assume wants are identical. Without such identity, the concept of choice, and therewith the criterion of equivalence basic to the theory, loses all meaning.

This gives rise to two questions: First, since there are no objective criteria for the identity of wants, what are the opinions expressed in the literature concerning the limitations this indispensable condition imposes upon the applicability of the theory referred to? Second, does it follow that even the concept of equal well-being as between persons whose wants are different becomes meaningless?

There are wide differences of opinion concerning the conditions in which the assumption of identical wants can be granted. Most writers agree that, for people living in the same locality, the identity can be admitted as between two distinct periods, provided the interval is short, say, a few years.[3] But R. G. D. Allen, for instance, insists that the concept of a change in the cost of living "is in no way a collective one, nor can it refer to the same, or different individuals at different times or in different markets".[4] A. L. Bowley, at the other extreme, would test the relevance of the assumption by general statistics of consumption "which show for England that the changes in habit were only slightly different between 1904 and 1927".[5]

Despite these variations, writers are unanimous in holding that the assumption of identical wants *cannot* be granted when living in different countries is compared. Any attempt to apply the theory in international comparisons of real incomes should therefore be ruled out.

But does the absence of identical wants destroy also the very meaning of the problem before us? J. R. Hicks, for example, has

[2] For a convenient summary of the various contributions, see Ragnar Frisch, 'The Problem of Index Numbers', *Econometrica*, Vol. 4, pp. 1–38.

[3] For quotations see my *International Comparison of Food Costs* (International Labour Office, Geneva, 1934), pp. 87 ff.

[4] 'On the Marginal Utility of Money and Its Application', *Economica*, May 1933, p. 201.

[5] 'Notes on Index Numbers', *Economic Journal*, XXXVIII (1928), 223–4.

argued: "if the assumption of constant wants cannot be granted, the question whether a consumer is better off in one situation or in the other loses all economic meaning".[6] In strict logic, of course, there is no way of disagreeing with this contention. In strict logic, therefore, the problem at hand ought to be abandoned as essentially insoluble, at least from the economist's viewpoint.

Yet, if we heeded Hicks, we would be implying that the Executive Committee of this distinguished Conference in requesting a paper on international comparisons of the cost of living were setting an insoluble problem. If only for that reason, another course seems preferable.

It makes indeed perfectly good sense to inquire whether people in different countries are or are not equally well off. The question is so reasonable that attempts to answer it have repeatedly been undertaken by several organizations reputed to be responsible and sane. To quote a few outstanding examples, the British Board of Trade in the first decade of this century compared the cost of living in Britain with the cost of living in Germany, France, Belgium, and the United States;[7] the Unilever Corporation (Lever Bros., Ltd.) estimated in 1930 the comparative cost of living in a considerable number of countries;[8] and the International Labour Office in 1931 compared the cost of living in fourteen European cities (including Barcelona, Istanbul, and Helsinki) with the cost of living in Detroit.[9]

We thus have the incredible spectacle of a government office, an international authority, and an international corporation all indulging in answers to a question that, strictly speaking, has "no economic meaning".

This discrepancy between theory and practice disturbed Vilfredo Pareto a half century ago: "no rigorous comparison of satisfaction is possible between different individuals"; on the other hand, "we daily compare the welfare of different persons

[6] 'The Valuation of Social Income', *Economica*, May 1940, p. 107.
[7] See the official British publications Cd. 3864 (1908), Cd. 4032 (1908), Cd. 4512 (1909), Cd. 5065 (1910) and Cd. 5609 (1911).
[8] For a summary of the investigation see *The Economist* (London), November 1930.
[9] *A Contribution to the International Comparison of Costs of Living*, Studies and Reports, Series N, 17 (International Labour Office, Geneva, 1932).

living in the same or even in different communities"; it would
"run counter to common sense, and be manifestly absurd" if we
concluded that such comparisons rest upon mere illusion.[10]

Like Pareto, we shall take the position that, while rigorously no
comparison is possible when wants differ, there seems to be room
for admitting the existence of intermediate states between identity
and non-identity of wants. If this is granted, there seems to be no
reason to deny the logical possibility of comparing the welfare of
individuals. At the same time we must not expect our results to be
of equal quality in all cases. Whatever the method, the reliability
of the outcome will depend not only on it but also on the position
the case occupies on the long scale between complete identity and
complete non-identity of tastes and wants. It may well be that
certain comparisons may turn out to be impossible. The best we
can hope for are results that will not be unreasonable. In all cases,
checks of consistency will have to be applied before one method
is adopted.

2   EXTRANEOUS STANDARDS OF EVALUATION

Before coming to grips with our problem, one more point must be
cleared up. The recent development in the scientific analysis of
diets has led to the establishment of standards of valuation. Work
of a similar kind has been done with respect to shelter. But how-
ever sound such standards may be from a nutritional or hygienic
point of view, they should not be used as criteria of 'equivalence'
in the sense relevant to cost of living comparisons. There is no
relation between, say, the nutritional 'adequacy' of a given diet
and the economic welfare of those who consume it. We should not
forget that most families eat to please themselves, not nutrition-
ists. It is, of course, quite possible that only ignorance prevents
many from following diets prescribed by nutritionists and that, if
properly informed, they would be glad to oblige. But if consumers
fail to conform (although their income might be sufficient to do
so), whether deliberately or from lack of knowledge, we surely
would be wrong to conclude that they are worse off than if they
did. We just do not know.

[10] *Cours d'Economie Politique* (Lausanne, 1897), Vol. 2, p. 47; see also the discussion
on pp. 48–51.

For example, Hazel K. Stiebeling and Esther K. Phipard found in their important work on the nutritional adequacy of diets according to income (or more exactly per capita food expenditures) that in the Pacific region, of 190 families analyzed, 21 spent on food about $1.50, and 37 about $4.00, per person per week.[11] In the former group one family's diet was classed as 'good', i.e., as deserving the highest commendation from the nutritional standpoint. In the latter group, 28 families failed to make the 'good' grade. A strict application of the adequacy standards would lead to the conclusion that, although spending nearly three times as much on food, the 28 families were deriving less satisfaction from eating than the one family whose expenditure was very modest. It would clearly be difficult to accept such a result.

The use of extraneous standards of evaluation just discussed is only a special case of the practice of 'equivalent substitution' which, though equally reprehensible, cannot always be avoided in cost of living comparisons, where consumption habits differ markedly. It is, of course, very tempting to consider the Englishman's tea as equivalent to the American's coffee, and Scandinavian rye bread to French wheat bread, or even fuel in cold climates to refrigeration in warm climates. But, however reasonable such 'equivalences', they represent judgments on others' satisfaction, and to that extent, they are wholly unwarranted and arbitrary. Especially when important interests are involved, such substitutions can be used to influence the result of the comparison. If at all possible, it is safest to adhere to the principle emphatically put forth by Gottfried Haberler: let consumers themselves by their behavior indicate their preferences.[12]

3   METHODS USED AND PROPOSED BY VARIOUS AUTHORS

What has been said about the inapplicability of index theories that postulate identical wants, and the inadmissibility of substitutions, actually disposes of the methods used in all the better known statistical attempts at international cost of living comparisons. Of the investigations already mentioned, that by the Unilever Com-

[11] *Diets of Families of Employed Wage Earners and Clerical Workers in Cities* (Department of Agriculture, Circular 507, Jan. 1939), p. 74.
[12] *Der Sinn der Indexzahlen* (Tuebingen, 1927), pp. 81–3.

pany must be disregarded for lack of information concerning methods. The published summary gives only results.

The widely known book by Colin Clark, *The Conditions of Economic Progress*, can be mentioned only in passing, for the same reason. Clark is very vague about the methods and data used in establishing comparisons between real incomes in a vast number of countries. At any rate, while his results may be excellent, and in many cases are very plausible, his contribution to the methodological aspects of the problem is not helpful in the present discussion.

The inquiry carried out by the British Board of Trade before World War I relied mainly upon a comparison of food prices and rents in England and other countries. In estimating food price levels, it used what has come to be known as Irving Fisher's 'ideal' formula which, however, cannot, in the light of later development in the economic theory of index numbers, be accepted unless applied to situations where tastes and wants are identical. The relative level of rents was measured by using equality in the number of rooms as a standard of equivalence, although the report on Germany noted that rooms were generally both larger and higher than in England. As a form of 'equivalent substitution', this procedure was bound to yield arbitrary, however reasonable, results.

The International Labour Office similarly used a mixture of direct price comparison and 'equivalent substitution', the former with respect to food, the latter concerning all other elements— shelter, clothing, fuel, and miscellaneous items. Moreover, being an attempt to determine incomes in Europe equivalent to those received by American workers in Detroit, it referred to a very specific level of real income which was in most cases far higher than that usually enjoyed by persons belonging to the same social and occupational classes in Europe. Since there is no reason why the relative cost of living should be independent of income, this investigation would not be particularly suitable for purposes of national income comparisons.

Because of the methods, neither of these two large investigations can make serious claims to accuracy—which, of course, in no way prevents their results from being essentially acceptable on grounds of common sense plausibility.

It might well be argued that, in view of the wide margin of error to which data on individual, and even more so, aggregate incomes are subject, cost of living comparisons of the kind mentioned are quite sufficient, despite the theoretical objections that may be leveled against them. And had more such inquiries been made, or could they easily be made, they might indeed be the best practical course to take. The difficulty is that investigations of this kind call for the collection of special data, with the help of a considerable staff, and entail heavy expenditure. If a sufficient number of countries are to be covered, only a well endowed international organization could reasonably undertake such a task.

Therefore the feasibility of comparisons that use data either already available or more easily procured remains to be considered. Unfortunately, only three methods that come under these headings seem to have been proposed.

First we may mention a device proposed by J. M. Keynes as the "direct method of comparing incomes of similar persons".[13] "When an Englishman is offered an appointment in Australia, . . . and is wondering what the money-income he will get is going to be 'worth', . . . he does not usually consult any of the official index-numbers, . . . he asks a friend who is acquainted with the conditions of life in the two places." Keynes even claims superiority for this kind of comparison when the character of expenditure differs widely and a substantial proportion is of the nonstandardized sort an index cannot cover. "In such cases . . . the direct method may yield a result materially different from the indirect (i.e., statistical) method and yet be nearer the truth."

This method which, Keynes observes, has been "entirely discarded by statisticians" (mainly, we may note, because it is not statistical) is here mentioned quite seriously. It certainly deserves attention, if only as a check on statistical computations. And it may well be that, arbitrary and vague as it seems, it is the only feasible solution in cases otherwise desperate. At any rate, whenever we call a specific result 'plausible', we have the results of Keynes' "direct method" in the back of our mind.

Of methods that use statistical materials of a kind available for

[13] *A Treatise on Money* (London, 1930), I, 100-2.

several countries, there are only two: one proposed by Ragnar Frisch, the other by myself. In an early contribution to the subject, Frisch sets out to find an economic parameter, defined only by the prices and quantities in one situation, or only by the prices and quantities in the other situation (so that no direct price or quantity comparisons are involved), that could be used as a criterion of equivalence for pairs of incomes in the two situations.[14] As such a parameter he suggests the flexibility of the marginal utility curve of money. After considerable refinement in the theory underlying it, he repeated his proposal in 1936, when he suggested, though cautiously, that the method might be applied even if the indifference maps and goods were different. But, he added, the association of such 'equivalent' incomes would then have merely heuristic value (see note 2).

The principle of Frisch's method is to choose some pure number, independent of units of measurement but closely connected in any given situation with the level of well-being, and to take the equality of this indicator of welfare for incomes belonging to different situations as the criterion of equivalence.

With the principle or the theory of Frisch's method, I have no quarrel. The forceful originality that characterizes all his writings cannot help impressing anyone who takes the trouble to study them. But anyone who has ever attempted to apply his method to statistical data cannot escape the impression that the number of unavoidable arbitrary decisions involved in numerous successive interpolations and smoothing processes is excessive.[15] It would, in my view, take more than ordinary courage to put the results to practical use. At any rate, a great deal more experimentation is required before Frisch's method can be accepted.

4   THE DISSIMILARITY METHOD

In describing my own method, I wish first to apologize to those who may have heard me at meetings of the Econometric Society. To members of this Conference who did not, I must confess that, after the many years that have elapsed since I first developed it,

[14] *New Methods of Measuring Marginal Utility* (Tuebingen, 1932), pp. 72 ff.
[15] In a small mimeographed pamphlet, dated July 1937, Frisch gave detailed practical instructions for the computation of his parameter.

I feel quite humble about it. Not that I have found any flaw in it. On the contrary, I still regard it as a fairly good approximate solution of a difficult problem, despite its lack of theoretical foundation. But, with the passage of time, I have become more convinced that in every economic problem too much is non-economic, and too much nonquantitative, for any method to be fool-proof.

With this general reservation, I now proceed to discuss, first, the method itself, then a few applications.

Let q_0', q_0'', etc. and q_1', q_1'', etc. stand for the quantities of a number of commodities consumed by two consumption units (families of given type or 'equivalent adult males'). Write p_0', p_0'', etc. and p', p_1'', etc. for the corresponding prices. Then, for any given article,

$$\frac{q_1}{q_0} - \frac{\Sigma\left(\frac{q_1}{q_0}\right) q_0 p_0}{\Sigma q_0 p_0}$$

is the deviation of its quantity ratio from the weighted average of the quantity ratios of all articles.

Taking each such deviation relatively to the weighted average from which it is measured, we write

$$\frac{q_1}{q_0} \cdot \frac{\Sigma q_0 p_0}{\Sigma q_1 p_0} - 1.$$

Finally, a weighted average of these relative deviations is obtained, without respect to sign, again using the same weights. The result:

$$D = \Sigma \left| \frac{q_1}{q_0} \frac{\Sigma q_0 p_0}{\Sigma q_1 p_0} - 1 \right| \frac{q_0 p_0}{\Sigma q_0 p_0} = \Sigma \left| \frac{q_1 p_0}{\Sigma q_1 p_0} - \frac{q_0 p_0}{\Sigma q_0 p_0} \right|$$

may be looked upon as a measure of dissimilarity in the structure of consumption as between 0 and 1.[16] The value of D ranges from

[16] An analogous measure may be obtained, for the same pair of units, using I instead of O as base:

$$D' = \Sigma \left| \frac{q_0 p_1}{\Sigma q_0 p_1} - \frac{q_1 p_1}{\Sigma q_1 p_1} \right|.$$

The numerical difference between D and D' is usually small, and their variations are in the same direction. When the same prices apply to the two consumption units compared, D is equal to D'.

nil (in case all the q_0's and the q_1's are proportional) to 2 (in case there are no common articles of consumption in the two budgets).

Table 1 gives the values of D calculated for all possible combinations from the per capita consumption of eleven important foodstuffs by five income groups of German workers' families in 1927–28. In the first row, the values of D rise from zero to 0.352. In the third row, they first fall from 0.195 to zero, then rise to 0.194.

Now suppose D is used for comparing consumption at different income levels in one situation with consumption at different income levels in another situation. For example, Table 2 gives the values of D obtained in the systematic intercomparison of the same five German income groups in 1927–28 with six income groups in 1906–07. It no longer contains any cases of zero-dissimilarity. But as in Table 1, the values of D in the first row rise continuously from left to right. Inspection of further rows shows that, apart from minor irregularities, for increasing incomes of 1927–28, D diminishes, reaches a minimum, then rises again. The same holds true for the columns: the minima occur with the same income combinations as in the rows. As in Table 1, the minima (whose values fluctuate narrowly between 0.290 and 0.317) shift to the right as the five income groups in 1927–28 are compared with higher income groups in 1906–07.

Tables 1 and 2 lend themselves to the following interpretation. If it can be assumed that all the families considered in any one situation have identical wants, live in the same milieu, and are confronted by the same prices, differences in their consumption must be due to differences in their real incomes. It will be observed (Table 1) that D becomes zero when the difference in (real) incomes is nil, and increases with the widening difference (in either direction) between a given income and another income, pertaining to the same situation, being compared with it.

When a given income in one situation is compared with a series of incomes in another (Table 2), the movement of D resembles that observed for comparisons within the same situation. Opposite any one income in one situation D is at a minimum value in the other situation. But that value is no longer zero. It represents a

minimum of dissimilarity among any two families' consumptions which belong, one to the one, the other to the other situation considered. That minimum of dissimilarity is due to differences in

TABLE 1

Values of D Based on the Consumption of Eleven Foodstuffs Obtained in the Comparison of Five Income Groups, German Workers' Families, 1927–1928

Annual income (Rm.) per person	Annual income (Rm.) per person				
	631	708	782	888	1,026
631	0.000	0.120	0.195	0.274	0.352
708	0.120	0.000	0.082	0.160	0.264
782	0.195	0.082	0.000	0.083	0.194
888	0.274	0.160	0.083	0.000	0.113
1,026	0.352	0.264	0.194	0.113	0.000

Die Lebenshaltung von 2,000 Arbeiter-, Angestellten- und Beamtenhaushaltungen, Einzelschriften zur Statistik des Deutschen Reichs, 22 (Berlin, 1932). The classification in the original is according to income per 'man', but the average incomes were recomputed per person for comparison with the prewar inquiry used in Table 2.

TABLE 2

Values of D Based on the Consumption of Eleven Foodstuffs Obtained in the Comparison of Six Income Groups, 1906–1907, with Five Income Groups, 1927–1928, German Workers' Families

Annual income (M.) per person 1906–07	Annual income (Rm.) per person, 1927–28				
	631	708	782	888	1,026
214	0.290	0.375	0.394	0.455	0.473
282	0.306	0.313	0.304	0.350	0.368
319	0.351	0.334	0.317	0.325	0.343
375	0.386	0.343	0.321	0.318	0.315
407	0.396	0.359	0.342	0.339	0.338
449	0.404	0.371	0.355	0.351	0.355

Same source as in Table 1 for Germany, 1927–28. For Germany, 1906–07, *Cost of Living in German Towns, Report of an Enquiry by the Board of Trade,* Cd. 4032 (1908).

wants, milieu, and relative prices as between the two situations. D increases, in both the rows and the columns of Table 2, with the income distance from the minimum value. Keeping in mind our assumption that within each situation the differences in consumption structure are due exclusively to differences in real income, and granting that the effects upon the value of D of different

wants, etc. on the one hand, and of different real incomes on the
other, are roughly additive, we can infer that the difference be-
tween the real incomes that correspond to the two incomes oppo-
site which the minimum D value occurs is as small as possible.
Such pairs of incomes may therefore be considered approximately
equivalent in satisfaction.

The statistical results of intersituation comparisons may be rep-
resented in three dimensions, plotting the money incomes in one
situation (say, the M_0's) in one horizontal direction, the money
incomes in the second situation (the M_1's) horizontally in the other
direction, and the values of D vertically. A surface may then be
obtained by interpolating between the observed points. *Whenever
this surface exhibits a 'valley' with an even bottom running obliquely
through the M_0–M_1 field, it may, by the foregoing reasoning, be in-
ferred that the ratio of the horizontal coordinates of any point on the
bottom of the valley is a cost of living index applying to a particular
level of real psychic income.*

TABLE 3

Values of D Based on the Consumption of Twelve Foodstuffs Obtained
in the Comparison of Four Income Groups in Poland, 1927, with
Five Income Groups in Estonia, 1925, Workers' Families

Annual income (zl.) per 'man', Poland, 1927	Annual income (Emk.) per 'man', Estonia, 1925				
	28,628	36,407	46,098	56,348	73,508
499	0.517	0.632	0.672	0.711	0.816
730	0.347	0.462	0.501	0.535	0.639
1,023	0.275	0.377	0.416	0.448	0.533
1,503	0.279	0.230	0.235	0.222	0.312

Budgets des familles ouvrières en 1925, Bureau Central de Statistique de l'Estonie,
Tallinn (s.d.), and *Budgets des familles ouvrières, 1927*, Statistique de la Pologne,
tome XL, facicule 1, Varsovie, 1930.

5 APPLICATIONS OF THE DISSIMILARITY METHOD

While the values of D in Table 2 are for situations that are only
temporally different, Tables 3–5 and their graphical representation
in Charts 1–3 show the results of interlocal comparisons. Chart 1
and Table 3 relate to workers' families in Estonia (1925) and
Poland (1927). A small section of the valley of equivalence can be

seen, indicating that the lowest Estonian incomes are just equivalent to the highest Polish incomes. The second comparison (Table 4 and Chart 2) is between the same Estonian incomes and workers' families in Finland (1921). Again a valley appears, indicating that the Finnish incomes are equivalent to the higher Estonian incomes.

TABLE 4

Values of D Based on the Consumption of Sixteen Foodstuffs Obtained in the Comparison of Four Income Groups in Finland, 1921, with Five Income Groups in Estonia, 1925, Workers' Families

Annual income (Fmk.) per 'man' Finland, 1921	Annual income (Emk.) per 'man', Estonia, 1925				
	28,628	36,407	46,098	56,348	73,508
4,530	0.884	0.781	0.743	0.712	0.725
5,973	0.914	0.805	0.747	0.705	0.691
7,280	0.923	0.813	0.756	0.706	0.683
9,812	0.930	0.820	0.763	0.706	0.633

For Estonia, same sources as for Table 3. For Finland, *Levnadkostnaderna under bokfoeringsperioden*, 1920–21, Finlands Officiella Statistik XXXII, Sociala Special-undersökningar V (Helsinki, 1925).

TABLE 5   .

Values of D Based on the Consumption of Ten Foodstuffs Obtained in the Comparison of Four Income Groups in Poland, 1927, with Four Income Groups in Finland, 1921, Workers' Families

Annual income (zl.) per 'man', Poland, 1927	Annual income (Fmk.) per 'man', Finland, 1921			
	9,812	7,880	5,973	4,530
499	1.051	1.033	0.955	0.902
730	0.849	0.831	0.753	0.701
1,023	0.814	0.786	0.713	0.693
1,503	*0.596*	0.645	0.601	0.578

See notes to Tables 3 and 4. To facilitate comparison with Chart 3, the Finnish incomes are given in decreasing order from left to right. The italicized value relates to the comparison of only 9 families in Finland and 17 families in Poland. As it is probably too low, it was smoothed in the chart.

According to the interpolation in the charts, the highest Polish income would seem to be approximately equivalent to the second lowest Estonian income, while the lowest Finnish income appears to be equivalent to an Estonian income just a little lower than the fourth one in the table. Consequently, the whole range of real in-

come considered in Poland should lie below the lower end of the range of real income considered in Finland.
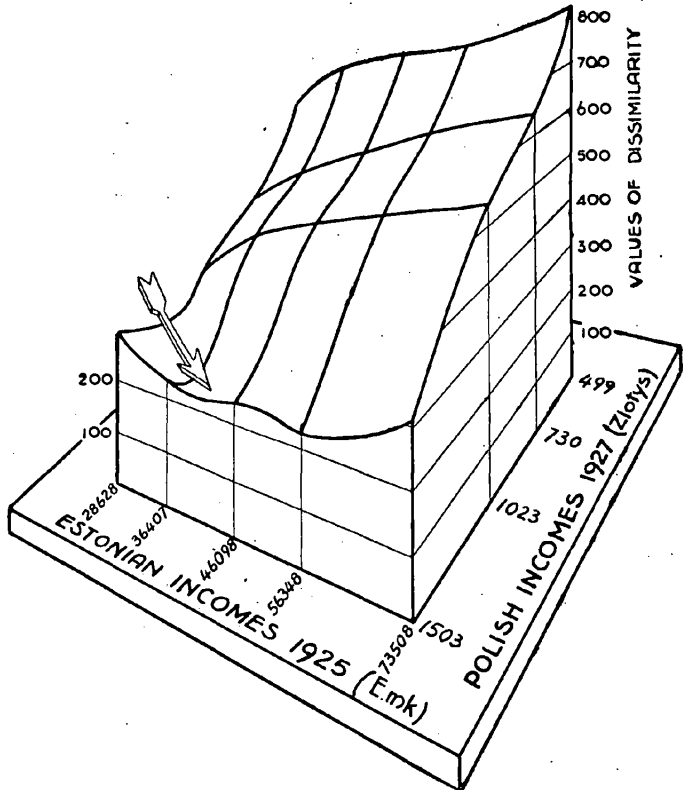

Chart 1

This inference is borne out by Table 5 and Chart 3, the direct comparison of the Polish and Finnish materials. In fact, no valley at all appears. The movement of the values of D is such as to situate the valley outside the region covered, in the direction of lower Finnish or higher Polish incomes.

The following 'equivalences' are given by way of example. The incomes are for different years. The dollar amounts are not equivalent United States incomes, but merely represent the dollar values of the various incomes according to the exchange rates in

900
800
700
600
500
400
300
200
100

700
600
500
400
300
200
100

VALUES OF DISSIMILARITY

28628 36407 46098 56348 73508 9812

ESTONIAN INCOMES 1925 (Emk)

7880 5973 4530
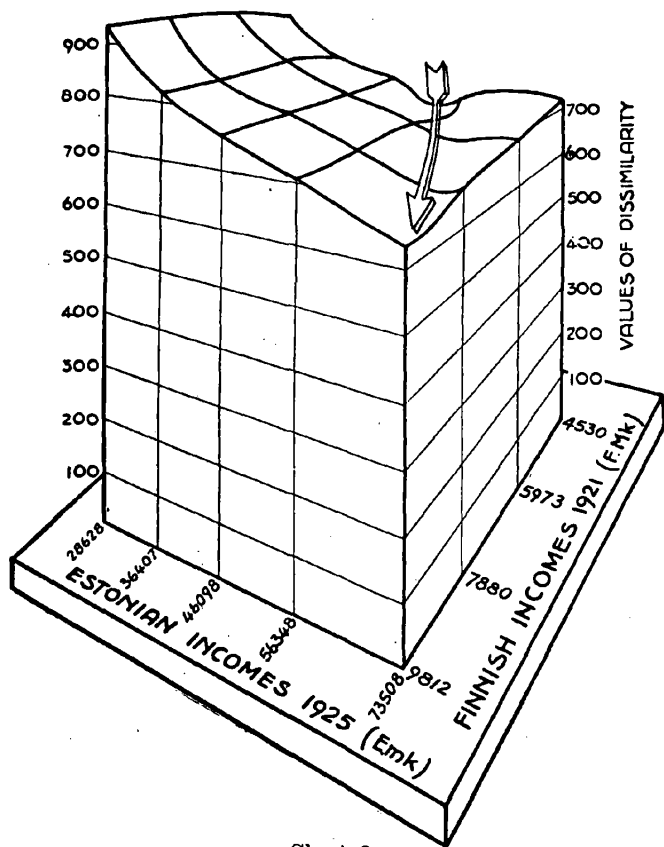
FINNISH INCOMES 1921 (F.Mk)

Chart 2

the respective years. All incomes are per 'man' and per year. From these 'equivalences' and from others that might be interpolated within the ranges of incomes covered, cost of living indexes can be constructed.

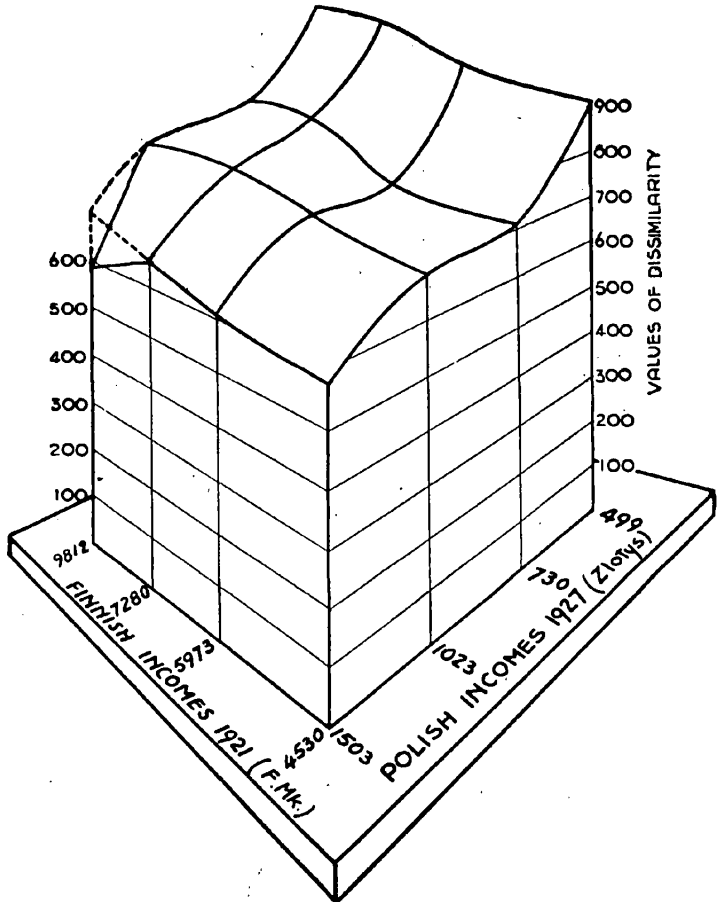| Poland, 1927		Estonia, 1925		Finland, 1921	
Zl.	U. S. $	Emk.	U. S. $	Fmk.	U. S. $
1,100	124.2	29,000	78.0		
1,500	169.3	42,000	112.9		
		51,000	137.1	4,500	169.8
		62,000	166.7	6,000	226.4
		74,000	198.9	9,800	369.8

Chart 3

These results are notable not only for their consistency but par-
ticularly because at least two of the situations compared manifest
very strong dissimilarity in the consumption of major foodstuffs,
and thus presumably in wants, milieu, and relative prices. Whereas
the minima of dissimilarity are surprisingly small between Es-
tonia and Poland (about 0.230, while they were about 0.300 in
the comparison between prewar and postwar Germany), D takes
quite exceptionally large values in the comparison between Es-

tonia and Finland. No method based on the assumption of identical wants could therefore be applied in this case.[17]

TABLE 6

Values of D Based on the Consumption of Nineteen Foodstuffs Obtained in the Comparison of Four Income Groups of Native-Born American Workers' Families with Four Income Groups of Scandinavian Immigrant Families, United States, 1909

Annual income ($) per person	Scandinavian immigrant families: annual income ($) per person			
	161	191	217	225
105	0.298	0.339	0.399	0.326
160	0.230	0.284	0.341	0.317
196	0.195	0.262	0.278	0.280
223	0.190	0.261	0.271	0.231

Cost of Living in American Towns, Report of an Enquiry by the Board of Trade, Cd. 5609 (1911).

TABLE 7

Values of D Based on the Consumption of Seventeen Foodstuffs Obtained in the Comparison of Four Income Groups of Swedish Workers' Families in Sweden, 1914, with Four Income Groups of Scandinavian Immigrant Families in the United States, 1909

Annual income (Kr.) per person, Sweden, 1914	Scandinavian immigrant families in the U. S.: annual income ($) per person, 1909			
	161	191	217	225
258	0.540	0.515	0.525	0.559
393	0.509	0.472	0.473	0.514
533	0.506	0.457	0.446	0.486
842	0.513	0.468	0.436	0.476

For the United States, see note to Table 6. For Sweden, *Levnadkostnaderna i Sverige 1913-1914, Del I: Utredningens Huvudresultat*, av K. Socialstyrelsen (Stockholm, 1921).

The relevance of the assumption of identical wants within each situation may be illustrated by another application.

Table 6 gives the values of D obtained in the comparison of four income groups of American families and four income groups

[17] The foregoing exposition relies heavily on my article 'A General Method for the Comparison of Prices of Living', in the *Review of Economic Studies*, IV (June 1937), 205–14.

of Scandinavian immigrant families in 1909. The data were taken from the investigation the British Board of Trade carried out for the purpose of comparing real wages in England and in the United States. The classifications of the American budget materials, especially collected, distinguish different ethnic immigrant groups.
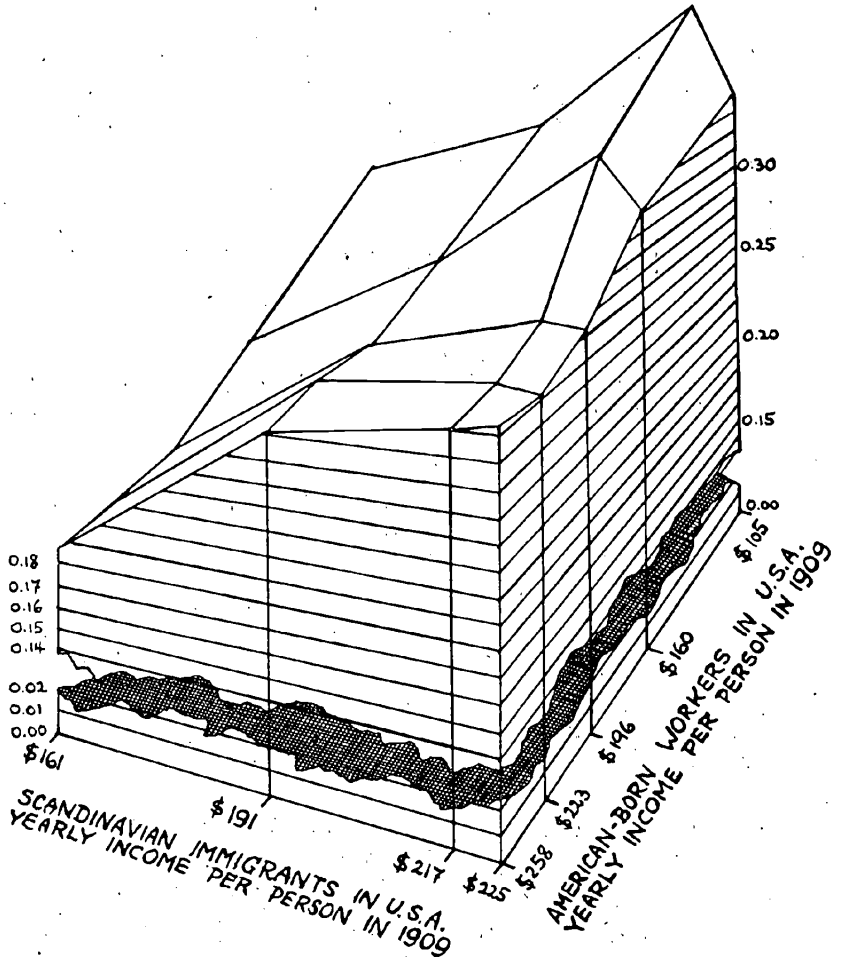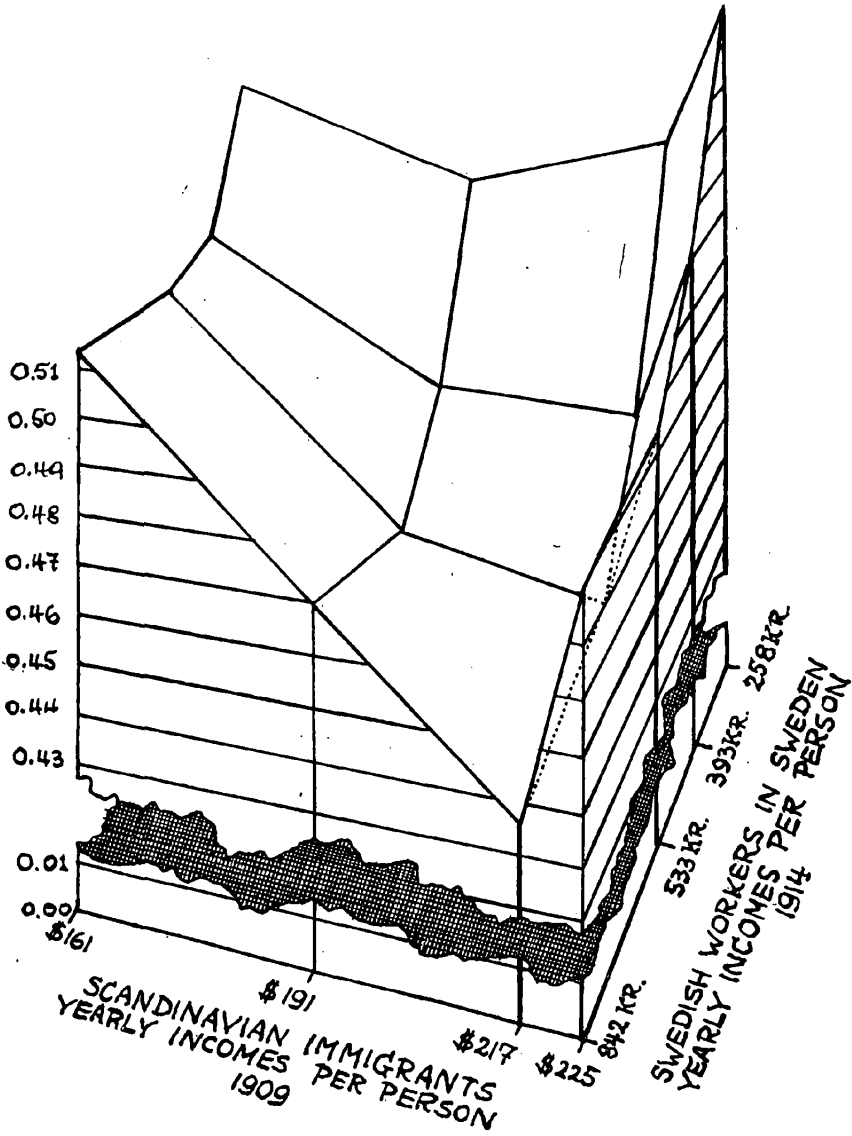

Chart 4

Table 7 compares the same Scandinavian immigrants with

Chart 5

Swedish families in Sweden in 1914. The Swedish material is classified according to income per 'man'.

D behaves in a fundamentally different way in Charts 4 and 5

and in Charts 1–3. The consumption of the lowest income group among the immigrants is more similar to that of the four groups of native American than the consumption in the second and third immigrant income groups (see Chart 4). At the same time, the lowest immigrant income group is less similar to any of the native Swedish income groups than the second and third immigrant groups. In the fourth income group of immigrant families the movement of D is reversed in both cases.

The failure of any 'valley' to appear is clearly due to differences in the degree of assimilation of the immigrants which also seems to depend on income. While at the lowest income level the immigrants are by the pressure of relative prices compelled to consume more or less like native Americans, they use their increased income to 'cultivate Swedishness'. Beyond a certain income, assimilation really begins. It is an open question whether it is the higher income (usually associated with a longer stay in the United States) that gives rise to, or at least accompanies, greater assimilation, or whether better assimilation is rewarded by a higher income.

TABLE 8

Values of D Based on the Consumption of Seventeen Foodstuffs Obtained in the Comparison of Four Income Groups in Sweden, 1914, with Four Income Groups of Native-Born Americans, 1909, Workers' Families

Annual income (Kr.) per person, Sweden, 1914	Annual income ($) per person, U. S., 1909			
	105	160	196	223
258	0.674	0.679	0.689	0.694
393	0.669	0.677	0.665	0.659
533	0.680	0.687	0.661	0.652
842	0.714	0.709	0.686	0.675

See notes to Tables 6 and 7.

Table 8, comparing native Americans with native Swedes, fails to bring out more than a hint of a valley, probably because the principle of classification in the American material fails to isolate the net influence of income upon consumption. This defect in classification, while strong enough to destroy the influence

upon D of real income, is not powerful enough to eliminate the effect of different wants. Thus similarity if not identity of wants within each situation is an essential condition for success in the application of the dissimilarity method.[18]

Finally, as an example of a less satisfactory result, consider the comparison of five incomes in Germany with five incomes in Mexico City, (Table 9). The 'valley of equivalence' hardly appears and can at best be perceived in a rough correspondence between the incomes of RM 1,000–1,200 and $ (Mex.) 289, and of RM 1,200–1,500 and $ (Mex.) 352. The Mexican data seem to be somewhat unreliable in bringing out the dependence of the structure of consumption upon the level of income. Actually, they were obtained from data on expenditure only, classified according to family income. Hence, they reflect the influence of family size as well as income on consumption.

TABLE 9

Values of D Based on the Consumption of Twelve Foodstuffs Obtained in the Comparison of Five Income Groups in Mexico City, Summer 1934, with Five Income Groups in Germany, 1927–1928, Workers' Families

Annual income ($ Mex.) per 'man', Mexico, 1934	Annual income (Rm.) per 'man', Germany, 1927–28				
	Less than 800	800 to 1,000	1,000 to 1,200	1,200 to 1,500	1,500 and over
181	0.640	0.570	0.600	0.623	0.639
233	0.619	0.532	0.504	0.523	0.550
289	0.607	0.511	0.471	0.480	0.504
352	0.683	0.565	0.505	0.461	0.508
429	0.658	0.549	0.485	0.475	0.522

For Germany, same source as for Table 1. For Mexico, 'Cómo se alimentan los obreros an la Ciudad de Mexico', by Federico Bach, *U.G.B.*, *Revista de Cultura Moderna*, January 1936, pp. 25 ff.

Nevertheless, as a very rough guess, it would appear that, at the income level of about $ (Mex.) 300 per unit in the summer of 1934, the equivalence with Germany (1927–28) was about $ (Mex.) 1 = RM 4. This guess is here put forth very tentatively indeed, since it is entirely without any support of the kind supplied by

[18] For a more detailed discussion of the immigrants' case, see 'The Reaction of Consumers to Changes in Prices and Income, a Quantitative Study in Immigrants' Behavior', *Econometrica*, II (1934), 59–72.

Keynes' 'direct method'. I do not possess the slightest experience in conditions of Mexican life.

6   APPRAISAL OF THE DISSIMILARITY METHOD

In reviewing the dissimilarity method, it may be convenient to list its strong and its weak points.

It commends itself, first, by the fact that there is no need to include the whole consumption set in the computations, provided the sector considered is sufficiently varied to include commodities whose income elasticities differ strongly. Since few family budget investigations give data on quantities consumed for commodities other than food, this is an important point.

Secondly, no assumption is involved as to similarity of wants, etc. in the different situations. Such dissimilarity as there is between the two situations will be reflected in the minimum values of D which should, ideally, be the same at each minimum point, irrespective of income.

Thirdly, it is not assumed that the cost of living ratio is the same for all income levels. The method puts no limit on differences in the relative cost of living depending on the level of real income, except so far as the expectation of a 'valley' presumes some continuity in the function connecting the cost of living ratios with income.

Finally, the method does not yield results in all cases, but may very well break down. No valley at all may appear, and the values of D may fluctuate erratically. I emphasize this point since the results yielded may be compared with certain forms of indexes which are actually so chosen as to yield results necessarily attaining a pre-arranged degree of formal perfection. If the results of a fixed-weight index satisfy the circular test, that should be no particular reason for rejoicing. If, on the other hand, results are found of the kind presented above in the three-cornered comparison Poland-Estonia-Finland, that is certainly significant.

The method is weak on several other points. First, it calls for the association of prices and quantities in different situations, i.e., for the identification of the price of a commodity in one country with that of the corresponding quantity of the 'same' commodity

in the other. The same name may, of course, hide great variations in the commodity concerned. It might, of course, be argued (and actually has been by Gini in conversation with the writer) that the form of D should be, say,

$$D = \Sigma \left| \frac{q_1 p_1}{\Sigma q_1 p_1} - \frac{q_0 p_0}{\Sigma q_0 p_0} \right|.$$

This is in fact what D becomes when only one price system is involved, e.g., in the comparison of different income levels within one country, where the p_0's and the p_1's are the same. Yet, while the direct association of a p_1 with a q_0 could thus be avoided, the meaning of the measure would change from one of dissimilarity in the structure of consumption into one of dissimilarity in the structure of expenditure. It would also make possible the inclusion of the whole budget, since only data on expenditures, no longer on quantities, would be needed. It might be worth while experimenting with Gini's suggestion.

Secondly, the method has no foundation in the theory of consumer's equilibrium. It is essentially empirical, and attempts to explain, in terms of accepted economic theory, why D should behave as it does have been unsuccessful.

Thirdly, the method depends, perhaps excessively, on the quality of the materials. Obviously, the data will have to bring out the *net* dependence of consumption on income alone. The material should therefore be classified, not according to family income, but according to income per 'man'. This, of course, raises the whole vexed problem of how to eliminate the influence of family type and size. I do not think it can be solved for all purposes merely by having separate classifications for several family types.[19] This method might be applied to any number of family types, by considering each as a different 'situation'. In fact, the very unit-of-consumption question might thus conceivably be clarified, if not

[19] In practical work I have found it both necessary and sufficient to apply some adjustment, however rough (say, counting all persons 16 years and over as one unit, all others as half a unit), to approximate the net relation between income and consumption. On this point, I heartily agree with Mr. Vickrey in his 'Resource Distribution Patterns and the Classification of Families', *Studies in Income and Wealth, Volume Ten.*

solved.[20] Unfortunately, the Bureau of Labor Statistics has con-
fined its food consumption tables for the Consumer Purchases
Study to showing average consumption for all family types com-
bined according to family income—which is also the reason this
country was included in the above applications only once, and in
that case, British data on American food consumption had to be
used.[21]

In the fourth place, the method assumed that all consumers
within a situation would consume identical quantities of every
commodity if they only had the same real income. And there is, of
course, considerable evidence that they do not. The spread around
the average quantity consumed at any income level, however
narrowly defined, has statistically been found to be considerable.
It is conceivable that, through an appropriate choice of the
consumption unit (or family type), the spread may be narrowed.
But some spread will certainly remain, even at the lower levels of
real income. With increasing income, family consumption is
certain to become more erratic, until at a certain level of comfort,
all systematic connection between income and consumption disap-
pears. For this reason, the measurement of real income becomes
more difficult the higher the real income levels to be compared.
The latter point is, of course, not peculiar to the dissimilarity
method, though its results, being more highly sensitive to the
quality of the material, will be more apparently affected. As far
as the lower income levels are concerned, the lack of unanimity in
consumption is not very serious, since the assumption involved
could easily be restated in terms of averages. Nevertheless, this

[20] Such experiments would, however, be very time-consuming, because of the vast
amount of calculation required. Moreover, unless Gini's suggestion were adopted,
which I am not prepared to do without substantial experimentation, the BLS ma-
terials of the Consumer Purchases Study could not be used, since family types
have not been distinguished in the tables showing food consumption according to
family income.

[21] I may be forgiven if I take this opportunity to express again the wish that the
original family schedules of future, and past, family budget inquiries in this
country be published, if only for a manageable sample of, say, 2,000 or 3,000 fami-
lies, in addition to whatever classifications are supplied. Such materials would
be highly relevant for many purposes at present unknown. Among other things,
it would permit explorations along the line of the following point mentioned in the
text,

whole matter of deviations from average behavior should be more fully explored on its own merits. However, one will have to wait for the publication of the necessary material for individual families before useful work can be done in this direction.

Although no rationing schemes were in operation during the periods to which the family budget data above utilized relate, the influence of such arrangements should at least be mentioned. From a purely formal point of view, rationing, if it is partial and if unrationed goods are also included in the calculations, may well be understood as part of the want system of the population concerned. If it were general and egalitarian, then our method would quite simply, and correctly, indicate that consumption structure and real income are independent of money income.

To continue our list of weak points, we may mention as the principal drawback of the dissimilarity method that it is not simple and straightforward. Unlike a direct comparison of prices, it does not appeal to common sense. I wonder if anyone in charge of responsible political decision could be convinced of its soundness to the point of basing action on the results obtained by it.

In conclusion, therefore, this method should, in my opinion, by no means be neglected. At the same time, it would be hazardous to base national income comparisons exclusively on the results it may yield. In no case should these results be accepted without at least the kind of evidence Keynes proposed for his 'direct method'. In the end, if the necessary apparatus for the collection and digestion of data were at hand or could easily be set up, direct price comparisons, supplemented by common sense substitutions, such as were applied by the British Board of Trade and the International Labour Office, are probably preferable, despite the theoretical objections to which they give rise. But in the absence of such comprehensive investigations, the dissimilarity method can at least be a reasonable substitute.

7   SPECIAL PROBLEMS IN DEALING WITH NATIONAL INCOME AGGREGATES

The results of international cost of living comparisons by the dissimilarity method will usually apply to different points of time in different countries, i.e., to the periods that happen to be

covered by the budget materials. If simultaneous international comparisons are desired, intertemporal price comparisons can be used to adjust the interlocal cost of living ratio to a common date. If the incomes to be compared and the interlocal index ratios were rigorously exact, a great deal would have to be said on the problem of intertemporal cost of living measurement. In particular, it would in that case be worth while to investigate ways , and means of applying the theories referred to in Section 1. Obviously, constancy of wants can more easily be postulated when prices are to be compared over time in the same place. But the main obstacle to the application of those theories, apart from serious problems in measuring the 'price' and the 'quantity' for many items in even an ordinary family's expenditure, is the absence of continuous family budget investigations. It is their absence that compels the use of 'base-year' weights in the calculation of all current official indexes, and therefore leads to systematic overstatement of the rise, and understatement of the fall, in the cost of living.

Moreover, rationing leads to an increase in the cost of living that is not reflected in the regulated prices but is no less real than an actual price increase. The overstatement due to base-year weights and the understatement due to rationing cannot be properly assessed.

Finally, it is equally uncertain to what degree current indexes should be corrected for changes in quality that occurred everywhere during the war—which may be a factor of great quantitative significance.

As a result, intertemporal indexes are more or less inaccurate. But all that can be done is to hope that their inaccuracy will not be in the same direction as the inaccuracies due to imperfections in the national income figures and in the interlocal cost of living ratios.

International cost of living comparisons are merely the first step in the comparison of real national incomes. When it comes to adding individual real incomes into national real incomes, or rather to deflating aggregates of individual money incomes, we are, of course, again confronted with the problem of the incom-

mensurability of different peoples' welfare. But, having already sinned in this respect as much as I have, I do not propose to raise this question again.[22]

There is one closely allied point, however, I would like to introduce. As I have repeatedly said, the cost of living ratios may be different at different levels of real income, both interlocally and intertemporally. They may also be different at any given level of money income, for, say, urban and rural populations, or even manual and nonmanual workers in the same place. To the extent that the ratios are substantially at variance, some account ought to be taken of this variance in the comparison of national incomes that are aggregates of individual incomes widely different in size.

Here again, we may quote one recent writer in violent dissent. In *A Treatise on Money*, Keynes considered (pp. 97–8) a situation where the cost of living (or, as he prefers to look at it, the purchasing power of money) might have changed, or, in the interlocal case, be different, in varying proportions for different levels of real income. By how much, in this case, has it changed for the community as a whole? He says: "There is, in my opinion, no satisfactory answer to this question—for the reason that we can give no meaning to a numerical comparison between the purchasing power of money to a poor man and its purchasing power to a rich man, the two things being, so to speak, in different dimensions." As a way out, he proposed (pp. 98–9) "to neglect those ranges in which comparatively few persons are to be found, and to say that the change in purchasing power for the community as a whole lies *between* the largest and the smallest change shown when the changes for those ranges of real-income, which include the bulk of the population, are arranged in order of magnitude."

While Keynes' solution is certainly acceptable, we may again indulge in a quotation from a less modern writer. In a little-known paper, 'The Dimensions of Happiness', Bentham wrote: " 'Tis in vain to talk of adding quantities which after the addition will

[22] I am not sure which proverb is appropriate to the situation: "C'est seulement le premier pas qui coûte" or "L'appétit vient en mangeant"?

continue distinct as they were before, one man's happiness will never be another man's happiness: a gain to one man is no gain to another. . . . This addibility of the happiness of different subjects, however when rigorously considered it may appear fictitious, is a postulatum without the allowance of which all political reasoning is at a stand."[23]

Whether we follow Bentham, lest 'all political reasoning be at a stand", or heed Keynes' advice, we shall surely in either case want more than mere data on aggregates of income. At least some rough subdivision of the national income into size groups, chosen so as to correspond to the different international cost of living ratios, will be required.[24] And this in turn raises another much debated problem—the unit with reference to which an income should be stated. Should the income groups be in terms of income per family, per person, or per man? It raises also the question of the units in which the frequencies should be put. Shall we want to know how many families, or persons, or 'men' each income group contains? Depending on the units adopted to express both the class intervals and the frequencies—choosing one by no means entails choosing the other—the results may differ substantially.[25]

I do not propose to discuss these alternatives, but merely wish to draw attention to their existence. Obviously, the questions they raise are not only very properly within the scope of this Conference, but also bring out how closely the whole field of income and wealth is intertwined. No part of it can be considered without soon being confronted with a series of other problems pertaining to it.

[23] Quoted by Elie Halévy, *La Formation du Radicalisme Philosophique* (Paris, 1904), III, 481, note 55.

[24] If, in order to avoid the adoption of a scale for measuring units of consumption, the various family types are to be kept separate at each income level, the amount of additional information required would become well nigh prohibitive. It is also humorous to note that the family type classification, adopted originally, I suppose, because of the incomparability of different persons' wants, would in the present connection, as it fatally does in many others, multiply substantially the number of times that very same incomparability would have to be disregarded.

[25] The degree of 'inequality' of a size distribution of incomes also depends heavily on the choice of units—a fact of which many writers on the subject, particularly Pareto, seem to be entirely unaware.

Only one other point is made here. The usual deflation procedure takes no account of saving. As a result, deflated incomes show changes in the volume of commodities that could be bought with them if the entire income were spent. But while this fiction makes good sense for an individual consumer, it becomes less meaningful when greater numbers of consumers are considered. If the community as a whole actually spent its income completely, prices, with given supplies, would be higher. This merely draws attention to the fact that national income, if it is to measure current welfare, should include only such goods and services as are currently consumable. Net investment, which is conventionally included in national income, does not simultaneously add to the flow of consumables, and should therefore be excluded from the computation of *real* national incomes. This procedure has the advantage of excluding such doubtful 'investment' as the government may be led to undertake in the pursuit of 'full employment' when the prejudice against 'the government going into business' is strong.

International comparisons of real national incomes call not only for materials on family budgets, but also, if the last two points are to be met, for more detailed data on money income than are at present available for most countries.

COMMENT

ABRAM BERGSON

My comments bear on the following questions raised or suggested by Mr. Staehle's paper: (1) the conceptual question that is raised by differences in tastes; (2) the applicability of the 'conventional' method of calculating cost of living indexes when tastes are different; (3) the conditions under which the 'conventional' method yields accurate results; (4) the question of the general applicability of the 'dissimilarity' method; (5) the conditions under which the 'dissimilarity' method yields accurate results.

1

With regard to the first question, I wish only to make the following suggestion. I wonder whether it would not be advisable here simply to acknowledge that in *theory* there are as many valid measures of the change in the cost of living as there are taste patterns. Or rather, we should say there are as many *sets* of valid measures, since for each given taste pattern (i.e., indifference map) there is a set of valid measures of the change in the cost of living, one measure for each level of real income. To acknowledge that the measure of the change in living costs depends upon tastes simply represents a further elaboration of the theory of index numbers, entirely analogous to the one that already took place when the dependence of the measure of the change in living costs on the level of real income was recognized.

Logically, the dependence of the measure of the cost of living on tastes is the corollary of our acceptance of individual preferences as the standard of value. Only if some other standard is used can we obtain a valid measure of the change in the cost of living that does not depend upon individual tastes.

At the risk of being set down as a heretic in these matters, I wish to urge that the foregoing formulation of the problem be adopted as an alternative to the one currently in vogue, according to which comparisons of the cost of living when tastes differ are said to be meaningless "in strict logic". This, I gather, is Mr. Staehle's view, though happily he does not allow it to interfere with his practical work in the field. My own view is that once one has acknowledged that in theory the measure depends upon

tastes, that there are many valid measures when tastes differ, one has said all one can say; to go further and say that the comparisons are meaningless is not only not particularly illuminating but, I fear, likely to be misleading. One might as well say that comparisons are meaningless when tastes are identical and real incomes different. If the measure could be determined for each set of tastes dealt with and each level of real income, obviously this would be a very considerable accomplishment. Very interesting and meaningful propositions certainly could be constructed if such indexes were at hand.

2

Granting this, I am not sure that I can agree with Mr. Staehle that the 'conventional' method of calculating the change in the cost of living is inapplicable when tastes differ. I am referring to the method whereby A's collection of goods is valued at the prices confronting B, and vice versa, to establish how much each would have to pay for his collection if he had to pay the other's prices. The difference in the cost of living in the two price situations is measured by comparing what each consumer would have to pay in the other's price situation with what he actually pays in his own. If tastes differ, this method inevitably leads to two measures of the difference in living costs, one from the viewpoint of A's tastes, the other from the viewpoint of B's tastes. This plurality of measures, however, in no way discredits the 'conventional' method. To repeat, if tastes differ, the plurality of measures is logically appropriate.

Even if tastes are the same, as has been noted, one is ordinarily confronted with an analogous situation, that is, with two measures of the change in the cost of living. This is because of differences in the real income of A and B. No one suggests that the 'conventional method' is inapplicable merely because real incomes differ.

3

The 'conventional' method has the virtue that the conditions under which it yields 100 percent accurate results can be stated. This is a virtue Mr. Staehle does not claim for his 'dissimilarity'

method, though my impression is that analogous conditions can be established for it too.   •

The 'conventional' method has a very long history, but as far as I am aware, no one has yet troubled to state precisely the conditions under which its results are completely accurate. I think it may be profitable to dwell a moment on this question.

Suppose A's money income were increased (decreased) just sufficiently to enable him to buy his old collection of goods at B's prices. It is a familiar proposition of index number theory that in actuality A would be better off than before, since when faced with B's prices he might economize on goods that were relatively more expensive than they were in his own price situation and buy more goods that were relatively cheaper than in his own price situation. To the extent that A would be better off, the 'conventional' method overstates the increase (understates the decrease) in the cost of living to A. But, evidently, this situation is excluded in the following case: when the commodities A consumes are completely complementary, that is, when A consumes goods in proportions that are fixed with respect to *relative price changes*. In this case, if A is fully compensated for any change in the total cost of his old collection of goods due to price changes, he will consume the different commodities in exactly the same quantities (as well as proportions) as formerly.[1]

In accord with the accepted definition of the change in the cost of living, as stated by Mr. Staehle, the problem of measurement is solved if one can determine for A the collection of goods that would yield him the same real income as his old collection and that he would wish to consume if confronted with B's prices. In

[1] The case of complete complementarity is analyzed in many mathematical studies of consumers' behavior. It need only be observed here that each indifference curve is represented geometrically by two straight lines at right angles, that the consumers' equilibrium position must be at the focus of one or another of the right-angle indifference curves (regardless of the price situation), and that there is only one expenditure line or expansion path (the line connecting the different foci) representing the change in the budget position due to changes in money income. This last contrasts with the situation where consumers' goods are more or less competitive with one another, in which case there is a different expenditure line for each price situation.

the case of complete complementarity, obviously the new collection of goods is the same as the old. The change in the cost of living is thus accurately measured by the 'conventional' method.

If both A's and B's consumption are characterized by complementarity, the 'conventional' method yields two measures of the change in the cost of living, each of which is entirely accurate: one represents the change in the cost of living from the viewpoint of A's tastes and real income, the other from the viewpoint of B's tastes and real income. I think this is worth stressing. The usual practice of referring to the two results obtained by the 'conventional' method as constituting limits for the change in the cost of living is misleading so far as it implies that there is some unique, intermediate, 'true' measure that is more accurate than either limit. In the case studied here, each measure obtained by the 'conventional' method is entirely valid in its own right.[2]

If consumption is characterized by complementarity and in addition tastes are the same, A's and B's budget position will both lie on the same expenditure line (expansion path).[3] However, the 'conventional' method will still yield two measures of the cost of living: one from the viewpoint of A's real income, the other from the viewpoint of B's real income. Both are completely accurate.

Only if A consumes goods in the same proportion (not necessarily the same absolute quantities) as B do the two measures merge and yield a unique measure of the change in the cost of living. This is the so-called case of expenditure proportionality.

Complete *complementarity* and *expenditure proportionality* are of course not one and the same thing. In *complementarity*, the structure of consumption is not affected by *relative price* changes, but may be affected by changes in *real income*. In *expenditure proportionality* the structure of consumption is not affected by changes in *real income*, but may be affected by changes in *relative*

[2] Without referring to the conditions under which the 'conventional' method yields accurate results, W. W. Leontief in 'Composite Commodities and the Problem of Index Numbers', *Econometrica*, Jan. 1936, has called attention to the plurality of correct measures of the price level.

[3] See note 1 above.

prices. If both conditions obtain, the structure of consumption is not affected by changes in either.[4]

I have commented on the 'conventional' method in some detail mainly in order to clarify its logical foundations, about which it seems to me there are many misconceptions. The foregoing discussion suggests also the need for further empirical work in this field, in particular on the question concerning the degree to which consumers' behavior in the real world approximates the case of complete complementarity. Budget studies could probably shed some light on this question. My own hunch is that the approximation, *in terms of broad categories* of consumption, may not be bad; that consumers, in determining the total amount of food they buy, housing space they rent, etc., may not be as sensitive to relative price changes as we are often inclined to assume. There is already abundant evidence, however, that as a rule the case of expenditure proportionality does not hold.

<div align="center">4</div>

To acknowledge that the change in the cost of living depends upon tastes is not in any sense to question the usefulness of the method devised by Mr. Staehle, which yields only one measure even though tastes differ. One must simply interpret Mr. Staehle's index as being at one and the same time an approximation to the change in the cost of living from the viewpoint of A and of B. How good an approximation it might be would depend on, among other things, how different tastes are.

With respect to Mr. Staehle's interesting method, I wish to raise a question about one central feature. As I understand it, the relative similarity in the proportions in which commodities are consumed is taken as the test of whether real incomes of different households facing different price situations are equivalent. This test evidently leaves open the awkward possibility that Household A will be said to be as well off as Household B, even though B is consuming more of everything, or of practically

---

[4] In this case there is only one expenditure line (expansion path) for all price situations, a straight line through the origin.

everything, than A, provided only that their consumption is in more or less similar proportions.

CHART 1
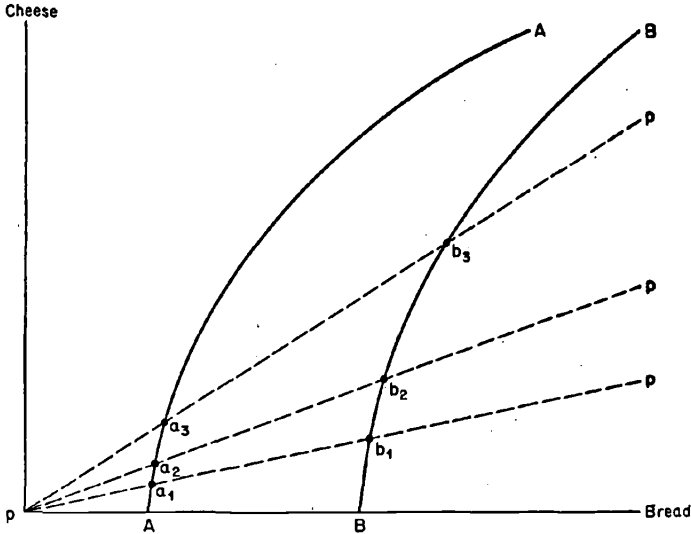
AA: Expenditure line (expansion path) of A households, facing price situation O.
BB: Expenditure line (expansion path) of B households, facing price situation 1.
a_1, a_2, a_3: Budget positions of households A_1, A_2, A_3.
b_1, b_2, b_3: Budget positions of households B_1, B_2, B_3.
Lines pp: Loci of budget positions for which the structure of consumption is identical and Staehle's D is O.

Chart 1 provides an extreme example of the sort of case I have in mind. The value of Mr. Staehle's D would be a minimum and in fact zero in the following cases:

a) For the pair of households a_1b_1 among comparisons of a_1 with b_1, b_2, and b_3;

b) For the pair of households a_2b_2, among comparisons of a_2 with b_1, b_2, and b_3;

c) For the pair of households a_3b_3, among comparisons of a_3 with b_1, b_2, and b_3.

Evidently, these minimum values would represent a 'valley' of the type to which Mr. Staehle refers. In this situation, *one would be compelled to conclude from Mr. Staehle's method that the*

*real income of the households paired in each case was the same, even
though the B households were all consuming much more of everything
than the A Households.*

It is of course not necessary to suppose that as between the A
and B households there are pairs of the sort just considered for
which the structure of consumption is identical. Mr. Staehle's
method encounters the same difficulties if this situation is only
approximated.

5

What is the probability that any such situation as the one
envisaged will ever confront us? It seems to me that under two
assumptions this sort of situation could be excluded:
a) If the tastes of the households facing one price situation are
similar to those facing the other;
b) If consumption is characterized by complete complementarity
in the sense referred to a moment ago.

In this case, evidently, all the observed budget positions would
tend to cluster about a single expenditure line, rather than about
two expenditure lines as appears in Chart 1 (see note 1). In this
case Mr. Staehle's D test would work without fail—the pairs of
households for which D is a relative minimum and very small
would be consuming goods in the same quantities as well as in
similar proportions.

It is subject, however, to one qualification. The 'dissimilarity'
test would still break down in the case of expenditure propor-
tionality. Here D would be zero for any pair of households, even
though their real incomes are, in fact, very different.

The conditions I have stated are sufficient conditions; if they
hold, the 'dissimilarity' method works. The interesting question is
whether they are also necessary. My impression is that they are,
in the sense that the more nearly each is satisfied, the more
nearly accurate Mr. Staehle's method is likely to be. Clearly, if
tastes differ radically, anything can happen; a situation compa-
rable to that in Chart 1 is quite conceivable. If tastes are the same
but the various goods consumed are highly competitive, a dif-
ference in the price situation may give rise to two situations, in

neither of which Mr. Staehle's method would work satisfactorily. On the one hand, the structure of consumption might be very different for households having the same real income (with the different budget positions being situated on the same indifference curve). Here D would be large, even though real income is the same. On the other hand, the structure of consumption might be similar even though real incomes are very different (as in Chart 1). Here D would be very small, even though real income is different.

If the foregoing is correct, a serious question would arise whether the 'dissimilarity' method is in any way superior to or indeed can even be considered on a par with the 'conventional' method. As indicated above, the 'conventional' method supplies us with a satisfactory approximation if the assumption of fixed proportions holds approximately. It is *not* necessary to assume also that tastes are similar. Whether tastes are similar or not would determine whether one gets one or two answers concerning the change in the cost of living for a given level of real income.

With respect to the method of Frisch to which Mr. Staehle refers, perhaps it is permissible for me to call attention to a study I carried out some years ago.[5] According to my analysis, the measure Frisch takes as an index of the level of real income turns out under his own assumptions to be constant for all levels of real income. For this reason, it would seem to me, Frisch's measure cannot very well be used for the purpose Mr. Staehle's D is intended to serve; that is, to select households with equivalent real incomes.

DOROTHY S. BRADY AND ELEANOR M. SNYDER

Mr. Staehle's method for determining differences in the cost of living between two situations depends upon locating the income groups that differ least in consumption pattern. The procedure pairs the income groups in the two situations for which the relative distributions of the cost of goods and services at fixed prices are most similar.

The usefulness of the method depends primarily on the fre-

[5] 'Real Income, Expenditure Proportionality . . .', *Review of Economic Studies*, Oct. 1936.

quency with which it will produce a unique pairing of the income
groups in the two situations. Examination of all possible compar-
isons indicates that the method clearly identifies pairs of income
groups only when the patterns of consumption in the two situa-
tions are very similar. As the differences between the consumption
patterns increase, the likelihood of ambiguity in the comparison
becomes very great. The more narrowly defined the items and the
greater their number, the less likely are consumption patterns,
as defined by this method, to be similar.

To visualize the probabilities, an inductive process is called for.
If only two items are compared, the value of D for the various
values of the ratios x_1 and x_2 in the second situation, compared
with a specific set (a, b) in the first situation, are given by

$$D = |x_1 - a| + |x_2 - b|$$
$$= 2|x_1 - a|$$

since $x_1 + x_2 = 1.00$ and $a + b = 1.00$.

D, accordingly, can be represented by a graph in relation to
the values of x_1, as two lines intersecting on the x_1 axis at $x_1 =$
a. Since D is minimum at $x_1 = a$, the method yields a unique com-
parison if the series of ratios x_1 includes the value $x_1 = a$. If the
series of ratios x_1 does not include the value a, there is no solution.

For the case of three items, the surface of values of D is given
in the contour map, Chart 1, which shows the values of D com-
paring a specific set of ratios in the first situation, $x_1 = 30$, x_2
$= 20$, $x_3 = 50$, with the range of possibilities in the second situa-
tion, in relation to the coordinates x_2 and x_3. Clearly D has a min-
imum value at the identical point, $x_2 = 20$, $x_3 = 50$. The number
of paths through this surface with distinct minima is very small
compared with all possibilities. The paths followed by consump-
tion ratios from the lowest income group to the highest are,
however, also a subset of the total. The method, therefore, de-
pends upon the degree to which the paths characteristic of con-
sumption data and the paths having minima correspond.

The characteristic ratios in consumption data for three groups
are of two patterns. First, one ratio tends to be constant over the

income scale, one ratio increases, and one decreases. In the ex-
ample, when x_1 is constant, and $x_2 + x_3 = k$, the paths followed
by the ratios are along the diagonals. If the ratios are along the
diagonal through the point of comparison, the alternatives are

CHART 1

Contour Map of the Values of D
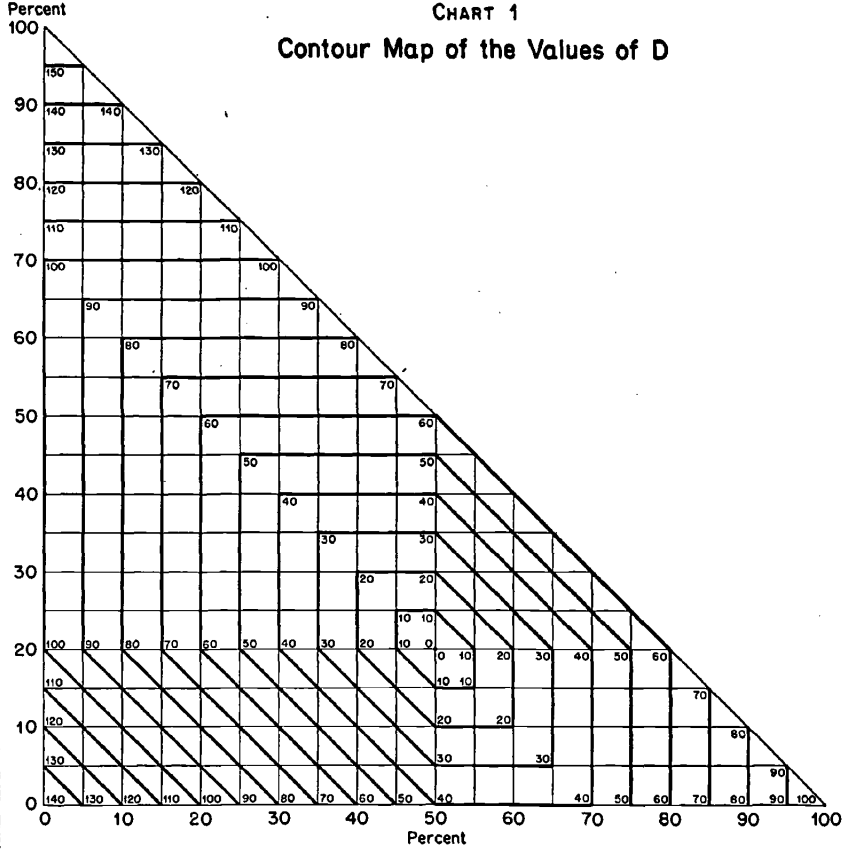
the same as in the case of two commodities. The ratios either
include the point of comparison or they do not. There is a mini-
mum value of D only if the identical sets of ratios appear in both
situations.

On diagonal paths close to the point of comparison, a 'valley'
appears in the values of D, which has a wider and wider floor as

the distance from the identical comparison increases. This 'valley' cannot be said to define the equivalent incomes in the two situations without appeal to some other criterion. Changes in the ratios along these 'valley' floors may cover a very wide range of incomes. If the point of comparison (30, 20, 50) was the pattern of families averaging $1,250 in the first situation, the 'valley' floor in the second situation might be $1,500 in width, say from $1,500 to $3,000 covering the income groups having the lowest (and the same) value of D.

At greater distances from the point of comparison, the values of D may decrease or increase or may coincide with long ranges of constant values. Ratios characteristic of consumption data are not likely to range from 0 to 100 on separate items, but ordinarily change within a range of .25 or .30.

In the other pattern characteristic of consumption ratios, one ratio decreases and the others increase. In this case the path of the consumption point resembles segments of a hyperbola. If on the map, x_2 is taken as the decreasing ratio and x_3 as the increasing ratio, these hyperbolic paths are convex to the axes. Near the point of comparison, these paths will be associated with values of D that have a distinct minimum. At greater distances the paths are not likely to yield a minimum D except in the very unusual case when individual ratios cover a wide range of values.

The analog of this map for the case of four commodities can be visualized as a contour diagram in three dimensions. The contours for constant values of D are shells formed by planes parallel to the coordinate axes and diagonal to them. The possibilities of paths yielding minimum values of D are similar to the case of three commodities but the likelihood of paths with constant D is increased.

In view of these considerations, the method can be expected to be successful only when the differences in the consumption patterns in the two situations are small. Small differences tend to appear in comparative studies when the commodity classification is very broad. The relative stability of the percentage distribution of total expenditures among such broad catagories as food, housing, and clothing has been noted ever since the time of

Engel. The relative distribution of total cost at fixed prices would obviously show a similar stability in most usual types of comparison.

Accordingly, in comparing two situations the procedure may yield an unambiguous pairing of income groups when a broad commodity classification is used but may or may not have a solution when a more detailed commodity classification is used. Frequently the narrower commodity classification will lead to a matching that differs from the correspondence established by using the broader classification. An example of this case is given by a hypothetical comparison.

Commodity	Relative Distribution of Cost			
	$1,000–1,500	$1,500–2,000	$2,000–2,500	$2,500–3,000
Situation '0'				
Bread, total	.6	.5	.4	.3
rye	.4	.3	.2	.1
white	.2	.2	.2	.2
Meat, total	.4	.5	.6	.7
beef	.1	.1	.1	.1
veal	.3	.4	.5	.6
	$500–1,000	$1,000–1,500	$1,500–2,000	$2,000–2,500
Situation '1'				
Bread, total	.6	.5	.4	.3
rye	.5	.4	.3	.2
white	.1	.1	.1	.1
Meat, total	.4	.5	.6	.7
beef	.2	.2	.2	.2
veal	.2	.3	.4	.5

VALUES OF D BASED ON 'BREAD' AND 'MEAT'

Incomes	Incomes, Situation '0'			
	$1,000–1,500	$1,500–2,000	$2,000–2,500	$2,500–3,000
Situation '1'				
$500–1,000	0	.2	.4	.6
1,000–1,500	.2	0	.2	.4
1,500–2,000	.4	.2	0	.2
2,000–2,500	.6	.4	.2	0

VALUES OF D BASED ON 'RYE BREAD', 'WHITE BREAD', 'BEEF', AND 'VEAL'

Incomes	Incomes, Situation '0'			
	$1,000–1,500	$1,500–2,000	$2,000–2,500	$2,500–3,000
Situation '1'				
$500–1,000	.4	.6	.8	1.0
1,000–1,500	.2	.4	.6	.8
1,500–2,000	.4	.2	.4	.6
2,000–2,500	.6	.4	.2	.4

The importance of the commodity classification for the success of the method can be emphasized by considering the case of a single variant in consumptions in two situations. Suppose in the two situations the foods consumed were the same so that a unique correspondence could be established between income groups in pairs but the principal beverage differed. If there are two beverages and both are included in the calculation, the procedure will probably not yield a distinct solution. This dilemma too can be shown by hypothetical figures.

Commodity	Relative Distribution of Cost			
	$1,000–1,500	$1,500–2,000	$2,000–2,500	$2,500–3,000
Situation '0'				
Foods	.7	.6	.5	.4
Beverages, total	.3	.4	.5	.6
coffee	.2	.3	.4	.5
wine	.1	.1	.1	.1
	500–1,000	1,000–1,500	1,500–2,000	2,000–2,500
Situation '1'				
Foods	.7	.6	.5	.4
Beverages, total	.3	.4	.5	.6
coffee	.1	.1	.1	.1
wine	.2	.3	.4	.5

VALUES OF D BASED ON 'FOOD' AND 'BEVERAGES'

Incomes	Incomes in Situation '0'			
	$1,000–1,500	$1,500–2,000	2,000–2,500	$2,500–3,000
Situation '1'				
$500–1,000	.0	.2	.4	.6
1,000–1,500	.2	0	.2	.4
1,500–2,000	.4	.2	0	.2
2,000–2,500	.6	.4	.4	0

VALUES OF D BASED ON 'FOOD', 'COFFEE', AND 'WINE'

Incomes	Incomes in Situation '0'			
	$1,000–1,500	$1,500–2,000	$2,000–2,500	$2,500–3,000
Situation '1'				
$500–1,000	.2	.4	.6	.8
1,000–1,500	.4	.4	.6	.8
1,500–2,000	.6	.6	.6	.8
2,000–2,500	.8	.8	.8	.8

It thus appears that, to be successful, the procedure must rely on judgments with respect to the equivalence of the Englishman's tea and the American's coffee. It may be argued that these difficulties can be avoided by confining comparisons to countries with similar consumption patterns and by chaining to effect the comparisons for those with obviously different consumption patterns. The degree of success in the procedure depends, however, upon the degree of similarity in consumption patterns, where similarity is very narrowly defined. The differences between the patterns of consumption in the northern and the southern regions of the United States are too great in terms of commodity detail to yield a successful pairing of income groups by Mr. Staehle's procedure, even though southern and northern families both consume 'Wheaties', 'hot dogs', and 'cokes'. The values of D based on 10 foods indicate the tendency toward a constant over a long range of

Incomes of Southern Families	Incomes of Northern Families	
	$1,000–1,500	$1,500–2,000
Under $500	.51	.46
500–1,000	.42	.38
1,000–1,500	.37	.32
1,500–2,000	.35	.31
2,000–3,000	.34	.32
3,000–5,000	.35	.32
5,000 and over	.44	.39

income. Even in terms of commodity groups as broad as meat, cereals, and vegetables the comparison of the North and the South by this procedure does not yield a unique correspondence of pairs of income groups.

The procedure is likewise not successful for this degree of detail

when more similar consumption patterns are sought, for example, in large cities in the North. Successful matching even when consumption is as similar as in American cities in the same region is apparently possible only when the procedure is applied to very broad commodity groups. The use of broad commodity groups is not really a significant change in the basic concept of Mr. Staehle's procedure. It means simply that the differences in cost ratios would be calculated not for items but for groups of items:

$$D = \sum_{i=1}^{n} \left| \frac{\Sigma q_{i1} p_{i0}}{\Sigma q_1 p_0} - \frac{\Sigma q_{i0} p_{i0}}{\Sigma q_0 p_0} \right|$$

where $i = 1, 2, 3 \cdots n$ designates the commodity group. The use of groups in the formula does not mean that the comparison cannot be based on a manageable number of items selected as representative of the total. Within each group some items can be repriced to give an index of the price differences between the two situations. These indexes can then be applied to the total expenditures to obtain the cost of the goods purchased in the second situation at prices prevailing in the first.

In formulating the results of the calculation, the relatives should be based on the actual expenditures for consumer goods and services in the two situations. If in all income groups expenditures were less than or at most equal to income, 'savings' could be included as an item in the comparison. Almost all studies of family expenditures, however, have shown average deficits in the lowest income groups when expenditures exceed income. If the final relatives are based on the paired incomes, the results might be contradictory in terms of other information on the relative cost of living in the two situations.

<div align="center">MORRIS A. COPELAND</div>

In writing his note on methods of handling the problems involved in comparing real national incomes Staehle appears to have been preoccupied with that attenuated modern version of the marginal utility theory of individual choice, the 'indifference map' theory, and to have drawn his criteria for judging methods of making inter-country income comparisons largely from this theory.

I think the basic question raised by Staehle's paper is whether we should confine our efforts at inter-country comparisons of real income to methods that commend themselves because they can be rationalized in terms of the indifference map theory. If so, the limitations on what can be done in the way of such inter-country comparisons are very serious, more serious even than Staehle makes explicit.

Before attempting to summarize these limitations, it may be well to inquire why the indifference map theory should be considered in connection with inter-country income comparisons at all. Staehle admits that this theory requires an *'identity of wants'* that cannot reasonably be assumed as between countries.

Staehle does not state his reason for his preoccupation with the indifference map theory; his language suggests that he regards this theory of individual choice as generally, perhaps universally, accepted. But if this theory of individual choice is to be assigned preeminence as a basis for judging methods of inter-country income comparisons, such preeminence cannot be justified on the ground that the 'indifference map' theory is generally accepted. If it were, it surely should be generally accepted by those whose business it is to study individual behavior, viz., psychologists. But this is far from the case. In fact, it is difficult to find psychologists who espouse this theory of individual choice.

Now as to the limitations Staehle imposes upon inter-country income comparisons. Let us first take the ones he makes explicit:

1) In his first paragraph he rules out what he designates the "volume of production" or "bulk" approach.

2) He does not deal with total national income—still less with gross national product—for he excludes investment or capital formation.

3) He rejects what he calls "extraneous standards" of comparison such as standards of nutritional and housing adequacy. Such standards would seem pertinent to welfare, but Staehle characterizes them as arbitrary.

4) His method does not provide an over-all exchange rate or inter-country deflator for consumer expenditures. Rather it provides a set of exchange rates or deflators, one for each matched

income level. His method provides no clear-cut answer to the question, 'How summarize this set of income-level exchange rates into an over-all average for purposes of an *inter-country* as distinguished from an *inter-income-level* comparison'?

But there are further limitations on his method that are not made fully explicit. It is not an accident that his illustrations are so largely confined to foods. He notes that his method requires that, for his q's, inter-country quality differences must be small, and that q's must be available by consuming-unit income classes for which the consuming-unit is comparable as between the two countries. In view of these requirements I submit the following as a minimum list of further limitations on Staehle's method:

5) His method is unlikely to help much in inter-country comparisons of housing. Because of quality differences as between countries the problem of 'equivalent substitution' looms large. But the same is to be said of other broad areas of consumption, particularly most services. Even for clothing, fuel, furniture, tobacco, beverages, miscellaneous personal effects, and some foods the problem of 'equivalent substitution' is highly important. An inter-country comparison that largely avoids 'equivalent substitutions' will cover only a small fraction of consumption.

6) His method is inapplicable to that area of consumption in which government is the purchaser.

7) In comparing the farm consumption of home produced food for two countries his method gives no clue concerning what prices to use.

8) It is not clear how he would take account of changes in consumers' stocks when such changes are substantial, as they have been in recent years. Yet logically a welfare approach would seem to require him to face this question.

All this suggests a very limited usefulness for Staehle's method, particularly when he tells us that the definition of consuming-unit employed in most of the data available for the United States has made it difficult for him to use his method on these data.

But Staehle would apparently claim a wider usefulness for his method than is suggested here, for he tells us: "It commends itself, first, by the fact that there is no need to include the whole

consumption set in the computations, provided the sector considered is sufficiently varied to include commodities whose income elasticities differ strongly. Since few family budget investigations give data on quantities consumed for commodities other than food, this is an important point." He seems to imply that one can safely draw from data on foods conclusions about exchange rates or inter-country deflators suitable for making real income comparisons for consumption categories other than foods.

In this connection I would like to call attention to the paper on international income and product comparisons presented last year by Jacobson, Clyman, and myself. We discussed the technical problems of comparison encountered in preparing the report to the Combined Production and Resources Board on 'The Impact of the War on Civilian Consumption in the United Kingdom, the United States and Canada.'[1] This report employed the "volume of production" or "bulk" approach which Staehle has ruled out. Nonetheless, I think its findings are pertinent to the question, Is it safe to draw conclusions about exchange rates for other areas of consumption from *exchange rates based on food statistics?* In our paper implicit United Kingdom-United States exchange rates or interspatial deflators were presented for various consumption categories for 1938–39 and for two major components of the wartime gross national product for 1943. The prewar exchange rate applicable to food was $3.75 per pound; for all measured consumption categories taken together it was $4.67. The wartime exchange rates cited were $6.67 per pound for munitions production and $8.67 per pound for the services of the armed forces. These findings suggest that it is unsafe and unwise to base international comparisons of national income on data for prices and quantities of foods alone.

Moreover, as I noted above, when consumers' stocks change markedly it is important to take this fact into account in comparing international consumption levels, and Staehle has not explained how his method would deal with such a situation. If the "bulk" approach is applied to an inter-country comparison of

[1] *Studies in Income and Wealth, Volume Ten,* Part III (1947).

gross national product, this problem of consumers' stocks can be dodged for purposes of comparing the over-all product levels. The "bulk" approach affords also some basis for dealing with this problem of changing stocks in comparing consumption levels.

Second, Staehle refers to international cost of living comparisons made by the Board of Trade, by Lever Brothers, and by the ILO, and characterizes them as "answers to a question that, strictly speaking, has 'no economic meaning' ". I suspect he arrived at this characterization by confusing two quite separate propositions: (1) The measurements cannot be rationalized in terms of the theory of indifference maps, (2) The measurements have no economic meaning. Interspatial consumption price indexes may not be susceptible to rationalization in terms of the esoteric theory of indifference maps but it does not follow that they have no economic meaning. Fortunately, Staehle seems loath to accept the logic of his own reasoning. But he does not attempt to indicate what economic meaning might attach to such price indexes. I suggest that the meaning is obvious. Such indexes may be useful as interspatial deflators for purposes of obtaining an inter-country bulk comparison. Analogous indexes were so used in the Combined Production and Resources Board report. Incidentally R. D. G. Allen, whom Staehle cites in connection with his 'have no economic meaning' proposition, was one of the two United Kingdom representatives on the combined committee responsible for the Combined Production and Resources Board report.

One minor point deserves mention. Staehle says: "The official cost of living indexes published in most countries, being computed by . . . fixed 'base-year' weights . . . actually overstate the rise and understate the fall in the 'true' cost of living." It should not be necessary today to point out that, under conditions of forced up-trading, a fixed-weight price index may understate the rise in the 'true' cost of living.

The extremely limited applicability of Staehle's method does not mean that we cannot make objective inter-country comparisons of gross national product and of consumption levels. Indeed, I would urge that we attempt such comparisons, using the "bulk" approach.

In our paper on the "bulk" approach we indicated that the problem remaining in providing such over-all comparisons between the United Kingdom and the United States is not a major one and that it should shortly be possible to provide such comparisons on a fairly firm objective basis. By a firm objective basis I think we should mean that competent workers in the two countries can agree on the physical volume relationships.

The type of comparison attempted in the Combined Production and Resources Board report is undoubtedly too detailed and too elaborate to be undertaken annually for any large number of countries. Moreover, as we pointed out, the difficulties of inter-country comparisons are greatly enhanced when the countries are characterized by wide cultural differences. But I think that, short of comparisons as detailed as that made in the Combined Production and Resources Board report, a great deal can be done toward providing approximate inter-country income comparisons. I have in mind comparisons that make a judicious use of price and physical volume indexes together with national income estimates and budget studies, each expressed in its own country's monetary unit. The possibility of developing such approximate measurements constitutes an area there is urgent need to explore in the years immediately ahead. I think we should not be deterred by Staehle but should proceed boldly to explore it.

WILLIAM VICKREY

Mr. Staehle's ingenious device would command more confidence if it could be shown that in the more usual cases where the price index method of comparison is reasonably reliable, it yields results that differ from the normal price index method by an amount that lies within the probable range of error of the price index method. It should be fairly easy to apply the method to a comparison of budgets in the United States in 1935–36 and in 1941, for example, as data are available for both periods.

However, I suspect that even here there may be substantial sources of error or at least bias that will be difficult to discover and allow for. For example, if there is any substantial basis in fact for Veblen's thesis that consumption patterns tend to be

handed down from the leisure class to the underlying population through emulation, then even if there were no change in prices from one period to another, one would expect Staehle's method to yield consumption patterns of the later period bearing a maximum similarity to those of slightly higher respective income levels in the earlier period; in short, to indicate a downward trend in the 'cost of living' even though prices did not change. It is difficult to see how an influence of this sort can be allowed for, and there may be others not as patent. Nor is the phenomenon entirely confined to comparisons in the same place at different times: Hollywood patterns may be slower in pervading Europe than the United States, and the influence of the cultural leadership of a given country may make itself felt with varying rapidity in other lands.

Mr. Staehle's technique is intriguing but it requires much more extensive validation before much reliance can be placed on it.