Parallel Session 7A
Friday, 29 August, a.m.

Paper Prepared for the 30th General Conference of
The International Association for Research in Income and Wealth

Portoroz, Slovenia, August 24-30, 2008

Output and Outcome – Measuring the Production of Non-market Services

Paul Schreyer

For additional information please contact:

Name	Paul Schreyer
Affiliation	OECD
Full mailing address	2, rue André Pascal, 75116 Paris Cedex, France
Email address	Paul.Schreyer@oecd.org

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

OUTPUT AND OUTCOME – MEASURING THE PRODUCTION OF NON-MARKET SERVICES

Paul Schreyer
OECD Statistics Directorate*
Paul.Schreyer@OECD.org

Version of 22nd June 2008

Abstract

We consider the notions of output and outcome in the measurement of health and education services. In principle, the national accounts production boundary encompasses outputs but not outcomes. However, we show that although output measures are different from measures of outcome, they are not independent from each other. This is most obvious when it comes to quality adjustment of outputs where implicit or explicit information about outcomes is needed. The paper discusses some quality adjustment methods, including a hedonic approach in a context of non-market services.

*Opinions voiced in this document reflect the views of the author and not necessarily those of the Organisation for Economic Co-operation and Development or its Member countries.

TABLE OF CONTENTS

1. Introduction

Although much effort is spent on measuring the value of GDP at current prices, an even more important objective of the National Accounts is to derive a measure of the growth of GDP and its components *in volume*. This implies decomposing current price measures into a price and a volume component. For complex services such as education and health, and moreover, in a context where there are no economically significant prices, this is a difficult task. Traditional methodologies have relied on measuring the volumes or the prices of inputs to obtain a measure of the volume or price of outputs. Quantity and quality of outputs are not identified consequently it is not possible to capture any productivity change in the production of health and education services. Productivity growth exists when more or better output can be produced with the same resources. As there is much evidence of changing quality of output in health and education services, ignoring productivity changes means foregoing important information for analysts and policy makers and measuring volume growth inaccurately.

The discussion about the deficiencies of input-based methods is by no means new but has resurfaced in the recent past. Eurostat (2001) stated in principle the desirability of applying output-based measures to non-market services. In the United Kingdom, the topic was taken up by the Atkinson (2005) report that was widely discussed inside and outside the U.K. The measurement of services output and productivity has been a longstanding topic of interest in the United States, with several new publications recently (Triplett and Bosworth 2004, Abraham and Mackie 2006). Over the past two years, the OECD has also looked at the topic and organised several workshops[1]. One of the objectives of the work at the OECD is to come up with a handbook that provides guidance on the measurement of volume output of health and education services[2] and the present paper draws on this work. Views about the feasibility to implement different measures in the official national accounts are not unanimous but it is clear that there is significant demand for research and a need to get a better handle on the measurement of outputs, inputs and productivity of major economic activities such as health and education that are largely or partly characterised by non-market production.

It will be argued later in this paper that the distinction between market and non-market producers is not very relevant for many of the statistical issues at hand. While it is clear that market production provides a basis for the measurement of production at current prices whereas the value of non-market production is typically estimated by summing up costs, most of the tricky issues associated with the measurement of volume output or with the measurement of price changes apply to both market and non-market producers. In particular, the difficulties to keep track of quality change and of entering and exiting products are present independently of the institutional affiliation of the producing unit. These problems are associated with the increasingly complex nature of modern services, for example in health care, and not with the question whether these services are provided under market conditions or not. In consequence, and despite the title of this paper, many issues of output measurement are relevant for the activity as a whole, and not only under conditions of non-market provision.

A good deal of this paper will therefore be devoted to the question of quality adjustment because it is mainly in this context that the notion of outcome enters the picture and has to be distinguished from the notion of output. However, a second conclusion of this paper is that the story does not end here – in

[1] For documentation see http://www.oecd.org/document/47/0,3343,en_2825_495684_37733615_1_1_1_1,00.html.

[2] A first and yet incomplete draft of this handbook can be found on the website for the 2007 meeting of the OECD National Accounts Working Party: http://www.oecd.org/document/42/0,3343,en_2649_34245_38677418_1_1_1_1,00.html.

concept as in practice, it is very difficult to conceive a notion of output that is independent from the notion of outcome – hence we have two distinct, but related concepts.

The principle objective of our efforts here is to shed light on the measurement of volume measures of output for complex and economically important service activities and we shall frame our discussion with examples from health and education, although they may apply more broadly. Terminology is important in this context and we therefore start by stating our definition of 'output' with a view to distinguishing it from 'outcome'. For completeness, we also define 'inputs' and 'activities'.

2. Terminology

Our point of reference is a production framework, i.e., an economic unit that transforms volumes of inputs into volumes of outputs. ***Inputs***, the goods and services to be transformed comprise labour services, capital services and intermediate inputs. Inputs are combined and transformed by way of a production technology. ***Outputs*** are suitably differentiated counts of actions or activities (in the case of services), and counts of physical units (in the case of goods).

Outcome is a state that is valued by consumers – a functioning car, the state of health, the level of knowledge etc. Outcomes are influenced by many factors, and one of them may be the level of outputs. For example, the state of health (an outcome) is a function of medical care (output of the health industry), peoples' lifestyles and the natural environment. Often, outcomes manifest themselves with a considerable lag to the provision of output as would be the case of long-term effects on human health. For all these reasons, outcomes are different from outputs. In principle, the production boundary as defined in the national accounts encompasses outputs, but not outcome. A superficial conclusion would therefore be that the national accounts statistician does not have to worry about outcomes, only outputs.

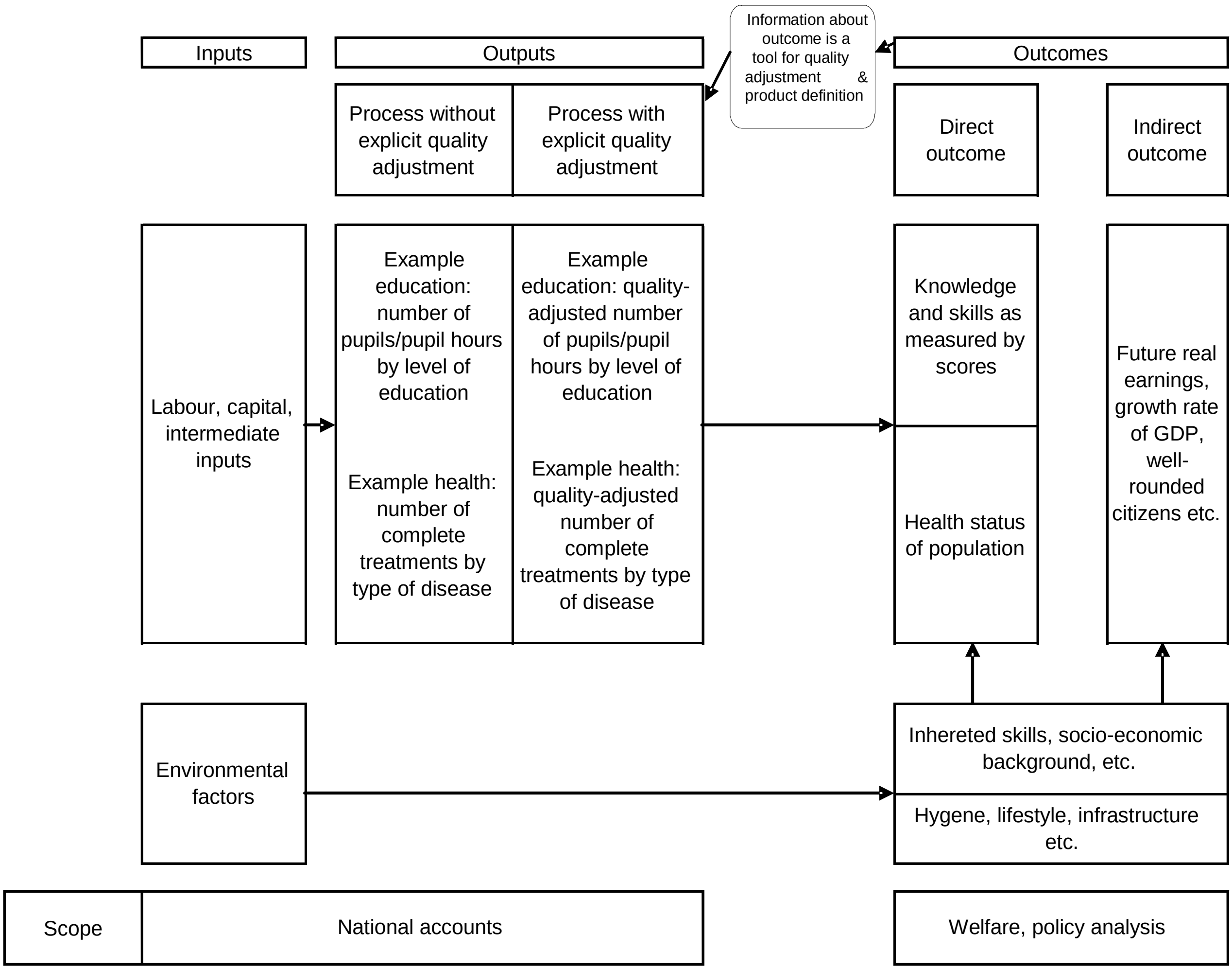
However, things are more complicated. While outcomes are different from outputs, they are not independent. One of the conclusions of the present paper will in fact be that it is virtually impossible - in particular for health and education services - to define quality-adjusted outputs without invoking outcomes one way or the other.

For the forthcoming discussion, it is useful to refine the definition of outputs and outcomes in two ways. First, outputs are broken down into two components: ***activities or processes*** and the ***quality adjustment*** applied to them. Processes are observable and countable actions by which services are delivered although their characteristics may change over time. For education, a typical process measure is the number of pupils or the number of pupil hours taught in a particular grade. For health, a typical process measure is the number of treatments of a particular disease such as hip replacements. A second refinement consists in distinguishing between ***direct*** and ***indirect outcomes***. Direct outcomes are closer to the act of service provision than indirect outcomes although neither is a measure of service nor output itself. For example, in the case of education, a direct outcome is the state of knowledge of pupils, estimated by scores or degrees. The indirect outcome associated with education is employment possibilities and enhanced real earnings due to better human capital. Indirect outcomes associated with health services are fewer working days lost due to diseases, or individual well-being. These distinctions between inputs, activities/processes, quality adjustments and direct and indirect outcomes are shown in Figure 1 below. Figure 1 also depicts the scope of national accounts measures which are defined via the production boundary. However, as has been mentioned above, there may be a need to bring considerations of outcome into the measurement of output to capture the quality of services - this link is also indicated in the figure and will be a central topic in the discussion in the sections to follow.

Different meanings of outcome

"Outcome" has been used in different ways in the relevant literature on non-market 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 how useful a particular definition is for the purpose at hand. The present paper follows in the line of Eurostat (2001) and the Atkinson Review (2005), and employs the term 'outcome' in the sense of the national accounts literature.

Figure 1: Inputs, outputs and outcomes



3. Competitive markets, no quality change in products

The consumer side

Our discussion starts with a simple market model of producers and consumers. For the moment, we take it that products are well-defined and transactions on the market are observable. On the demand side, consumers purchase the goods and services supplied by producers. Standard economic theory attributes a utility function to consumers where *utility* depends on the quantity of goods and services consumed. The utility function indicates how the consumer appreciates (in unobserved 'utils') the quantity of products purchased.

To fix ideas, let households' utility function be $U^t = U(H^t)$, where U^t stands for the level of utility in period t which depends positively on H^t, a state of the world that consumers value. H^t corresponds to our notion of direct outcome. In general, H^t will be vector-valued because there are many different states valued by consumers – for example, different aspect of the status of health, the level of knowledge or the state of the natural environment. 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. We 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. In an application to health care, Berndt, Cutler, Frank, Griliches, Newhouse and Triplett (1998) distinguish between medical care ('output' in our terminology), the state of health ('direct 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.

If we follow this idea, outcomes depend on N different services consumed by households, and we shall label quantities of these services in period t as $\mathbf{y^t} = [y_1^t, y_2^t, y_3^t, \ldots y_N^t]$. Importantly, the outcome variable H^t depends not only on services $\mathbf{y^t}$ that are purchased or obtained from producers but also on a host of other factors, $\mathbf{\Omega}^t$. Examples include households' behaviour with regard to smoking, alcohol consumption or physical exercise when H^t stands for the health status. Or $\mathbf{\Omega}^t$ could stand for student's efforts or natural giftedness when H^t stands for educational attainment and the state of knowledge of the population. We shall come back to these 'environmental' or 'conditional' factors $\mathbf{\Omega}^t$ in the discussion on quality change. For the moment, we formulate our 'outcome' function and insert it into the utility function defined earlier:

(1) $H^t = H(\mathbf{y^t}, \mathbf{\Omega}^t)$ and therefore

$U^t = U[H(\mathbf{y^t}, \mathbf{\Omega}^t)] = V[\mathbf{y^t}, \mathbf{\Omega}^t]$.

Households demand services to the point where prices equal the marginal utility generated by these services. In money terms, this gives $p_i^t = [\delta U/\delta y_i^t]/\lambda^t$. Here, $1/\lambda^t$ is the marginal utility of income that is needed to convert 'utils', the unobserved units of utility, into currency units. It is not difficult to see that the marginal contribution of a good or a service to utility is in fact a composite term, namely the product of the marginal contribution of the product to outcome, $\delta H/\delta y_i^t$, and the marginal contribution of outcome to utility, $\delta U/\delta y_i^t = [\delta U/\delta H][\delta H/\delta y_i^t]$.

Although the above discussion does not give us any immediate insights into measurement, there is one point of importance for our discussion on output and outcome that emerges from the simple consumer demand relation above: the marginal contribution to outcome for a particular product i, $\delta H/\delta y_i^t$, relative to the marginal contribution of another product j, equals the price ratio of the two products:

(2) $[\delta H/\delta y_i^t]/ [\delta H/\delta y_j^t] = p_i^t/p_j^t$.

The implication of this relation is that there is a conceptual link between outcomes and classifications for (consumer) goods and services: in principle, product items should be grouped into a category i if they generate similar contributions to outcome – because then consumers will attach the same value, i.e., the same price to them. Product items they should be grouped into a different product category if their contribution to outcome is different. Note that in our simple presentation here we make the implicit assumption that there is exactly one homogenous unit y_i^t and one price p_i^t. In practice, this is not the case and similar types of products have to be grouped. In the context of an elementary consumer price index, for example, the p_i^t's would be the un-weighted averages of individual items within product group i. Similarly, the y_i^t's for a volume index could be thought of as un-weighted average number of individual items for product type *i*. How individual products ('items') are grouped is a question that has to be answered with respect to the purpose of the price or volume index. Above, it was mentioned that from a consumer perspective, the criterion for grouping individual items is that they potentially offer the same contribution to outcomes, i.e., they satisfy the same or similar consumer needs. Put differently, they are substitutes from a consumer perspective. Conversely, if different items are not interchangeable from a consumer perspective, they should be treated as different products. In the presence of quality change or new and disappearing items, the question of grouping items becomes important. But the point to retain is that the organisation of price or quantity measurement, in particular how products are classified and stratified, cannot proceed without some reference to outcome, if one wants to bring in a consumer perspective.

The producer side

We now turn to producers and take it that their production technology can be represented by a cost function[3] that shows the minimum costs required during a given period to produce a quantity of N products $\mathbf{y^t}=[y_1^t, y_2^t, y_3^t, \ldots y_N^t]$, for a given set of input prices and for a given technology. In the case of health service producers, a particular product could consist of a (complete) treatment of a particular disease, in the case of education it could be the hours of teaching provided for a particular grade. To keep the exposition tractable, it will be assumed that there is a technology set S_i^t for each product that links output y_i^t to a set of M inputs, $\mathbf{x_i^t} =[x_{i1}^t, x_{i2}^t, \ldots x_{iM}^t]$. Inputs are purchased at the prices $\mathbf{w^t}=[w_1^t, w_2^t, \ldots w_M^t]$ on factor markets where producers are price takers. A cost function for product i can then be written as

(3) $C_i^t=C_i^t(y_i, \mathbf{w}) = \min\{\mathbf{w_i} \cdot \mathbf{x} : (y_i, \mathbf{w}) \text{ is contained in } S_i^t(y_i, \mathbf{w})\};\ i=1,\ldots N.$

A unit or average cost function for output type i is defined as $c_i^t(y_i,\mathbf{w_i}) \equiv C_i^t(y_i,\mathbf{w})/y_i$ and measures the costs per unit of output y_i during period t. With constant returns to scale which we shall assume for convenience here, unit costs are independent of the level of output, and average costs equal marginal costs. When producers minimize costs, it follows that total costs for a particular product are equal to minimum costs:

(4) $C_i^t(y_i^t,\mathbf{w^t}) = c_i^t(\mathbf{w^t})y_i^t = \mathbf{w^t} \cdot \mathbf{x_i^t}.$

We follow Diewert (2008) and define the producer's or sector's total cost function C^t as the sum of cost functions for different products. This will be helpful in the definition of index numbers below:

(5) $C^t(\mathbf{y^t},\mathbf{w^t}) = \sum_i^N C_i^t(y_i^t,\mathbf{w^t}) = \Sigma_i^N c_i^t(\mathbf{w^t})y_i^t = \mathbf{w^t} \cdot \mathbf{x_i^t}.$

Under constant returns to scale, and in a competitive market, producers provide services at the point where prices equal marginal costs. More generally, under market conditions, excess profits are competed away and prices will equal average costs so that:

(6) $p_i^t = c_i^t(\mathbf{w_i^t}).$

[3] For more information about cost functions see Shephard (1970) or Diewert (1974).

When the consumer and the producer side are combined, market equilibrium is characterized by

(7) $[\delta U/\delta H][\delta H/\delta y_i^t]/\lambda^t = p_i^t = c_i^t(\mathbf{w_i^t})$.

Thus, marginal utility from consumption and marginal costs of production are equal in equilibrium. There is of course nothing new about this statement which can be found in any micro-economics textbook. It is nonetheless useful because we see that in equilibrium, consumer valuation and producer valuation of a good or of a service coincide at the margin when there is a market-clearing price. The implication is that when weights are needed to aggregate across products, there is no need to invoke either a consumer or a producer perspective – the value of market transactions is all that is needed and it combines the two sides of the market[4]. Also, in the absence of quality change and as long as there are no new or exiting products, it would appear that there is little to worry about the distinction between output and outcome – market prices jointly value the marginal contribution of consumption to outcome and also reflect marginal production costs.

4. Non-market production, no quality change

Having established the competitive case with a well-defined set of products that does not change over time, we shall now direct attention to the case of non-market production. When goods or services are provided by non-market producers, they are provided at a price that does not cover costs and which may even be zero. In this case, the price at which products are transacted loses its significance as an indicator of marginal or average costs. Nor is the price necessarily linked to a utility-maximising quantity of consumer demand, as was the case in the competitive environment. Therefore, our convenient link between the producer and the consumer side, established in the first section, breaks down:

(8) $[\delta U/\delta y_i^t]/\lambda^t \neq c_i^t(\mathbf{w^t})$

An immediate consequence from this situation is that the well-established body of literature on the theory of producer price indices, notably Fisher and Shell (1972), Archibald (1977), IMF et al. (2004) no longer applies. The theory of the output price index relies on revenue functions for producers and stipulates revenue-maximising behaviour, given a set of market prices for producers' outputs. In a non-market environment[5], this is not a useful assumption, and consequently, revenue functions cannot be used as a conceptual basis for output price indices.

[4] This is a simplification. In practice, and in the presence of transport costs or taxes, there is no unique market price and a distinction has to be made between different valuations. For example, from a consumer perspective, a valuation at purchasers prices is appropriate, which is inclusive of taxes and transportation margins. From a producer perspective, a valuation at basic prices would be more appropriate, which excludes for example taxes payable and subsidies receivable in conjunction with production or sale. The statement in the text is also a simplification in the sense that output price indices for producers and input price indices for consumers are typically compiled at different levels of aggregation and on the basis of different classifications. This may also lead to differences in weights between output price indices for producers and input price indices for consumers. However, even when aggregation happens differently in the producer and in the consumer case, the price at the lowest level of aggregation at which the transaction takes place, is a market price and reflects the joint influences of producers and consumers.

[5] Note that despite the fact that our discussion has been couched in terms of non-market producers, it carries over to the more general case of regulated industries. For example, Lawrence and Diewert (2006), measure the quantity index of output for New Zealand electricity utilities with a cost-based index because there are no meaningful revenue shares or prices for the three types of outputs identified for utilities: throughput of electricity, system line capacity and connections.

However, measurement can be based on *cost-based or quasi prices*. When transaction prices are significantly below cost or zero, it is customary in the national accounts to measure the money value of output as the sum of costs. One could also say that output is valued at quasi prices. They are those (unobserved) prices that emulate a competitive situation where prices equal average costs per product. With unit costs at hand, they can be treated *as if they were prices*:

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

If we maintain the (courageous) assumption[6] that non-market producers are cost-minimising units, then minimum costs equal actual costs or $c_i^t(\mathbf{w^t})y_i^t = \mathbf{w^t} \cdot \mathbf{x_i^t}$, and it follows that

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

Expression (10) states the obvious, namely that with cost-based 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*[7]. What is important for the purpose at hand is the fact that this equality of inputs and outputs in value does *not* imply equality of inputs and outputs in volume or quantity. If this were the case, our efforts to derive volume measures of output that are separate from volume measures of inputs would be put in question and with it any attempt to measure productivity change in non-market production.

The main difference between cost-based prices of outputs 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 or the costs for one year of schooling) whereas the latter correspond to the *costs per unit of input* (such as wages per hour of a nurse or the salary of a teacher).

Diewert (2008) shows formally how a cost-based volume index of output can be defined. He proposes a family of cost-based output quantity indices and focuses on the Laspeyres (Q_L), on the Paasche (Q_P) and on the Fisher (Q_F) case. In line with the economic approach towards index numbers, Diewert 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:

(11)
$$Q_L = C^0(\mathbf{y}^1, \mathbf{w}^0) / C^0(\mathbf{y}^0, \mathbf{w}^0) = \Sigma_i^N c_i^0 y_i^1 / \Sigma_i^N c_i^0 y_i^0$$
$$Q_P = C^1(\mathbf{y}^1, \mathbf{w}^1) / C^1(\mathbf{y}^0, \mathbf{w}^1) = \Sigma_i^N c_i^1 y_i^1 / \Sigma_i^N c_i^1 y_i^0$$
$$Q_F = [Q_L Q_P]^{1/2}.$$

[6] The main advantage of this assumption is that it allows a conceptually clean identification of productivity change with technical change. Suppose that cost changes are broken down into changes in the quantity of output, changes in the prices of inputs and productivity change as in Diewert (2008). Then, one of the conditions that the measured productivity change is a reflection of technical change (or a shift of the cost function) is that the assumption of cost-minimising producers holds, If this assumption is relaxed, the measured productivity change can encompass several elements, in particular a movement towards or away from the efficiency frontier and a movement of the cost function itself. As our present concern is with the measurement of output and only indirectly with the measurement of productivity, we have maintained the assumption of cost-minimisation for reasons of simplicity. Giving up this assumption would complicate the exposition and would require more discussion by way of applicable index number formulae. See also Balk (1998) for productivity measurement when producers are not acting as cost minimisers.

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

For a number of practical reasons, we prefer working with a ‘price’ rather than a quantity index à la Diewert for non-market producers and then deflate total costs by a price index. To this end, we construct an *indirect index of quasi prices* by dividing total costs by the volume index of output:

$$(12) \qquad P_L = [C^1(\mathbf{y}^1,\mathbf{w}^1)/C^0(\mathbf{y}^0, \mathbf{w}^0)]/Q_P \quad = \Sigma_i^N c_i^1 y_i^0/\Sigma_i^N c_i^0 y_i^0$$

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

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

Note a useful interpretation of this quasi-price index that is obtained by re-writing the Laspeyres or Paasche version in expression (12). For example, after inserting the theoretical expression for Q_P into the first line of (12), P_L can be presented as the product of two terms:

$$\begin{aligned}(13) \qquad P_L &= [C^1(\mathbf{y}^1,\mathbf{w}^1)/C^0(\mathbf{y}^0, \mathbf{w}^0)]/Q_P \\ &= [C^1(\mathbf{y}^1,\mathbf{w}^1)/C^0(\mathbf{y}^0, \mathbf{w}^0)]/[C^1(\mathbf{y}^1,\mathbf{w}^1)/C^1(\mathbf{y}^0,\mathbf{w}^1)] \\ &= [C^1(\mathbf{y}^0,\mathbf{w}^1)/C^0(\mathbf{y}^0, \mathbf{w}^0)] \\ &= [C^1(\mathbf{y}^0,\mathbf{w}^1)/C^1(\mathbf{y}^0, \mathbf{w}^0)][C^1(\mathbf{y}^0,\mathbf{w}^0)/C^0(\mathbf{y}^0, \mathbf{w}^0)]\end{aligned}$$

The first term in the last line of (13) is an 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 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: we recall that in a market situation, and under competition, 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 present non-market case, the quasi-price index for outputs plays a similar role as the market price index for outputs in a market situation. If unit costs rise less rapidly than input prices, there has been productivity change.

In fact, the measurement of productivity as a shift in the cost function is a well-established methodology. Whenever there are situations of imperfect competition, or non-constant returns to scale in production, there are at least two ways of measuring productivity change – as a shift of the production possibility curve or a shift of the cost function[8]. The two approaches do not generally yield the same result but there is no strong a-priori reason to prefer one over the other. We conclude that the cost-based productivity measure is a fully valid measure of technical change for non-market producers.

But (13) also shows that despite that 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. But 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. The output perspective remains a proxy only insofar as it brings in no direct consumer valuation – unit costs are not a product of the interplay between producers and consumers as in the market case but only driven by the supply side. However, some elements of consumer valuation enter the picture when it comes to quality adjustment and this is the topic of the section to follow.

[8] 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.

5. Non-market production and quality change

Direct outcome, stratification and implicit quality adjustment

There is an extensive literature on how to deal with quality change in existing products, with the exit of old products and with the entry of new products if one wants to compile price or quantity indices. Quality change counts among the most serious measurement issues in estimating price indices. Early references include Stone (1956), Griliches (1971), and more recent ones IMF et al. (2004), as well as Triplett (2006). The reader is referred to these volumes for a complete discussion. The task in this section is twofold: discuss how the measurement of quality change relates to outputs and outcomes, and propose a method for quality adjustment for the non-market case. To start, recall some key principles and methods that are followed in measuring quality change.

> "Agencies that estimate price indexes employ, near universally, one fundamental methodological principle. The agency chooses a sample of sellers […] and of products. It collects a price in the initial period for each of the products selected. Then, at some second period, it collects the price for exactly the same product, from the same seller, that was selected in the initial period. The price index is computed by matching the price for the second period with the initial price, observation by observation or 'model by model' as it is often somewhat inaccurately called." (Triplett 2006, p.15)

One technique to deal with quality change in products is thus 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 in education are establishments that provide different services in addition to education, such as boarding schools as opposed to day-time schools or hospitals with different levels of non-medical services. The price of a particular treatment would then be followed inside the group of establishments of a particular type. This is a way of controlling for quality characteristics and to track a price for a constant quality service.

Note, however, that grouping also relies on an important assumption: to show a price or quantity movement that is representative of a product group, the price or quantity movements of those products that *are* matched have to be a good indicator of the price or quantity movements of those products that are *not* matched – in particular products that are newly entering the market. Price or quantity changes that arise outside the sample of matched products are ignored.

The non-market case on which our attention is focused here shows also the importance of choosing the right level of aggregation where matching takes place. And it is again considerations of direct outcome that govern this choice. Take the case that is described in Box 1. Two medical procedures are considered, of different unit costs, but equally interesting from the consumer's viewpoint – both procedures treat the same medical problem equally well. In other words, the contribution of each treatment to outcome, from a consumer perspective, is the same. Treating each procedure as a separate product, i.e., setting the elementary level for the construction of the price or volume index below each procedure can lead to counter-intuitive results: more of the cheaper but equally helpful treatment translates into a decline in the volume of output because, in a non-market context, the new procedure only gradually replaces the old procedure. When both procedures are treated as providing the same service and at the elementary level both procedures are treated as the same product, some of the bias is eliminated. But no judgement about substitutability can be made without invoking at least implicitly, some judgement about outcomes. By definition, (complete) substitutability of services implies that they are equally appreciated by consumers, in other words, they generate the same outcome.

We conclude that even in a situation where matching between products is perfect, and no explicit quality adjustment may be needed, *some* reasoning about outcome is in place, if only to group substitutable products together in one stratum.

But matching is rarely perfect. Matching of the quasi prices or volumes of non-market services such as health and education services is unlikely to control completely for particular characteristics associated with the provision of these services. Or a representative service may change its characteristics, akin to a product that price collectors are no more able to find in a particular outlet and that has to be replaced with a new product. These are the instances where explicit quality adjustment comes into play, of which more later.

Indirect outcome and quality adjustment through re-definition of products

Jorgenson and Fraumeni (1989, 1992) were the first to apply a human capital approach to the measurement of output of the education sector of the United States. Their approach constitutes a particular way of accounting for quality change and despite the fact that the human capital approach cannot easily be generalised to other services, it is worth putting the work by Fraumeni and Jorgenson into the context of our present discussion. The approach has a clear theoretical foundation and has stood the test of empirical implementation.

At the core of the human capital approach towards measuring education output lies the idea that educational services are investment flows that add to human capital. Private households, it is assumed, demand educational services to the point where the marginal costs of an extra year of education (in the form of fees or income foregone, for example) equal the marginal benefits from education. The latter are measured as the discounted differentials in lifetime income due to the additional year of schooling. This supposes in turn that wages on labour markets correspond to the marginal productivity of workers – in this way, the relative level of human capital and consequently the relative level of worker's marginal productivity can be gauged via relative additions to discounted life income of workers. The quantity measures that enter the calculation are a set of student enrolment numbers, stratified by various criteria such as level of education, type of studies, gender and so forth. At this, lowest level of aggregation, quantity measures for the education services provided are not different from the unadjusted quantity measures that formed the body of our discussion earlier and which would be matched over time to control for quality change. However, under the human capital approach, the same quantities are now quantities of investment goods and need to be matched by investment goods prices for aggregation, not by unit costs as was the case earlier.

Let there be I different education services and let the price of the educational investment for each particular type of education in period t be $p_{edu,i}{}^{t}$ (i=1,2…I). The value of $p_{edu,i}{}^{t}$ is not a market value that is readily observed but has to be imputed. Fraumeni and Jorgenson (1989) show how to implement such computations empirically. Let the quantity of the educational investment be $y_i{}^t$, so that the value of education services is $\Sigma_i{}^I p_i{}^t y_i{}^t$. The Laspeyres, Paasche and Fisher-type volume index of education services between periods 1 and 0 are then:

$$
\begin{aligned}
(14) \qquad Q_{L,edu} &= \Sigma_i{}^I p_{edu,i}{}^0 y_i{}^1 / \Sigma_i{}^I p_{edu,i}{}^0 y_i{}^0 \\
Q_{P,edu} &= \Sigma_i{}^I p_{edu,i}{}^1 y_i{}^1 / \Sigma_i{}^I p_{edu,i}{}^1 y_i{}^0 \\
Q_{F,edu} &= [Q_{L,edu}\, Q_{P,edu}]^{1/2}.
\end{aligned}
$$

The different y_i's are thus valued with the corresponding returns on extra education. When the relative valuation of the different types of education changes, the volume index of education services changes. For example, if the returns to one extra year of tertiary education rise quicker than the returns to an extra year of secondary education, and if more students undergo the extra training, the resulting change in the volume index would not only capture an increase in years of schooling but also reflect the fact that the average quality of human capital has risen. In this sense, the human capital approach adjusts volume measures of education output for changes in the quality of human capital. This is achieved by re-defining education output as the production of an investment good and by imputing the price of this investment good.

Note, however, that the human capital approach assumes that at the lowest level of aggregation, the quantities y_i^t are not subject to quality change – one hour of a particular course taught this semester equals one hour of the same course taught the following semester.

How does the human capital approach relate to outcome? There is no formal link here except that the quantity of human capital should in some way be related to the stock of knowledge of humans having undergone education. But in terms of the underlying economic theory, the human capital approach does not really need a variable H as in our formulation earlier. Human capital is only a means to insure wage income on the labour market. Households' utility depends on the goods and services that can be purchased with these labour incomes, but there is no need for an explicit introduction of the stock of knowledge as an argument in the utility function. Outcome only enters in a monetary sense, as the reward on the labour market. In terms of the terminology introduced at the beginning of this paper, we are dealing with indirect outcome.

In summary, the human capital approach is an example of how certain (important) aspects of quality change can be captured by re-defining education output as an investment good and by relying on labour market outcomes for its valuation[9]. But it is also a fact that not all aspects of quality can be captured. Further, the human capital approach has to rely on some strong assumptions. In particular, it has to be true that observed differences in remuneration on labour markets are attributable to differences in educational attainment. Finally, there is no immediate way of applying a similar approach to other areas of non-market services such as health or general government. Further, future incomes are not the only motivation for education - learning per se and the nature of future work can give satisfaction. This aspect cannot be captured with discounted lifetime incomes.

Outcome and explicit quality adjustment

If matching is insufficient to control for key characteristics of service provision, other, explicit, techniques have to be invoked to account for quality change. In general, the quality of a product can be expressed by the quantity of its characteristics. Quality change can then be captured by the change in characteristics. Similarly, price changes in products can be attributed to pure price changes and to those

[9] We note several accounting issues that would arise if the human capital approach were implemented in the national accounts. First, there is a question about the scope of inputs to and outputs from the production of educational investment goods. For example, Jorgenson and Fraumeni (1989) consider students' time as one of the inputs into the production process. On the output side, they value not only future income from labour market activities but they also value non-market labour. Other authors, for example Ervik, Holmoy and Haegeland (2003) have argued that in a national accounts context, students' time should not be part of inputs. They also exclude non-market labour from the output computations. This can lead to significant differences in the measured value of production of the education sector. A second issue is that the value of the education output at current prices under the human capital approach (which could be seen as a social gross production) is not necessarily equal to the value of inputs. This implies a kind of gross operating surplus for non-market producers that is hard to deal with in an established accounting framework. Ervik, Holmoy and Haegeland (2003) estimate the human capital-based output for the Norvegian education sector to be around NOK 77 billion in 1995, compared to the cost-based national accounts estimate of about NOK 14 billion.

price changes that reflect changes in product characteristics. This is the approach followed by hedonic price indices[10] that are now well established among statistical agencies.

Quality adjustments require the identification of a set of characteristics such as the speed, engine size or equipment for a car or the processor speed for a computer. 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'. An important result of the hedonic model is that it allows the identification of characteristics and provides a market valuation of each one[11]. Market valuation, in turn, is a convenient way of aggregating across characteristics because everything is expressed in a single monetary unit.

At first glance, hedonic regressions would appear to be badly suited for a non-market environment, given that hedonic methods are all about extracting 'market' information from examining the link between market prices and product characteristics. At a second glance, however, the basic principles of hedonic regressions *would* appear to be adaptable for a non-market environment, albeit with a different interpretation.

To demonstrate this idea, we shall augment the simple producer model above by product characteristics. Thus, for every product y_i^t (i=1,2,…N) there is a vector of n(i) characteristics $\mathbf{z_i^t} \equiv [z_{i1}^t, z_{i2}^t, \ldots z_{in(i)}]$ that qualifies this product. Characteristics play a double role:

First, they help dealing with heterogeneity. Truly homogenous products are rare if they exist at all, and every type of product contains a more or less important element of heterogeneity. For example, if health product y_1 is "treatment of a stroke", this will encompass strokes of different severity, and patients of different age suffering from a stroke. If old patients require more intense treatment than young patients or if more severe strokes necessitate more intense care than less severe strokes, there are in fact different services involved. To some degree, this can be accommodated by stratification and matching (see above) but only up to a point. Additional heterogeneity is best captured by identifying, through knowledge of the product, those characteristics that make one service distinct from another. The variable z_{11} could thus be 'age of patient', the variable z_{12} 'degree of severity of stroke' and so on.

10 See Griliches (1988) for an introduction and Triplett (2006) for a comprehensive discussion.

11 Rosen (1974) demonstrated that in general, those characteristics of a product will show up in the function that are valued by consumers *and* that have cost implications for producers. Triplett (2006) writes on this: "It is well-established – but still not sufficiently understood – that the functional form of [the hedonic equation] cannot be derived from the form of the utility function or of the production function. Neither does [the hedonic equation] represent a 'reduced form' of supply and demand functions derived from [the utility and the production functions] as the term is conventionally used. Establishing these results requires consideration of buyer and seller behaviour towards characteristics]" (page 231).

Box 1. Why narrow specifications of products may not always be sufficient to capture quality change

A straight forward technique of dealing with quality change in a price or in a volume index is to match models, i.e., to compare only prices or quantities of products that are tightly specified. In other words, products are treated as different products whenever their characteristics are different. The more specific the characteristics of a particular product, the less likely it is that a modification of the product goes unnoticed and that a change in quality is not recognised as such. Such implicit quality-adjustment is well adapted when the set of products observed is stable and when it is representative for the universe of products. It may, however, be insufficient, when products change, when there are substitution processes between new and old products and when there are no markets or when existing markets operate imperfectly. This is best illustrated by way of an example. We use a quantity index here but the same points could also be made by way of a price index that is subsequently used to deflate values.

Suppose there are two treatments for a disease, traditional surgery and laser treatment, and assume that laser treatment is introduced in period 1. In addition, as may well be the case, the unit cost of laser treatment is lower than the unit cost of traditional surgery. The total number of interventions remains the same.

	Traditional surgery			Laser surgery		
	Period 0	Period 1	Period 2	Period 0	Period 1	Period 2
Unit cost	100	100	100	-	90	90
Number of interventions	50	40	5	0	10	45
Total cost	5000	4000	500	0	900	4050

Now consider a *matched-model approach* towards calculating a volume change from period 0 to period 1. In the simplest case, the volume index is given by the quantity changes in the two treatments, each weighted by the cost share it occupies in period 0. As laser surgery does not yet exist in period 0, it receives a zero weight so that the volume index of treatments is simply the change in the number of traditional surgery, or (40/50-1)=-20%. Between periods 1 and 2, the corresponding volume index equals $[s_T(5/40)+s_L(45/10)]-1=-7.1\%$ where $s_T=82\%$ and $s_L=18\%$ are the period 1 cost shares of the traditional and of the laser treatment respectively. This approach treats the two treatment as different products and the sharp drop in the total volume index in period 1 reflects the 'new goods' problem that arises when new products enter the sample that cannot be compared with quantities in the base period. Note also the assumption implicit in this model: consumer valuation of the two products is captured by the relative unit costs, so if laser surgery is cheaper than traditional surgery, this method implicitly quality-adjusts *downward* the quantity of laser surgery when it is combined with traditional surgery. In a perfectly operating market, the price of the traditional treatment would see an instantaneous downward adjustment, bringing consumer valuation of the two processes in line but in the practice of health services provision, in particular in a non-market context, this would appear an unrealistic assumption.

A different result arises when it is considered that the two treatments are perfect substitutes, i.e., that they are in fact the same product. In this case, no cost weighting is applied between the two treatments – and the number of treatments is simply added up. As there are 50 interventions in every period, the result is a volume index that shows zero growth and a declining price index, reflecting the drop in average unit costs of treatment.

The previous method is justified if it is plausible that consumers are indifferent about the two treatments. If this is not the case, and they prefer laser over traditional surgery because the former is less intrusive or requires fewer days of recovery, an explicit quality-adjustment is needed. Such an adjustment can be applied to the quantity measures, either by scaling up the quantity of laser treatments or by scaling down the quantity of traditional treatments. Whichever way this is done, the implication is always that one treatment is expressed in equivalents of the other treatment, and the ratio should in some way reflect consumer preferences. Alternatively, prices or unit costs could be rescaled before constructing a price index. Suppose the adjustment factor is 1.1 – each laser treatment is the equivalent of 1.1 traditional treatments. Then, expressed in 'traditional surgery-equivalents', the number of treatments is 50 in period 0, 40+10*1.1=51 in period 1 and 5+45*1.1=54.5 in period 2. The resulting volume index is +2% in period 1 and +6.9% in period 2. Obviously, the difficulty lies in determining the adjustment factor which should (i) reflect consumer preferences; (ii) be uni-dimensional.

Second, characteristics help dealing with quality differences. In the case of health services, for example, observations on the same type of treatment may come from different institutions for which there exists additional explicit information on characteristics that are considered attributes of quality from a consumer perspective, for example the in-hospital waiting time after hip fracture[12]. Waiting time is an

[12] This is an example from the OECD Health Care Quality Indicators, see Garcia Armesto, Lapetra, Wei and Kelley (2007).

example of a direct measure of a quality characteristic. Other characteristics may only be captured indirectly. For example, the treatment of a particular disease may or may not include the use of a scanner. If there are reasons to believe that the use of certain types of modern equipment increases the quality of diagnosis, this is a relevant variable to be picked up. Yet another type of characteristic may have to do with secondary products or amenities provided by the unit that produces services. For example, the same medical service can be provided in hospitals of very different classes: multiple versus single-patient rooms, quality of meals, equipment of rooms with phones and TV sets, etc. The total average costs per medical act from two different establishments cannot be compared unless one controls for these characteristics. A similar case can be made for education establishments with regards to the provision of meals in school or the leisure activities on offer.

A very important characteristic for the health and education area is the number and quality of personnel per patient or per student. Class size, for example, has long been quoted as an important quality characteristic of schools. Waiting time, if not measured directly, can probably be tracked well by the number of nurses or doctors per patient and is an obvious quality characteristic for medical services. However, we have to be careful in introducing input measures lest there be a danger of falling back into measuring the volume of outputs by the volume of inputs. And this is exactly what we had set out to get away from in the first place! Which characteristics are admissible? And how should they be introduced?

In a market situation, characteristics should be price-determining. In the first place this means they are valued by consumers but it also implies that there is a cost associated with the provision of characteristics, otherwise they would have no impact on prices. By analogy then, in the non-market case, the choice of characteristics should be such that they are primarily relevant for consumers, i.e., they are utility-enhancing. Characteristics that are only or predominantly relevant for producers such as whether catering is carried out in-house or whether it is outsourced, should not be candidates for characteristics. There is no single, universal rule here and the choice of characteristics, as in all hedonic regression studies, is a matter of knowledge of the product. The choice of characteristics brings us back to the notion of outcome. Identification of relevant characteristics cannot proceed without recourse to outcome – only outcome considerations from a *consumer perspective* will help getting on with the choice of quality characteristics to be included in the hedonic regression in a non-market context.

It is worth noting that this is not an issue specific to non-market services although it may be more accentuated in this context. Empirical difficulties with the choice of characteristics are also an important issue for market-provided products, such as computers in the following reference (Triplett 2006):

> "In the end, economic theory does not specify the characteristics. Choosing the characteristics requires marketing or engineering or other information about the product and what buyers want to do with it." (page 167).

The difference to the non-market case is that in a market context, the hedonic regression provides a market valuation of the characteristic in question, i.e., a valuation that arises as the intersection between consumers and producers. In the non-market context, the valuation is confined to the cost side. It could be seen as the marginal valuation of the characteristic by the non-market institution or by the government through its choice of allocating resources to different characteristics.

Having discussed the nature of characteristics $\mathbf{z_i}^t$, we shall now link them to the cost function that constitutes the measurement framework. Recall that for every product i, there was a unit cost function. y_i^t remains the right measure for the volume output but we now have to allow for the fact that y_i^t may be not directly observable. We shall call the observable, but not quality-adjusted measure of output ζ_i^t. For example, ζ_i^t could be the total number of stroke treatments in different hospitals, irrespective of the age of patients or of the severity of cases treated. To model quality change one can say that the quality-adjusted, but unobserved output y_i^t equals a measure of unadjusted but observable output ζ_i^t , multiplied by a quality-adjustment function g_i^t which in turn depends on the characteristics $\mathbf{z_i^t}$ that attach to each product.

Characteristics should be defined in such a way that a bigger quantity of characteristics is associated with higher quality.

(15) $y_i^t = \zeta_i^t\, g_i^t(\mathbf{z_i^t})$.

Total costs for a non-market product should remain invariant to how they are split up into a unit cost and a quantity component. This requires that

(16) $c_i^t(\mathbf{w^t}) y_i^t = u_i^t \zeta_i^t$,

where $u_i^t \equiv c_i^t y_i^t / \zeta_i^t$ has been defined as the unadjusted unit cost for product i. In combination with (15) one obtains

$$
\begin{aligned}
(17)\qquad u_i^t &= c_i^t(\mathbf{w}^t) y_i^t / \zeta_i^t \\
&= c_i^t(\mathbf{w}^t)\, g_i^t(\mathbf{z_i}^t), \text{ or, taking logs,} \\
\ln u_i^t &= \ln c_i^t(\mathbf{w}^t) + \ln g_i^t(\mathbf{z_i}^t).
\end{aligned}
$$

The last expression is a hedonic function[13]. Quality-unadjusted measures of unit costs of non-market output are broken down into a term that reflects the true, quality-adjusted unit costs and a term that captures the effects of characteristics. When set up as a hedonic regression, the relationship (17) constitutes an empirical avenue towards correcting for differences in quality or more general, for differences in characteristics between different observations within the same type of product. There is a vast literature on hedonic regressions and price measurement[14] and we shall not expand on it here beyond providing a sketch for an implementation method.

In a non-market context, one can think about $\{u_{ik}^t\}$ as a set of observations on unadjusted unit costs for a particular product, for example costs per cataract treatment in periods t=0,1 from a cross-section of hospitals, where each observation is indexed by $k=1,2,\ldots K_i$. A hedonic regression that uses the dummy variable method then takes the following form:

(18) $\ln u_{ik}^t \;=\; \beta_0 + \alpha D^t + \sum_n \beta_{in} \ln z_{ink}^t + \varepsilon_k^t$.

Here, the hedonic equation in the last line of (17) has been transformed into an implementable form by setting the quality-adjusted unit costs $\ln c_i^t$ to equal the regression parameters $\beta_0 + \alpha D^t$ where D^t is a dummy variable that equals 1 in period t=1and that equals zero in period t=0. The quality adjustment term $\ln g_i^t(\mathbf{z_i^t})$ has been specified as $\sum_\mathbf{n} \beta_{in} \ln z_{ink}^t$ and a statistical error term ε_k^t has been added. With suitable econometric methods, parameter estimates of the various coefficients are obtained. Of particular interest here is the parameter α that comes with the time dummy[15] variable. In antilog form, it represents the quality-adjusted index[16] of unit costs for product i:

[13] "A hedonic function is a relation between prices of varieties or models of heterogenous goods or services and the quantities of characteristics contained in them." (Triplett 2006, page 229).

[14] Court (1939) is frequently credited to have been one of the first researchers to use a hedonic function for purposes of price measurement. Good introductions and overviews of the topic can be found in Triplett (2006), ILO, IMF, OECD, UN-ECE and World Bank (2004), and ILO, IMF, OECD, UN-ECE, Eurostat and World Bank (2004).

[15] To simplify notation, we made no distinction between theoretical parameter values β_0, α etc. and estimated parameter values as they are obtained from the regression.

[16] Triplett and McDonald (1977) showed that the specification of the hedonic regression implies a certain index number formula in the aggregation of the individual observations. For the double-log specification used here, one gets $\exp(\alpha)=[\Pi_k (u_{ik}^1)^{1/K(1)} / \Pi_k (u_{ik}^0)^{1/K(0)}]/QA$. QA is a quality adjustment term of the following form $QA=\exp\{\sum_n \beta_{in}[\sum_k \ln z_{ink}^1/K^1 - \sum_k \ln z_{ink}^0/K^0]\}$. This is a term that captures differences in characteristics

$$(19) \qquad \ln c_i^1 - \ln c_i^0 = (\beta_0 + \alpha) - \beta_0 = \alpha.$$

Hedonic coefficients are thus a tool for estimating the quality-adjusted rate of change in unit costs. In a market context, hedonic coefficients reflect the valuation of characteristics by the interaction of consumers and producers (Rosen 1974). When competition is imperfect, interpreting coefficients becomes more difficult because in addition to consumer preferences and producer costs, the degree of market power affects the size and possibly the sign of coefficients (Pakes 2003). In a non-market setting, there is no direct market interaction between consumers and producers. Resources are allocated by producers, more or less in response to demand but the valuation of characteristics is a cost valuation. Thus, coefficients in the hedonic equation above provide an indication about the technology of supply of characteristics and its cost structure. Supply in a non-market context does not interact with demand. In such a situation, regression coefficients are indicative of costs or producer valuation. However, despite the fact that the *valuation* of characteristics is a cost valuation, the *choice* of characteristics should largely reflect a consumer perspective as was explained earlier.

The above sketch of a hedonic method has not been tested and conveniently ignores the many empirical problems that its implementation would likely have to face. Data requirements are probably enormous and a core question, the choice of characteristics, makes this approach a challenging task. Nonetheless, exploratory studies would be helpful to assess feasibility and usefulness.

Output as the marginal contribution to direct outcome

One consequence of (8) was the inapplicability of the theory of output price indices. We solved this by reverting to cost-based or quasi prices. In so doing, we implicitly signed up to another decision on how to deal with the inequality of marginal utility and marginal costs: cost-based output prices (or cost-based output volumes à la Diewert (2008)) imply that weights are cost shares of goods and services, and not utility shares. When it came to quality change, we stuck with a cost-based approach but applied explicit quality adjustment to unit costs, and implicitly, to volume measures of output. Outcomes played an indirect role only in the sense that they helped stratifying activity counts and choosing the set of characteristics for quality adjustment.

An alternative way of constructing volume and price indices of output consists of directly invoking the consumer perspective and constructing measures of output that are based on a product's marginal contribution to outcome. Atkinson (2005) puts this forward as one of the methods of measuring the output of non-market producers. In our notation, this would mean tracking movements of output through observation of movements of outcomes caused by the change in output, i.e., $\delta H/\delta y_i^t$. One obvious advantage of this procedure is that – at least in concept – it provides a solution to the problem of quality adjustment. If all that consumers care about is direct outcome, and we are able to trace outcome and measure its change then there is no need to explicitly quality-adjust measures of production.

Under an outcome-based approach, the volume of output is taken as the marginal contribution of the product to outcome as perceived by the consumer. Consider first a non-market producer, whose value of production has been set to equal costs. With this constraint, total costs would equal the total value of each product's contribution to outcome:

$$(20) \qquad \text{Total value of production}_i^t = u_i^t \zeta_i^t = (\delta U/\delta H^t)/\lambda^t \, (\delta H^t/\delta y_i^t) y_i^t.$$

between the two periods, where each difference gets weighted with the corresponding regression coefficient. In the absence of changes in characteristics between the two periods, and no change in the number of observations, QA equals unity.

As before, we allow for quality change in the provision of goods or services. Thus, a direct comparison of unadjusted unit costs over time, u_i^1/u_i^0 would give rise to a biased price index and a comparison of unadjusted quantities ζ_i^1/ζ_i^0 would give rise to a biased volume index because the quality of the observed units ζ_i^t changes over time. At the same time, all that can be directly observed are the unadjusted unit costs along with the unadjusted quantities. An outcome-based approach towards measuring prices and volumes deals with quality change in the following way: the money value of the contribution to utility that a change in outcome generates, is taken as the 'true' price $p_{W,i}^t \equiv (\delta U/\delta H^t)/\lambda^t$. On the volume side, the marginal contribution of the good or service to outcome would constitute the 'true' quantity $y_{W,i}^t \equiv (\delta H^t/\delta y_i^t)y_i^t$. By way of example from health services, $p_{W,i}^t$ would correspond to the net gains in welfare from an additional life-year and $y_{W,i}^t$ would correspond to the number of life-years gained from medical service of type *i*. Note that when the usual assumption of declining marginal utility applies, the true price of health services $p_{W,i}^t$ declines with an improved state of health, as $\delta^2U/\delta H^2<0$, everything else being equal. Although not directly observable from transactions on the market, these 'true' prices and quantities can in principle be estimated.

In the area of health care, there is an increasing number of studies that have used measures of direct outcome to value health care output[17]. For example, Cutler, McClellan, Newhouse and Remler (1998) pursue this research avenue and derive a price index for heart attack treatment. Conceptually, their cost-of-living index is comparable to the change in $p_{W,i}^t$. It is measured as (one minus) the net benefit from a change in heart attack treatment. If the benefits from medical practice changes are greater than the costs, consumers are better off and the cost of living declines. Empirically, then, the authors measure $p_{W,i}^1/p_{W,i}^0-1$ via estimating the net benefits from improvement in health due to the treatment of heart attacks. Given an estimate for $p_{W,i}^1/p_{W,i}^0$ and given the constraint (11), we can work out the implied quality adjustment factor to the observed quantities:

$$(21) \qquad y_{W,i}^1/\, y_{W,i}^0 = \quad [\zeta_i^1/\zeta_i^0][(u_i^1/u_i^0)/(p_{W,i}^1/p_{W,i}^0)].$$

This shows that the index of quality-adjusted volumes of medical services $y_{W,i}^1/\, y_{W,i}^0$ corresponds to the index of unadjusted quantities ζ_i^1/ζ_i^0 times a quality adjustment factor $(u_i^1/u_i^0)/(p_{W,i}^1/p_{W,i}^0)$ that is just the ratio between the index of unadjusted unit costs and the cost of living index[18].

It is useful to compare this result with the hedonic adjustment described in the previous section. Expressions (17) and (19) taken together yield a formally similar quality adjustment where the regression coefficient α replaces the cost of living index $p_{W,i}^1/p_{W,i}^0$:

$$(22) \qquad y_{,i}^1/\, y_{,i}^0 \quad = \quad [\zeta_i^1/\zeta_i^0][(u_i^1/u_i^0)/\exp(\alpha)].$$

Despite the formal similarity, there is of course no reason why the two adjustments should be equivalent empirically. The quality adjustment factor in (21) is based on a non-parametric calculation that relies on a one-dimensional measure – outcome - that is supposed to subsume all quality aspects of production. The quality adjustment in (22), on the other hand, is based on a parametric calculation that uses multi-dimensional quality characteristics, to which cost weights are applied. The link to outcome remains indirect in that only the choice of characteristics reflects considerations of outcome but no direct measure of outcome enters the set of characteristics.

---

[17] See also the volume edited by Berndt and Cutler (2001) for other examples of new medical care price indices.

[18] Triplett (1998), referring to the cost-of-living adjustment by Cutler and al. (1998) points out that there is great reluctance to put a monetary value on lives saved or extended and yet such a valuation is necessary to measure the price change $p_{W,i}^t/p_{W,i}^{t-1}$. He suggests the direct use of cost-effectiveness studies in health care by using an adjustment factor that is based on Quality-Adjusted Life Years (QALY).

Outcome-based measures of health services: a UK example

The U.K Department of Health (DH) estimated outcome-based measures of production in two areas where there is evidence of health gain being significantly higher than cost; one of them is prescription of statins.

"Statins reduce cholesterol, which can block arteries, and so reduce the risk of heart disease and stroke. DH assesses the health gain from statin therapy in terms of the risk of coronary heart disease (CHD) for different groups, based on analysis of the Health Survey for England (HSE) in 1998 and 2003, which contained questions and clinical measures relevant to CHD. The estimates of health gain take account of HSE evidence on preexisting CHD and stroke, diabetes (which increases risk of CHD), total cholesterol, HDL cholesterol, blood pressure, smoking status, age and sex. This approach makes it possible to estimate the health gain from statins as used in the population, making use of evidence from clinical trials on effectiveness of statins for different groups. DH estimates that statin therapy in 2003 added 77,000 life years, compared with no therapy, for the 1.9 million patients who took the drug. It estimates that each prescription has a marginal benefit of 0.0038 life years. If each life year is valued at £30,000, the value of each prescription is £115. This compares with £14 as the unit cost in 2005/06. Further Developments updates previous work and estimates that growth in health care output would be higher by 0.7 – 0.9 percentage points a year, on different assumptions if a value weight replaced the cost weight." (Lee 2008)

There are many practical considerations before any of these approaches can be implemented and output measures that rely on the marginal contribution to outcome such as in (21) are a long way from being regular features in statistical programmes. Note the following empirical issues.

It is not obvious how to identify the marginal impact of output to outcome. As was mentioned earlier in this paper, outcomes (for example the state of health) are affected by many factors, not only by the good or service under consideration. Thus, the effect of health care on the state of health should not be affected by any other factors that influence consumer outcome such as the lifestyle of patients. It is empirically difficult to control for these other influences on outcome.

A more subtle issue is the following: a measure of the contribution to outcomes should reflect the normal, or expected effect of the activity whose output is to be measured. 'Normal', 'average' or 'expected' effects should be considered rather than ex-post, individual effects. This distinction arises from in the context of service provision where the consumer is typically actively involved in the provision of the product. It would seem that a measure of service production should not be influenced by the *individual* capacity of the consumer to make use of these services. For example, the same teaching activity performed on two different students, should be measured as the same quantity of teaching services towards each student, even if it turns out that one of them benefits less from teaching than the other. Or the same medical treatment, applied to two different persons with the same disease (and similar patient characteristics) should be measured as equal quantities of medical services. Unless, that is, the two persons come from different groups of patients where 'groups' is understood as sets of patients with characteristics that require a-priori different services, for example young and old persons for particular treatments. It is easy to see that making this difference in practice will be difficult and sometimes impossible. This limits the applicability of output methods that rely on the direct contribution to outcomes.

6. Conclusions

This paper looked at the notions of output and outcome in the context of service activities with market and non-market producers. We started by defining the various notions, in particular inputs, processes/activities, outputs, and outcomes. The following main conclusions were derived:

- Even in the simplest market model without any quality change in products, output measurement may require *some*, often implicit, reference to outcome when products are grouped and classified. Thus, output and outcome are different, but they are not independent of each other.

- With non-market producers, the equality of marginal utility and unit cost ratios breaks down. The money value of output is determined as the sum of costs of inputs, but it is possible to construct meaningful indices of cost-based volumes and quasi prices along with indices of volumes and prices of inputs. Akin to the simple market case, cost-based indices are not void of implicit information about utility and outcome – typically, this enters via the product classification on the basis of which aggregation takes place.

- In the presence of quality change, all existing methods require implicit or explicit information or reasoning about outcome. The simplest and most widely-used way of controlling for quality change – matching or stratification – requires implicit statements about the similarity of product in their capacity to generate direct outcome when products are grouped together. A second method uses explicit information on product characteristics to adjust unit costs for differences in characteristics. The choice of characteristics cannot proceed without considering outcome.

- Problems of quality adjustment arise whether services are provided by market or non-market producers. The fact that there is an observable market transaction in one period and another market transaction in the next does not imply that they are comparable – otherwise, price statistician would be in the convenient situation of only having to deal with quality adjustment for non-market production. This is of course not the case. Although the distinction between market and non-market producers is useful and has consequences for measuring the current price value of output, it loses most of its significance when we come the thorny part of measuring output – the treatment of quality change and the treatment of exiting and new products.

- A pragmatic approach will be called for to proceed with services measurement. In particular, there is no reason to approach every type of service with the same method for quality adjustment – methodologies should be robust but will also have to be reflective of data availability and transparency for users.

- Measuring output for complex services is difficult but the conclusion should not be that it is simply too difficult to do anything. Health and education account for a too large and growing part of the economy to ignore output measurement for them. It may take a while before consensual and internationally comparable methods are agreed upon but active research and data development is vital to achieve this objective.

REFERENCES

Abraham, Katharine G. and Christopher 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.

Archibald R. B. (1975); "On the Theory of Industrial Price Measurement: Input Price Indexes"; *Working Paper No 48*, U.S. Bureau of Labor Statistics, Washington D.C.

Archibald, R. B. (1977); "On the Theory of Industrial Price Measurement: Output Price Indexes"; Annals of Economic and Social Measurement 6, pp. 57-72.

Atkinson Review (2005); *Final Report: Measurement of Government Output and Productivity for the National Accounts*; Palgrave McMillan.

Balk, Bert M. (1998); *Industrial Price, Quantity and Productivity Indices, The Micro-economic Theory and an Applicati*on, Kluwer Academic Publishers.

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

Berndt, Ernst R., and David M. Cutler (editors) (2001); *Medical Care Output and Productivity*; Vol. 62, NBER Studies in Income and Wealth, University of Chicago Press.

Cutler, David M., Mark McClellan, Joseph P. Newhouse and Dahlia Remler (1998); "Are Medical Prices Declining? Evidence from Heart Attack Treatment"; *Quarterly Journal of Economics*, Vol. CXII, November, pp. 991-1023.

Diewert, W. Erwin (2008); "The Measurement of Nonmarket Sector Outputs and Inputs Using Cost Weights"; *Discussion Paper 08-03, January 17, 2008; Department of Economics, University of British Columbia*, available at http://www.econ.ubc.ca/diewert/dp0803.pdf.

Diewert, W. Erwin (1980); "Aggregation Problems in the Measurement of Capital"; in D. Usher (ed.); *The Measurement of Capital*, University of Chicago Press, Chicago and London.

Diewert, W. Erwin (1974); "Application of duality theory"; in M. D. Intriligator and D. A. Kendrick (eds.) *Frontiers of Quantitative Economics,* Vol. II, Amsterdam, North Holland.

Diewert, W. Erwin and Alice O. Nakamura (2007); "The Measurement of Productivity for Nations"; in: *Handbook of Econometrics*, Volume 6A; © 2007 Elsevier B.V.

Ervik, Astrid Oline, Erling Holmøy and Torbjørn Hægeland (2003); "A Theory-Based Measure of the Output of the Education Sector"; *Statistics Norway Research Department Discussion Paper* No. 353, July.

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

Fisher, Franklin M. and Karl Shell (1972); *The Economic Theory of Price Indices*; Academic Press, New York and London.

Gold, Marthe R., Joanna E. Siegel, Louise B. Russel and Milton C. Weinstein (eds.) (1996); *Cost effectiveness in health and medicine*; New York, Oxford University Press.

Griliches, Zvi (ed.) (1971); *Price Indexes and Quality Change: Studies in New Methods of Measurement*; Cambridge MA, Harvard University Press.

International Monetary Fund, International Labour Organisation, Organisation for Economic Co-operation and Development, United Nations Commission for Europe, The World Bank (2004); *Producer Price Index Manual, Theory and Practice*, Washington, D.C.

Jorgenson, Dale W. and Barbara Fraumeni (1989); "The Accumulation of Human and Nonhuman Capital, 1948-84"; in Lipsey, R. E. and H. S. Tice (eds.), The Measurement of Saving, Investment and

Wealth; National Bureau of Economic Research Studies in Income and Wealth, Vol. 52, pp. 227-282, University of Chicago Press.

Jorgenson, Dale W. and Barbara Fraumeni (1992): "Investment in Education and U.S. Economic Growth"; *Scandinavian Journal of Economics,* 94, supplement pp. 51-70.

Könus, A. A. (1924); "The Problem of the True Index of the Cost of Living", translated in *Econometrica* 7 1939, 10-29.

Lawrence, Dennis and Erwin W. Diewert (2006), "Regulating Electricity Networks: The ABS of Setting X in New Zealand", in: Chapter 8 in *Performance Measurement and Regulation of Network Utilities*, T. Coelli and D. Lawrence (eds.), Cheltenham: Edward Elgar Publishing, pp. 207-241.

Lee, Phillip (2008); *Public Service Productivity: Health Care; UK Office for National Statistics*; available at: http://www.statistics.gov.uk/cci/article.asp?ID=1922.

Nordhaus, William D., "The Health of Nations: The Contribution of Improved Health to Living Standards" (February 2002). *Cowles Foundation Discussion Paper No. 1355*. Available at SSRN: http://ssrn.com/abstract=302789.

Garcia Armesto, Sandra, Maria Luisa Gil Lapetra, Lihan Wei, Edward Kelley (2007); *Health Care Quality Indicators Project 2006: Data Collection Update Report*; OECD Health Working Papers 29.

Pakes, Ariel (2003); "A Reconsideration of Hedonic Price Indexes with an Application to PCs"; *American Economic Review*, 93(5); pp. 1578-96.

Rosen, Sherwin (1974); "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition"; *Journal of Political Economy*; 82(1), pp. 34-55.

Shepard, R. W. (1970); *Theory of Cost and Production Functions*, Princeton University Press.

Stone, Richard (1956); *Quantity and Price Indexes in National Accounts*; Organisation for European Economic Cooperation, Paris.

Triplett, Jack E. (1983); "Concepts of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource-Cost Debate"; in M. F. Foss (ed.), *The U.S. National Income and Product Accounts: Selected Topics*; University of Chicago Press, Chicago and London.

Triplett, Jack E. (2006); *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*; OECD Paris.

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

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