Session Number: 3
Session Title: Measurement of government output
Session Organizer(s): David Caplan, Tim Smeeding
Session Chair: David Caplan

Paper Prepared for the 29th General Conference of
The International Association for Research in Income and Wealth

Joensuu, Finland, August 20 – 26, 2006

International Comparisons of Non-Market Services How to improve

Paul Konijn, Peeter Leetmaa and Silke Stapel, Eurostat

 For additional information please contact:

Paul Konijn
Eurostat, L-2920 Luxembourg
paulus.konijn@ec.europa.eu
Fax: +352-4301-33029
Tel: +352-4301-33438

This paper is posted on the following websites: http://www.iariw.org

1. Introduction

Non-market services are regarded as being among the "comparison-resistant" components within the Eurostat/OECD PPP (*Purchasing Power Parity*) Program, which estimates volume and price differences across countries. Their share of GDP is significant, but the sources and methods applied do not bring satisfactory results. Like for comparisons over time, 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 program assumes as practical compromise – however unrealistically - equal productivity levels of civil servants across countries.

In the national accounts of EU Member States, input methods used for measuring the volume growth of individual non-market services (in particular health and education services) are in the process of being replaced by methods that attempt to directly measure the growth in the output produced. It is therefore natural to investigate whether similar methods can be used for cross-country comparisons.

To that end, Eurostat, in co-operation with OECD, has set up a Task Force to carry out these investigations. 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.

In this paper, we describe the current approach with its weaknesses and past attempts to allow for productivity differences across countries. We'll look at the conceptual and practical issues arising from using direct output measures, without wishing to pre-empt the work of the Task Force.

The paper draws to a large extent on the Eurostat/OECD PPP Methodological Manual which can be consulted at
http://forum.europa.eu.int/Public/irc/dsis/palojpi/library?l=/methodological_papers/ppp_manual&vm=detailed&sb=Title.

2. The current method

The Eurostat/OECD PPP program currently approaches the government activity measurement from the input side. 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". There are 9 basic headings[1] for which the PPPs are calculated based on the collected government compensation data:

Health:

- Physicians
- Nurses and other medical staff
- Non-medical staff

Education:

- Pre-primary and primary education
- Secondary education
- Post-secondary non-tertiary education
- Tertiary education

Collective services:

- Compensation of employees n.e.c.
- Compensation of employees - Defence

PPPs for these basic headings are calculated using the compensation of employees that countries are required to report 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 employees working in these three government services. 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 terms of the kind of work done. The military occupations are specified as NATO ranks and their country equivalents.

The compensation of employees for the selected occupations is defined as comprising gross salaries and wages – that is, the basic salary or wage plus other cash payments over and above the basic salary or wage - employers' actual social contributions and imputed social contributions. The definition is consistent with that of the SNA93/ESA95 except for the exclusion from gross salaries and wages of overtime payments and benefits in cash and kind not related to the salary scale. These two deviations from ESA95/SNA93 have been introduced mostly for comparability reasons.

[1] The basic heading is the lowest level of the PPP expenditure classification for which a PPP is calculated. Basic headings are in principle defined as a group of similar well-defined goods or services. They are the "building blocks" out of which all aggregate PPPs are established.

The compensation of employees to be reported for the selected occupations should not be extracted from government payrolls. Instead, the compensation of employees should be derived by a less representative but more comparable approach that involves working from government salary scales directly. Many participating countries find it difficult, if not impossible, to implement this approach. These countries resort to extracting the compensation of employees from government payrolls or other statistical sources.

The compensation of employees reported for each selected occupation must be annual. The compensation of employees should also be the national average taking into account the discrepancies in compensation which may arise both between various levels of government – that is, between central, regional, state and local governments. Often there are no national salary scales available and, some approximations, e.g. a weighted average, should be used.

Countries are required to report for each of the selected occupations the compensation of employees, the standard number of hours worked per week, the holiday entitlement and number of public holidays falling on working days during the reference year. The norm applied is 1710 hours – that is, 52 working weeks, each of 38 hours duration, less 7 working weeks (or 35 working days) of paid leave and public holidays. The data on standard working hours and holiday entitlement are used to standardise the compensation of employees across countries.

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.

3. Similar market services

The survey on compensation of employees is used in the first place to deflate the part of government final consumption expenditure that relates to the government-produced (i.e. non-market) services. For the same services, there is usually also a share of market production, paid for by households and/or the government (as social transfers in kind). Until now, for these services the deflator for non-market services is used as proxy PPP.

Eurostat and OECD are trying to improve on this by collecting actual prices for 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 needs 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 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.

4. Failure to account for the productivity differences between the countries

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.

Differences in productivity are being disguised as price differences. Hence, the volumes of output of non-market producers in countries whose cost of inputs are relatively low are being overestimated and the volumes of output of non-market producers in countries whose cost of inputs are relatively high are being underestimated. The implicit assumption in this statement is that low-cost countries are producing less efficiently or with a lower quality output than high-cost countries. This seems a plausible assumption but there is no evidence for it as long as outputs are not actually measured.

To see if we can show this effect in reality, table 1 gives volume indices per capita for 2004 for a number of country groupings. For confidentiality reasons we cannot show individual countries' data.

Table 1: Volume indices per capita EU25 = 100

2004	GDP	Actual individual consumption			Collective consumption
		Total	Health	Education	
EU25	100	100	100	100	100
Old MS	109	109	107	101	104
New MS	55	55	63	96	81
Countries with GDP per capita volume index > 125					
average	155	122	119	132	122
Countries with GDP per capita volume index > 100, < 125					
average	116	112	112	121	109
Countries with GDP per capita volume index > 75, < 100					
average	86	85	77	101	94
Countries with GDP per capita volume index > 50, < 75					
average	63	62	68	99	96
Countries with GDP per capita volume index < 50					
average	38	41	39	83	58

Source: Eurostat

The countries are divided into 5 groups with different income levels. The first column gives the (unweighted) average volume indices of GDP for the countries within each group. The following three columns show the average volume indices for total actual consumption of individual services and actual consumption of health and education services respectively. The last column gives the volume indices for the collective consumption of government.

It can be seen that the volume indices for total actual consumption are very close to those for GDP, with the exception of the richest group of countries where the GDP index is – for other reasons - very high in one specific country. For health, the volume indices are also rather close, but for education we notice that the poorer countries have relatively high volume indices (compare for example old and new Member States). The same is true, but to a lesser extent, for collective consumption.

Table 2 shows the corresponding relative price level indices (which are to a large extent determined by the salary levels of government employees). Here we note that the price level indices for health, education and collective services have a wider spread than the indices for GDP and total actual consumption. In high income countries, the difference

between the price level and the EU25 price level is larger than for the overall price level, and the opposite is true for low income countries. The effect is strongest for education. Indeed, the costs of inputs (mainly salaries of teachers etc.) in the poorest countries are at a level of about a quarter of the average EU25 level.

Table 2: Relative price level indices **EU25 = 100**

2004	GDP	Actual individual consumption			Collective consumption
		Total	Health	Education	
EU25	100	100	100	100	100
Old MS	105	105	107	112	109
New MS	53	51	40	33	42
Countries with GDP per capita volume index > 125					
average	119	123	126	145	131
Countries with GDP per capita volume index > 100, < 125					
average	112	115	119	124	116
Countries with GDP per capita volume index > 75, < 100					
average	83	83	78	79	82
Countries with GDP per capita volume index > 50, < 75					
average	62	61	50	49	51
Countries with GDP per capita volume index < 50					
average	46	45	33	23	31

Source: Eurostat

The relatively high per capita volume indices for education and collective services for poorer countries seem indeed to indicate that the "productivity" effect exists, i.e. these volume indices would be lower if we could actually take productivity and quality differences into account.

5. Past attempts to allow for productivity differences

One solution would be to complement the input-price approach with adjustments for differences in productivity. This was the strategy adopted by Statistics Austria when it organised the comparisons covering eastern and central European countries for the European Comparison Programme. An adjustment was made for differences in labour productivity. It was based on the assumption that productivity differences between the non-market producers of countries were roughly equal to the productivity differences between their market producers. The market sector was defined as all economic activities except agriculture and non-market services. Productivity was defined as value added per employee. And nominal market-sector value added was converted to real market-sector

value added with the PPPs for that part of final expenditure on GDP that was generated by market activities.

Table 3 shows two sets of volume indices for GDP per capita for the thirteen transition economies that constituted "Group II" in 1996. Both sets of indices were obtained by the input-price approach. The first set are the indices before being adjusted for differences in labour productivity. The second set are the indices after being adjusted. The percent differences between the two sets are large and, if the adjusted indices are to be believed, large enough to be of concern.

Table 3: Per capita volume indices for GDP before and after productivity adjustments

countries	GDP per capita volume indices		Per cent difference
	Input prices	Productivity adjusted	
Austria	100.0	100.0	---
Slovenia	59.7	53.6	-10.2
Russia	30.4	26.1	-14.2
Estonia	30.2	24.8	-17.9
Romania	29.8	25.0	-16.1
Croatia	28.4	23.8	-16.2
Lithuania	26.0	21.4	-17.6
Belarus	23.3	19.3	-17.3
Latvia	22.8	18.9	-17.1
Bulgaria	22.6	18.9	-16.2
Macedonia	19.1	14.1	-26.2
Ukraine	15.0	11.6	-22.6
Albania	13.0	11.0	-15.4
Moldova	9.5	7.3	-23.0

Source: PPP Manual, box B in chapter 5.

EU Member States and OECD Member Countries have not been in favour of such an adjustment being employed in Eurostat-OECD comparisons. Besides considering the approach too simple for their economies, they question its underlying assumption, arguing that productivity differences between countries are likely to be smaller in the non-market sector than they are in the market sector because the productivity of non-market producers is less dependent on capital inputs than the productivity of market producers. More generally, they argue against productivity adjustments because they are inevitably based on assumptions which cannot be verified without a genuine measurement of output. And, if output could be measured, then input methods would not be necessary.

Many outputs of non-market producers of individual services are equivalent to those of market producers of individual services. An alternative to productivity adjustments would be to value these outputs of the non-market sector at the same market prices as their counterparts in the market sector and to calculate PPPs for non-market services accordingly. But even if a non-market service is identical to a market service, there is no reason to believe that the production of the service by non-market producers is as efficient as that of market producers. Generally, non-market producers are under no

constraint to maximise profits or minimise losses and it cannot be assumed that they employ inputs with a view to equating average real wages and marginal product. Their production, in other words, may be less efficient and more costly than market production and have correspondingly higher shadow prices. The cost differences between market and non-market producers are likely to be greater in some countries than another. To use the prices of market output to calculate PPPs for non-market output would also be making an assumption, albeit implicit, about the productivity of market and non-market producers.

To use the PPPs for individual consumption expenditure by households as proxies have also been proposed as an alternative to the input-price approach. Two arguments have been advanced for adopting this alternative. The first is that, as the purpose of Eurostat-OECD comparisons is primarily to compare economic welfare, it is justifiable to use the PPPs for household expenditure as proxy PPPs for government expenditure because it provides a measure of consumption forgone. The second is that the PPPs for household expenditure are better estimates of the PPPs for government expenditure than the PPPs obtained using the input-price approach.

Neither of these arguments has found support among EU Member States and OECD Member Countries. The first argument is based on the wrong premise. The purpose of Eurostat-OECD comparisons is to compare real GDP across participating countries. In so far as GDP per capita is a measure of economic welfare, its purpose is to compare economic welfare. But it is output that is being compared, not consumption or consumption forgone. The second argument is difficult to justify theoretically and there is little empirical evidence to support it. While use of a proxy PPP appears to remove some of the counter-intuitive results of the input-price approach, it produces other results that are equally counter-intuitive.

6. 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* (2001) and Commission Decision 2002/990. The latter specifies a deadline of 2006 for the removal of input methods for individual non-market services. The main argument for preferring output methods to input methods lies in the impossibility to analyse government performance or productivity when output is not directly measured. This fact was already recognized in SNA93 and ESA95, and indeed long before that (see e.g. Hill (1975)), but these two manuals still left the choice between input and output methods to the countries. It was felt that this situation did not guarantee sufficient comparability across countries, which was one of the reasons for developing the above-mentioned Handbook.

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, at least 10 countries will have implemented output measures for education by the end of 2006. Two more will follow in 2007. Of the remaining 3, the implementation date is at the time of writing this paper unknown. Of the 10 new Member States, 4 countries will have implemented by the end of 2006, one has indicated 2007, the others are yet unclear[2]. Detailed results of this questionnaire will be presented to the OECD/ONS workshop on non-market services in London in October 2006 (see http://www.oecd.org/document/34/0,2340,en_2649_33715_36450978_1_1_1_1,00.html for information).

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 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 more difficult to implement output measures for health than for education. Not all responses to the questionnaire have been received, but there will be fewer countries ready for health by 2006 than for education.

[2] 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.

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).

7. Potential direct output measures for spatial comparisons on education and health

The question naturally arises whether similar direct output measures can also be used in the PPP context. As mentioned in the introduction, a special Task Force has been set up by Eurostat in co-operation with OECD to investigate the possibilities.

The Task Force has met twice so far. As a starting point, education services were considered as this area is considered to be somewhat less complicated compared to measurement of health services. The OECD, in the framework of a project on government output measurement, is so far the most active participant in developing proposals for direct output measures.

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 two crucial issues as mentioned above in the context of intertemporal measures have to be faced here. First, it has to be determined which part of differences in outcome of education (as for example measured by PISA) is attributable to the services provided. Secondly, 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.

One important aspect that has to be considered when deciding on new methods 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.

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 consultations. This kind of information is collected in many countries through household surveys. It has to be investigated if this information can reliably be used for cross-country comparisons.

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 to be based on input price data from a salary survey.

8. Way ahead

The problems outlined in this article with the methods currently employed to establish cross-country comparisons of non-market services leave no doubt that much can be improved. As mentioned, work is well underway, 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 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. From information collected by the Task Force it became clear that countries are increasingly using payroll information and statistical sources, with which it is more difficult to follow the existing guidelines.

In the longer term, hopefully genuinely output based methods can be developed and implemented. The respective Task Force work will develop the first ideas outlined under point 7 above, both conceptually and by trying them out with real numbers. It is clear though that it will be a challenging task.

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.

T.P. Hill (1975) *Price and volume measures for non-market services*, Statistical Office of the European Communities, OS/5/75-E

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