MEASURING HEALTH SERVICES IN THE NATIONAL ACCOUNTS: AN INTERNATIONAL PERSPECTIVE

Paul Schreyer* (OECD) and Matilde Mas (IVIE and University of Valencia)

Draft 3rd October 2013

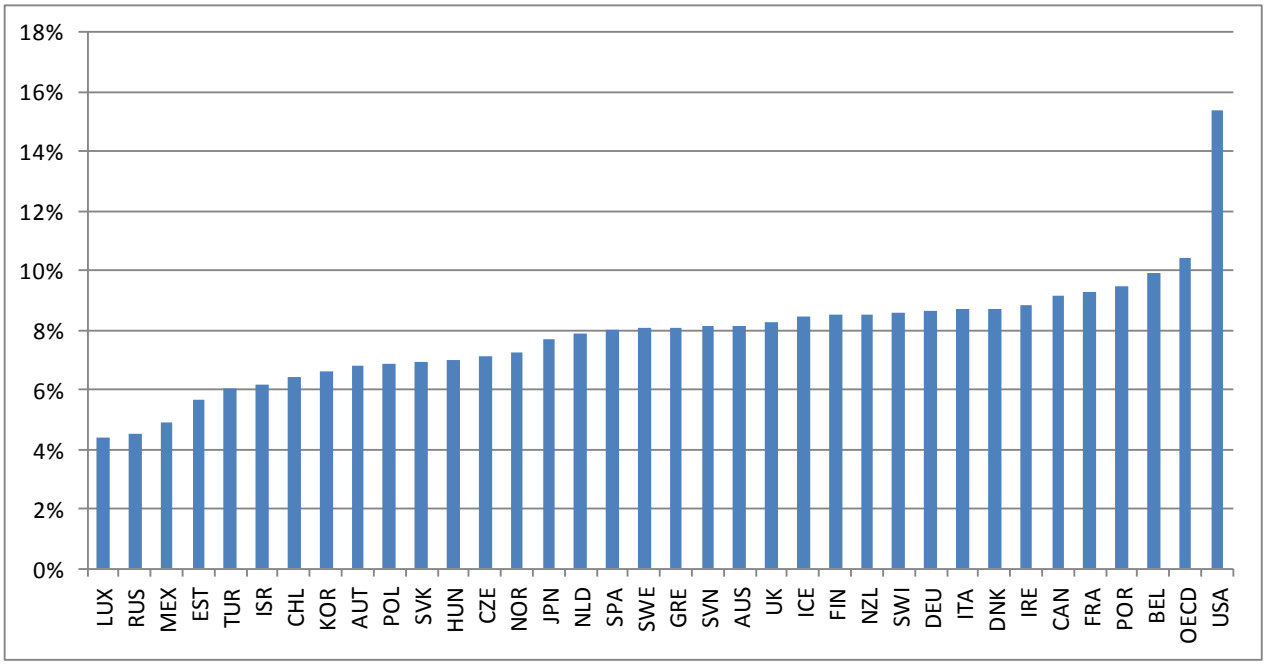
Paper prepared for the NBER/CRIW Conference on Evaluating Health Care Spending: Progress and Gaps in the Health Care Statistical Infrastructure on October 18 and 19, 2013, Washington DC.

*This paper reflects the view of the authors and not necessarily those of the OECD or its Member countries.

1. Introduction

In 2011, domestic demand for health services accounted for an average of 11% of GDP in OECD countries, as an item of household demand second only to housing. At the same time, variations between countries are significant, ranging from a modest 4% in Luxembourg to a sizable 15% in the United States. Such differences within a fairly homogenous set of countries immediately raise a number of questions: are we comparing like with like? And if so, are differences in the value of health services due to differences in prices or to differences in the volume of health services provided? A similar question arises when comparing the evolution of health expenditure within a country over time: how much of an increase in expenditure has occurred because of more services delivered and how much has occurred because of services having become more expensive? This paper aims at exploring the issue of measuring health services and the break-down of expenditures between prices and volumes from an international perspective. It will ask whether health services are defined in the same way across countries and whether statistical offices apply similar methods to undertake a price-volume split when nominal expenditures are tracked over time. The paper will also present new inter-country comparisons of the volume of health services consumed, based on an approach recently put in place by the OECD and Eurostat.

Figure 1 Domestic health expenditure as a percentage of GDP, 2011, current prices



Source: derived from *OECD Annual National Accounts* 2013.

Figure 1 is more complex to construct than meets the eye. Indeed, its construction reflects a number of measurement issues that are specific to health services. The first specificity is that unlike, say a haircut, health services are not necessarily the object of transactions between two parties. Most countries' health systems operate under a private or public insurance system and the price for the service is often negotiated between the insurer and the health care provider rather than between the patient and the health care provider. Payments or reimbursements by health insurers are counted as consumer expenditure in the national accounts so require an imputation. A second specificity is that government may provide health services directly to individuals with only a nominal fee or no fee involved at all. Such social transfers in kind do not figure among consumer expenditures. International comparisons of health expenditure are thus best based on a measure of individual health services that sums up expenditure by patients and the value of

the in-kind services provided by government. Such in-kind services need to be identified and valued. Figure 1 reflects such a valuation and shows total health expenditures whether incurred by patients (or their insurance companies) or whether provided by government. The third specificity is that health care providing units[1] are more often non-market producers than in other industries. This distinction entails a different accounting treatment, at least in the way the value of health services at current prices is measured: whereas the value of sales constitutes output for market producers, the value of output for non-market producers is measured as the sum of production costs[2]. The distinction between market and non-market producers is also important from the perspective of assessing efficiency in the provision of health services: market and non-market producers may take their decision on the quantities (and prices charged) following different objective functions. Differences in health care productivity performance may be associated with the share of non-market versus market producers and provide useful insights from international comparisons. Finally, the measurement of the volume of health services (as opposed to health expenditure) is tricky: rapid progress in medical technology, and complex services bring out many of the measurement challenges that statisticians face when developing price indices and volume measures in the national accounts.

The discussion about the measurement of health and education services is by no means new. Nearly forty years ago, Peter Hill (1975) developed a set of principles and guidance for measuring health, education and collective government services. More recently, the debate has resurfaced. Eurostat (2001) stated the desirability of applying output-based measures to non-market services. In the United Kingdom, the topic was taken up by the widely-discussed Atkinson Review (2005). The measurement of services output and productivity has also been a longstanding topic of interest in the United States, with a series of publications including Triplett (2001), Cutler and Berndt (2001), Triplett and Bosworth (2004) and Abraham and Mackie (2006). Health services in particular have been the subject of research on cost-effectiveness and productivity (Cutler, Rosen and Vijan 2006 or Rosen and Cutler 2007). A recent overview of concepts and quality adjustments of measures of health and education services can be found in Schreyer (2012) and Schreyer (2010).

This paper will only provide partial answers to these issues. Its aim is to provide an international perspective on the measurement of health care in the national accounts. Section 2 takes a look at the international accounting conventions for health services, as spelled out in the 2008 System of National Accounts (2008 SNA). Section 3 reviews relevant national accounts practices in a broad selection of OECD countries. Section 4 turns from inter-temporal to inter-spatial comparisons and reports on recent efforts by the OECD to construct internationally comparable measures of the price levels and volumes of health care services. Section 5 concludes by summing up the key measurement tasks ahead.

2. What the SNA has to say about measuring health services

Current price measures

The national accountant's task of measuring production begins with identifying the units that produce health services and distinguishing between market and non-market producers. Market producers sell their output at prices that are economically significant. Thus, for market health services, the value of output in current prices can be measured by the value of sales of these services. However, health provision is among the most common examples of services provided by government free of charge or at prices which are not

[1] Statistical information on health providers can be found in Section Q, Division 86 of the *International Standard Industrial Classification of all Economic Activities* (ISIC) Rev. 4, which includes Hospital Services; Medical and dental practices; and other human health services providers.

[2] As will be discussed below, the costs recognised by the SNA are incomplete as only depreciation is recognised as capital costs.

economically significant and thus constitute non-market output. A price which is not economically significant is deliberately fixed well below the equilibrium price that would clear the market. The SNA defines it as a price which has little or no influence over how much the producer is willing to supply and which has only a marginal influence on the quantities demanded.

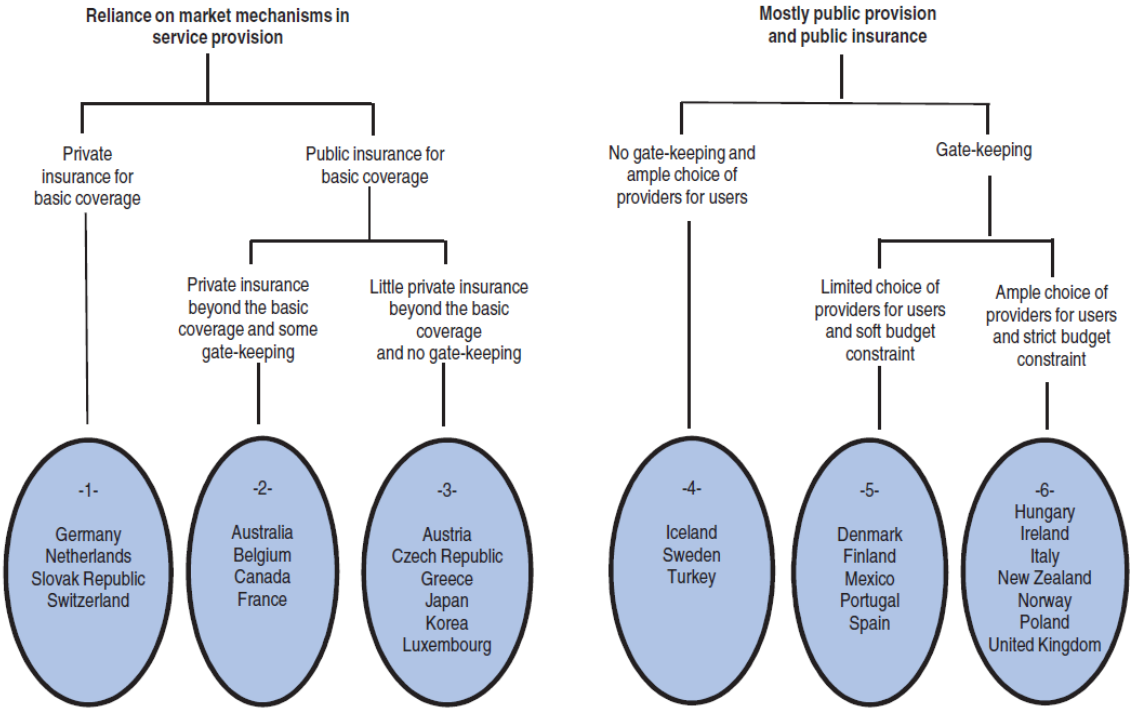
There are differences in country practices to identify the economic significance of prices. For instance, the European System of Accounts (ESA 1995) considers, for practical reasons, that a price is not economically significant if it covers less than half of the costs of producing the service. Neither the 2008 SNA nor its predecessor, the 1993 SNA have specified a particular level of cost coverage which complicates international comparisons of market and non-market provision. Whatever the exact rule, valuation of output is based on adding the costs incurred in production; namely the sum of:

- Intermediate consumption (the goods and services used up in producing the service)
- Compensation of employees (costs of doctors, nurses, etc…)
- Consumption of fixed capital[1] (depreciation of hospital buildings, of medical equipment etc.)
- Other taxes, less subsidies, on production.

Note that, according to the 2008 SNA, capital costs for non-market producers are solely measured as the value of depreciation, thus ignoring that part of costs of capital services that reflect the opportunity costs of capital and revaluation. The main reason for this convention lies in the fact that any such imputation directly affects GDP and national income and that there is a broad spectrum of possible imputations. That said, Jorgenson and Landefeld (2006), Jorgenson and Yun (2001) and OECD (2009) show alternatives for dealing with this complication. From the perspective of productivity measurement, the asymmetric treatment of assets used in market and in non-market production results in an incomplete estimate of capital inputs and in an asymmetric treatment of the same asset, depending on the sector affiliation of the asset owner (Jorgenson and Schreyer 2013). For analytical applications it may therefore be considered useful to deviate from the national accounts convention. An example for such an application is Mas, Pérez and Uriel (2006) who examine the contribution of infrastructure capital, largely held by government entities, to economic growth in Spain and who apply a complete user cost expression to public capital. We conclude that a breakdown between market and non-market production in the publication of national accounts data would be of significant interest to analysts.

A further complication arises in health provision measurement due to the existence of insurance schemes of different scope and variations. Unlike other services that are directly transacted between the supplier and the consumer, health service transactions often occur between three parties: health service supplier, the consumer and public or private insurance schemes. The consequence is that transacted payments between the supplier and the consumer are not necessarily indicative of the price of the health service. Institutions vary greatly between countries as shown in Figure 2. Any international comparison of health care expenditures, say in proportion to GDP needs therefore to be based on measures reflecting full costs in health care provision, whether they accrue to patients, private providers or government. This is indeed the approach pursued by the OECD-Eurostat Programme on Purchasing Power Parities (Koechlin, Lorenzoni and Schreyer 2010) where the value of actual individual consumption of health care is deflated with international price indices to arrive at volume comparisons of per capita consumption of health services between countries.

Figure 2 Institutions in health care provision in OECD countries



Source: Joumart, Höller, André and Nicq (2010).

Volumes

Market and non-market producers. The current value of health services, if provided by non-market producers, is always valued at cost in the national accounts. Thus, the value of inputs equals the value of outputs. At the same time, this does not mean that the volume of outputs cannot be distinguished from the inputs used to produce it. Changes in productivity may occur in all fields of production, including the production of non-market services[3]. Volume measurement is thus inherently different from the measurement of values, also in the case of non-market producers. However, volume measurement of the services provided by non-market producers is not inherently different from volume measurement of the services provided by market producers. This was first pointed out by Hill (1975):

> "It is proposed as a matter of principle that the basic methodology used to measure changes in the volume of real output should always be the same irrespective of whether a service is provided on a market or on a non-market basis. This is not to say that the actual numerical measures would not be affected by whether the service is market or non-market, because different weighting systems would be involved, but at least the methods of measurement should be conceptually similar" (page 19).

[3] See 2008 SNA, Paragraph 15.116.

Schreyer (2010) confirms this principle but points out that in practice, there has been a tendency to create separate volume indices for market and non-market production[4]. Traditionally, volume output measures for non-market producers have been based on volume measures of inputs with the implication of assuming zero productivity change and the risk of inadequately capturing changes in living standards and macro-economic productivity. A number of possibilities exist for deriving output-based volume measures of health services.

In a market-based health system where there is information on market prices, expenditure on the treatment of a disease can be deflated by a disease-specific price index to arrive at a volume output measure of the disease. For example, Berndt et al (2000) have estimated a price index for heart attacks and this index can be used to deflate disease-specific expenditures. This is similar to what happens in other market sectors in the economy where volume output measurement is accomplished by dividing data on revenues or sales by a price index.

In some countries, hospitals and other providers of medical services are considered market producers because they receive economically significant revenues from reimbursement schemes that, on average, cover their costs. In such cases, a 'quasi price' index consists of average revenues per treatment. One notes, however, that reimbursement schemes are themselves based on cost so that the differentiation between costs and revenues is blurred. Also, the fact that there are revenues does not imply that there is a competitive market where prices necessarily carry signals about consumer preferences.

In some instances, it may also be possible to draw on market price information for purposes of deflating values of non-market production. A potential candidate is the medical services part of the Consumer Price Index. However, care has to be exerted to make sure that the CPI is representative for the deflation of the non-market production. In particular, (i) the services supplied by the market provider have to be sufficiently similar to those supplied by the non-market provider; (ii) the scope of the CPI has to match the scope of non-market production. This may not be the case when the CPI is designed to reflect prices for out-of-pocket expenditures and when consumers only pay part of the full price for the medical good or service. In this case, the CPI is not an appropriate tool for deflation of non-market production which relies on a concept of measuring production at its full cost.

Alternatively, direct volume indices can be constructed. A direct volume index is the weighted average of the volume indices of different types of treatments, where the cost share of each type of treatment constitutes the weight. Berndt et al. (p.173) suggest that "real output of medical care could be formed from cost of disease accounts by counting quantities of medical procedures (the number of heart bypass operations, say, or of appendectomies, or of influenza shots), and weighing each procedure by its cost." Although there are some differences between a direct volume index and a volume index derived at by deflation (such as index number formulae, timeliness of data), the basic idea remains the same – volume measures of outputs are sought, as opposed to volume measures of inputs.

Outputs and Outcomes. A key distinction in this context is between inputs, outputs and outcomes. The 2008 SNA makes this distinction as follows:

> "Taking health services as an example, input is defined as the labour input of medical and non-medical staff, the drugs, the electricity and other inputs purchased […] These resources are used in the activity of primary care and in hospital activities, such as a general practitioner making an

[4] Perhaps slightly confusing, the 2008 SNA recommends a 'volume output method' for volume measurement of health services (Paragraph 15.118) but anchors this recommendation in a discussion on non-market output. This may create the impression that the volume output method is specific to non-market producers which it is not.

examination, the carrying out of a heart operation and other activities designed to benefit the individual patient. The benefits to the patient constitute the output associated with these activities. Finally, there is the health outcome, which may depend on a number of factors apart from the output of health care, such as whether or not the person gives up smoking" (Paragraph 15.120).

From a national accounts perspective, the target measure for the production of health services is outputs, not outcomes. This distinction is more difficult than meets the eye, however. First, the SNA reference to output as 'benefits to the patient' is best understood as the marginal contribution of health care activities to health outcomes, controlling for all other factors influencing outcomes. This means that the notion of outputs does not exist independently of outcomes. A similar conclusion (Schreyer 2012) arises in the context of quality adjustment (see below). Berndt et al. (1998) distinguish between medical care ('output' in our terminology), the state of health ('outcome' in our terminology) and utility. They envisage a relationship whereby utility depends, among other variables, on the state of health and where the state of health is itself dependent on health care services, on the environment, lifestyle etc.). Thus, a health care activity with a higher composite quality than another health care activity could be identified as such if it contributes more to health outcome than the alternative activity[5].

In practice, output of health service providers in the national accounts is increasingly operationalised via disease-based measures of health service provision, more or less in line with the OECD guidance on the matter: "In the case of diseases, our central notion in defining health care services is the ***treatment of a disease or medical services to prevent a disease***. Volume measures of output are then disease-based measures. Ideally, in the case of a treatment, the unit of output would capture *complete treatments*, and would take into account quality change in the provision of treatments. This measurement of health care output would then be able to differentiate among price, quantity and quality changes." (Schreyer 2010, p.73). When disease-based measures are introduced, they tend to be applied to both market and non-market producers of health services. This does not apply to those general government institutions that are part of the health sector at large (such as Ministries of Health) but not part of the providing industry. Nearly universally, the volume of general government output is measured via the volume of its inputs.

Weights. Another conceptual question concerns the choice of weights to aggregate across different types of outputs. For non-market production, prices, if they exist, are not a meaningful tool to aggregate. However, measurement can be based on *unit costs or quasi prices*. They are those (unobserved) 'prices' that emulate a competitive situation where prices equal average costs per product. Unit costs are observable and can be treated *as if they were prices*. Diewert (2011, 2012) and Schreyer (2012) discuss the question of weights extensively but for the purpose at hand it suffices to remind us that unit cost weights are a legitimate way of aggregating across non-market services that can subsequently be applied to obtain productivity measures.

Consider the treatment of disease i that is characterised by a unit cost function $c_i^t(\mathbf{w}^t)$ where $\mathbf{w}^t$ is a vector of input prices such as doctors' wages or user costs of hospital equipment. As c_i^t is a cost function, it represents minimum costs necessary to carry out the treatment at hand. Quasi prices are then simply defined to equal unit costs:

[5] Things are further complicated in practice. First, as Berndt et al. (1998) point out, there is an issue of lags: the state of health may be affected by medical care and by other factors with a lag so that utility derived from the state of health occurs at a different date from when medical services are provided. Second, there may also be a trade-off between immediate utility derived from consumption (say a fatty diet) and long-term disutility from reduced health status. This complicates formalisation of consumer behaviour but is secondary to the issue at hand, namely the measurement of health services.

(1) $p_i^t \equiv c_i^t(\mathbf{w^t})$

If minimum costs equal actual costs one has $c_i^t(\mathbf{w^t})y_i^t = \mathbf{w^t \cdot x_i^t}$, where y_i^t is the number of treatments of type i and $\mathbf{x}_i^t$ is the quantity vector of inputs that corresponds to $\mathbf{w}^t$:

(2) $p_i^t y_i^t \equiv c_i^t(\mathbf{w^t}).y_i^t = \mathbf{w^t \cdot x_i^t}$.

Expression (2) states the obvious, namely that with quasi prices, the value output of product i equals the value of inputs used in production of product i. This is the way non-market output is valued in the *System of National Accounts*[6]. However, as pointed out earlier, equality of inputs and outputs in value does *not* imply equality of inputs and outputs in volume or quantity.

The main difference between cost-based prices of outputs ('quasi prices') and prices of inputs is that the former correspond to *costs per unit of output* (such as the costs for one treatment of a heart attack) whereas the latter correspond to the *costs per unit of input* (such as wages per hour of a nurse).

Diewert (2008) shows formally how a cost-based volume index of output can be defined. He defines the Laspeyres version of a cost-based output quantity index as the (hypothetical) total cost $C^0(\mathbf{y}^1, \mathbf{w}^0)$ of producing the output vector $\mathbf{y}^1$ of period 1 under the conditions of period 0 technology and input prices, divided by the actual costs of period 0, $C^0(\mathbf{y}^0, \mathbf{w}^0)$. Similarly, he defines a Paasche type index as the actual costs of period 1, $C^1(\mathbf{y}^1,\mathbf{w}^1)$, divided by the hypothetical costs $C^1(\mathbf{y}^0, \mathbf{w}^1)$ that would have been incurred, had the products of period 0 been produced in period 1, under the technological constraints of period 1 and given period 1 input prices:

(3) $Q_L = C^0(\mathbf{y}^1, \mathbf{w}^0)/ C^0(\mathbf{y}^0, \mathbf{w}^0) = \Sigma_i^N c_i^0 y_i^1 / \Sigma_i^N c_i^0 y_i^0$

$Q_P = C^1(\mathbf{y}^1, \mathbf{w}^1)/ C^1(\mathbf{y}^0, \mathbf{w}^1) = \Sigma_i^N c_i^1 y_i^1 / \Sigma_i^N c_i^1 y_i^0$

$Q_F = [Q_L Q_P]^{1/2}$.

The same reasoning can be applied to quasi prices and an *indirect index of quasi prices* constructed by dividing total costs by the volume index of output:

(4) $P_L = [C^1(\mathbf{y}^1,\mathbf{w}^1)/C^0(\mathbf{y}^0, \mathbf{w}^0)]/Q_P = \Sigma_i^N c_i^1 y_i^0 / \Sigma_i^N c_i^0 y_i^0$

$P_P = [C^1(\mathbf{y^1},\mathbf{w^1})/C^0(\mathbf{y^0},\mathbf{w^0})]/Q_L = \Sigma_i^N c_i^1 y_i^1 / \Sigma_i^N c_i^0 y_i^1$

$P_F = [P_L P_P]^{1/2}$.

A useful interpretation of this quasi-price index can be obtained by re-writing the Laspeyres or Paasche version in expression (4). For example, after inserting the theoretical expression for Q_P into the first line of (4), P_L can be presented as the product of two terms:

(5) $P_L = [C^1(\mathbf{y}^1,\mathbf{w}^1)/C^0(\mathbf{y}^0, \mathbf{w}^0)]/Q_P$

$= [C^1(\mathbf{y}^1,\mathbf{w}^1)/C^0(\mathbf{y}^0, \mathbf{w}^0)]/[C^1(\mathbf{y}^1,\mathbf{w}^1)/C^1(\mathbf{y^0},\mathbf{w}^1)]$

[6] For a genesis of the treatment of non-market production in the national accounts and the many issues associated with it, see Vanoli (2002).

$$= [C^1(\mathbf{y}^0,\mathbf{w}^1)/C^0(\mathbf{y}^0,\mathbf{w}^0)]$$

$$= [C^1(\mathbf{y}^0,\mathbf{w}^1)/C^1(\mathbf{y}^0,\mathbf{w}^0)][C^1(\mathbf{y}^0,\mathbf{w}^0)/C^0(\mathbf{y}^0,\mathbf{w}^0)]$$

The first term in the last line of (5) is an economic index of input prices: costs are compared between two situations, with technology and the level of output held fixed but input prices are allowed to vary. The second term in the same line is an inverted productivity index: for a given reference output and input prices, changes in minimum costs between the periods are compared. Similar transformations could be applied to P_P and then combined with P_L to yield a decomposition of P_F, but there is no need to present them here. The main point can easily be explained with the decomposition of P_L only: in a market situation, a productivity index equals an input price index divided by an (output) price index: if output prices rise less rapidly than input prices, this implies productivity improvements. In the non-market case, the quasi-price index for outputs plays a similar role as the output price index in a market situation. If quasi prices (unit costs) rise less rapidly than input prices, there has been productivity change.

The measurement of productivity as a shift in the cost function is a well-established methodology[7] and we conclude that the cost-weighted measure of outputs is a fully valid measure output that also qualifies for productivity comparisons. Despite the fact that much of the discussion about non-market producers has been by way of costs, we *are* lending an output perspective to our calculations: unit costs or quasi prices are productivity-adjusted input prices and the productivity adjustment marks the movement from an input perspective towards an output perspective in measuring non-market activity. This is not always well understood, because costs are rightly seen as input-related variables. The above makes it clear that considering costs per unit *of output* differentiates an output perspective from considering costs per unit *of input*, i.e., the input perspective. However, the cost-based measures of output remain incomplete insofar as they invoke no direct element of consumer valuation – unit costs are not a product of the interplay between producers and consumers as in the market case. Unit costs are only reflective of the supply side.

Quality change. An unrealistic assumption in the model above is the unchanged set of products between two periods. In reality, the quality of products changes over time, certain products disappear from the market and new products emerge. These changes constitute not only a major practical challenge for statisticians; they also have consequences for theoretical considerations about output and utility. The distinction between new products and quality change[8] will be ignored here but a few general points about quality adjustment[9] of prices or quantities will be noted.

One technique to deal with quality change in products is to group them such that only products of the same specification are compared over time or in space. Such grouping or matching ensures that only prices or quantities of products of the same or very similar quality are compared. The idea is that products of different quality are treated as different products. Examples for such grouping are medical services provided by hospitals with different levels of non-medical services. Also, when the nature of the service changes due to certain consumer characteristics, grouping may be necessary. For example, an elderly patient suffering from the same disease as a young patient may need more care due to longer time to recover. This may result in higher expenditures for the group of older patients. Note, that capturing quality differences through grouping and matching the groups over time relies on an important assumption: the price or quantity movements of those products that are matched have to be a good indicator of the price or

[7] Balk (1998) provides a full treatment of the various productivity measures. In his terminology, our measure of technical change would be labelled a 'dual input based technical change index' (page 58). Diewert and Nakamura (2007) also discuss dual, cost based measures of productivity change.

[8] For a discussion see for example ILO et al. (2004).

[9] For an in-depth treatment of quality adjustment in price measurement see Triplett (2006).

quantity movements of those products that are not matched – in particular products that are newly entering the market. Also, all other price or quantity changes that arise outside the sample of matched products are ignored.

A more sophisticated way of grouping is with hedonic regression techniques[10] that help controlling for characteristics of treatments and patients. For instance, Berndt et al. (2001) use patient characteristics, information on different types of depression, variables on medication and the like to estimate a hedonic price model for the treatment of depression; the idea being to isolate those price changes that are due to changes in characteristics from those price changes that constitute 'inflation'. However, in situations of non-market production, the applicability of hedonic techniques is more limited or at least more complex (Schreyer 2012).

Yet another way to tackle quality change in medical care is to start from the observation that consumers attach utility to a good or to a service because it affects outcome, i.e., a particular state that they value and which can be measured. One could also say that outcome is an intermediate step between consumption and utility and this is indeed the way it has been treated in the literature. Thus, one possibility to deal with quality adjustment and aggregation is to subsume several characteristics into a single indicator that reflects the *contribution of the product to outcome*. For example, in the case of price indices for health care, Triplett (1999) suggests quality-adjusted life years (QALYs) as a single dimensional measure that could be used for the quality-adjustment of different treatments within a product group. The point is to derive a single indicator that serves as a reasonable summary of a true, multi-dimensional set of quality characteristics valued by consumers when purchasing health services. Careful judgement needs to be applied in the choice of such a measure. In particular, it should not be affected by any other factors that influence consumer outcome (e.g., socio-economic background of students or lifestyle of patients).

Box 1. The meanings of 'outcome'

Outcome has been used in different ways in the relevant literature on health services. Two usages are common:

In the health care literature, 'outcome' is typically defined as the resulting change in health status that is directly attributable to the health care received. Triplett (2001) indicates this usage in the cost-effectiveness literature and quotes Gold et al. (1996) who define a health outcome as the end result of a medical intervention, or the change in health status associated with the intervention over some evaluation period or over the patient's lifetime. Employed in this sense, some authors suggest that the 'output' of the health care industry be measured by 'outcome'.

Among national accountants, 'outcome' is typically used to describe a state that consumers value, for example the health status without necessarily relating the change in this state to the medical intervention. For example, Eurostat (2001) gives as examples of "outcome indicators" the level of education of the population, life expectancy, or the level of crime. Atkinson (2005) has the same usage of the word. Understood in this sense, outcome in itself cannot be a useful way to measure output or the effectiveness of the health or education system. In terms of national accounts semantics, the 'marginal contribution of the health care industry to outcome' is the equivalent to the notion of 'outcome' as used in the health care literature.

As long as a particular definition is used consistently, the substance of the argument is of course unaffected and the only question is the usefulness of one definition or the other. As the note follows in the line of Eurostat (2001) and the Atkinson Review (2005), it also employs the term 'outcome' in the sense of the national accounts literature.

[10] See Triplett (2006) for a comprehensive discussion.

3. Overview of country practices – comparisons in time

In this section we take an international perspective and address the issue of how health services are measured in countries' national accounts in practice. Schreyer (2010, table 4.4) provided an overview for thirty OECD countries, plus a more detailed analysis for six European countries: Austria, Denmark, Germany, Netherlands, Norway and United Kingdom. The first task addressed in this section is updating the information for the set of countries. Table 1 below reflects a few updates but a more extensive process of updating is presently being launched through the OECD's Working Party on National Accounts in 2014. Consequently, for the time being, we mainly rely on existing information from Schreyer (2010) and some more recent and specific examples for Germany, Spain, Hungary and the United Kingdom that have been investigated as part of the European Union's INDESCER project (Goerlich et al. 2012, Huttl et al. 2011) as well as a research project by Statistics Canada (Gu and Morin forthcoming).

Residential care. Note important differences between areas of health care. The above, conceptual, discussion was framed with 'a treatment' in mind and led to endorsing a disease-based approach towards measuring health care services. While the disease-based approach is no doubt useful for hospital services, it may be less evident when it comes to the broader set of health care institutions. In particular, residential care activities are different in nature from hospital and medical practice activities and account for sizable shares of overall health expenditure. It is difficult to conceptualise the correct measure of output of residential care and typically, one will be led back to a measure of inputs or number of days in residential care, possibly differentiated by intensity of care. Certainly in practice, these are the measures most frequently found.

Pathway through institutions. Another issue, potentially important, is whether treatments can be observed throughout the pathways of health care institutions. For instance, a treatment may start as an inpatient treatment in a hospital and continue as outpatient treatment. In most countries, tracking treatments in this way is not possible. As a consequence, the effects of shifts between inpatient and outpatient treatments on volume measures of health care may be lost or obscured.

Table 1. Overview of country practices in the volume measurement of health services

		Hospital activities		Residential care activities	Medical and dental practice activities		Other human health activities
Country	**Status**	**Acute hospitals**	**Mental health and substance abuse hospitals; Specialised hospitals**		**Doctor services**	**Dental services**	
Austria	Implemented, data since 2001	Deflation with index based on unit costs per treatment by DRGs, cost weights	Deflation with index based on unit costs per treatment by DRGs, cost weights	Number of occupant days, weighted by revenues, no quality adjustment	Number of treatments weighted by revenues, no quality adjustments	64 indices based on fees per single service item paid by the social security, weighted by revenues	Deflation by HCPI
Australia	Implemented	Direct volume index based on DRGs, cost weights	NA	Number of cases by level of care weighted by subsidy rates	Number of services weighted by fees charged	Number of services weighted by cost	NA
Belgium	Implemented in 2009, data available since 1995	All hospitals are market producers; Direct volume index, based on DRGs, cost weights	Number of occupant days by level of care, weighted by income by category of hospital services	Number of occupant days by level of care, weighted by income by category of hospital services	Number of consultations, use of regulated price of services	Number of consultations, use of regulated price of services	Number of consultations, use of regulated price of services
Canada	Implemented	Deflation with input price index	NA	NA	NA	NA	NA
	Planned	Exploratory work (Gu and Morin 2013)		NA	NA	NA	NA
Czech Republic	Implemented	Deflation with index based on daily rates for hospital			Number of treatments	Number of treatments	CPI - component
Denmark	Implemented	Deflation with index based on unit costs per treatment by DRGs, cost weights	Deflation with index based on unit costs per discharge by diagnostic group, cost weights	Deflation with index based on unit cost per patient by type of care, cost weights	Deflation - CPI component	Deflation with index based on unit cost per patient by 2 types of care, cost weights	
Finland	Implemented data available since 2000	Volume index based on DRGs, cost weights	Number of day care days	Number of day care days	Number of consultations by type of consultation (17)	Number of consultations by type of consultation (3)	
France	Implemented, data available since 1998	Volume index based on DRGs, cost weights	Volume index based on DRGs, cost weights	Volume index based on DRGs, cost weights	Deflation - CPI component	Deflation - CPI component	Deflation - CPI component

		Hospital activities		Residential care activities	Medical and dental practice activities		Other human health activities
Country	**Status**	**Acute hospitals**	**Mental health and substance abuse hospitals; Specialised hospitals**		**Doctor services**	**Dental services**	
Germany	Implemented data available since 2006	All hospitals are market producers; Deflation with index based on unit costs per inpatient treatment by groups of DRGs, cost weights+ explicit quality adjustment	Number of day care days or number of treatments, cost weights	Number of persons at the end of the year , cost weights by care level	Deflation – unit value for medical/dental services (statutory) and CPI component(private)		Deflation - CPI component
Greece	Implemented	Number of day care days	Number of day care days	Number of day care days	Deflation - CPI component	Deflation - CPI component	Deflation - CPI component
Hungary	Implemented data available since 2001	Volume indices based on DRGs weighted by quasi prices	Volume indices based on DRGs weighted by quasi prices	Number of visits	Number of consultations	Number of scores	Number of treatments on basis of services provided
Iceland	Implemented	Deflation with input price index	NA	NA	NA	NA	NA
	Planned	NA	NA	NA	NA	NA	NA
Ireland	Implemented	Deflation with input price index	NA	NA	NA	NA	NA
	Planned	NA	NA	NA	NA	NA	NA
Italy	Implemented data available since 2000	Volume indices based on DRGs, weighted by costs	Volume indices based on DRGs, weighted by costs	Volume indices based on DRGs, weighted by costs	Number of prescriptions	Deflation - CPI component	Deflation - CPI component
Japan	Implemented	Market - CPI component	Market - CPI component	Market - CPI component	Market - CPI component	Market - CPI component	Market - CPI component
Korea	Implemented	Market - CPI component	Market - CPI component	Market - CPI component	Market - CPI component	Market - CPI component	Market - CPI component
Luxembourg	Implemented data available since 2000	Deflation - CPI component	Deflation - CPI component	Number of day care days or number of cases by level of care for non market (cost weighted, no quality adjustments); Deflation - CPI component for market	Number of consultations or treatments for non market (cost weighted, no quality adjustments); Deflation - CPI component for market		

		Hospital activities		Residential care activities	Medical and dental practice activities		Other human health activities
Country	**Status**	**Acute hospitals**	**Mental health and substance abuse hospitals; Specialised hospitals**		**Doctor services**	**Dental services**	
Netherlands	Implemented	Direct volume index based on ICDs by age and discharge numbers + share in day care days as weight	Direct volume indicators based on days of treatments, days of hospitalization and hours of delivered care	Deflation - CPI component (CTG Tariff)	Deflation - CPI component (CTG Tariff)	Deflation - CPI component (CTG Tariff)	
New Zealand	Implemented	Government (non-market) hospitals: Composite volume index based on DRGs, cost weighted; patient discharge and bed-night numbers. Private market: deflation – CPI component	Combined with acute hospitals	Number of employee hours worked	Deflation - CPI component	Deflation - CPI component	Deflation - CPI component
Norway	Implemented	Direct volume index based on DRGs, cost weighted	Number of day care days by levels of care	Number of day care days	Deflation - CPI component	Deflation - CPI component	
Portugal	Implemented	Direct volume index based on DRGs; use of regulated price by DRGs (quasi price)	Direct volume index based on DRGs; use of regulated price by DRGs (quasi price)	Not applicable	Direct volume index based on number of consultations, use of regulated price (quasi price)	Not applicable	Not applicable

		Hospital activities		Residential care activities	Medical and dental practice activities		Other human health activities
Country	**Status**	**Acute hospitals**	**Mental health and substance abuse hospitals; Specialised hospitals**		**Doctor services**	**Dental services**	
Sweden	Implemented data available since 2003	Direct volume index based on DRGs, cost weights	Direct volume index based on number of days of care by level of care	Direct volume index based on number of days of care by level of care	Direct volume index based on number of consultations, cost weighted	Direct volume index based on number of consultations, cost weighted	Number of consultations or treatments
Switzerland	Implemented	Deflation with input price index	NA	NA	NA	NA	NA
United Kingdom	Implemented. Data from 1995. England and Northern Ireland	Direct volume index based on HRGs, cost weights	Direct volume index based on HRGs, cost weights	Proxied by growth in hospital activities (only includes health-related residential care activities)	Direct volume index based on number of consultations, cost weighted	1995-2006 : Direct volume index based on number of treatments, cost weighted. From 2006: proxied by growth in hospital activities.	Proxied by growth in hospital activities
United States	Implemented	Deflation - use of relevant component of CPI/PPI	Deflation - use of relevant component of CPI/PPI	Deflation - use of relevant component of CPI/PPI	Deflation - use of relevant component of CPI/PPI	Deflation - use of relevant component of CPI/PPI	Deflation - use of relevant component of CPI/PPI
	Planned	Direct volume index based on DRGs, cost weights	NA	NA	NA	NA	NA

Source: adapted from Schreyer (2010).

Table 1 calls for several observations:

- There are still significant differences in the methods used to measure the volume of hospital services. For instance, to date, the United States, Canada, Mexico, Chile, Japan, and Korea are employing input-based volume measures; Australia, New Zealand, and many EU countries use output-based measures. At the same time, there are many shadings to the output based measures and indeed, it is not always clear whether certain methods do qualify as input-based or output-based, for example the number of hospital days[11]. More information is also required to pass a judgement on the nature of those output measurements that are based on relevant CPI or PPI components. Do these components reflect full prices? How have they been valued?

- Where output-based methods for hospital care have been chosen, these tend to rely on DRGs or hospital discharge information and thus share the characteristic of a disease-

[11] This is the case of Greece that has been placed under the "Deflation with input price index" heading in table 1

based measure. For reasons mentioned earlier, there is also great similarity in countries' approaches towards measuring residential care activities.

- It is tremendously difficult to make a statement about the degree of international comparability of measures of hospital services based on the above Table. While it is obvious that methods vary between countries, this does not necessarily imply significant problems of comparability of results. Comparability is often quoted as one of the advantages of traditional, input-based measures for health services. However, as there is no reason to believe that the bias induced by input-based methods (instead of output-based measures) is the same across countries, reverting to input-based computations would not really solve the problem of comparability. One avenue to gain insight into the comparability of output-based measures is currently pursued by the OECD: as standardised data for spatial comparisons of health prices is progressively collected (see Koechlin, Lorenzoni and Schreyer 2010), it may be possible to use this information to also construct temporal indices of health care services that would then serve as a counter-fact to national methods.

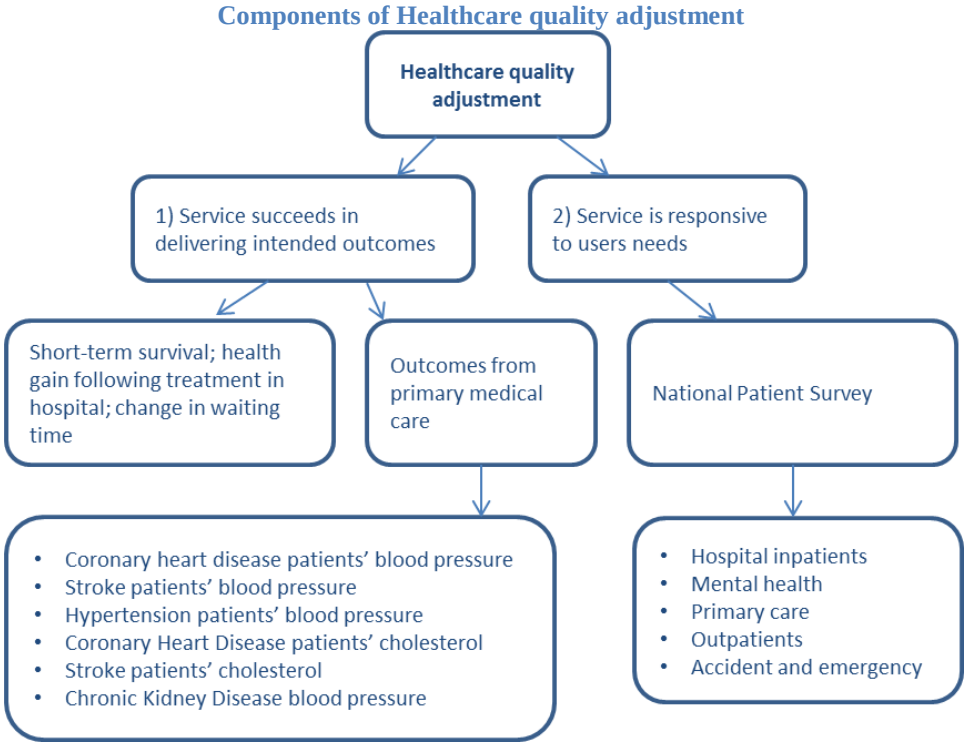
Quality Adjustment in Practice. Of the various methods to quality-adjust volume or price indices of health care, the vast majority of OECD countries has relied on stratification and matching. A good example is Finland whose approach towards quality adjustment is clearly rooted in stratification. Statistics Finland aims at capturing quality change by classifying medical services into strictly homogeneous quality groups of products. Statistics Finland considers that outcome is not a concept in national accounts, and correcting for changes in outcome introduces a normative element that is not in line with the positive approach of national accounts. From a practical angle, Statistics Finland considers that outcome-based quality corrections might offer too little and arrive too late for decision makers. Experimental work with explicit quality adjustment has been pursued by the U.K. Office of National Statistics (see Box) but is scarce otherwise. Eurostat, the Statistical Office of the European Union, has even advised against the use of explicit quality adjustment procedures on the grounds that if explicit methods are used by some EU countries but not by others or if the quality adjustment methods used are very different, this would undermine comparability of volume measures of health care in the European national accounts.

Explicit quality adjustments – United Kingdom.

The U.K. Office for National Statistics (ONS) is the statistical office among OECD countries that has gone furthest in investigating and advancing the measurement of volume health services and government services more generally. One of the triggers for this activity was the *Atkinson Review* (2005) commissioned by the British Government and work carried out for the UK Department of Health by the University of York and the National Institute of Economics and Social Research, NIESR (Dawson et al. 2005). However, at present, the quality adjustments remain exploratory and have not been reflected in the U.K. National Accounts. The explicit quality adjustment procedure is developed by the Centre for Health Economics (CHE) at York University (CHE 2005) and the Department of Health (DH 2005, 2007). The method was implemented using data for England and an assumption is made that the rest of the UK follows the same trend. The quality adjustments take account of some aspects of quality that are not readily captured by disease-based activity measures. The adjustments reflect two dimensions of quality (see Figure): (i) the extent to which the service succeeds in delivering its intended outcomes; and *(ii)* the extent to which the service is responsive to users´ needs.

In practice, the first dimension accounts for at least 99.5% of total quality adjustment. It consists of two composite measures, (i) short-term survival rates, health gain following treatment in hospital and change in waiting times, and (ii) outcomes from primary medical care. According to the ONS (2011), in 2009 quality adjusted output was 7.1% greater than quantity (unadjusted output). From 2001 to 2009, quality adjustments added an average of 0.5 percentage points (pp) a year to output growth. The main contribution to quality change came from survival, health gain and waiting times, which improved by an annual average of 0.66 pp from 2001-02 to 2008-09. Smaller contributions come from primary care and responsiveness to users´ needs, with an annual coverage improvement of 0.07 pp and 0.01 pp respectively over the same period. Finally, quality change rose from 0.4 pp in 2007-08 to 1.11 pp in 2008-09. This came almost entirely from an improvement in 30 day survival rates following treatment and a reduction in waiting times

was the main reason for an increase in quality in 2009.



Source: ONS (2011 pg. 12)

4. Price levels and volumes of health services – comparisons in space

While the measurement of the evolution of health services in a particular country is of considerable interest, so is the comparison of the level of health services in different countries at a particular point in time. For example, Figure 1 showed levels of health expenditure as a share of GDP across countries with marked differences. What policy makers and analysts would like to understand is whether these differences in expenditure reflect more or less health services or higher or lower prices for these services in the various countries. This requires a spatial price index of health services that permits breaking down nominal expenditures into a price and volume component. The spatial price index comes in the form of a health-specific Purchasing Power Parity (PPP).

PPPs are regularly measured for all components of GDP[12]. Despite a long tradition of work in the area, the task remains challenging. Three main problems have to be addressed in the measurement of PPPs. The first is to identify products that are comparable across countries. This can be complicated because products are not identical, because there are differences in quality or because products simply do not exist in all countries. The second issue is to ensure representativeness of products: whatever price is compared, it has to be the price of a product that is widely and typically purchased in each country. The third issue arises when there is a product, but no meaningful market price for comparison. Issues one and two arise in the comparison of all prices, issue three arises in the comparison of products that are produced and delivered outside markets. In many countries, health services count among these products.

[12]. For a full description of the methods used, the reader is referred to Eurostat-OECD (2006, 2013).

When goods or services are supplied by a non-market producer the prices charged to consumers are significantly below the price that a market producer would charge. In some cases, the price may even be zero. It would make no sense to compare such prices charged to patients or consumers across countries as they reflect administrative decisions and not the value of products. A recent pilot study by the OECD (Koechlin, Lorenzoni and Schreyer 2010) compares *quasi prices* across countries. In direct analogy to the temporal indices of quasi prices (see above), this deals with the issue of absent market prices in health provision. In what follows, we briefly report on these results pointing out that work is progressing in the area to move from a pilot stage to full, period implementation and to a broader scope than hospital services.

The products: case types. For the study at hand, products were defined through *case types*. These refer to classes of hospital services that are similar from a clinical perspective. For instance, 'heart failure' constitutes one case type. Each case type is further specified so as to compare similar occurrences of diseases. In the case of heart failure, the indication is given that 'no operating room procedure is performed'. This leads to greater homogeneity of case types also in terms of their consumption of resources. 29 in-patient[13] case types were identified[14] based on the following criteria. The case types should:

- represent common procedures or diagnoses;
- account for a significant percentage of hospital expenditures;
- represent procedures which are likely to be the principal procedure within one hospitalisation (for surgical case types); and
- represent well-identified conditions (for medical case types).

The valuation: quasi-prices. It is rare that case types can be directly valued through free-standing costing studies and clinical trials. A more promising avenue is to use secondary data sets available through health administrations and national insurance funds for purposes of reimbursement and health financing. The administrative data sets provide quasi-prices, encompassing both negotiated prices and administered prices. The former are established through independent negotiations between purchasers/third party payers and providers, and are not necessarily directly tied to the cost of care. While there may be differences between negotiated and administered regimes (Castelli, 2007; Triplett, 2003), the general principle for compilation of quasi prices is that at a minimum they are reflective of the full set of costs, compatible with costs as defined in the national accounts (see above).

Results. One key result of a comparison of hospital quasi prices is an index of comparative price levels for medical services. By way of example, the Table below shows results from the OECD pilot study for different types of inpatient hospital services.

[13] Akin to temporal price and volume indices, we note that the explicit distinction between inpatient and outpatient case types implies that inpatient and outpatient services are considered different products. While plausible in some ways, this also means that the methodology is not able to capture price differences that are due to the fact that an inpatient treatment has been substituted by an outpatient treatment or vice versa. At this point it is not possible to quantify the extent of this possible bias.

[14] See Koechlin, Lorenzoni and Schreyer (2010) for a full list. The selection was based on a list of inpatient case vignettes (Huber, 2007), on a proposal by the OECD Expert group on procedures under the Hospital Data Project (Smedby, 2007), and on the list that is currently used at the OECD for Health Data collection (OECD, 2012).

Comparative price levels for hospital services and GDP, 2007

	AUS	CAN	FIN	FRA	ITA	ISR	KOR	POR	SLV	SWE	USA	Group
Inpatient Medical Services	122	125	91	140	158	60	37	90	65	112	173	100
Inpatient Surgical Services	124	113	99	114	132	65	66	81	56	116	163	100
Total Inpatient hospital services	123	113	98	121	140	62	57	85	59	114	164	100
GDP	104	101	118	112	103	120	73	83	79	121	90	100
Reference: per capita real GDP	115	118	108	99	95	82	81	69	81	113	142	100

** Source: Koechlin, Lorenzoni and Schreyer (2010).*

Results were compiled for 12 countries. They are expressed as indices, with the average for the group of countries set to equal 100. PPPs were computed so as to be invariant to the choice of the base country. Computation started with the United States as reference country, then comparative price levels (CPLs) were derived by dividing PPPs by market exchange rates, and the average of the group was calculated as the geometric mean of the CPLs of the different countries. This average was then set to equal 100 and each country's CPL expressed in relation to it. CPLs provide a measure of the difference in price levels between countries by indicating – for a given category or aggregate – the number of units of the common currency needed to buy the same volume of the category or aggregate. In our example, there is no common currency as such and results should be interpreted looking at the relativities between countries rather than looking at absolute levels. For example, the figures in the table should be read as follows: in 2007, price levels for total inpatient hospital services in the United States stood at 163 % of the average price level of the group of countries and were therefore nearly 44 % (163 compared to 113) higher than in Canada.

Main findings, generally in line with evidence from other sources, include (i) hospital services in the United States are significantly more costly than in the other countries considered in this study. In particular, price levels in Korea and Israel are only around 60% of the average of all countries; (ii) for the 12 countries under consideration, price level differences cannot be explained by differences in the average length of stay – rather, high-priced countries also exhibit high prices per day of hospitalisation.

The above results are a first step towards more systematic and broad-based measurement of spatial price and volume indices for health services. The methodology needs further refinement, and a second-best approach for countries where the available data does not allow following the standard approach. Also, the methodology has to be expanded to cover PPPs for the services of mental health and speciality hospitals, nursing and residential care facilities. The objective is to translate PPP results into volume measures of health services. This requires a set of expenditure data from the national accounts that are consistent with the present framework for health PPPs. Such consistency (for example with regard to classifications) is important otherwise deflating health expenditure with health PPPs will give rise to biased measures of the volumes of health services across countries. These and other developments are presently undertaken by the OECD and Eurostat.

6. Conclusions

This paper provided a national accounts perspective to the measurement of health service provision. It spelled out some of the key concepts and looked at practices in a number of OECD countries. A new

approach towards cross-country comparisons of price and volume measures of health services was also presented. Key messages and conclusions are:

- While the measurement of the *value* of production of non-market producers is necessarily different from the measurement of the value of production of market producers (sum of costs for the former, revenues for the latter), the measurement of the *volume* of production may and indeed should follow the same method. There is increasing recognition that for many purposes, a disease-based approach towards output measurement is the right way forward;

- Information on the precise treatment in national accounts of institutional units involved in health care provision is scattered and incomplete. In particular, there are gaps in the information on market versus non-market producers although this constitutes an analytically relevant distinction. It is not always clear whether methodologies for volume output figures differ between market and non-market producers (and, among non-market producers between general government and non-profit institutions serving households).

- It is tremendously difficult to make a statement about the degree of international comparability of measures of hospital services. While methods vary between countries, this does not necessarily imply significant problems of comparability of results. Comparability is often quoted as one of the advantages of traditional, input-based measures for health services. However, as there is no reason to believe that the bias induced by input-based methods (instead of output-based measures) is the same across countries, reverting to input-based computations would not really solve the problem of comparability**;**

- A new approach towards comparing volumes of health services internationally has been developed in the context of the Eurostat-OECD Purchasing Power Parity Programme. As evidence from this approach accumulates over several years, it is planned to construct time series of health service provision which will provide a new point of comparison with the existing national accounts data and advance the discussion on future developments in the measurement of health services nationally and internationally.

REFERENCES

Abraham, K. G. and Ch. Mackie (2006);"A Framework for Nonmarket Accounting"; in: Jorgenson, D., J. S. Landefeld, W. Nordhaus (eds.), *A New Architecture for the U.S. National Account*, NBER, Studies in Income and Wealth, University of Chicago Press.

Atkinson, Anthony (2005): *The Atkinson Review: Final Report. Measurement of Government Output and Productivity for the National Accounts,* Palgrave Macmillan, Basingstoke.

Berndt, E. R., D. M. Cutler, R. G. Frank, Z. Griliches, J. P. Newhouse and J. E. Triplett (1998); "Price Indexes for Medical Care Goods and Services: an Overview of Measurement Issues"; *NBER Working Paper Series*, No 6817.

Berndt, E.R., Cutler, D.M., Frank, R.G., Griliches, Z., Newhouse, J.P. and Triplett, J.E. (2000) "Medical care price and output" in Culyer, A.J. and Newhouse, J.P. (Eds.) *Handbook in Health Economics* vol.1a., Elsevier.

Berndt, E.R., Cutler, D.M., Frank, R.G., Griliches, Z., Newhouse, J.P. and Triplett, J.E. (2001) "Price Indexes for Medical Goods and Services: An Overview of Measurement Issues" in Cutler, D.M. and Berndt, E.R. (eds.).

Castelli, A., *et al*. (2007), "Improving the measurement of health systems output growth", *Health Economics*; Vol. 16: pp. 1091-1107.

Chessa, Antonio G. and Foske J, Kleima (2006): "The Dutch experience in measuring health output and labour productivity", *Discussion paper 06002*, Statistics Netherlands.

CHS (2005): *Developing New Approaches to measuring health care output and productivity*, York University and National Institute of Economics and Social Research.

Cutler, D. M. and E. R. Berndt (eds.) (2001); *Medical Care Output and Productivity*; NBER Studies in Income and Wealth, Volume 62; University of Chicago Press.

Cutler, D. M., A. B. Rosen, S. Vijan (2006); "The Value of Medical Spending in the United States, 1960–2000"; *New England Journal of Medicine*; 355: pp. 920-927.

Dawson, D., Gravelle, H., O´Mahony, M., Street, A., Weale, M., Castelli, A., Jacobs, R., Loveridge, P., Martin, S., Stevens, P. and L. Stokes (2005): "Developing new approaches to measuring NHS outputs and productivity", Final Report to the Department of Health, Centre for Health Economics (York) and National Institute of Economic and Social Research.

Department of Health (2005); *Healthcare output and productivity. Accounting for quality change*, December.

Department of Health (2007); *Review of Data Sources and Methodology for the Calculations of Hospital output in the NHS*.

Diewert, W.E. (2012); "The measurement of productivity in the nonmarket sector"; *Journal of Productivity Analysis* (2012) 37:pp. 217–229.

Diewert, W. E. (2011); "Measuring productivity in the public sector: some conceptual problems"; *Journal of Productivity Analysis* 36; pp.177–191.

Eurostat-OECD (2006), *Methodological Manual on Purchasing Power Parities*, OECD, Paris.

Eurostat-OECD (2013), *Methodological Manual on Purchasing Power Parities*, OECD, Paris.

Eurostat (2001); *Handbook on Price and Volume Measures in National Accounts*, European Communities, Luxembourg.

Goerlich, F., J. Peréz, A. Huttl, M. O´Mahony, E. Schulz and L. Stokes (2012): "Health capital: an application for Germany, Hungary, Spain and the UK", http://indicser.com/images/dp26_goerlich_et_al.pdf

Gu, Wulong and Stéphane Morin (2013); "Experimental Measures of Output and Productivity in the Canadian Hospital Sector"; in Jorgenson, D.W., S. Landefeld and P. Schreyer (eds.); *Measuring Economic Progress and Economic Sustainability*; NBER Book Series Studies in Income and Wealth, forthcoming.

Hautakangas, Sami and Jani Heikkinen (2010): "Measurement of non-market output in local and central government in Finland" Statistics Finland.

Hill, T. P. (1975); *Price and Volume Measures for Non-Market Services*; Report to the Statistical Office of the European Communities, Brussels.

Huber, M. (2007), "International comparison of prices and volumes in health care among OECD countries." European Center for Social Welfare Policy and Research, paper presented at the first meeting of the Task Force for the Development of Health-specific Purchasing Power Parities, Paris, 8 June.

Huttl, A, M.Mas, A.Nagy, G.Okem, M.O´Mahony, E.Schulz and L.Stokes (2011): "Measuring the productivity of the Healthcare sector: Theory and Implementation" http://indicser.com/images/RP5_Huttl_etal.pdf

Joumard, I., P. Hoeller, C. André and C. Nicq (2010), *Health Care Systems: Efficiency and Policy Settings*, OECD Publishing, Paris

Jorgenson, Dale W. and J. Steven Landefeld, "Blueprint for Expanded and Integrated U.S. Accounts: Review, Assessment, and Next Steps," in Jorgenson, Landefeld, and Nordhaus (eds.), 13-112, 2006.

Jorgenson, Dale W. and Paul Schreyer (2013); "Industry-level Productivity Measurement and the 2008 System of National Accounts"; *Review of Income and Wealth*, 59, no 2 June; pp.185-211.

Jorgenson, Dale W. and Kun-Young Yun, *Lifting the Burden: Tax Reform, the Cost of Capital, and U.S. Economic Growth*, Cambridge, MA, The MIT Press, 2001.

Koechlin, Francette, Luca Lorenzoni and Paul Schreyer (2010); "Comparing Price Levels of Hospital Services Across Countries: Results of a Pilot Study"; *OECD Health Working Paper No 53*.

Malizia, Raffaele (2009): "The valuation of General Government output at macro-economic level" paper presented to the International Workshop on Government output and productivity measurement: lessons from the international experience, Rome, April 23-24.

Mas, Matilde, Francisco Pérez and Ezequiel Uriel (2006), "Capital Stock in Spain, 1964-2002. New Estimates", *Growth, Capital and New Technologies*, M. Mas and P. Schreyer (eds.) Fundación BBVA, Madrid.

OECD (2011). *Health at a Glance 2011: OECD Indicators*. Paris.

OECD (2009), *Measuring Capital: Revised Manual*, Paris.

ONS (2010): *United Kingdom National Accounts- The Blue Book*, from 2000 to 2009, Office for National Statistics.

ONS (2011): "Public Service Output, Inputs and Productivity: Healthcare. Correction Notice", *Office for National Statistics*, UK, April 7

ONS (2012): "Quality adjustment of public service health output: current method", *Office for National Statistics*, UK April.

Rosen, A. B. and D. Cutler (2007); "Measuring Medical Care Productivity A Proposal for U.S. National Health Accounts"; *Survey of Current Business* June 2007, 54-58.

Schreyer, Paul (2010): "Towards Measurement the Volume Output of Education and Health Services. A Handbook", *OECD Statistics Working papers*, 2010/02, Paris.

Schreyer, Paul (2012); "Output, Outcome and Quality Adjustment in Measuring Health and Education Services"; *Review of Income and Wealth* 58, no 2 June; pp. 257-278.

Schreyer, Paul and Dirk Pilat (2001): "Measurement of Aggregate and Industry-Level Productivity", *OECD Economic Studies* No. 33.

Smedby, B. (2007); "A selected list of hospital procedures for international comparison"; Report on the work of the expert group on procedures under the HDP2 project. Presentation at the OECD Health Data National Correspondents meeting, Paris 9-10 October 2007.

Statistics Netherlands (2007): I*nventory of sources and methods for price and volume measures in the Dutch National Accounts* (revised 6-11-2007); http://www.cbs.nl/NR/rdonlyres/AE3A6C16-0523-4C4E-88E9-DC7567CD34DA/0/ExternRapportEurostatokt2007Finaleversie.pdf

Triplett, J.E. (2003), "Integrating Cost-of-disease Studies into Purchasing Power Parities", *A Disease-Based Comparison of Health Systems*, OECD Paris

Triplett, J.E. and B. P. Bosworth (2004); *Productivity in the U.S. Services Sector: New Sources of Economic Growth*; The Brookings Institution, Washington D.C.

Vanoli, A. (2002); *Une Histoire de la Comptabilité Nationale*; Paris, La Découverte.

Zannoni, Silvia (2009): “Non market health care service in Italy”, paper presented to the International Workshop on Government Output and Productivity Measurement: Lessons from the International Experience, Rome, April 23-24.