United Nations Statistical Office

000869 13 X

1

PRICE AND QUANTITY INDICES OF SERVICES

New York, March 1988

I. INTRODUCTION

1. The methodology for calculating price and quantity indices for services is one of the most controversial areas of the statistics of services and of price and quantity index calculation in general. Methodological problems are growing in importance and are becoming increasingly troublesome to national statistical offices. This paper is a review of the general issues related to service indices and is based on information currently available to the Statistical Office of the United Nations (UNSO). A more comprehensive review is planned after more information will be available to the UNSO.

2. It is a well-known fact that indices for services are far less accurate than indices for goods. In earlier days, when less was demanded from national accounting and especially when services constituted only a relatively small portion of the gross domestic product, not very much attention was paid to the calculation of service indices. Various methods were developed. However, it was recognized that for services (at least for some services), indices were of low quality and, in some cases even their acceptability was put in question. Nevertheless, this did not stop the regular computation of the indices and constant price aggregates, since it was considered that service indices were computed not so much for their own sake as for the sake of the total GDP indices; for these latter the inaccuracies of the service indices caused relatively smaller error only. For instance, even if the quantity indices for the public administration are not sufficiently reliable to draw analytical conclusions for the output and productivity of these services (e.g. to compare them with those of other activities) they can be applied quite well in the computation of the gross domestic product constant price data (e.g. for determining whether the rate of growth in the total economy increases or decreases).

3. While this thesis may have been quite true in a period when services constituted a modest part of the economy, even in this early period a number of difficulties hade already been encountered. Since many producers of statistics did not specify sufficiently the limitations of the service indices, these indices, once published, were applied by many users as if they were of the same quality as other indices. For instance some users compared the productivity indices of the public administration (derived from the published national accounting data) with the productivity indices of manufacturing or agriculture and drew conclusions from these comparisons which were highly questionable if one knew the methods by which the service indices were computed.

4. The real amplification of the service indices problem came, however, as service rapidly increased in importance in the national economies. In the last few decades the share of services in the gross domestic product substantially increased in practically all countries; in many developed countries it exceeded 50 per cent and in some it even reached 70 per cent. Is the above thesis still acceptable in these circumstances? Can a gross domestic product quantity index still be considered reliable if the quantity indices of its large constituent parts are not reliable?

5. Although the problem is not so severe as it may look from the 50 and 70 per cent references made above, it is still a big problem. It should be admitted that for some services (e.g. for transport, communication, laundries, hairdressers, a number of other personal services) the methodological problems encountered are not too different from those in the goods producing industries; thus, the quality of the quantity and price indices in these fields is not inferior to those of other indices. Nevertheless, the share of the service areas where problems are big (public administration, defense, health, education, a number of business services, etc.) is still relatively large and troubling to the constant price computations of the national accounts summary aggregates. This is one of the reasons why many countries attach high priority to the development of the methodology of the service indices, since they reckon that if no improvement can be achieved in this field, sooner or later the quality of all summary growth rates (quantity indices) will become questionable. And, of course, in spite of the inherent limitations mentioned, the improvement in the computation of the service indices can substantially contribute to the knowledge on what happens in the field of services in the national and international economies.

6. The present paper first tries to explain the reasons for the difficulties encountered in respect of the service indices (section II). Then a review is given of the various methods applied in different practices, of the background assumptions on which they are built and of the sources of inaccuracies (section III). Some selected methodological issues are considered next (section IV) and finally, suggestions are made in respect of ways for possible improvements (section V). In general, only the problem areas will be considered here; transport, communication, and other services for which no special problems arise, will not be dealt with.

II. THE REASONS FOR THE DIFFICULTIES WITH THE SERVICE INDICES

7. The problems with the service indices do not stem from one but from several factors. Some service areas are affected by just one of these factors, others by several. The present list (in paras. 8 to 10 below) does not claim to be exhaustive; nevertheless, an attempt is made to enumerate the main causes for many service indices to suffer from deficiencies.

8. Many services are not sold on the market and thus do not have prices which can be used to express their relative importance. Such is the case with public administration, defence, and in many countries with health, education, and research (or at least a large part of it). As a consequence, in these fields the market does not provide its beneficial role as measuring scale (via market prices), thus, some surrogate measuring scales have to be constructed, based on this or that assumption. The most common assumption accepted for the non-market services is that output moves proportionally with inputs which, in

many cases, is not an unreasonable assumption; however, measuring output on the basis of inputs (as it is done also in current price national accounting) disregards the fact, that there may be productivity differences between the different activities and that productivity (output/input ratio) may also change from one period to the other.

9. Many services have a unique character in the sense that each given service differs in respect of some quality characteristics from all other services of the same kind. Thus, either there are no appropriate units of measurement by which the performances of these services can be aggregated or, one can make such aggregations but at the cost of disregarding some quality characteristics. Health, public administration and research are typical of this category, but many of the business services also belong here. (For instance, there are no two entirely identical surgery operations, or criminal investigations or lawyer's services.)

10. The properties described in the two preceding paragraphs are the most important limitations for the service indices. There are, however, a number of additional problems causing difficulties for quantity and price index computations of various services.

(a) For some services, the changes in the quality are very large; computer services may be the most typical example of this. This causes frequent discontinuities (incomparabilities); the same service that was provided in the base period does not exist any more in the current period and, conversely, the service which is provided in the current period did not exist in the base period. Productivity changes are also very rapid; therefore, the assumption that "output moves proportionally with input" is hardly valid.

(b) Some services have a preventive character, that is, they do not so much to produce something as prevent something. Defense, police, some type of health services are examples of these. The problem with the output measurement in these cases is not that we do not have appropriate units of measurement, but that we simply do not have anything to measure.

(c) In many cases the effect obtained as a result of some services is not solely the result of the service provided but also the result of something else, and it is very difficult to separate the various effects from each other. For example, the result of teaching is the increase in knowledge (measured, e.g. by number of certificates); however, in improving the knowledge not only does the teacher (who, indeed, provides services) participate but so does the student, and it is hardly possible to separate the teacher effect from the student effect;

(d) For some services, quality differences can be judged only on the basis of subjective criteria. Cultural services (cinema, theater, etc.) are the typical examples of these. The consequence is that in most cases these quality differences are ignored.

(e) For some services, the customer (puchaser of the service) cannot appreciate the quality, and makes his preferences on the basis of some secondary, tertiary criteria. For instance, patients in general cannot judge the efficiency of the doctor (of the treatment he proposes) and therefore make their preferences on the basis of criteria which are less relevant, e.g. how nicely the doctor handles his patients, etc.

(f) Some services consist of a composite item, of a margin, and here, even conceptually, it is difficult to determine what is the quantity change and what is the price change. Trade is one example. Should the quantity index of the trade margin be proportional with the quantity index of the turnover (which seems to be the practice in most countries) or should one recognize that shifts in the market composition or commodity composition also may have a quantity effect on the trade margin. For instance, if the trader discovers a less expensive supply price and by this increases his trade margin, is this a quantity or a price change? Or if he/she increases the share in the turnover of the commodities with relatively higher margin while the whole turnover (quantity) remains constant, is the corresponding increase in the value of trade margin a pure price increase or is there room also to recognize some quantity changes in this case? Insurance may be another example for this type of problem.

(g) Some services are provided both on a market basis (e.g. financed by households) and on a free-of-charge basis (e.g. financed by government). In some cases it is difficult to judge e.g., whether it would be better to treat an appendectomy financed by the household and an appendectomy financed by the government as one single product representing the same quantity or as two different products, representing different quantities.

As a consequence of all the limitations enumerated above, for many of the service branches there are no genuine quantity and price indices computed, only some approximations of them. Approximations are not necessarily poor in accuracy; however, they are based on some asssumptions. The basic question determining the quality of the service indices, is therefore, how valid the underlying assumptions are.

III. THE VARIOUS APPROXIMATIONS APPLIED IN THE CALCULATION OF SERVICE INDICES

11. At present the UNSO possesses only partial information on the methods (approximations) applied in the computations of service indices in the national practices. 1/ Nevertheless, on the basis of this information it is possible to distinguish among some main approximation methods. On the basis of more detailed information to be collected from national statistical offices it will be possible to better elaborate the data presented in Table 1.

12. Table 1 concentrates on the main features of the approximations applied, without going into much detail. The denominations in this table are not necessarily identical with the terms used in national practices (or in earlier international recommendations). The assumptions applied and the sources of inaccuracies listed in the table are only those specific to the given method; common assumptions applied or common sources of errors (e.g. inaccuracies caused by the poor quality of basic data or sampling errors) are not mentioned. The examples are illustrative and do not imply any recommendations for a given service branch.

13. As can be seen from the table, the sources of inaccuracies are very different from one type of approximation to another. Therefore, the measures to be applied in trying to improve the quality of the service indices may also differ from method to method.

1/ The most detailed information in this respect is available in an OECD study: Measurement of Value Added at Constant Prices in Service Activities, Paris, 1987.

TABLE 1

APPROXIMATIONS USED FOR SERVICE QUANTITY AND PRICE INDICES

No.	Denomination	Assumption applied	Method of computation	Source of inaccuracies	Examples
1.	Rough input measure a) price approach b) quantity approach	Output moves with input (productivity does not change).	a) Deflation of input elements b) Extrapolation of input elements	Productivity changes are entirely ignored.	Public administration value added quantity moves with number of persons engaged in the administration, or public administration "prices" move with the average wages.
2.	Stratified input measure a) price approach b) quantity approach	Within strata output moves with input (productivity does not change).	a) Deflation of input elements (by strata) b) Extrapolation of input elements (by strata)	Productivity changes within the strata are ignored.	Health value added quantity index is a weighted average of the changes in the number of doctors and of the changes in the number of nurses.
3.	Rough or stratified input measure with productivity change assumption.	Output moves with input, adjusted for estimated productivity changes.	As in 1 or 2, but corrected with estimated productivity change indices.	Actual productivity change may be different from the estimated one.	As in 1 or 2, but it is assumed that labour productivity in the given service branch increases each year, e.g. 1 per cent (or to the same extent as in some other branches).
4.	Rough output measures.	Output moves with rough output type indicators.	Extrapolation by output indicators.	Quality changes are ignored.	Output (value added) quantity index of medical centers is taken equal with the changes in the number of patients treated.
5.	Benefit type measures.	Output moves with some benefit type indicators.	Extrapolation by benefit type indicators.	The effect of some external factors is incorporated into the output index. Quality changes may also be ignored.	The output quantity index of an educational establishment (e.g. university) is equal with changes in numbers of the diplomas (certificates) issued.
6.	Borrowed quantity or price indices.	Quantities or prices move in the same way as in some other branches.	Deflation by borrowed price indices or extrapolation by borrowed quantity indices.	Q or P indices may move differently in the two branches.	The price index of computer services is taken equal with the price index of another branch or with the overall wholesale price index.

IV. SELECTED ISSUES

A. Extrapolation versus deflation, double versus single indicator method

14. Constant price values for the current period can be obtained in practice in one of two ways: either by multiplying the base year constant price value by a quantity index or by dividing the current year current value by a price index. The first method is generally referred to as extrapolation, the second as deflation.

15. In determining value added (the difference between gross output and intermediate consumption) in constant prices another distinction between the two methods is worth mentioning. Both deflation and extrapolation can be made separately for gross output and intermediate consumption (and then the constant price value added can be determined as the difference between the constant price gross output and the constant price intermediate consumption) or the value added itself is deflated or extrapolated. The first is generally referred to as double indicator (double deflation, double extrapolation) method, the second as single indicator (single deflation, single extrapolation) method. A combination of the extrapolation and deflation methods can also be applied in the case of the double indicator method: e.g. gross output can be extrapolated while intermediate consumption is deflated.

16. Although these distinctions are general problems of the constant price calculations, they deserve a special scrutiny in the case of services. As to the choice between the extrapolation and deflation methods, the general theory of constant price computations gives some preference to the latter method 1/. This is also true with services, since the general rule that price relatives display less variation than quantity relatives (and, therefore, a representative price index has a smaller sampling error than a representative quantity index taken from the same sample) is also valid for most of the services. However, for a number of services other circumstances should also be taken into account which may reverse the relative preferences. The most important of these other factors which may lead to opting for the extrapolation method are the following:

1/ See, e.g. Manual on National Accounts at Constant Prices, United Nations Statistical Papers Series M, No. 64.

(i) services which are not sold on the market (like public administration, education, etc.) have no real prices, and, therefore, deflation (by cost components) may be more complicated and less accurate than extrapolation;

(ii) even for some marketed services, price and value data for other than benchmark years may be incomplete or not even exist, while some quantity information (to be used for extrapolation) may be more complete and reliable.

17. As to choosing between the double indicator and single indicator methods, the former is to be preferred, based strictly on theoretically considerations, since the implicit assumptions behind the single indicator method, i.e. that the intermediate consumption/gross output ratio does not change substantially and that the price indices of gross output and intermediate consumption do not differ much, are not sufficiently justified in all instances. However, in many cases no sufficient information is available for the application of the double indicator method, and in cases where the above assumptions are likely to hold, or the relative share of the intermediate consumption is relatively low, the single indicator method may provide a good approximation of the constant price data required.

B. Output versus input approaches

18. Output indices can be constructed directly (based on output type measures like quantities of services provided) or approximated by input indices (based on input type measures like persons employed, wages paid, materials used in connection with the services provided). Any input type approximation assumes constant input/output relationship, i.e. that productivity does not change between the two periods compared. For instance, if the quantity index of the public administration services is determined on the basis of the quantities of persons employed (for sake of simplicity let us disregard the other input items, i.e. intermediate consumption and consumption of fixed capital) it is assumed that each person produces the same amount of output in the current period as in the base period.

19. Input approximations can be applied not only in the area of non-marketed services (where even in current price accounting input is substituted for output) but also for some marketed services where the extrapolation method is applied and where appropriate output type extrapolation indicators are either not available or are only very crude. For instance, in a number of countries the base year value added of financial services is extrapolated on the basis of some labour input indicator, in most cases on the basis of number of

persons employed. Here again, constant productivity is assumed, i.e. that one person employed in the financial sphere provides the same quantity of services in the current period as in the base period.

20. Conceptually, output measures are, no doubt, superior to input measures, since the constant productivity assumption in the service activities is hardly justified. This shortcoming of the input type measure has been recognized for a long time, and earlier recommendations and guidelines proposed shifts in favour of output type measures. However, this problem is not so simple that it can be reduced to the simple statement that whenever feasible the output type approach should be applied. The productivity distortion of the input approach is not necessarily larger than the distortion of the output approach in cases when only relatively crude output type indicators, which disregard substantial quality differences, can be applied. And, as will be shown below, there are some ways to reduce the extent of the productivity distortion. Thus, the preference to be given to the output approaches should be interpreted only with some caution.

C. The productivity adjustment

21. The main shortcoming of the input type approach is its productivity distortion stemming from the implicit constant productivity assumption. It is understandable, therefore, that one of the main areas of experimentation in improving the service indices is just of how one can get rid of the constant productivity assumption. If input type indices were corrected by productivity indices they would become genuine output indices.

22. A number of countries, in their national accounting, apply some kind of productivity adjustments, i.e. they modify their input type indices by some estimated productivity changes. Some countries assume that productivity in the given service activities (where the input approach is applied) change at the same rate as productivity in the rest of the economy. Other countries adjust their input based quantity indices mechanically by an arbitrarily selected annual productivity increase rate (which varies between 0.5 and 2 per cent). Neither of these methods is sufficiently justified; however, it may well be true that any of these assumptions is still better than the zero productivity change assumption.

23. The crucial question is whether or not the actual increase in productivity can be estimated on the basis of some factual information. In some views these estimations should be based on some empirical studies, e.g. how many fewer staff who are better equipped and/or more experienced are needed to provide the same services as those in the base period. In other

views the estimations should be based on econometric type studies, i.e. on the relationship between the capital equipment/manpower ratio and the productivity. So far only limited experience is available on the feasibility of this type of estimation.

24. It should be noted that one component of productivity changes, namely, the effect of the shifts in the skill composition of the labour force can relatively easily be incorporated into the quantity indices. To attain this, only appropriate stratifications are to be applied. More details on this will be given in the next sub-section.

D. The stratifications in the index compilations

25. Stratification, in general, improves the accuracy of the indices by reducing their sampling error. This applies equally for all types of quantity and price indices. In the case of input type quantity indices, however, stratification may have an important additional advantage: it can incorporate into the overall index that part of the productivity changes which stem from the shifts in the composition of the strata (manpower). If the overall quantity index is compiled as an average of the strata quantity indices, and if these latter ones are weighted by values which can be considered as good proxies for the productivity differences (between the strata), the improvements or deteriorations in the inter-strata staff composition (from the point of view of the productivity differences) will be reflected in the overall index. For instance, if doctors and nurses are treated as two strata, and the average earning differences between doctors and nurses are considered as relatively good proxies for expressing productivity differences between these two categories, the overall index will reflect the productivity changes stemming from shifts in the doctor/nurse composition of the health personnel. However, this index will not reflect those productivity changes which occur within the particular strata, e.g. which stem from the fact that doctors are producing more than before.

26. The deeper the stratification is, the larger is the portion of the productivity changes that can be caught and reflected by the overall index. For instance, if doctors are not only separated from nurses but highly qualified doctors and moderately qualified doctors are separated also into two strata, then the overall quantity index will also reflect those productivity changes which stem from the shifts between highly qualified and moderately qualified doctors.

27. The advantages of the stratification, in general, cannot be denied. There are, however, two limitations to be taken into account. The first is a

practical one: how far the basic data needed are available in practice. Or, to put it in another way, is the improvement in the accuracy of the indices worth the cost of the additional information to be collected for distinguishing the various strata? The second problem relates to the mentioned underlying assumption: how far is it justified to assume that earning differences closely approximate productivity differences? There are, therefore, some limits in this type of improvement of the input type indices. Nevertheless, it seems that in many countries there is some room for improving the quantity indices of services by deeper stratifications than are presently applied in the practice.

V. WAYS OF IMPROVING THE QUALITY OF SERVICE INDICES

28. Most countries are not satisfied with the present quality of their service quantity and price indices and are making various efforts in trying to improve them. Although it is generally recognized that service indices - owing to their very nature - will always have limitations, a large number of national statistical offices think that some improvements would be feasible.

29. Some of the possible improvements have been dealt with already in the preceding sections; these will be mentioned only briefly. There are, however, a number of further possibilities of improving the quality of the service indices. The list in this section does not claim to be exhaustive, and can be extended and better elaborated as soon as more detailed information on national practices and plans become available.

30. One of the main methods of improving the service quantity and price indices is the detailing of stratification. This relates to the input type approximations, either in a shift from method 1 (of the table above) to method 2 or 3, or by making the stratification used for method 2 or 3 more detailed. The more strata are distinguished in a given area (e.g. in health or public administration in respect of employment), the better the productivity changes in that area are taken into account, since with this stratification the inter-strata productivity differences are taken into account in the computation while the intra-strata productivity differences are not. The problems of the stratification were described in paras. 25-27 above.

31. Direct estimation of productivity differences is the second method of improving the indices. This relates to approximation 3 in Table 1. The problems of the productivity adjustment were dealt with in paras. 21 - 24 above.

32. The third group of improvements relates to output type measures (item A in Table 1). Output type indices can be improved by using more homogeneous

units. For instance, instead of using surgery operations as units, one could subdivide the operations into groups (e.g. appendectomies, heart surgeries, colonectomies, or simply into categories such as minor and major operations). Then each type of operation could be used as a quantity indicator for the extrapolation, with some weights (e.g. man hours required for a given type of operation) applied to approximate the quantity differences among the various types of operations.

33. Further possible improvements can be achieved

(a) by switching from the single indicator method to the double indicator method where it can be assumed that the intermediate consumption/gross output ratio changes substantially, or that prices for intermediate inputs and outputs move differently;

(b) by switching from extrapolation to deflation where basic data available would permit it and where it can be assumed that in this way the error caused by the incomplete observation can be reduced;

(c) by selecting more precisely defined specifications as price representatives, in order to avoid the "unit value bias", i.e. the distortion caused by the fact that the average quality of the price representative may change substantially. In some fields, e.g. for repairs or for lawyers' services - where practically all services are "unique products" - it may be expedient to select standard specifications (even if in reality they do not exist) with exactly the same characteristics, and to price them by special enquiries (e.g. to ask a number of lawyers how much they would charge for a service in a divorce case with some given detailed characteristics, and to use these as prices for the price index computation);

(d) by using more refined units where the input approach is applied, e.g., by using man hours worked instead of number of persons employed.

34. There is one conclusion that already emerges quite clearly at this stage in the effort to improve service quantity and price indices: one should be very flexible in this exercise, it would not pay off to strive for strict uniformity. Conditions and circumstances differ to a large extent from service to service, and one has to adapt the selection of the method to be applied to these conditions. It would not be expedient, e.g. to press for using the input approach in education just because the input approach is applied in public administration, if it turns out that for education in the given country an output type approach promises better results. Neither would it be expedient to renounce the application of productivity change estimations

in a given service branch because in other branches it is not feasible to do the same. Uniformity is needed only on the conceptual level, requiring that the content of the quantity and price indices should be the same. However, no uniformity is needed in respect of the method of approximation which is to be applied. In each service branch the method of approximation should be applied which promises the most accurate results, and this method may differ from branch to branch.

35. Neither would it be expedient to strive for uniformity internationally, in that, e.g. the same approach should be recommended for a given branch for all countries, since circumstances in this branch may differ form country to country.