

THE MEASUREMENT OF REAL OUTPUT OF PUBLIC SECTOR SERVICES

BY REINO T. HJERPPE*

Recently there has been discussion concerning the renewal of the volume measurements of public sector services. This renewal has been proposed e.g. in the recent United Nations Draft Manual on Public Sector Statistics. In the present paper we discuss some theoretical and practical problems connected with this renewal. According to some preliminary calculations concerning the Finnish educational sector, the new methodology might lead to a considerable revision of figures of output and labour productivity in the public sector. The revisions are of such a quantity that they might cause significant changes in the measurement of the volume of the total gross domestic product. This is a fact which may still require reflection before the new methodology is generally introduced, even though the revisions as such may be highly desirable from several aspects.

INTRODUCTION

Traditionally the volume of public sector services has been calculated mainly on the basis of labour inputs. There is considerable similarity in the methodology in different countries (Hill, 1977a). Justification for this has been that the value of the services of the public sector are at least equal to the costs of their production (Hicks, 1940),¹ or that many of the services are of such a nature that labour productivity does not change significantly in them. The most important practical reason is, however, that it has been considered difficult to generate the necessary output data due to conceptual or data gathering problems. All these ideas can be criticized as not quite tenable.

In the recent U.N. Draft Manual on Public Sector Statistics (DMPSS, 1979) there is a recommendation to introduce a new methodology for output measurements. According to this proposed methodology data on the quantities of the services produced will be combined with data on uses of these services in order to get a volume measure of public sector services. This can be achieved by recording the transactions between a producer and a user of a service.

It is suggested in the DMPSS that volume series should be established for government final consumption, capital formation, consumption of fixed capital, value added, intermediate output and gross output. For government services such as health and education which involve transactions between producers and users, gross output in constant prices should be measured by the number of transactions that take place (DMPSS, p. 83). It is essential in this that the extent of the use of the services should be a basis for volume measures.

*An earlier version of this paper was presented at the 16th General Conference of the International Association for Research in Income and Wealth, Pörschach, Austria, August 19-26, 1979. The paper was prepared in connection with a research project led by Olavi E. Niitamo, Director of the Central Statistical Office of Finland and financed by the Academy of Finland. For final preparation of this paper a fellowship from the American Council of Learned Societies is gratefully acknowledged. Prof Victor Fuchs and Mr Laszlo Drechsler have provided valuable comments and Mr Jouko Kaartinen from the Central Statistical Office of Finland has been very helpful in the preparation of the paper.

¹The fact that the parliament has approved the expenditures can be used as a proof of this assertion, although this may not be quite convincing.

On the other hand collective services (mainly public administration and defence) do not involve such transactions and they are recommended to be measured by the amount of work performed (DMPSS, p. 85).

There are still a number of problems in this new methodology. The most essential question is whether the proposed methodology is better than the present one and what are the possible criteria for this superiority? In the following we shall briefly consider reasons why the volume should be measured independently of inputs and the problems in doing this and in interpreting the results. The new methodology will be preliminarily evaluated with some empirical calculations concerning the Finnish educational sector.

THE NEED FOR VOLUME MEASUREMENTS

There are a number of reasons to try to improve the volume measurements in the public sector. Some of them are listed in the following:

- Since in many countries the relative size of the public sector has expanded, a considerable proportion of national resources are allocated to public production. Therefore it is more important than earlier to know how efficiently these resources are used and what is happening to public sector output. Activities of the public sector have often been criticized as inefficient without any reference to indicators of this inefficiency. Of course one has to recognize that there are some aspects of public sector activities which are difficult or perhaps even impossible to measure.
- The underlying reason for output measurements is that they somehow relate to the concept of welfare, even though national accountants may have tried to deny this sometimes in the face of severe criticism of present output concepts (Saunders, 1977). Since aggregate welfare cannot be measured directly, the measurements have to be based on some convention. The problem is to select among the different conventional measures the one which is considered the best.
- Some theories of public sector growth start explicitly from the assumption that productivity will not increase (e.g. Baumol's model can be interpreted like this; see Baumol, 1968, also Skolka, 1977). If this kind of theory happens to be correct it has many interesting implications. Productivity measurements can be used to test these kinds of growth theories.
- Since services provided by the government are often similar to those provided by the private sector or have only been transferred from private to public production it is natural to ask whether the volume measurements should differ in the public sector from the general procedures used to calculate private sector output. The main point here is, however, the treatment of external effects which will be discussed below.
- The development of planning systems in the public sector has given rise to all sorts of output measurements in the public sector. The problem often is that these measurements may not have been based on any common idea.

The development and uses of planning data has not been in this sense systematic. One may think, however, that useful results could also be exploited in national accounting.

For example in Finland there have been several experiments to develop output (or performance) measurements in different services in the state planning system. Attempts have been made in the areas of health, social services and education. Some of these measures have appeared in the government documents reviewing the past activities of the government. The Ministry of Finance has given a task to the Central Statistical Office to develop a system of economic and activity statistics for the public sector. In this work the main emphasis so far has been on the development and coordination of economic statistics. The work is still in its initial stages. It may be mentioned that attitudes toward the development of output measurements vary in different branches of the government reflecting, perhaps, personal views. Some offices may think that it only involves the interests of the Ministry of Finance and therefore they are not eager to develop new measurements. This may be a practical aspect which also has to be considered in the development of the system.

There are also considerable differences between different branches of the government in the ease with which one can develop output measurements. In some parts of health services it may not be too difficult to record the number of transactions whereas difficulties are encountered particularly in areas dealing more with general administrative tasks.

BASIC THEORETICAL QUESTIONS

There are a number of basic theoretical questions in all volume measurements which are also relevant in this case.² First, one may ask whether the basic purpose is to measure output from the point of view of welfare or from the point of view of production possibilities. (This is a division suggested already by Hicks, 1940.) In the first case we are looking at the relevant utility functions and indifference curves and use as weights the appropriate marginal rates of substitution in consumption. In the latter case we are interested in the situation of the production frontier of the economy and in asking what is happening to this frontier. The relevant weights are marginal rates of transformation in production activities.

The second question relates to the interpretation of the real output indices. If the purpose is to look at the matter from the point of view of welfare, the question is what assumptions are made about individual utility functions and how these are assumed to be aggregated to permit interpretation in terms of the whole nation.

The third question relates to the weighting system. Although some of the questions (like whether to use consumer prices or factor costs) may already have been solved in answering the first question, there still remains a separate problem whether to use market valuations or some other pricing system, such as shadow prices. By using market prices we usually make an assumption of optimality of the price system. But if we have reasons (often well grounded) to believe that market prices are not optimal then the question remains whether we can generate a better weighting system. Of course in non-market services no market prices for output exist and therefore we usually resort to cost figures as weights. But the costs can be marginal or average and we have to decide between them. In order to use any cost

²Setting out these questions draws heavily on the classification provided recently by Sen (1979).

figures as weights at all we have also to make assumptions on the relationship between costs and shadow output prices. With no cost figures available these must be imputed.

The fourth question concerns income distribution. Are there ways to incorporate income distribution in aggregate measures of output? These are questions which have been discussed in recent welfare theory and in the public choice literature. It is clear that the usual assumption of optimal income distribution behind output indices cannot be very realistic.

Even though one cannot possibly answer all these questions unequivocally it is, however, useful to be as clear as possible on these matters in order to make the appropriate reservations concerning the practical calculations and their interpretations.

DEFINITION OF OUTPUT

In the past discussions of the concept of output considerable variations have appeared. These can be illustrated with the help of a production function. Any administrative or public process can be defined as a production process with a production function

$$Q = f(R),$$

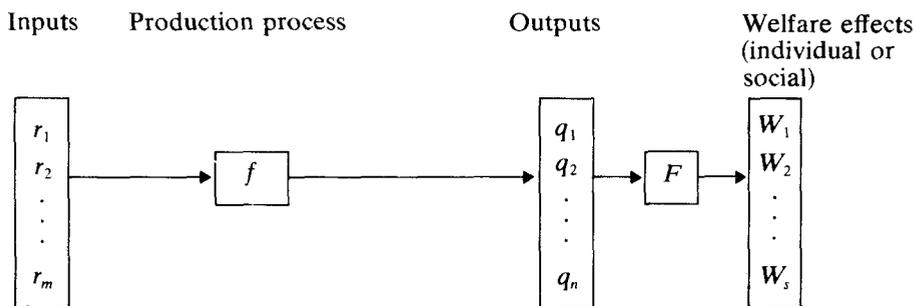
where Q is a vector of outputs, R is a vector of inputs and f describes the transformation process which may exhibit various properties of continuity etc. In order to have a practical interpretation one has to define operationally the elements of Q and in order to get an aggregate output measure one has to define relevant weights for the elements of Q .

A usual idea is that inputs can be easily measured. This may not, however, be the case. One can just think of measuring the labour input of the students in education or the inputs used by households in the care of the elderly. The main problem is, however, that outputs often seem either to be nonexistent or they are difficult to define. Sometimes Q has been interpreted as an input vector in a more general function transforming it into the welfare space. We have therefore another transformation

$$W = F(Q)$$

where the vector W describes the final welfare effects of Q .

This can be described also as a figure:



If we exclude the possibility of measuring welfare (individual or social) in cardinal terms, we can look e.g. at the goals and targets of an activity in order to have measures or indicators of W . As an example of this may be mentioned a paper by Hurst (1977) in a previous IARIW meeting. (There are also several other papers where Q is looked at as an intermediate input in producing W .)

On the other hand Hill (1977) takes a position, which is also reflected in the DMPSS recommendations, that one should not try to measure W but only Q . This corresponds more closely to the usual way of measuring output in private services.

The welfare effects can be measured on different levels and aspects, as has been pointed out e.g. by Nestor Terleckyj some time ago (1971). For example the immediate program objective for vaccination can be the prevention of an epidemic. This may have an effect on the general level of health, and better health leads again to a higher level of welfare. Depending on the purposes we can try to measure the effects of vaccination on these various levels.

It may be useful to compare this to the social indicator approach. The purpose of social indicators is (according to *Towards a System of Social and Demographic Statistics* (1975)) that they are connected with some area of social concern or they may just satisfy thirst for learning or understanding or they may serve some activity. This definition of social indicators is rather broad and it gives possibilities for a flexible approach. This flexibility is also a weakness: it does not contain a theory of how to form or construct social indicators. It would be desirable to specify whose learning, understanding or activity is in question or what is the social concern area and what are the criteria of selecting a particular area. By defining social indicators in such a flexible way it is clear that they cannot form a basis for aggregate output measurements in national accounting.

One may inquire whether modern cost-benefit analysis would supply methods of measuring output of public services. Cost-benefit analysis (CBA) is nowadays well over 40 years old and so roughly is national accounting (NA).

The basic idea in the CBA approach is of course different from that in national accounting in the sense that CBA is usually concerned with planning of alternative future projects, whereas NA will record the results of past activities. One of the central themes in CBA has been whether it is possible to express the value of a project with one cost-benefit indicator.

There are opinions that present possibilities expressing the value of public projects by one indicator are more limited than earlier (Dorfman, 1978). The first reason for this is that the nature of public projects has changed: earlier they were concerned more with economic type projects such as agricultural developments, construction of roads, prevention of floods etc. Nowadays public projects concern areas like health, education, environment etc. where results can be much less clearly stated. Another reason is that evaluation of a project is in the end a political decision and therefore it is not possible to describe the value of a project with one number.

The third reason is that any project has some bearing on income distribution: the benefits are not distributed evenly and somebody may lose even in generally good projects.

The fourth aspect is that there are regional benefits and costs involved. The importance of regional aspects has in recent times increased in many countries.

These reasons make it necessary to supplement quantitative cost-benefit estimates with qualitative descriptions. This can be seen as a reflection of the fact that the social welfare function is composed of at least two components: an economic one and a social or environmental one.

Some analysts are of the opinion that at least in principle income distribution effects can be combined by using some explicit social choice function. Opinions in this matter are not, however, unanimous. It seems clear that the CBA methodology does not provide a solution to the measurement problem.

A common feature in CBA, in planning, programming and budgeting systems and in national accounting is that these all try to evaluate output or results from user's side. The national accounts are to serve several purposes and therefore measurements should be sufficiently general. As long as we try to form national product aggregates it is obvious that the methodology of social indicators or cost-benefit analyses is not appropriate, since one has to get somehow a weighted aggregate measure which those systems do not provide.

CLASSIFICATION PROBLEMS

In the past there has been discussion of how one should divide government output into final consumption, capital formation and intermediate products. According to the current United Nations System of National Accounts (SNA) all government output which is not sold to consumers is defined as final product (either government consumption or capital formation). When discussing this Hicks (1948) and Kuznets (1948) have been of the opinion that some part of the product should be considered as intermediate. Recently, e.g. Leffler (1978) has again raised the question of the intermediate nature of some government outputs.

Government products are intermediate by nature in three ways: (1) Government can produce services for private enterprises (e.g. various economic services, some research etc.) which are preconditions or infrastructure of the whole production system in the economy. (2) Some products (like roads) are used by consumers in their consumption process as inputs (in producing travelling). In this case we have conceptually a Lancasterian (1966) or Beckerian (1965) type household production function (see also Sandmo, 1973). Household activities are not, as is well known, considered as production in the present SNA. (3) There are "intermediate" type expenditures inside the government sector itself. Government is producing a number of services for its own use (planning services, information, preparations of laws, research etc.). The purpose of these services is to increase the efficiency of government operations. If these lead to better social policies we can think of these services as collective goods (Head (1962)). It is quite plausible that for example health and education require more administrative services than earlier. From the point of view of welfare interpretation it would be desirable to delineate these services as "intermediate government services."

The making of a distinction between final and intermediate services has been objected to as impractical but on the other hand there have also been arguments (as exemplified already by Hicks, 1948, p. 164) that it is practically quite possible to make this distinction. We could also point out that the mere size of the

government is such that it would be useful to have a more detailed analysis of the nature of services than what is now generally provided.

The SNA takes a somewhat intermediate position on this problem since it recommends the separation from public consumption of the part which can be allocated to individuals. It would also be useful to consider separately intermediate type services which are used only by enterprises. This is relevant in the analysis of the welfare implications of government expenditures and in international comparisons.

CLASSIFICATION OF GOVERNMENT SERVICES FROM THE POINT OF VIEW OF VOLUME MEASUREMENT

There are no classifications of services available for the purpose of volume measurements. It is possible that without such a classification various practices will develop in different countries. Classifying the expenditures by the functions of government is not quite suitable for these purposes. One of the possible tasks for international cooperation would therefore be to look for a relevant service classification for the purpose of volume measurement.

EXTERNAL EFFECTS

Government services which involve transactions are of individual type and they usually are also services which include external effects.³ The purpose of government in producing these services is often that it can best take care of externalities connected with these services. A typical example is compulsory education which is necessary for the production system to prevail in its present form and in order to develop it. Other examples are health services where treatments bring benefits also to other individuals e.g. preventive vaccination. The essential point is that the value of these services is not necessarily reflected in their costs. Therefore transactions may be assigned too low a value. It is difficult to find a general solution to this. In some cases, e.g. basic education, the social benefits and costs cannot be measured with a reasonable accuracy. On the other hand, in some other cases it might be feasible to make the necessary adjustments. This would lead to the use of shadow prices. There is a critical argument by McKean (1968) about selecting shadow prices as an alternative value system since it might easily lead to arbitrariness. This is where future research could be done in the new methodology and it would be useful to consider cases where externalities can be measured in a reliable way.

DETERMINATION OF WEIGHTS

Since non-market activities have no market prices to use as weights and there are no obvious shadow price alternatives available, production costs are usually used as the required weights.

³External effects can, of course, be related to both private and public production. Externalities have been used, however, to justify public production in some cases.

It should be pointed out that the relevant prices in this case are not necessarily the average unit prices. If we think of an analogy to market goods it is usually assumed that market prices reflect marginal costs (provided there is Pareto-optimal allocation in the economy). There is a point where marginal and average costs are equal and it is of course possible to assume that this situation is relevant for non-market goods. This would be the case with no fixed costs and proportional increase of costs to output.

Going from market goods to non-market goods we could use the analogy and require the use of marginal costs as weights. These should be, in principle, marginal social costs and not marginal private costs. Even this principle has a limited application in a so-called second best economy. If a Pareto-optimum (first best) condition is not possible to achieve, marginal cost pricing is not optimal in any case (this has been established e.g. in Baumol and Bradford, 1970). The implementation of this principle leads, however, to rather strong information requirements concerning the structure of the economy and it is therefore generally doubtful whether second best pricing is a practical way to construct the weights at all.

Due to these informational requirements there may be cases where the marginal cost pricing principle is still the best one (this so-called third best situation has been examined by Ng, 1977).

For the case of pure collective goods we have a well known result by Samuelson (1954, 1969) based on the Wickseil-Lindahl voluntary exchange model which also requires marginal cost pricing for collective goods. This would also require that the collective goods economy is in a voluntary exchange equilibrium, which is a very stringent assumption considering the way collective goods are practically allocated and financed. This result again gives only a theoretical basis for weighting but we have no knowledge to what extent it is applicable in a real world situation. There may be no good alternative to using approximate average costs as weights, but one should be aware of the limitations of this type of weighting system.

THE QUALITY OF PUBLIC SERVICES

One important and difficult problem is connected with the measurement of the quality of the services. For example the use of computers in making diagnoses, research and say, by police in a criminal case, has made analyses much easier compared to the past. Many computations which have been practically impossible earlier are nowadays quite easy and standard procedures. These are clear indications of the rise of both quality and productivity.

In similar ways the quality of, say, a student hour in education may have changed for the following reasons

- better abilities of the students (due to e.g. better education in earlier stages of educational process or better environment at home)
- development of the qualifications of teachers
- increase in knowledge in the subject matter due to new scientific results
- changes in the sizes of student groups and classes
- progress in the methodology of teaching (pedagogy).

These appear relevant problems when we look at the empirical figures later. It is quite possible that the quality of an average lecture has improved during the past twenty years or so (there may be opposite factors, too). If this is the case, the number of student hours may be a poor indicator of the volume of educational services.

Theoretically we could try to develop methodologies which would decompose services into different characteristics as has been suggested in the case of durable goods. The theoretical basis for this could be for example the theory developed by Lancaster (1966). According to this, consumers are interested not in the goods themselves but in the various characteristics of the services which they can use to generate utilities for themselves (by using a process which has been called a household production function). The difficult thing in this approach is that there is no obvious way in practice to achieve the required decomposition of services to desired characteristics.

It may be possible to take the increase in capital into account in the calculations to make corrections in the productivity measures (see the examples on computers above). This does not, however, provide an exact measure of the increase in the productivity and may lead to arbitrary results.

If there is separate information on prices it is possible to use this by deflating expenditure series and compare these to the quantity data. The differences between the deflated series and the quantity series give in principle an index of quality change. This kind of methodology has been used e.g. by Usher (1975). For non-market services the problem lies in the absence of the relevant price data.⁴

INCOME DISTRIBUTION

The usual assumption in index theory is to assume that income distribution is optimal. The obvious counter argument to this assumption is of course that one does not believe in this optimality in practice.

The solution to this dilemma is to modify the weighting system so that it would reflect one's opinion of the desirable income distribution. In recent theoretical discussions it has been pointed out that, for example, using the shares of average consumers as weights in the consumer price index implies a value judgement (that it is the consumption of an average consumer which is important to follow, Mullbauer, 1976). One could ask what is the intrinsic value of using these weights. Why do we not use the weights of a poor consumer in the consumer price index? The answer is that we have made a value judgement when selecting our weighting system. Therefore the correct way to incorporate the income distribution to our index calculations would be to use socially desirable weights.⁵ It is obvious that these weights are subjective and are based on a subjective social welfare function.

⁴The use of price index deflation has been recommended in a private discussion by Prof. Victor Fuchs. The reason for this is that it may be easier to collect price and cost data on a sample basis than output data.

⁵Sen (1979) has suggested the use of so-called named goods, which means that a good going to different persons should be considered as different goods receiving different prices depending on the social valuation of the consumption of these persons.

One simple way to incorporate the income distribution in the calculation of the volume indexes would be the following. Suppose x is some measure of the inequality of the income distribution (say Dalton's, Gini's or Theil's).

If Q is our normal volume index we could compute "income distribution corrected" index by Qx . If for example we have a Gini index G , suppose it takes two different values G_1 and G_2 in two different income distribution situations. We can get $Q(1 - G_1) \geq Q(1 - G_2)$ whenever $G_1 \leq G_2$ (which means that the income distribution is more equal in the first case). (For this see Sen (1979) and the criticism of Hammond (1979)).⁶ One problem in this is that Gini's index does not take into account the form of the income distribution which may be relevant in evaluating the social preferences of different distributions.

WORK AS A FINAL PRODUCT

More speculative questions are connected with treating work as a consumption good. The classical treatment is to consider work as a negative utility. However there are nowadays a number of sociological studies which indicate that work may be an important thing in itself by giving possibilities for social contacts, job satisfaction etc. It has even been argued that the rapid increase in public services is due to the fact that people prefer to work in government jobs (Pollard, 1979). This could be an interesting factor in explaining the phenomenon which would be otherwise interpreted as a decline in productivity. But even though this idea could be taken seriously there are limited possibilities to integrate it in volume calculations.

Above we have discussed a number of theoretical problems which are connected with the construction of volume indices for government services. The list could be continued still, for example, to the issues of the changes in the tastes of the government. (If we take the tastes of politicians as relevant, then what happens to tastes when after elections the government changes, e.g. from left to right? What happens in this case to our welfare interpretation of the volume index?) We shall not, however, pursue more of these issues here.

SOME EMPIRICAL FINDINGS

So far very little experience is available reflecting the methodology of the new volume measurements.

As an empirical example we look at some figures on the educational sector in Finland. In the new methodology it is suggested that the volume should be measured as a number of student hours appropriately weighted.

Of course, we first have to define the educational institutions. That would include institutions established primarily for educational purposes. Therefore for example on-the-job education would not be included.

If we are interested in the production process of the educational institutions we have the following production function

$$x = f(L, K, R)$$

⁶Hammond's criticism of Sen's procedure is directed to the narrow nature of the implied utility functions.

where x is educational output, L is labour input, K is capital input and R materials.

In many applications x is identified with a student achievement or the level of knowledge of something else measuring the general level of education. (See e.g. a recent study by McGuckin and Winkler (1979).) If we now assume that x is merely student hours we have to think that this is just an input in the process which generates the level of education. Therefore the level of education is a function, say

$$S = S(x, A, E, J, Y, P, H)$$

where

S = the level of knowledge or education

A = inherited ability to learn

E = desire or motivation to learn

J = on-the-job education

Y = private education (at home etc.)

P = personal experiences

H = learning by doing

(It is certainly not easy to define all these variables operationally.)

By looking at the situation in this way we can see that educational institutions do not produce the level of education which is a result of another production process by a student himself.⁷ The national accountant would include as production only that process in the educational institutions and the relevant output would be in this case, e.g. student hours.

The output of educational services consists of student hours in lectures, of personal advising, and of participation in tests and examinations. The number of student hours can be achieved by multiplying the given lecture hours by the number of participating students. In a case when the average number of hours per student does not change in time, the number of students will qualify as an (unweighted) measure of volume of student hours.⁸

The available data in Finland consists only of the number of students in different branches of education. There are no exact data on student hours nor on the amount of student counseling by teachers. We have used therefore the number of students as an approximate measure of student hours. The different branches of education consist of the primary school, the secondary school, vocational schools, higher educational institutions and universities. These series are compared to the published national accounting data in the following table.

According to the national accounts the volume of educational services has increased by an average 4.2 percent annually during 1960–75. The labour input has increased by 3.8 percent. Due to the estimation methods there is practically no increase in productivity (the increase in labour productivity is by assumption

⁷As regards the weights they could be for example the value of alternative use of student time spent in the classroom. This would lead to different weights for different students and information on this could be hard to get.

⁸The difference between the two series is due to the volume of imputed rents for the school buildings.

zero).⁹ Taking the number of students as an output indicator reveals that the annual increase in output is only 0.4 percent. Accordingly the productivity of labour has decreased on the average by 3.4 percent annually. There is a considerable difference here from the figures in the national accounts.

It is quite true that these figures reflect intentional efforts to decrease class sizes. The educational authorities have obviously assumed that this increases the quality of education.

TABLE
INDICES OF EDUCATIONAL OUTPUT AND LABOUR INPUT IN FINLAND, 1960-75
(1964 = 100)

Series from the National Accounts					
	Volume Index of Educational Services (1)	Government Consumption in Education (Volume) (2)	Labour Input (3)	Number of Students in Educational Institutions (4)	Number of Students/Labour Input (4)/(3)
1960	80	80.4	83.3	98.6	118.4
1961	84	86.0	86.4	99.4	115.0
1962	90	95.3	92.0	99.9	108.6
1963	96	97.5	97.4	100.6	103.3
1964	100	100.0	100.0	100.0	100.0
1965	104	102.8	104.3	100.3	96.2
1966	107	107.3	105.4	100.6	95.4
1967	111	112.7	108.0	101.1	93.6
1968	116	118.5	111.9	101.4	90.6
1969	121	122.8	116.4	101.0	86.8
1970	126	128.6	119.8	101.7	85.0
1971	131	135.6	125.3	102.0	81.4
1972	136	144.8	129.3	102.5	79.3
1973	142	152.5	135.4	103.4	76.4
1974	150	163.7	145.2	104.3	71.8
1975	154	175.5	147.2	105.0	71.4
Total increase 1960-75	92.5%	118.3%	76.7%	6.5%	-60.2%
Average annual growth rate	4.2%	5.2%	3.8%	0.42%	-3.38%

Whether the result achieved is a sensible measure of output of educational services at all depends on what has generally happened to the quality of educational services. Here we can refer to the possible reasons for an increase in quality, that were discussed above. There are difficulties in measuring quality. For

⁹The students demand educational services e.g. in order to increase their future earnings. For this purpose they attend (after qualifying) those lectures which they consider most relevant for their purposes (they may acquire relevant "characteristics" from the educational process in order to "produce" relevant knowledge themselves).

example the student ratings used in evaluating high school graduates may have changed in time and they do not necessarily reflect correctly the achievement levels of the students. The requirement levels of examinations may have also changed. Without due reservations for the quality changes, the figures may give an erroneous picture about what has happened. It may also be that the general administrative tasks of the teachers have increased.

The calculated figures would indicate that of the total increase in labour input in the educational sector only a little less than ten per cent is due to the increase in the number of students. The rest has to be allocated to other factors.

If we suppose that the development in other branches of public services would have been the same (a heroic assumption!)—the growth of their volume being zero instead of four percent—that would decrease the growth rate of the total gross domestic product by approximately half a percent (the weight of these services being about 15 percent). Of course in practice there could be considerable differences between different service sectors.

SUMMARY

In principle the recommendation for the new methodology for volume calculations as suggested e.g. in DMPSS is interesting and should be welcomed. There is considerable interest in the measurement of government output independently of inputs and these efforts should be encouraged. This paper has discussed the assumptions, implications and problems connected with this new methodology. Most of these problems are well known by national accountants and it is quite understandable that they have been reluctant to accept alternatives to the present methodology. The very rough empirical figures also indicate—and this is perhaps most interesting—that the new methodology may in fact lead to considerable revisions of the present volume measures. When adopting a new method we should be, however, convinced about the superiority of the new method. This paper has not tried to yield this but this is not to say that this could not be achieved by devoting more effort to the measurements. One of the most important things which would require some clarification is certainly the quality measurements of public services.

We recognize that the final motivation for all output measurements in the framework of national accounts is to relate these measurements somehow to the concept of social welfare, but we cannot measure welfare directly. Therefore the problem of output measurements can be solved only by convention. This applies to volume measurements in the public sector as well. The question to be solved then is, which conventional methods lead to the most meaningful results.

International efforts could be devoted to the development of the new methodology. In addition, the need for a suitable product classification should also be discussed.

REFERENCES

- [1] Baumol, William (1967), The Macroeconomics of Unbalanced Growth, *American Economic Review*, June.
- [2] Baumol, William, and Bradford, David (1970), Optimal Departures From Marginal Cost Pricing, *American Economic Review*, Vol. 60, pp. 265–283.

- [3] Becker, Gary S. (1965), A Theory of the Allocation of Time, *The Economic Journal*, September.
- [4] Dorfman, Robert (1978), Forty Years of Cost-Benefit Analysis, in *Econometric Contributions to Public Policy*, Richard Stone and William Peterson (eds.), Proceedings of a Conference held by the International Economic Association, MacMillan.
- [5] Draft Manual on Public Sector Statistics (1979), United Nations, ST/ESA/STAT. 85/Rev. 1, February 1979.
- [6] Hammond, Peter (1978), Economic Welfare with Rank Order Price Weighting, *Review of Economic Studies*, June pp. 381–384.
- [7] Head, J. G. (1962), Public Goods and Public Policy, *Public Finance*, No. 3, pp. 205–208.
- [8] Hicks, J. R. (1940), The Valuation of Social Income, *Economica*, May 1940.
- [9] Hicks, J. R. (1948), The Valuation of the Social Income—A Comment on Professor Kuznets' Reflections, *Economica*, August 1948.
- [10] Hill, T. P. (1977), On Goods and Services, *Review of Income and Wealth*, December.
- [11] Hill, T. P. (1977a), Price and Volume Measures for Non-market Services (mimeographed).
- [12] Hurst, J. W. (1977), Sectoral Input, Output and Productivity for Health Services Illustrated for Maternity Services in England 1964–1974, IARIW Conference, York, England (mimeographed).
- [13] Kuznets, Simon (1948), On the Valuation of Social Income—Reflections on Professor Hicks' Article, *Economica*, February.
- [14] Leffler, Keith B. (1978), Government Output and National Income Estimates, in *Public Policies in Open Economics*, Karl Brunner and Allan H. Meltzer (eds.), Carnegie–Rochester Conference Series on Public Policy, North-Holland.
- [15] Lancaster, Kelvin J. (1966), A New Approach to Consumer Theory, *Journal of Political Economy*, April.
- [16] McKean, Ronald N. (1968), The Use of Shadow Prices, in *Problems in Public Expenditure Analysis*, Samuel B. Chase Jr. (ed.). The Brookings Institution.
- [17] McGuckin, Robert H. and Winkler, Donald R. (1979), University Resources in the Production of Education, *The Review of Economics and Statistics*, No. 2, May.
- [18] Mullbauer, John (1976), Economics and the Representative Consumer, in *Private and Enlarged Consumption*, Edited by L. Solari and J.-N. duPasquier, North-Holland.
- [19] Ng, Yew-Kwang (1977), Towards a Theory of Third-Best, *Public Finance*, No. 1.
- [20] Pollard, Sidney (1979), The Rise of the Service Industries and White-Collar Employment, in *Post-Industrial Society*, Bo Gustafsson (ed.), Groom Helm, London.
- [21] Samuelson, Paul (1954), The Pure Theory of Public Expenditures, *Review of Economics and Statistics*, pp. 387–389, November.
- [22] Samuelson, Paul (1969), Pure Theory of Public Expenditures and Taxation, in J. Margolis and H. Guitton (eds.) *Public Economics*, New York: St. Martin's.
- [23] Sandmo, Agnar (1973), Public Goods and the Technology of Consumption, *Review of Economic Studies*, 40, 517–528.
- [24] Saunders, C. T., Welfare and the GNP, in *Modeling for Government and Business*, Essays in Honor of Prof Dr P. J. Verdoorn, Martinus Nijhoff Social Sciences Division, Leiden, 1977.
- [25] Sen, Amartya (1979), The Welfare Basis of Real Income Comparisons: A Survey, *Journal of Economic Literature*, March.
- [26] Skolka, Jiri (1977), Unbalanced Productivity Growth and the Growth of Public Services, *Journal of Public Economics*, May.
- [27] A System of National Accounts and Supporting Tables (1968), United Nations, Studies in Methods, Series F, No. 2, Rev. 3, New York.
- [28] Terleckyj, Nestor E. (1971), The Measurement of Government Output, paper presented at the IARIW Conference at Ronneby, Sweden, 1971 (mimeographed).
- [29] Towards a System of Social and Demographic Statistics (1975), United Nations, Studies in Methods, Series F, No. 18.
- [30] Usher, Dan (1975), Measuring Real Consumption from Quantity Data, Canada 1935–1968, in *Household Production and Consumption*, Nestor E. Terleckyj (ed.), National Bureau of Economic Research.