6

SOME PROBLEMS IN THE MEASUREMENT OF REAL NATIONAL INCOME

By J. L. Nicholson
(Chief Statistician, United Kingdom Central Statistical Office)

The Truth is one and incapable of contradiction;
All knowledge that conflicts with itself is Poetic Fiction.
—W. H. Auden.

I. THE GENERAL SETTING

BEFORE discussing problems of method, one ought to be clear about aims. One is hardly in a position to tackle the numerous and various problems that are peculiar to the measurement of real national income until one knows the purposes for which the results are intended to be used. This is the kind of proposition with which few people would disagree but which, as Oscar Wilde said about someone's face, once seen never remembered.

Since estimates of real national income may have several possible objects, one should not expect to find perfectly general solutions to many of the problems. The correct, or most appropriate, solution for one purpose may often differ from the correct, or most appropriate, solution for some other purpose. There would be general agreement that the particular purpose of the estimates may affect the choice of weights, as well as the definition of what is included within the boundaries of economic activity. But it may also influence the best method of making use of the available data; and one should even bear in mind the possibility that, in certain cases, it may affect the choice of indicators.

The purposes of estimates of the real national income or product seem to fall into two main categories. Such estimates may be needed in connection with

(*a*) problems concerned with the potential or actual welfare or satisfaction derived by consumers and purchasers of final products; or
(*b*) problems of productivity, and the effective utilization of resources, the emphasis in this case being on the producer, not the consumer.

It must be admitted that the usefulness of estimates of the

real national income, in either connection, is very limited. They do not, certainly not without modification, and perhaps cannot provide general quantitative measures either of economic welfare, or of total output in the productivity sense. The estimates usually correspond with, or derive from, some generally accepted definition of the national income, which in turn implies a definition of economic activity. Not all economic activities need contribute, while many activities that are not regarded as economic can contribute, to total production or economic welfare. According to the specific or general purposes in mind, the concepts of productivity and welfare can also be defined in different ways. But the real national product does at least provide a starting point from which a measure, or index, of one kind or the other may be derived; and can itself, therefore, be given a flavour – to put it no higher – either of productivity or of welfare. So much would, perhaps, be fairly widely admitted. But the practical implications of this dichotomy need to be considered rather carefully. And we must be prepared for some seemingly paradoxical results.

It is worth, first, briefly reconsidering the *rationale* of national income aggregation. In order to obtain an aggregate of all the goods and services which together comprise the total national product, each item must be valued at a certain price, which – begging a few questions for the moment – is the price at which it changes hands. Some common unit of measurement is obviously necessary. But what is the justification of valuing each item at this particular price? We would naturally use these prices if we merely wanted to describe the actual exchanges taking place at a particular time, or to study the interrelationships of different parts of the economic system. But there must be more to it than this. What is it that the exchange price is intended to measure? From the point of view of the consumer the answer must, I think, be that the distribution of his expenditure over different commodities is assumed to be such that the marginal utility of each purchase is proportional to its price; that, in other words, the last shilling spent on any one item yields, broadly speaking, the same satisfaction as the last shilling spent on any other item. From the producer's point of view the answer must, I imagine, be that the distribution of factors of production is assumed to be such that the marginal product of each is proportional to its price.

If this is the *rationale* of the usual process of aggregation it may be worth recalling the conditions to which it applies. It may be noticed at once that, if the aggregate is to be given a welfare connotation, different goods and services must be valued at their market prices, including the effects of indirect taxes and subsidies; but that, if the aggregate is to be interpreted in a productivity sense, the valuation of final goods and services must exclude indirect taxes and subsidies levied after the final stage of production – otherwise the proportionality conditions will not hold good. It may be noted, in passing, that the marginal products and the costs of different factors will be proportional only if the costs are those actually incurred in purchasing one factor rather than another. An entrepreneur who is deciding on how much of different factors and materials (which embody factors used at earlier stages) to purchase will have regard to their market prices, including any indirect taxes, and will compare these costs with the value of the product, excluding any taxes that are levied on the product. In the aggregate, therefore, all indirect taxes (less subsidies) that are necessarily incurred in the process of production, i.e. all except those imposed after all the processes of production have been completed, must be included in the valuation of the national product. Marginal products are assumed, in other words, to be proportional to *factor costs*, not to factor *rewards*.[1] It is clear, therefore, that even in a closed system the two aggregates themselves will generally differ – by the net amount of any indirect taxes that are imposed after the final stage of production.[2]

Consider the welfare aspect first. The assumption that there is perfect competition in the markets for all consumer goods is perhaps not *too* wide of the mark. But the assumption about proportionality applies, of course, only at the margin. Although for each consumer the marginal utility of any item is supposed to be proportional to its price, its total utility certainly cannot be assumed to be proportional to his total expenditure on it. The first unit purchased may not give the same satisfaction as the tenth, or hundredth, or last unit purchased. It follows that two

[1] The term factor costs is commonly, and misleadingly, interpreted to mean factor rewards. A fuller discussion of this problem will be found in my article on 'National Income at Factor Cost or Market Prices', *Economic Journal*, June 1955.

[2] The only taxes of this kind, in the United Kingdom, are those which represent the price of a privilege, not of a good or service (e.g. marriage, dog and gun licences).

aggregates are not strictly comparable if they contain widely different quantities of several commodities. Furthermore, although the proportionality condition may hold good for a single consumer, it obviously cannot be assumed to hold good as between different consumers. For these two reasons, it is difficult to attach any meaning to an aggregate which covers all consumers, and all units consumed of every commodity. The most that it can do is to serve (i) as an index, for the purpose of arranging different aggregates of commodities in order; and (ii) as a starting point from which relatively small changes in the neighbourhood of that starting point can, on certain assumptions, be compared. For a single consumer whose tastes remain the same, small changes of apparently the same amount, measured from points which are not far apart, will have approximately the same significance, provided that substantial changes do not occur in the relative quantities of the different commodities comprising the aggregate. A change from 98 to 100 may be assumed to mean pretty much the same as a change from 100 to 102, and will perhaps have something in common with a change from 120 to 122, but cannot be directly compared with a change from 200 to 202. On the same assumptions, marginal changes in an aggregate covering all consumers may be compared in a similar way, and the aggregate can be used for purposes of ordering, provided further that there is no change in the distribution of incomes. The last qualification is important: the limited significance which can be attached to changes in the aggregate is likely to be very sensitive to changes in the distribution of incomes; indeed it is difficult to define what is meant by 'no change in the distribution of incomes' if the situation changes at all.

It is important to recognize that the significance of a national income aggregate, whether for purposes of ordering or for comparison of marginal changes, is bound up, firstly, with a given composition of this aggregate, although limited changes in composition are permissible, and secondly with a particular distribution of incomes. Such an aggregate is bound to reflect the distribution of incomes, not only because it is without body or substance unless a particular distribution of incomes is postulated, but also because the distribution of incomes affects the prices of different commodities and hence the weights used in calculating the index. These limitations apply equally to the

welfare and to the productivity aspects of national income.

If substantial changes take place in the distribution of (real) incomes, or if movements in the output of different commodities – especially commodities which are known to have largely different markets – are widely different, a general measure of changes in real income can have little meaning. In these circumstances we can hope to obtain valid indices only for particular classes of the community which have not experienced such changes. A general measure embracing everyone has validity only if everyone has experienced the same change in real income. It is obvious, indeed, that a general measure, covering all members of the community, involves some assessment of the relative importance of each member. Why should the rich (poor) be given greater weight simply because they are rich (poor)? And yet who is to say what should be the relative weights for different people? Should everybody – men and women, healthy and infirm, young babies and octogenarians with one foot in the grave, spinsters and profligates, scholars and fools – should they all be given the same weight (they are not at present) simply because no good reason can be found for giving them different weights? Clearly the assumptions made – and assumptions of this kind, whether stated or not, are often made – can very materially affect the results.

The interpretation of national income aggregates in terms of productivity rests on assumptions which are closely analogous to those already considered. All industries are assumed to enjoy perfect competition, in the sense that there are supposed to be no restrictions on the entry of new producers. It is also assumed that there is perfect competition between purchasers, as well as between sellers, of all factors of production. Proportionality is assumed between the price and the marginal product of every factor; but of course the marginal product of a factor, or combination of factors, may vary with the level of production: constant returns to scale are unlikely to prevail at all levels of output. It follows, again, that only marginal changes in the aggregate, measured from points in the same neighbourhood, are directly comparable; and that changes in the aggregate lose their significance when the changes in the different commodities comprising the aggregate are markedly different.

As a measure of changes in productivity the index is therefore bound to be influenced by the extent to which different industries

or factors enjoy monopolistic or monopsonistic positions. If the degree of monopoly were the same in all industries, or if changes in the output of different commodities were approximately the same, there would be need to worry. But when changes in the output of different industries show substantial variations, the inclusion of monopoly earnings in the weights for each industry is likely to distort the general picture. This type of distortion may not often be serious, but is likely to become important in periods of radical change.

All this sounds rather discouraging. But the position is not, perhaps, as bad as it appears from a recital of the underlying assumptions. It would be a very topsy-turvy world in which the assumptions were so wide of the mark as to destroy the value of all national income estimates. It is as well, nevertheless, to recognize their full limitations from the outset.

The qualifications that have been mentioned above apply, it must be remembered, not only to comparisons of the national product at different points of time within a given country, but also, and indeed with much greater force, to comparisons between the national products of different countries. To assume that competition is the general rule in a single country is one thing. But to assume a similar degree of competition between the products of industries, or between the factors of production, situated in countries that are separated by oceans, that have different climates, traditions, customs and, maybe, different industrial systems; and to assume even broad similarity in the tastes of consumers, or in the distribution of incomes, in different countries is to place a very severe strain on the imagination.

It follows from what has already been said that a measure of the change in the aggregate should be such that, if marginal quantities of some commodities, or factors of production, are substituted for marginal quantities of others, the value of the aggregate, whatever the point of view adopted, should be unaffected. This, again, shows that both the aggregate itself and a measure of the change may be different, according to whether we are thinking of productivity or of welfare. If, however, there are large-scale substitutions the whole method breaks down and it becomes impossible to measure the change at all accurately.

It is useful to remember these principles when a sudden

change occurs in the composition of the aggregate and it becomes necessary, unless all attempt at comparison is abandoned, to resort to some arbitrary method of bridging the gap. When, for instance, there is a fundamental change in the character of a product it is possible that a measure of the 'change' in output which seems appropriate when considering welfare may not be appropriate when considering productivity. Take the familiar example of beer. Suppose that water is added to beer, so that the number of bulk barrels increases relatively to the number of standard barrels. There has been a fundamental change in the quality of the product and it is necessary to find some measure of equivalence between the two kinds of beer. From the welfare point of view we need to decide how much of the new, weaker beer gives the same satisfaction as a given quantity of the old, stronger beer (or vice versa). From the point of view of productivity, we need to decide how much of the new, weaker beer could have been produced with the resources which went into producing the old, stronger beer (or vice versa). The answer in the one case may approximate to the number of standard barrels, and in the other to the number of bulk barrels. At all events the two measures may be different.

It is also worth considering, against this background, the suggestion that the units in which the various goods and services are measured should always be determined by reference to the market contract. At first glance this seems a sound enough general principle; but it can lead to rather curious results. Consider its application to the case of doctors. The productivity of doctors would then depend on the form of the institutional organization which arranges medical attention for patients – whether privately, or through a national health service. In an epidemic, when doctors in the health service are called on to do more work, the index would not show any rise in output. Even from the welfare point of view it seems unduly Erewhonian to argue that because the consumer, under a national health service, pays the same as before, therefore he obtains the same quantity of services as before. It seems more reasonable to assume, even from the point of view of welfare, that when there is an epidemic the public obtains a greater quantity of the doctors' services; and that the price in effect is reduced. Doctors will be prevented, during an epidemic, from doing the work they would otherwise be able to do in their spare time, which we may

suppose they spend in writing their autobiographies. The output of autobiographies will fall and, on the market contract principle, there will be no compensating increase in the output of medical services supplied under the national health scheme. Doctors may be more or less efficient at writing their autobiographies than at doctoring. But the proposition that their autobiographies are *far* more valuable than the extra medical services which they provide during an epidemic would not, I think, command general assent. Suppose that, in a very severe winter, consumers spend more than they usually do on clothing. No one suggests that the extra output, thus induced, should not be regarded as part of the national product.

Or take the case of life assurance. What the consumer purchases – life cover – is the same in each year that his policy is in force. But the work of the insurance company is much greater in the first year, when the policy is drawn up, and in the last year, when the claim is met, than in the intermediate years. This is, indeed, recognized in the indicators that are generally used to represent the output of the insurance industry, which take into account the number of new policies and the number of claims met, as well as the number of policies in force. The allocation of output between different years may also depend, therefore, on which of the two problems, to which we attach the labels productivity and welfare, we have in mind.[1]

There is a further widespread illusion which ought to be dispelled. The main test of a price or quantity index should be whether it stands up to the demands made on it in actual, practical problems; whether, that is to say, it will behave, in given conditions, in a way that satisfies the requirements which are inherent in the nature of the problem. It is sometimes taken for granted, without proof being offered, that a quantity index and a price index of the same group of commodities should be such that their product equals the (unique) ratio of total values. Yet the possession of this property is not, in practice, felt to be a necessary attribute of either price or quantity indices; nor is there any reason in logic why price and quantity indices should

[1] The distinction between a measure which relates to changes in productivity and one which relates to changes in welfare was, I believe, first adumbrated by Hicks in 'The Valuation of Social Income' (*Economica*, 1940). Though attacked by Kuznets (*Economica*, 1948), it was re-established by Little in *Economica*, 1949, and in 'A Critique of Welfare Economics' (Oxford University Press, 1950) and elaborated by Samuelson in 'Evaluation of Real National Income' (Oxford Economic Papers, 1950).

have this property.[1] The change in the value of a group of commodities can in fact be analysed into (i) the effect of changes in prices, (ii) the effect of changes in quantities, and (iii) the interaction between changes in prices and changes in quantities. There is no reason to assume that the price and quantity indices should be such as to make the interaction term equal to zero.

II. THE FUNDAMENTAL NATIONAL INCOME AGGREGATES

Something should be said, before proceeding further, about the main national income aggregates, and in particular about the effects of foreign trade. We are concerned with the following three categories:

(*A*) gross domestic product=consumption+ home investment +exports−imports;

(*B*) gross national income=gross domestic product+net (property) income from abroad;

(*C*) gross domestic resources=consumption+home investment.

Each of these categories is important and has its own uses. The difference between the gross national income and the gross domestic product is simply net income from abroad. This is all property income since it is part of the gross domestic product of some other country.

The difference between gross domestic product and gross domestic resources is

exports−imports=net investment abroad−net income from abroad+net unilateral transfers abroad.

The difference between gross national income and gross domestic resources is net foreign investment (i.e. net purchase of foreign assets)+net unilateral transfers abroad.

The importance of the first two aggregates is well recognized. The third – gross domestic resources – provides a measure of the total resources actually used during the year, either for investment or for consumption, and includes any resources

[1] This was one of the wholly arbitrary tests which Irving Fisher introduced in that lamentable example of misplaced energy, 'The Making of Index Numbers'. The book was properly castigated by Yule (*Journal of the Royal Statistical Society*, 1923, p. 424) and Bowley (*Economic Journal*, 1923, p. 90) on its first appearance, but the criticisms have since apparently been forgotten. It was pointed out, for instance, that if two sets of rectangular areas are compared, the average difference in length multiplied by the average difference in width does not equal the average difference in area, except under peculiar conditions.

obtained as income from abroad or from foreign disinvestment. Since it includes all imports and excludes all exports, it automatically reflects any changes in the terms of trade.

We can conceive of changes in both (*A*) and (*C*) being measured in real terms. But the measurement, in real terms, of changes in (*B*) presents a real conceptual difficulty. Income as such does not, strictly, have any real (physical) counterpart. Until income is spent, it cannot be identified with real goods and services. Of the three possible aspects of national income – product, income and expenditure – only two – product and expenditure – can be identified in real terms. Any attempt, therefore, to measure national income as such in real terms is bound to be very arbitrary. It involves, in particular, the deflation of net income from abroad, the conventional method being to use the index of the prices of imports. But there seems no good reason to use the prices of imports, in preference to the prices of exports or of domestic expenditure; income from abroad can be used either to finance additional imports or to reduce exports, or consequently to make use of more home-produced goods, or even to increase foreign investment. Since it is impossible to know how this income is used – it is not earmarked for any particular purpose – deflation by any of these indices is scarcely better than multiplying by the number of snakes in Ireland. Any adjustment for changes in the terms of trade is also bound to be arbitrary and conventional. The adjustment usually consists in replacing exports by imports of the same total (current) value, although some exports may represent property income payable abroad and some may be balanced by foreign investment.

The gross domestic product and gross domestic resources can each be valued either at constant market prices, or at constant factor costs. It may appear, at first sight, that the gross domestic product has more meaning if valued at constant factor costs; and that total domestic resources has more meaning if valued at constant market prices. The method generally followed for the valuation of imports at constant market prices – a necessary stage in estimating the gross domestic product at constant market prices – is indeed somewhat arbitrary when some imports are subject to heavy customs duties. Conventionally, customs duties are not included in the valuation of imports, on the grounds – presumably – that all taxation is a matter which is purely internal to the country concerned. And yet the actual

cost of those imports, measured against the possible alternative use of home-produced goods, clearly includes the customs duties. But, apart from this type of difficulty, there appears to be no reason, in principle, why the gross domestic product should not be valued at constant market prices, or gross domestic resources at constant factor costs. The distinction between product and resources need not be tied to the distinction between productivity and welfare.

III. ESTIMATION BY TWO ROUTES

In practice there are only two routes by which the gross domestic product (or gross domestic resources) can be estimated: one proceeding from the production side, the other from the expenditure side. Although the two routes should end up at the same aggregate, they proceed from quite different directions, at almost no point do they coincide, and detailed comparisons between the individual components of the two estimates are therefore, in the nature of the case, impossible.

The production method consists in summing the net contributions of all industries to the gross domestic product (so-called net output, including depreciation), valued at base year prices. All industries are included, irrespective of whether they produce raw materials or intermediate products for use by other industries, or goods for final use in investment or consumption, or – as often happens – some combination of these. The expenditure method, on the other hand, consists in estimating all the components of final expenditure on goods and services – consumption, government expenditure, investment in fixed capital and stocks, and exports – at constant prices. No attempt is made to determine the contributions of different industries to any particular item of final expenditure. If we had complete, accurate and detailed information about all transactions, the two methods should, in theory, lead to the same answer. But without this information – without in fact a full Leontief matrix for two or more years – any detailed reconciliation of the two sets of results is impossible.

There are, in fact, various reasons why the two methods must be *expected* to give different results. In the production method, we have to measure the change in the net contribution of each industry to the gross domestic product. In practice, the change in

net output is generally represented by a measure which relates to gross output, including the contributions of earlier stages of production. Any change, therefore, in the ratio of net to gross output would result in the measure being, to that extent, inaccurate. For example, if a firm decides to instal its own electricity generating plant, the resulting increase in its net output would not be reflected in the measure of gross output. Changes in the ratio of net to gross output can also be caused by changes in the amount or quality of workmanship incorporated in the product; by economy or extravagance in the use of materials; by changes in the technical processes of production; by changes in the degree of integration of industry not already reflected in the available statistics of gross output; or by changes in selling costs or payments for other services per unit of output.

On the expenditure side, also, it is often necessary to make assumptions and approximations. Arbitrary assumptions, for instance, usually have to be made about the proportions of certain items of consumption that are chargeable as business expenses, and which are not, therefore, part of personal consumption. Changes in quantities are sometimes measured directly and sometimes indirectly from money expenditure and estimated changes in prices. In either case it is seldom possible to be sure that full account has been taken of changes in the quality of the product; and there may also be changes in the amount, or quality, of services purchased with the product, the accurate measurement of which is virtually impossible. More important, there is usually no direct information about physical changes in stocks, which therefore have to be estimated indirectly, for the United Kingdom at least, from estimates of the value of stocks and rather precarious estimates of 'stock appreciation'. The estimate of the change in stocks is the least reliable component of our real expenditure estimates; more will be said about this later.

IV. ESTIMATES FOR THE UNITED KINGDOM

Estimates for the United Kingdom have now been made for each year since 1946 by both these methods. In 1952 some tentative estimates, using the expenditure method, were prepared by the Central Statistical Office and published in the

National Income Blue Book. At that time only preliminary investigations had been made. Further work has since been undertaken, a new set of estimates has been compiled by the expenditure method and a virtually independent set of estimates has been compiled by the production method. Both sets of estimates relate to the gross domestic product of the United Kingdom. Two estimates were at first made by each method, one at factor cost and the other at market prices; but the allocation of indirect taxes and subsidies between different industries, in the production method, was necessarily somewhat arbitrary and has not since been reattempted. All the estimates have been expressed in terms of 1948 prices, as this was the most recent year for which we had the full results of a Census of Production. The estimates have been published in a summarized form in recent issues of the Blue Book.[1]

Importance was attached to keeping the estimates derived by the two methods as independent as possible. Co-ordination has resulted in the elimination of discrepancies due to differences in definition (e.g. of what is included in the field of economic activity); the adoption as far as possible of the same principles for measuring changes in the output of final goods and services; the derivation of all base year figures and weights for both calculations from the same source, namely the estimates for 1948 given in the Blue Book; and the avoidance of arbitrary differences in the use or treatment of the available data. Where, for example, we were measuring changes in the same thing and there was a choice between two or more possible indicators (e.g. physical quantities and deflated values), the same method was used in both sets of estimates. And we have adhered to the definitions of subsidies and indirect taxes used in our national income estimates, which follow very closely the definitions which are now generally accepted in national income work. Accordingly none of the government services to industry (cf. p. 161 below) have been treated as subsidies. It is intended to provide a full detailed description of the derivation of the real product and expenditure estimates in due course.

The production method has for some time been applied to all the industries coming within the scope of the Index of Industrial Production; that is, mining, manufacturing industries generally,

[1] National Income and Expenditure 1946–53 (H.M.S.O., August, 1954) is the latest which has so far appeared.

building and (so-called) public utilities, which together account for about half the gross national product. The problems involved in measuring changes in output in these industries have been described and full details of the statistics used in compiling the Index have been given elsewhere.[1]

More difficult as well as more interesting problems arise in measuring the less tangible contributions of industries engaged in transport, distribution, the production of services and public administration and defence. The estimates for these branches of the economy were prepared in close collaboration with the Department of Applied Economics of Cambridge University and the methods used in their path-breaking work[2] have, with one or two important exceptions, been followed pretty closely. The general method was to use, wherever possible, a measure of the quantity of services rendered. The services of retail distribution, for example, are measured by the volume of turnover; national insurance is represented by the numbers insured and the numbers in receipt of benefit; hospital and similar services by the number of people receiving attention; courts of justice by the number of cases tried. Indicators of this type, representing changes in output, were used for the major part of local authority expenditure and for all central government expenditure on services to persons. Output indicators were used wherever possible, but in many cases no sensible measure of output was available; in such cases estimates of expenditure within the industry on wages, materials and other items of cost were deflated by appropriate price indices for each group of items. The latter method was applied to public administration and defence; and, in fact, to most of central government expenditure, as well as to certain items of local authority expenditure such as civil defence, child care, the fire service and expenditure on roads.

Where we were seeking indicators for the same industry we aimed to achieve consistency, as far as possible, between the production and expenditure methods; and in some cases the

[1] Studies in Official Statistics: No. 1. The Interim Index of Industrial Production, H.M.S.O., December, 1948, and No. 2. The Index of Industrial Production, H.M.S.O., October, 1952.

[2] cf. W. B. Reddaway, 'Movements in the Real Product of the United Kingdom, 1946–1949, *Journal of the Royal Statistical Society*, Vol. CXIII, 1950, and C. F. Carter, 'Index Numbers of the Real Product of the United Kingdom', *ibid*, Vol. CXV, 1952. For the more recent estimates, cf. A. A. Adams, 'The Real Product of the United Kingdom, 1946–1952', *London and Cambridge Economic Service Bulletin*, September 1953 (published in *The Times Review of Industry*).

same indicator has been used in both sets of estimates. But the most appropriate indicator for measuring changes in gross output (required in the expenditure method) is not necessarily the most appropriate indicator for measuring changes in net output (required in the production method). In principle, consistency ought to involve the use of different indicators[1] – and would do so more often in practice if the necessary data were available. The well-known Geary method is often useful to bear in mind, even when it cannot be applied.

It is obvious that many of the output indicators provide very crude measures of the volume of services rendered by the industries in question. But deflated figures of expenditure, a considerable part of which may consist of overhead expenses, will not in general provide a good measure of output. The number of persons benefiting from the services provides at least a first approximation to the kind of measure which is wanted. And in time it may be possible to devise better indicators, or to make use of additional information, as well as to extend the field which is represented by direct measures of output.

All in all the results of estimating the gross domestic product by the two routes, which are set out in the following table, agree remarkably well. The two sets of estimates labelled 'at 1948 factor costs' are in fact based on factor rewards and are so described in deference to the usual terminology (cf. p. 147 above). Fuller details of these estimates can be found in the 1954 National Income Blue Book.

Indices of the Gross Domestic Product of the United Kingdom at 1948 prices, 1948=100

	1946	1947	1949	1950	1951	1952	1953
Expenditure Method							
GDP at 1948 factor costs . . .	94.9	95.5	103.6	106.6	111.0	110.4	115.0
GDP at 1948 market prices . . .	95.9	96.1	103.1	105.7	110.1	109.8	114.3
Production Method							
GDP at 1948 factor costs . . .	94.6	95.5	104.5	108.6	111.4	110.1	115.0

[1] Except where an industry purchases nothing from other industries, so that its gross and net output are identical.

Up to 1950 the production method shows a larger increase than the expenditure method, and after 1951 it is the other way round. It is tempting to suggest, as a possible explanation of these divergencies, that there may be a tendency for the ratio of net to gross output to fall in a period such as 1946–50 when production is rising fairly rapidly, and to rise, e.g. because of economies in the use of materials or in selling costs, when production is relatively stable or falling. But to seek for explanations from among so many possibilities is not likely to be rewarding.

The remainder of this paper deals with some of the practical problems that have been encountered in the United Kingdom.

V. SOME SPECIAL PROBLEMS

1. *Rationing*

The Cambridge view is that rationing is part of the input of distribution. Their argument, as far as I understand it, is that, just as advertising stimulates sales, so rationing facilitates the distribution of rationed commodities – the implication being that if the government did not do so, distributors would be willing to pay for a rationing system – and that it should be treated, therefore, as a subsidy.[1] I wonder what retailers would say to this! Rationing can of course contribute towards a more equitable and even towards a more orderly distribution of commodities in short supply. But no subsidy is in fact paid to retailers; and indeed rationing is often accompanied by price control, which effectively limits the retailer's margin. Retailers are simply obliged, by law, to allocate so much to each customer; they do not obtain any extra reward for doing so. Rationing does not itself bring about a reduction in prices. Nor does it have any effect on the total quantity of goods made available.

It does, of course, influence the distribution of particular goods, making it more even than it would otherwise be. It could perhaps be regarded, therefore, as the cost of ensuring a more equal distribution of essential commodities than would otherwise result from the existing distribution of incomes. But the cost of operating a rationing scheme is obviously not the same as the resulting increase in real income, if indeed that can

[1] cf. W. B. Reddaway, 'Some Problems in the Measurement of Changes in the Real Geographical Product', *Income and Wealth, Series I*, p. 267 *et seq.* Mr. Reddaway admits the weakness of his argument in the case of sweets.

be measured. Moreover, as between two situations, one having an unequal distribution of incomes so that rationing is found to be necessary, the other having a more even distribution of incomes and not therefore feeling the same need for a rationing system, the Cambridge treatment places a higher value (at factor cost) on a given total supply of goods and services in the first situation than in the second.

It is also argued that rationing in effect improves the quality of distribution. But even if this be granted, it can hardly be maintained that the cost of the rationing scheme is a measure – even an approximate measure – of the extra quality obtained. The cases of the agricultural advisory service and of the employment exchanges are different since, although they are financed by the government, each of these services can be assumed to contribute to the output, either of a particular industry, or of industry in general. It is therefore reasonable to treat the cost of these services as subsidies to the industry, or industries, which benefit from them; although a similar treatment could be justified for many other items of government expenditure.[1]

Unlike most subsidies, rationing is not designed to benefit any particular industry or any specific class of the community. Unlike others, it does not result in a general reduction in prices from which everyone benefits in different degrees. It is intended, of course, to be for the benefit of the whole community, but so are a great many other branches of administration. The analogy with subsidies, therefore, seems to be tenuous in the extreme and rationing has not been treated as such in the present estimates.

2. *Utility Clothing*

In 1948, utility clothing and all footwear was subsidized, but certain specifications had to be met before clothing could be classified as utility, and non-utility clothing attracted purchase tax. From 1949 onwards there were no subsidies and in 1952 the utility scheme was abolished and replaced by the D-scheme which has a graduated purchase tax; the tax being levied when the goods are delivered by the wholesaler to the retailer. In estimating expenditure on clothing in 1949 and subsequent years

[1] The criterion which must be satisfied before this treatment can be justified is, in Mr. Reddaway's own words, that 'the service(s) in question must be essentially for the benefit of *industry*, not final consumers or the community as a whole' (*op. cit.*, his italics).

at 1948 factor costs, it is necessary to decide what clothing (if any) would qualify as utility, if the utility scheme and subsidies to clothing were still in operation. This is an example of the general problem of deciding whether goods produced at one date are the same goods as those produced at an earlier date, or whether they are different enough to constitute new goods. The method which was adopted depends on being able to obtain an index of clothing prices at factor costs. The procedure was briefly as follows.

For each main category of clothing expenditure a Paasche-type index of wholesale prices (excluding purchase tax) was obtained, an addition was made for the retail margin and an index thus constructed of retail prices excluding purchase tax; the index reflecting the subsidies in 1948. The index figure for 1948 was then increased by the proportion which total clothing subsidies in 1948 bore to total expenditure on clothing (at retail prices excluding tax) in 1948. The resulting indices, representing changes in factor costs since 1948, were used to deflate the estimated current expenditure on each category at current factor costs. The method implies that subsidies were intended to cause a general reduction in the prices of clothing by a certain proportion, which was broadly their effect from the point of view of the consumer. This seemed, looking back on it, the only thing to do since any subdivision, once the utility scheme had been abandoned, was practically impossible.[1]

3. *Electricity*

The question to be considered is: when is one unit of electricity to be regarded as the same commodity as another unit of electricity and when is it a different commodity? An answer to this question is needed, whether we want to measure changes in the consumption, or changes in the prices, of electricity. The two problems are connected and the following remarks apply to both. We must decide, first, what in principle should be done, irrespective of whether the required data are available or not; and consider afterwards how best to adapt the available data to our decision.

Under many of the arrangements governing the sale of electricity in this country, the price varies with the level of

[1] The treatment here very summarily described was suggested by Mrs. Muriel Venning.

consumption. The price per unit is constant within specified 'blocks' of consumption, but varies from one 'block' to another. There is generally an additional charge which represents overhead costs; and in industry the total charge commonly depends both on the maximum demand at any one time, and on the number of units consumed. The problem can to a certain extent be simplified by assuming that any one consumer purchases two commodities: (i) the facility to be able to consume electricity by simply turning on a switch, and (ii) the electricity itself. It is the second which causes difficulties.

As a general rule, if two commodities sell at different prices at the same date (and place), they are regarded as different commodities. But there are some exceptions to this rule, e.g. bulk purchase agreements, hire purchase schemes, clothing clubs, special discounts for favoured customers; but perhaps some of these are exceptions only in the sense that the necessary data for any other treatment are not available. Applied to electricity the rule would mean that all units within one 'block' of consumption would be regarded as one commodity, all within another 'block' as another commodity, and so on. The implications of this rule are not easy to accept. Is there, for instance, any reason why one kilowatt-hour consumed on the morning of Monday, November the 5th should be regarded as a different commodity from one kilowatt-hour consumed from the same connecting point on the morning of Monday, November the 12th? Surely not. The same rule applied to the price index would mean that when consumption increases and average payments per unit of electricity decline, the index would show no change provided the price per unit within a given block remained unchanged. The fact that the electricity authority has not *announced* any change in price does not seem a sufficient reason to assume that no change in price has occurred. In the case of goods which can be purchased in larger quantities at reduced prices, the concession reflects economies in the costs of production or distribution; furthermore the customer has to accept a large quantity all at once and, in doing so, accepts certain risks (the possibility of deterioration, or that he may not be able to use all of it). From both points of view, there is some justification for treating the sale of larger quantities as though they were a different commodity. But it cannot be said that either of these reasons applies to the case of electricity.

Thus there appears to be no good reason for distinguishing between one unit of electricity and another obtained from exactly the same point. Common sense suggests that the two units should be regarded as one and the same commodity. On this reasoning, the price can be taken as the average price per unit, over any given period, of the electricity obtained from a particular point. We have here, moreover, a workable solution. A complete analysis of electricity consumption into the different tariff 'blocks' is clearly unthinkable.

Having decided that all the electricity obtained from a particular source is one commodity, the only question that remains is, how far should we go in distinguishing between different sources of supply. The service involved in supplying a unit of electricity to, say, a motor car factory in Coventry is different from that involved in supplying a unit of electricity to a farmer in the Scottish Highlands; and the difference will not be entirely reflected in the charges which are intended to cover overheads. The two services should therefore be regarded as different commodities, just as coal in Oxford is a different commodity from coal in Newcastle. In a sense, indeed, the service obtained by any one consumer is different from that obtained by any other consumer. But it hardly seems necessary, and would certainly be impossible to collect separate figures for every meter in the country. It is sufficient to have separate figures for the major categories of consumers that have different price arrangements. Thus we ought to distinguish domestic consumers in Carlisle from domestic consumers in Canterbury, an aluminium factory in Birmingham from a cotton mill in Bolton, a naval shipbuilder in Glasgow from a trawler-builder in Hull. But how far we can go in distinguishing between different classes of consumer must, in the last resort, depend on the degree of detail in which the basic figures can be conveniently subdivided.

4. *Imports*

Price and volume indices of imports of merchandize are regularly compiled by the Board of Trade from the Trade and Navigation Accounts which differ, in regard to both coverage and timing, from the Balance of Payments statistics; and it is the latter which are used in the national income estimates. The Board of Trade indices of the prices and volume of imports are, in

addition, based on valuation at c.i.f. prices whereas for national income purposes insurance and freight have to be dealt with separately; only the part which is undertaken by foreign firms belongs to imports. Thus the Board of Trade indices are not in a form in which they can be directly used. For certain years detailed figures are available of the ratio of f.o.b. to c.i.f. prices. The volume indices were accordingly reweighted, and their coverage extended, to provide, as closely as may be, estimates of the value of imports of merchandize in each year at 1948 f.o.b. prices, having the same coverage as the Balance of Payments figures; similar estimates were made of the value of imports of merchandize at current f.o.b. prices. As the differences in timing in the two sources of information could not be satisfactorily dealt with, the price index implied by these two sets of estimates was then applied to the figures of the total value of imports of merchandize given in the Balance of Payments White Papers. For other items in the Balance of Payments (shipping, travel, military expenditure, etc.) price indices were estimated from such information as was available.

5. *Stocks and work in progress*

This item is the weakest link in the estimates obtained by the expenditure method. The volume of stocks and work in progress can show large variations from one year to another and, as there is considerable uncertainty in the estimates, any errors can have a quite marked effect on estimates of the gross domestic product.

The value of the physical change in stocks and work in progress is estimated as the difference between changes in the value of stocks and estimates of 'stock appreciation'. The estimates of 'stock appreciation' depend on a number of asumptions about the methods adopted by firms in valuing stocks and there is, it must be admitted, very little positive evidence for some of these assumptions. In particular, stocks are assumed to be valued on the first-in-first-out principle. Any firms which follow the last-in-first-out principle would have virtually no stock appreciation (unless there happened to be a reduction in physical stocks during the year, when it would be reckoned on the amount by which stocks were reduced).

Furthermore, in periods of rising prices – and it is more common, nowadays, for prices to rise than to fall – firms have an incentive to understate the value of stocks held at the end of

the year, since taxes are levied on profits including any increase in the value of stocks held since the beginning of the year. In periods of rising prices profits may therefore be underestimated and the estimates of 'stock appreciation' may be correspondingly too high. At any rate the estimates are more likely, in such circumstances, to be too high than too low and the estimates for recent years may, therefore, contain an upward bias. It so happens in fact that, if the initial estimates of 'stock appreciation' are reduced by a constant percentage each year, a rather closer agreement is found between the estimates of the gross domestic product for recent years obtained by the expenditure method and those obtained by the production method. On the assumption that the bias introduces a constant proportionate error into the initial estimates, a simple method of removing the bias is to reduce all the estimates of 'stock appreciation' by a given proportion x, and to determine x such that the sum of the squares of the differences between the two estimates of the gross domestic product over a period is a minimum (subject to the restriction $0<x<1$).

ACKNOWLEDGMENTS

Anyone who now attempts to estimate the real domestic product of the United Kingdom must feel greatly in debt to the pioneers of the Cambridge school, who were the first to make systematic estimates by the production method. I have had the benefit, in addition, of several discussions with Mr. W. B. Reddaway and Mr. A. A. Adams who generously made available their considerable fund of knowledge and experience. Much of the detailed work involved in preparing the estimates was carried out by statisticians in the Central Statistical Office and other government departments. I must, however, make it clear that I alone accept responsibility for the statements and views contained in this paper.