Session Number : 1
Session Title : Measuring non-market output, efficiency and productivity
Session Chair : Dr Christian KASTROP

Paper prepared for the joint OECD/ONS/Government of Norway workshop
”Measurement of non-market output in education and health”

London, Brunei Gallery, October 3 – 5, 2006

The national accounts perspective, time and spatial comparisons : the projects of Eurostat and OECD

Paul KONIJN, Eurostat, unit C1
Alain GALLAIS, OECD, STD/NAFS

For additional information, please contact :

Author name(s) : Paul KONIJN / Alain GALLAIS
Author address(es) : Eurostat, L-2920 Luxembourg, Luxembourg ;
OCDE Tour Europe, 33 place des Corolles La Défense 2, 92400 Courbevoie, FRANCE
Author E-mail(s) : paulus.konijn@ec.europa.eu / alain.gallais@oecd.org
Author fax(es) : (352) 4301 33029 / (33) 1 45 24 98 14
Author telephone(s) : (352) 4301 33438 / (33) 1 45 24 81 01

This paper is posted on the following website :
http://www.oecd.org/document/34/0,2340,en_2649_33715_36450978_1_1_1_1,00.html

Abstract

Non-market services could be regarded as being "comparison-resistant" services, in the context of intertemporal as well as interspatial GDP volume comparisons. For intertemporal comparisons, measuring the growth in the output and productivity of each country, the SNA 93 and the ESA 95 recommends the use of output based measures. EU Member States have a deadline of 2006 for the implementation of the Eurostat Handbook on Price and Volume Measures, implying the use of output measures for health and education services. Other OECD countries are less advanced. Overall there is necessity to exchange experience to ensure international comparability of growth and productivity measures.

But inter-country comparisons also greatly matters, especially for international organisations, in particular inter-country comparisons of the volume of health and education expenditures. Thus, Eurostat and OECD are now also investigating the use of output measures for the Eurostat/OECD Purchasing Power Parities (PPP) program, which estimates on an annual basis spatial price differentials.

The share of non-market services in GDP is significant, but the sources and methods applied currently in the PPP program do not bring satisfactory results. As for temporal comparisons, the problem in cross-country comparisons is related to capturing productivity differences. The current input price method (based on salaries of civil servants) applied in the PPP program assumes as practical compromise – however unrealistically - equal productivity levels of civil servants across countries. Eurostat and OECD have set up a Task Force – composed of NA and PPP experts - that is looking for improvements either from using output measures for interspatial comparisons or from improving the input method.

More generally , the OECD Committee on Statistics supported the launch of a "non-market" project based on the mobilisation of the specialised expert networks on Health and Education, in association with national accounts and price experts. In the temporal dimension, the concrete output of the project should be an OECD Statistical Manual to be published in end 2007, bringing more detailed international guidelines for the development of volume measures of non-market output, in particular for Education and Health, and reviewing the best practices. In the spatial dimension, the project consists of a proposal for an output approach for the above mentioned NA-PPP Task Force.

The complex nature of price level comparisons for health services requires, however, that such a project be carried out in close collaboration between health statisticians and health experts on the one hand and price statisticians and national accountants on the other hand. For the years 2007-08, and in cooperation with Eurostat, the OECD plans therefore to set up a Task Force, specifically devoted to health PPPs, whose mandate would cover both market and non-market production of health goods and services and who would take over from the current Eurostat/OECD Task Force on PPPs for non-market services, as regards the investigations on sources and methods for health, as soon as the new Task Force is created.

1. Introduction

Eurostat and OECD have in the recent past collaborated at many occasions in the work on the improvement of volume measures for non-market services (as well as on the price and volume measurement of other goods and services). This started in various Eurostat Task Forces on health, education and government services between 1998 and 2000 of which the results were included in the Eurostat *Handbook on Price and Volume Measures in National Accounts*, published in 2001. Currently, Eurostat and OECD are co-operating in a Task Force on spatial comparisons of non-market services, while the OECD in parallel is running a project that aims at providing more practical guidelines to countries for the measurement of health and education services. This paper gives some background on each of these developments.

In both intertemporal and interspatial comparisons, the volume of non-market services have traditionally been measured by input methods, i.e. methods that assume that changes in the volume of outputs can be estimated by looking at changes in the volume of inputs. The reasons for this practice were difficulties in determining what really the outputs of these services are, the non-availability of data on output of non-market services and difficulties in assessing differences in quality of the services over time and across countries.

These input methods were recognized to have serious drawbacks, the most important being that the basic assumption that outputs follow inputs does not have to hold: there may be changes in productivity over time and across countries. The share of non-market services in GDP being significant, this also has an impact on GDP volume measures.

The above mentioned Handbook on Price and Volume Measures in National Accounts concluded that there is no real alternative than to attempt to measure the output of at least health and education directly. This conclusion is in line with the recommendations of SNA93 and ESA95.

The Handbook is implemented by a Decision of the European Commission from 2002 (2002/990/EC). This Decision puts a deadline of end 2006 for the elimination of input methods for health and education services.

Eurostat and OECD are now also investigating the use of output measures for the Eurostat/OECD Purchasing Power Parity (PPP) program. To that end, a Task Force has been set up, composed of national accounts and PPP experts. The Task Force will look at possible improvements of the input price method as well as the feasibility of using direct output measurement for health and education. It should come up with recommendations by mid 2007.

2. Developments in national accounts

For GDP volume growth measures, EU Member States are introducing direct output methods for health and education services, in order to implement the Eurostat Handbook on Price and Volume Measures and Commission Decision 2002/990. Here is a brief summary of the recommendations of the Handbook regarding education, health and collective services.

For education services, the Handbook defines pupil/student-hours or pupil/student numbers, as the output *quantity* measure. These indicators should be measured at the most detailed level of types of education possible (using the costs per type of education as weights) and suitably adjusted for quality. The Handbook mentions three ways of finding information on *quality*:

– using outcome-based measures, such as examination results
– using direct quality information, for example from school inspections
– using indicators on the quality of inputs, for example pupil/teacher ratios. In this case, the Handbook says this should be based on a sound analysis of the relationship between the pupil/teacher ratio and the quality of classroom teaching.

In each case, there is an open question about how to combine the quantity and the quality measures. The Handbook is clear however that outcome indicators should not be used as output measures directly, but only as quality indicators. The problem with using e.g. examination results as quality indicator is to estimate which part of a change in exam results is due to changes in the quality of the teaching, and which part is due to other effects. This corresponds to what the Atkinson report calls "the attributable incremental contribution of the service to the outcome" (Atkinson (2005), "principle B").

According to provisional results from a recent questionnaire organised by Eurostat, of the "old" 15 Member States, 8 countries have already implemented output measures for education and 3 will follow by the end of 2006. Denmark has asked for a derogation until 2012. Other countries, including the new Member States[1], encounter some delays but are working on it. Detailed results of this questionnaire will be presented separately to this workshop during session 5.

For health services, the Handbook defines the number of complete treatments as the output *quantity* measure. This should be defined at a very high level of detail (cost-weighted) and of course suitably adjusted for quality. In practice, complete treatments are defined per type of health service, i.e. hospital services, general practitioners etc. For hospital services, numbers of treatments can be derived from DRG (Diagnosis Related Groups) systems, which are increasingly used in the countries. In this system, hospital stays are classified into medically meaningful groups that are as homogeneous as possible with regard to resource use. Concerning quality adjustments, the Handbook notes that

[1] This deadline of 2006 was decided in the 2002 Commission Decision, i.e. before the accession of 10 new Member States in May 2004. It is therefore not surprising that these countries in particular have difficulties meeting this deadline.

very detailed DRGs can capture changes in the treatment mix (for example new treatments) but changes in individual treatments remain difficult to measure. Like for education, the challenge is to separate outcome from output, or rather to determine which share of changes in outcome is due to changes in the quality of the output.

From the above-mentioned questionnaire, it can be seen that countries find it somewhat more difficult to implement output measures for health than for education. Of the "old" 15 Member States, 8 countries also have already implemented output measures for health and 1 will follow by the end of 2006. Denmark has asked for a derogation until 2012; other countries, including the new Member States, encounter some delays but are working on it. Detailed results of this questionnaire will be presented during session 2.

The Handbook allows the continued use of input methods for collective services, as it has not (yet?) been possible to define appropriate output measures (save a few exceptions such as social security offices (see ONS (2006) for an example of work in progress) and tax offices).

3. Developments in the Eurostat/OECD PPP program

The Eurostat/OECD PPP program currently approaches the government activity measurement from the input side[2]. Final expenditure of government on non-market services is broken down into input components as compensation of employees, consumption of fixed capital, intermediate consumption and taxes minus subsidies from which, in order to move from government output to final expenditure, the market sales as a lump sum are deducted. The national accounts (NA) values of each of these components are deflated with separate PPPs.

In fact, only compensation of employees is covered by a specific survey, called the "PPP survey of compensation of government employees". For other parts of government expenditure there is no specific data collection and proxy PPPs are applied, assuming that these proxies show the same price differences as when actual prices would have been collected. Those PPPs are "borrowed" from other parts of the comparison. For example, for intermediate consumption the PPPs of a large range of consumer goods and services have been aggregated in order to establish a proxy PPP for this component.

The above-mentioned survey collects the "price for labour input". Countries are required to report the compensation of employees for a selection of occupations. For each occupation, representative wages and social security contributions are to be provided for a "model" civil servant with a specific grade and family situation. The intention is to represent the various education and skill levels that are commonly to be found among government employees. Forty-six occupations are included in the current selection: forty civilian occupations and six military occupations. The civilian occupations are defined using job descriptions taken from ISCO-88. These descriptions specify the occupations in

[2] See the PPP Methodological Manual (Eurostat and OECD (2005)) for a full description of the methodology applied for non-market services.

terms of the kind of work done. The military occupations are specified as NATO ranks and their country equivalents.

It is worth mentioning that the current round of the International Comparison Program (ICP) uses the same method for government services as the Eurostat/OECD comparison. Back in the early 70's, in the so-called "Phases I and II" of the ICP, physical input measures like numbers of doctors, nurses and teachers were used. During Phase III, this was refined by making adjustments for capital inputs (as proxy for productivity) and the use of pupils in education. See Kravis, Heston and Summers (1982), chapter 5, for more details.

As mentioned above, the input price approach does not take into account differences in productivity in different countries. It assumes that non-market producers are equally efficient and that the same level of input will yield the same volume of output regardless of the country in which the non-market producer is operating. This assumption might be tenable when countries are at similar levels of income - as were most of the nine participating countries in the early 1970s when the input-price approach was first used by Eurostat. But when income levels vary to the extent they do among countries currently participating in Eurostat-OECD comparisons, the assumption is difficult to defend and the anomalies it gives rise to have at least to be recognised.

The problems outlined here with the methods currently employed to establish cross-country comparisons of non-market services leave no doubt that much can be improved. Work is well underway in the above-mentioned joint national accounts-PPP Task Force, following a two-stage strategy.

In the short term, the existing input methods will be improved and maybe alternative input methods can be proposed. The aim is to finalise the work on the input based approach by end-September 2006 and report the proposals, together with the information on the progress made on output based methods at the November 2006 Eurostat PPP Working Group meeting. The Task Force work on output based methods will continue after that in 2007.

The OECD is currently setting up a new Task Force that will investigate (market and non-market) health PPPs (see section 6 below). This Task Force should bring in experts on System of Health Accounts and on health prices. This new Task Force will continue the work undertaken by the current NA-PPP Task Force on health.

The proposals for improving the input based method go into two directions. Firstly, a simple revision to the current list of occupations is foreseen in order to remove some obsolete or less comparable occupations. Secondly, the comparability of the definitions and the consistency of the data with national accounts will be improved. This second direction includes bringing the rules and definitions more in line with the available data sources in the Member States.

In the longer term, hopefully genuinely output based methods can be developed and implemented. For education, the most obvious output *quantity* measure is a measure of pupil or student numbers broken down by type of education, weighted with the costs per pupil for each type of education. This kind of data is widely available in international databases, e.g. of Eurostat and the OECD (e.g. OECD (2005)). Using this information has the additional advantage that no additional data collection by individual countries is required.

The more difficult issue to deal with is to find international measures of the *quality* of education. For international comparisons it is imperative to make adjustments for quality, as it is clear that pupil or student numbers alone will not provide comparable measures. The differences in quality of education across countries are likely to be a lot higher than the differences in quality of education in one country across time.

Fortunately, there exist international studies that compare pupil attainment across countries. The most well-known is probably PISA (Program for International Student Assessment), steered by the OECD (see http://www.oecd.org/department/0,2688,en_2649_35845621_1_1_1_1_1,00.html for more information). PISA tests 15-year-olds for their knowledge and skills in the domains of reading, mathematics and science. It has been running for several years now, so that it becomes possible to better assess the reliability of its outcomes. Its results may shed light on the quality of (primary and) secondary education in various countries, but of course not on the quality of higher education.

The same crucial issues as mentioned above in the context of intertemporal measures have to be faced here. First, it has to be determined whether available measures of student performance such as PISA are good indicators of the differences in education outcome. Second, one has to find out which part of differences in outcome of education (as for example measured by PISA) is attributable to the services provided. Third, one has to find a way of combining the quantity and quality indicators. What weight should each of the two get? The Task Force will continue to investigate these matters both conceptually and empirically.

For health services, the situation is arguably more complicated. First of all, health services consist of a more heterogeneous set of services than education. It is composed of services of hospitals, general practitioners, specialists, dentists, paramedics, other kinds of institutions, etc. Each of these services has its own characteristics that determine the definition of its output, whereas for education the definition of output is the same for all types of education.

For hospital services, probably the largest component in most countries, the DRG system is mentioned above as a possible source of output *quantity* data for growth measures in the national accounts. A problem for cross-country comparisons is that the DRG systems in place in individual countries are not very well comparable, even if they are often derived from the same (US) system. Thus, it will be difficult to compare numbers of complete hospital treatments across countries. Perhaps it will be possible to construct certain higher aggregates of treatments that can be compared; this has to be investigated.

For general practitioners, medical specialists and dentists the output measures agreed for national accounts are based on numbers of (first or total) consultations. This kind of information is collected in many countries through household surveys or sometimes through administrative data. It has to be investigated if this information can reliably be used for cross-country comparisons. For certain standard or 'simple' medical services deflation based on price surveys with clear specifications remains an option.

As regards *quality* adjustments, there are no international quality comparisons for health like we have for education. There are some *outcome* indicators, like life expectancy. But with this kind of indicators it will be very difficult to determine the contribution of the health services to changes in the indicator.

A lot more research has to go into the available data and the conceptual issues related to the use of them.

For collective services, in the national accounts input methods continue to be used, due to the difficulties in defining measurable output. The same will be true in the PPP context. Thus, collective services will remain based on input price data from a salary survey.

4. Similar market services

One important aspect that has to be considered is the consistency between the methods applied for market services and for non-market services. This is important because the distribution between market and non-market is different across countries. Currently, generally speaking, for market services output prices are collected, while for non-market services input prices are used. If the baskets of services under market and under non-market are not the same in different countries, the results will not be comparable

Using output measures for non-market services will resolve this issue, at least conceptually. Suppose education was provided fully on a market basis. Each pupil would pay a market price for receiving teaching. Total expenditure on education would simply be the number of pupils times the average price paid. Deflating this expenditure with a market price index will result in a volume index that reflects the number of pupils. Hence, the implicit volume index for market education services is the same as the proposed volume index for non-market services. Therefore, PPP could compare education expenditures and volumes across countries without regard to the market / non market status of the provider provided that expenditure and volume data are comparable.

In the PPP program, until now, for market health and education services deflators for non-market services are used as proxy PPPs. Eurostat and OECD are trying to improve on this for health by collecting actual prices for privately produced medical services. A selection of services have been defined (e.g. a consult of a general practitioner, a consult of a cardiologist (with some accompanying services), a tooth extraction by a dentist, laboratory tests, etc.) and countries have attempted to collect prices for these services at the end of 2005. Eurostat and OECD are still in the process of validating the results, so we cannot elaborate on these data at this stage.

The biggest challenge for countries in this price collection is to find the full market price, i.e. the total price the producer finally receives either from the patient (or his/her private health insurance) and/or from the state. The full market price is needed to be able to properly calculate the volume of the services produced. Both household expenditure and government expenditure need to be deflated by the same full price, otherwise the volume would be double counted.

The PPP Manual (chapter 4, footnote 21) explains this with the following example:

Suppose that the quantity of a pharmaceutical product purchased is 1000 units and that the price per unit is 10 € of which households pay 2 € and government 8 €. In the national accounts, 2000 € will be recorded as household expenditure and 8000 € will be recorded as government expenditure. If the amounts actually paid – that is, 2 € by households and 8 € by government – are used to deflate these expenditures, it will seem that both households and government have each purchased 1000 units or 2000 units in total. But if the total amount paid – that is, 10 € – is used, households will appear to have purchased 200 units and government 800 units - a total of 1000 units.

From this example, it is clear that the volume of the total expenditure – in other words, the *actual final consumption* – is the most important indicator. The volumes for household expenditure (200) and government expenditure (800) are a fictitious, conventional distribution of the total volume. The example shows that the best indicator to use for cross-country comparisons of volumes of e.g. health and education is the actual final consumption. Household and government expenditure volume data are more difficult to compare across countries as they are affected by actual reimbursement practices of governments and by the share of market and non-market output.

The same is true in the intertemporal context of national accounts: the volume changes of household and government expenditure on health or education are rather meaningless indicators, as they are affected by changes in the shares households and governments pay in the total costs of the services consumed.

5. **OECD "non market" project**

In this context of national accounts reflections and European experimentations, the basic idea of the OECD "non market" project is to extend these researches to other countries, of course, but above all to other statisticians communities: the specialized experts networks of education, health and government productivity in particular. If national accounts are no longer to rely on a conventional zero productivity growth, and are to define a "real" output of non market services, i.e.: the real contribution of these services to important and sensitive outcomes, this matter should involve other statisticians and economists (even politicians, citizens?), reach if possible a theoretical consensus and build rather simple formulas with available data.

Such a proposal was approved by the OECD Committee on Statistics in 2005 (Agenda Item 5 – Productivity Measures for the Business and Non-market Sectors and International Comparisons of Productivity Levels), which consisted initially to organize in 2006 a seminar on the measurement of output of education and health.

As INSEE has accepted to assign on this project a full time expert (Alain Gallais) during 18 months (April 2006 -> September 2007), two “workshops” will be managed in fact, the first one in October 2006 in London (this workshop) and the second one in June 2007 in Paris. London was obviously chosen because of the deep impact of the Atkinson review and following UKCeMGA reports, that make the United Kingdom leading the work in this area. They propose several schemes/formulas in fact, still to be decided within the UK and even within the ONS. The ONS and the Government of Norway had the kindness to co-organize this meeting with the OECD. Paris will be a more classical place for an OECD meeting: this event should be held in La Muette.

The coverage of the project is both temporal and spatial comparisons. In the temporal dimension, the concrete output of the project should be an OECD Statistical Manual to be published by end 2007, bringing more detailed international guidelines for the development of volume measures of non-market output, in particular for education and health, and reviewing best practices. In the spatial dimension, the project consists of a proposal for an output approach, within the Eurostat NA-PPP Task Force on non-market services and the future OECD Health Task Force.

6. **OECD Health-specific Purchasing Power Parities (PPPs) project**

As a natural continuation of all these explorations there will be an OECD project on Health-specific PPPs. The aim of the project is to develop output-based PPPs for health goods and services. These data should fulfill two purposes: (i) provide a tool for the analysis of the volume of health expenditure in OECD and EU countries. This objective is complementary to other data projects in the field of health accounts with a view to improving the information base for health policy makers; (ii) provide input to the broader purpose of deriving economy-wide PPPs for international comparisons of volume GDP. The project is envisaged to be partly financed by an EU contribution to OECD work on health, and implemented in close co-operation with Eurostat.

The OECD would set up a PPP-SHA Task Force specifically devoted to health PPPs. The group’s mandate would cover both market and non-market production of health goods and services. The Task Force will take over from the current Eurostat/OECD Task Force on PPPs for non-market services, as regards the investigations on sources and methods for health, as soon as the Task Force is created. An important feature of the health PPP Task Force would be the joint membership of health specialists, price statisticians and national accountants. This seems indispensable in light of the technical specificity of the subject. This Task Force would work during 2007-2008 with a view to generating input for the next PPP survey on health prices in 2008.

There would be three components to this project:

- Development of concepts and methodology (outcome/output, quality, unit of measurement);
- Assessment of data availability and feasibility of data development;
- Formulation of joint OECD/Eurostat guidelines for data collection for the purpose of health PPPs.

References

T. Atkinson (2005) *Atkinson Review Final report - Measurement of Government Output and Productivity for the National Accounts*, Palgrave Macmillan, published on http://www.statistics.gov.uk/about/data/methodology/specific/PublicSector/Atkinson/final_report.asp

European Commission, IMF, OECD, UN and World Bank *System of National Accounts 1993*, published on http://unstats.un.org/unsd/sna1993/introduction.asp

European Commission (1996) *European System of Accounts 1995 (ESA95)*, published on http://forum.europa.eu.int/irc/dsis/nfaccount/info/data/esa95/esa95-new.htm

European Commission (2002) *Commission Decision of 17 December 2002 further clarifying Annex A to Council Regulation (EC) No 2223/96 as concerns the principles for measuring prices and volumes in national account*, published on http://eur-lex.europa.eu/LexUriServ/site/en/oj/2002/l_347/l_34720021220en00420059.pdf

Eurostat (2001) *Handbook on Price and Volume Measures*, published on http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1073,46587259&_dad=portal&_schema=PORTAL&p_product_code=KS-41-01-543

Eurostat/OECD (2005) *PPP Methodological Manual*, published on http://forum.europa.eu.int/Public/irc/dsis/palojpi/library?l=/methodological_papers/ppp_manual&vm=detailed&sb=Title.

I.B. Kravis, A. Heston and R. Summers (1982) *World Product and Income – International comparisons of real gross product*, John Hopkins University Press

OECD (2005) *Education at a glance 2005*, see http://www.oecd.org/document/34/0,2340,en_2825_495609_35289570_1_1_1_1,00.html

ONS (2006) *Public Service Productivity: Social Security Administration* published on http://www.statistics.gov.uk/articles/nojournal/1596.pdf