This draft: Nov 22, 2009

Health System Productivity

Jack E. Triplett
Brookings Institution[1]

Chapter 31 in *Oxford Handbook of Health Economics*
Edited by Sherry Glied and Peter Smith

I. Introduction.

Productivity is the ratio of outputs to inputs, and productivity growth is the growth of the ratio—that is, it is a shift in a production function. In concept, health system productivity differs little from productivity in any other industry or sector. Though the economics of medical care may in some respects appear unique, production in the sector is still described by a production function, a relation between medical care inputs and output. Productivity change in medical care is a shift in that medical care production function.

However, productivity in the medical care sector has behaved very differently from other industries, even other services industries: Measured productivity growth in medical care has typically been negative. Murray (1992) reported negative labor productivity growth in Swedish hospitals. Triplett and Bosworth (2004, 2007) found negative productivity growth in U.S. medical care between 1987 and 2005, at a rate of about one percent per year, a finding confirmed by Harper, et al. (2008). When the U.K. statistical office added the output of medical care to the country's national accounts, the negative productivity growth implied by the new measure provoked an outcry in Parliament and the appointment of a special commission on public sector services productivity measurement (Atkinson, 2005) to determine what was wrong.

Few industries have experienced more innovation, so medical care's negative productivity growth is highly suspect. Economists generally believe that measured productivity growth in the sector is biased downward because of difficulties in measuring medical care output accurately and also that measurement errors are pervasive in some of the inputs, particularly in pharmaceuticals and medical devices and in the high-tech portions of medical equipment. In all productivity measurement, the most essential tasks are getting the data right, which provides the agenda for section III of this chapter. Data on inputs and outputs—indeed, economic data generally—for the health care sector are much less well developed than for many other sectors of the economy, which is bizarre considering the size of health care in most industrialized countries and the importance of the sector. Many studies of medical care productivity employ data from the national accounts (which are described in Commission of the European Communities, et al., 1993), but microdata studies face the same measurement problems.

Medical care is not the only determinant of health. In section IV, I discuss some of the economic implications of the fact that medical care, though it is demanded to improve health, does not, *by itself*, produce health.

II. Concepts of Productivity Change

The conventional framework for any measurement of productivity rests on the production function, in this case for the output of medical care services. The inputs are conventionally notated as KLEMS: capital services, labor services (the vector of all labor inputs, from surgeons to janitors), energy, intermediate or purchased materials (which in this sector includes pharmaceuticals used in hospitals and clinics, stents, and so forth), and purchased services. That is:

---

[1] I appreciate valuable comments from Joseph P. Newhouse, Mary O'Mahoney, and the editors.

(1.1) medical services = f (K,L,E,M,S)

In equation (1.1) all the variables, including output, should be understood as vectors,[2] the elements of which, however, are usually aggregated for analysis.

The medical care production function in equation (1.1) implies:

(1.2) Productivity change (medical care) = growth of medical services / growth in f (•)

Equation (1.2) is known as "multifactor productivity" (MFP) growth. It is also called TFP (total factor productivity), which is a synonym.[3] Productivity change is usually interpreted as a measure of changing efficiency in production, or of technological change, though in practice measurement errors and other inconsistencies limit the validity of this interpretation.

A *partial* productivity measure is "labor productivity" (LP). LP growth is usually calculated as the growth in medical services over the growth in the labor input, which implies the growth of the average product of labor. Properly, LP growth in the medical care sector is:

(1.3) LP (medical care) = ∂ (medical services) / ∂ L,

using equation (1.1), that is, it is the growth of the marginal product of labor. I say little about LP in this chapter because MFP is the more comprehensive measure.

For an alternative perspective, recast the numerator in the productivity ratio to measure the welfare gain that an economic activity produces (Hulten, 2001, contrasts the two perspectives). In health *system* productivity, the welfare gain is the improvement in health; it may also encompass aspects of the patients' medical experiences, such as the painfulness of medical procedures, length of waiting times, pleasantness of their encounters with health system personnel, quality of food and amenities in hospitals and so forth. In the rest of this chapter we neglect for brevity elements of the patients' medical encounters, without suggesting that they are unimportant.

The alternative perspective requires an expanded notion of the inputs and resources that contribute to improved health. We posit a relation between a measure of health and the inputs that produce health, or to put it alternatively, the determinants of population health. Thus, in parallel with equation (1.1), we can write the "health production function":

(1.4) health = h (medical care, time, consumption, R&D, environment, etc).

The inputs in equation (1.4) include medical care services, the output in equation (1.1).[4] The non-medical inputs include time spent by people investing in their own health (for example, exercise), as well as

---

[2] So the usual mathematical convention properly would have: 0 = g (medical services, K, L, E, M, S).

[3] The term MFP was introduced in National Research Council (1979) to avoid the implication that equations such as (1.2) have necessarily enumerated all the inputs. *Measured* productivity change also reflects inputs that have not been accounted for, or not accurately or fully measured.

[4] If national accounts *industry* data are used to estimate equation (1.1), the output domain would normally include the public and private medical care sectors, and therefore the drugs and medical devices provided by those institutions. The domain would exclude, however, expenditures for drugs and medical devices purchased as final consumption by households (even though those expenditures are included in GDP), because the SNA excludes household production. Equation (1.1) could be made to match total health care spending in GDP if a household "industry" were added to the

consumption items that have health implications (some food, plus tobacco and alcohol, for example), research and development, the quality of the environment, and other influences.

This way of thinking about health—and the relation between medical care and health—was introduced into economics by Grossman (1972), but it is a formalization of relations long known in epidemiology. For example, McKeown (1976) showed that most long-term improvements in health could not be accounted for by medical interventions. Grossman explained that the loss of utility from abstaining from the consumption of pleasurable activities today that (like smoking) have deleterious long-term consequences for health was comparable to individuals' investing in health (chapter 11, by Bolin).

Equation (1.4) implies an alternative productivity measure, one for the production of health:

(1.5) MFP change (health) = change in health / change in h $(\cdot)$

In this case, productivity change can be interpreted as a measure of the efficiency of all of society's resources that are used to produce health. The resources include utility foregone in desisting from the consumption of unhealthy goods and from abandoning the pursuit of unhealthy lifestyles.

A major question in health economics is the value of national expenditures on medical care, which are increasing in nearly all countries (see chapter 15 by Chernew). The proper way to estimate the *productivity of the medical care resources used to improve health* is to make use of equation (1.4). This task implies a partial productivity measure, which can be expressed as (using Z as the partial measure of productivity, and substituting equation 1.1 into 1.4):

(1.6)
$$
\begin{aligned}
Z &= \partial\,(\text{health}) / \partial\ \text{medical care} \\
&= \partial\,(\text{health}) / \partial\,[\text{f (KLEMS)}]
\end{aligned}
$$

As with any partial derivative, the value of Z will depend on the values of the other variables in the equation, particularly in the case of health the life style, diet, and environmental variables.

Thus, a nation's health is not a straightforward function of its per capita expenditures on medical care. The contribution of medical services to improved health cannot be established without considering non-medical determinants of health.

For example, substantial U.S. expenditures for blood pressure medication intended to reduce the amount of salt in patients' bodies may just be offsetting the excessive salt added by the U.S. food processing industry, so there may be no net gain in health, relative to societies that consume less salt and fewer medications. And one cannot judge the performance of the U.K. National Health System by a naïve examination of trends in NHS expenditures and vital statistics. U.K. health may change because of the other variables in $h(\cdot)$, variables that perhaps have more influence on health than does the National Health Service. Philipson and Posner (2008) present another example when they suggest that the inverse relation between obesity and income is a consequence of a fall in the price of consuming calories and a rise in the cost of exercise (once a costless by-product of manual labor), which the higher-income groups can better afford: Expenditures on medical care then partly offset the negative health impacts of obesity.

One often hears statements such as: The U.S. medical care system must not be efficient or productive because the U.S. does not have the highest health level in the world, even though it has the

---

domain of equation (1.1), and households were treated as an additional sector that contributes to the production of medical care. There are other reasons for adding nonmarket production to the national accounts—see Jorgenson, Landefeld, and Nordhaus (2006).

highest per capita spending on health care. This is a non sequitur, because it does not take into account any other variables in the health function of equation (1.4). More healthful diets or lifestyles in Japan or Mediterranean countries may give inhabitants of those countries better health with lower inputs of medical care than in the U.S. Similarly, regressions of cross-country expenditures on medical care (or similar variables, such as consumption of pharmaceuticals) and levels of health are of little value unless they control for the dependence of health on health-determining factors other than medical care.

III. Estimating Medical Care Productivity

Equation (1.2) can be estimated in a number of ways. One option is econometric (see any standard textbook that covers production function estimation), or by some type of envelopment method (Fare et al, 1984). Because the form of the function f(•) implies an index number formula, a frequently-employed alternative is to estimate it by a ratio of index numbers (see section III.B). Hulten (2001) reviews the relative advantages of econometric and index number approaches and Schreyer (2001) presents a comprehensive review of methods for estimating productivity change. Regardless of the method of estimation—econometric or index numbers—the measurement issues are the same.

A. Estimating Productivity of the Medical Care Sector: Output.

The greatest difficulty in measuring medical care productivity is measuring output. Both conceptual and practical problems arise. Alternative approaches compete in the literature.

1. Treatment for an illness. A growing number of health economists have endorsed an episode of treatment for a disease or illness as the unit of output for measuring medical care. Authors who proposed, developed, or supported the treatment-based approach include Scitovsky (1964, 1967), Newhouse (1989, 1992), Cutler, McClellan, Newhouse, and Remler (1996), Berndt, Busch, and Frank (1999), Shapiro, Shapiro and Wilcox (2001), Berndt et al (2000), Cutler and Berndt, eds. (2001), Triplett (2001), and Dawson, et al. (2005). Atkinson (2005), in his review of methods for estimating medical care productivity writes: "Ideally, we should look at the whole course of treatment for an illness…." If the treatment is the unit, then the products of medical care are treatments classified by disease, so the approach demands grouping medical care expenditures by a disease classification, such as the International Classification of Diseases, published by the World Health Organization.

It is useful to state the rationale for the treatment-based approach. On the one hand, one can proceed from the consumer side. What people ultimately want from medical care is improved health. The medical care system pursues this through treatment of ailments and diseases. Doctors' appointments, hospital patient days, drugs—things have typically been treated as outputs in past measurements of medical care—are, from a consumer perspective, more accurately viewed as inputs used in the production of treatments, or (perhaps better) as intermediate stages to the ultimate goal of obtaining a treatment. Medical care is only one determinant of health, but that does not change the fact that the consumer is seeking better health when deciding to visit a health care establishment.

A second rationale starts from equation (1.4). The medical care system's contribution to health is:

(1.7) contribution to health = ∂ (health) / ∂ (medical care),

other variables in equation (1.4) constant. An increment to medical care is an intervention. Most interventions are disease specific (though some might directly affect overall health). The effect of an intervention, its contribution to health, also manifests itself on a disease-specific element of health. For example, one might want to determine the effect on health from treatment of a heart condition. The effect comes from either reduction in mortality from heart disease or reduction of disabling effects of the heart condition. Both the treatment and the effect (measured by QALE, for example[5]) are disease specific.

This rationale provides a concept for measuring the medical care *input* to the production of health, and one might object that measuring the *output* produced by the medical care sector should reflect an output concept. However, there is no very strong case for any alternative concept from the output side. In a famous exchange with Griliches, Gilbert (1961) contended that units such as a visit to a doctor or a day's stay in the hospital were the appropriate units because, he argued, those were the transaction units, the units for which medical care suppliers charged their customers.[6] Even if this were true in Gilbert's day, compensation schemes have increasingly been shifting toward Diagnostic Related Group (DRG) systems, where the unit is clearly the treatment of an illness, not a day in the hospital.

Implementing an illness-based output concept is greatly facilitated by the International Classification of Diseases (World Health Organisation, 1977), which provides a well-developed classification system that can be used for medical expenditures, and by DRG systems. DRG systems fit into the ICD system at the chapter level. DRG systems render collecting data by illness groups more feasible because providers already have the data in their records. However, DRG systems are not currently completely comparable across countries (Schreyer, 2009).

As well, much other medical information, information that an analyst would use in conjunction with medical care data and health accounts, is collected, tabulated, and organized on similar principles. Scientific advances also fit into the system: Research on new treatments takes place, obviously, at the level of a specific disease, so it can be fitted naturally into the ICD. This is important for developing methods to deal with treatment improvements in output measures for medical care (considered in a subsequent section).

Thus, using a disease-based classification system such as ICD for measuring medical care output makes the economic classifications line up with the classifications that are used for other scientific work, and with the classification used for payments. This is a great advantage.

Several reservations are routinely expressed. First, not all medical expenditures fit into a disease-based system. Much preventative care (routine physical examinations, for example) cannot be attributed to a chapter of the ICD. Some part of nursing home and other rehabilitative care may not be explicitly related to any particular illness, even though some medical event—a stroke, for example—may have precipitated the nursing home admission. Additionally, the problem of comorbidities must be allowed for in some way in assigning costs to chapters or subchapters of the ICD.

These frequently-voiced reservations must, however, be put into context: Cost-of-illness (COI) accounting has been carried out for many years, starting in the U.S. with Rice (1966)—see also Rice, Hodgson and Kopstein (1985). Hodgson and Cohen (1999) allocated 86 percent of U.S. medical expenditure to the ICD classification; most of the unallocated amounts were caused by data deficiencies and had nothing to do with comorbidities or other conceptual problems. Heijink et al. (2006) present data from COI accounts across a group of other countries. The challenge is not "Can it be done?" The challenge is to improve on what has been done in the past (improved methods for distributing the costs of comorbidities, for example).

It will also be necessary to shift the orientation from one-time snapshots of illness costs—the traditional focus—to time series, which are necessary for productivity analysis, and indeed for much other

[5] For a discussion of QALE and similar measures of health outcomes, see chapter ZZ.
[6] Griliches' response was repeated in his collected essays, among other places—Griliches (1988). The debate is best interpreted as an example (of which there are many) of the difference between output and input measurement concepts for economic data.

analysis of medical care. A start toward the appropriate time-series focus is Roehrig et al. (2009), which appears to be the first such study. Several previous studies compute rates of expenditure change by disease grouping by linking together COI estimates carried out independently for two or more years: Polder and Achterberg (2004) for the Netherlands, 1994 to 1999, Heijink et al. (2006), for several countries and pairs of dates, and Triplett (2001), who links a number of mental health studies in the U.S. Linking COI studies that have been done separately creates error because of methodological non-comparability in the base studies, so is inferior to a study that applies a consistent methodology through time. OECD (200x) recommends constructing time series.

Statistical agencies in several countries have adopted the cost-of-illness approach, which has made great improvements in the data. Among these, prominent milestones are the hospital price indexes in the U.S. Producer Price Indexes (Catron and Murphy, 1996), which have from the early 1990s collected the cost of a sample of diagnoses (extension of this concept to other medical care providers has been announced), the Eurostat Manual for services output in national accounts (Eurostat, 2001), which specified that medical care output should be estimated on a cost-of-illness basis, the North American Product Classification System (NAPCS, which guides collection of industry product data in the U.S., Canada, and Mexico), which adopted a version of the ICD as the "product codes" for medical care industries,[7] and the forthcoming OECD manual for health care and education output in national accounts (Schreyer, 2009).

Finally, researchers working on international comparisons of medical care costs have adopted cost-of-illness as the appropriate empirical framework (OECD, 2003; Busse, Schreyogg and Smith, 2007). Indeed, constructing data on expenditure trends at the cost-of-illness level provides essential information for any analysis of medical care costs. It is astonishing that so much discussion has taken place about the rising cost of medical care, in so many countries, with minimal information (or concern) about the illnesses whose medical care are experiencing rising costs. Generating data on trends in the cost of disease is important not just for productivity analysis, but also for the analysis of medical care costs.

Arraying medical expenditures by a disease classification is only the first step in implementing a treatment-based output concept. It is also necessary to separate trends in medical care expenditures into trends in quantities of treatments and prices or costs of treatments. Before discussing this in section III.B, I consider in the next subsection an alternative concept for medical care output.

2. Medical outcomes as the output measure: A treatment index or an outcomes index?

Equation (1.4) suggests valuing medical care interventions by their incremental contributions to health, that is, the contribution of each intervention is its medical outcome measure. If so, why form an index of treatments? Why not measure medical outcomes directly, disease-by-disease, and combine them into a weighted measure? Indeed, Dawson et al (2005) proposed exactly that—to ignore the intervention entirely and look only at its effect on health. Their preferred basic measure of the output of the National Health Service is a weighted index of QALYs, by disease classifications, where one QALY is valued at £30,000 (see their equations 12 and 111).[8]

Among reasons for preferring the medical outcome measure is the general knowledge that not all treatments are effective. Errors and mistakes, misprescriptions and misdiagnoses (patients still receive antibiotics for viral infections, for example), botched operations, and variance across areas in modes of treatment are well known. Interarea differences in medical practice may be errors or may be differences of opinion about best practice, but even if the latter, presumably more knowledge will eventually show that some treatments that were thought to be best practice were not. This evidence suggests that some interventions are not making a positive incremental contribution to health. For these cases, bypassing the

[7] Although it appears that only the U.S. will implement this portion of NAPCS at this time.
[8] A summary report by a subset of the same authors is Castelli et al (2007).

treatment measure would bring us closer to an output measure that measures the incremental contribution to health that the medical care system makes.

However, similar phenomena occur in other parts of the economy and are not adjusted out of the productivity measures for other sectors. Botched and inappropriate car repairs, for example, occur with some frequency; sometimes, they are corrected by the original repairer so the corrections do not result in new output, but sometimes the customer seeks out a new shop, so that repairing the botched job actually increases GDP (and is included in the numerator for the estimate of car repair productivity). We do not subtract such "redos" from GDP, even though they hardly contribute to consumers' welfare, nor do we adjust for defective manufactured products that are also not infrequently produced.

Whether appropriate or not, treatments are still produced in the medical care sector and they still use resources in the medical care sector. By that standard, they are outputs of the medical care sector.

Another reason for preferring treatments as the output measure for medical care arises from the distinction between output and welfare that is embodied in equations (1.1) and (1.4). Medical care services are an input in equation (1.4). The output is health. One never wants to measure an input by its output (nor an output by an input). If the output of the medical care sector were measured as a health outcome, and that measure then used as an input in the equation for the determinants of health, the possibility of productivity change in the health equation is largely eliminated by convention. One of the things we need to know is the productivity of the medical care sector in the production of health. To estimate that, the measure of medical sector output must be independent of the output of health. It must in principle be possible that the output effect of a change in an input quantity differs from the change in the input. [9]

A final, pragmatic, reason also has great weight: A health output measure that is based on disease treatments is grounded on a more precise statistic than an output measure that is based entirely on medical outcome measures. Measuring health output by treatments may not be that far along at present. Nevertheless, more information is available on expenditure by disease, on numbers of treatments by disease, and even on health care prices by disease than on medical outcome measures, and treatment information is inherently more concrete and therefore more precisely measured information. Measuring medical care output by treatments is not that different from the way we measure car repair in national accounts (Triplett, 2001), and can readily be understood within the usual framework of economic statistics. In contrast, even health economics professionals raise difficulties, both conceptual and practical, with existing medical outcomes measures (Meltzer, 2001). A sound measurement principle is to minimize the use of undeveloped and potentially controversial measures, using them only where they are necessary and not where alternatives exist.

In summary, measuring medical care output by health outcomes, though seemingly appealing, mixes concepts from two different types of measurement—the measure of production and productivity in the medical care sector and the measurement of the determinants of health. In the absence of a system for estimating equation (1.4)—that is, in the absence of a "health account"—it might be tempting to attain a partial measure of the impact of the medical care sector on the determinants of health by using health outcomes to estimate the output of medical care. But mixing concepts in an economic measurement always provokes confusion. Better to estimate both equations (1.1) and (1.4) than to intermingle and confound separate conceptual systems.

B. From Expenditures on Treatments to Quantities of Treatments.

---

[9] To avoid confusion, it is not inconsistent to make the quality adjustment for a changed treatment depend on the ratio of medical outcomes for the new and old treatments, suggested in section III.B.

Collecting expenditures on treatments, by disease classification, is necessary but not sufficient to measure output. From Section III.A.1, output for medical care is measured by the quantities of medical care treatments, which we will alternatively designate *medical care services*. Expenditures on, for example, heart attacks equal the price or cost of treating heart attacks times the quantity of medical care services for the treatment of heart attacks. Both price (medical care inflation) and quantity information are necessary, but the quantity of medical care services is the most crucial measure, because the services quantities are the outputs of the medical care sector, and they are one of the inputs into the production of health.

Two methodologies exist for obtaining estimates of quantity changes of medical treatments: deflation and direct quantity indexes.

1. Deflation. The first, "deflation," is the conventional methodology of national accounts: The change in expenditures for some category of goods and services is divided by a price index (a measure of inflation) for that category to obtain the quantity measure. For example, expenditures on treating heart attacks would be divided by a price index for treating heart attacks. Generally accepted procedures for deflation in national accounts are presented in the SNA handbook (Commission of the European Communities—Eurostat, et al, 1993, ch. 16).

Because medical care services are heterogeneous, both their price and their quantity are expressed as index numbers. Index numbers most commonly employed in productivity measurement are the Törnqvist index and the Fisher index (Caves, Christensen and Diewert, 1982), both of which are termed "superlative index numbers."[10]

Suppose the deflating price index is a Fisher index (the Fisher index number is the geometric mean of Paasche and Laspeyres index numbers). Then, for periods 0 and 1:

$$(1.8)\quad (\sum P_1Q_1/\sum P_0Q_0) / \{[\sum P_1Q_0/\sum P_0Q_0]\,[\sum P_1Q_1/\sum P_0Q_1]\}^{1/2} = \{[\sum P_0Q_1/\sum P_0Q_0]\,[\sum P_1Q_1/\sum P_1Q_0]\}^{1/2}$$

The terms in square brackets on the left hand side are Laspeyres and Paasche price indexes, and the term in curly brackets is the Fisher price index. Deflation yields, on the right-hand side, a Fisher quantity index number (in parallel, the geometric mean of Laspeyres and Paasche quantity indexes, each of which is shown inside a square bracket); the Fisher quantity index measures the change in the quantities of treatments for (in the example) heart attacks. To give it an alternative interpretation, the quantity index on the right-hand side of equation (1.8) shows the rate of growth in medical services for this medical ailment. The quantity index takes account of differences in severities and of variations in treatment; it is not just a count of the number of heart attack treatments.

The theoretically correct way to aggregate outputs is by marginal cost (MC) weights (Fisher, 1993). By the usual competitive assumption, MCs are proportional to prices, which are the weights in the quantity index (the right hand side of equation (1.8)). Thus, deflation preserves (approximately) the theoretical aggregation condition.

However, the assumption that prices and MCs are proportional is especially problematic in medical care, where many prices have remote connections to costs. Even in the United States, where prices charged for medical care treatments are routinely collected and medical care price indexes published (in the Producer

[10] The modern concept of the superlative index number was developed by Diewert (1976), though the term itself was in use much earlier. The Fisher index number appears in the text. The Törnqvist (quantity) index is: $I(T)_{01}) = \prod \{(q_1/q_0)^s$, where $s = (½\,[p_0q_0/\sum p_0q_0 + p_1q_1/\sum p_1q_1])$, for each of the ith quantities. In this expression, and in equation (1.8), the subscript i, designating the ith quantity, is suppressed.

Price Index), the usual national accounts deflation methodology is therefore questionable when applied to medical care, because the prices-MCs proportionality assumption is untenable. In other countries, government-provided medical care means a price index for medical care has no relevance. For this reason, medical care output in most countries' national accounts is not estimated by deflation.

2. Direct quantity index. The alternative method is to aggregate the quantities directly—that is, to compute directly the right-hand side of equation (1.8). As an example, one could compute an index of the quantity of heart attack treatments. This computation requires data on quantities of treatments, classified by case severity, type of treatment and so forth; the weights for the index are the costs (not the prices or consumer valuations) of the various treatments. The correct weight for each treatment is its marginal cost (MC); in practice, average cost (AC) is more likely to be obtainable, so AC provides an approximation to MC.

Several European and Oceanic countries are experimenting with direct quantity indexes of medical care. Schreyer (2009) lists: Australia, Finland, France, Germany, Netherlands, New Zealand, Norway, Sweden, and the UK. In some of these cases, however, the indexes are not computed on a disease basis, and not all of them are based on completed treatments.[11] In principle, countries with government-provided health care would have the treatment quantities and costs by disease in their health care management systems; in practice, most do not, either for obscure political or bureaucratic reasons (in a few countries) or because managing governmental health care systems has not been done very effectively.

3. When Treatment Characteristics Change. The major methodological problems with either the deflation method or the direct quantity index method arise because of (a) improved treatments, generally but not necessarily in the same facility, and (b) treatments that move between facility types or that cross industry lines, so that the equivalent treatment (measured by medical outcomes) is available at lower cost.

a) *Improved treatments*. When improved treatments for a disease are introduced, they create the infamous price index "quality change" problem. If the better treatment is introduced at higher cost, and the new one is compared directly with the old (either its cost or its quantity), it will show up inappropriately as medical care inflation, instead of (as it should) an increase in the quantity of medical services, unless some "quality adjustment" is made. Even if the new, improved treatment is cheaper (yet better) than the old (so that comparing the new and the old will record a drop in cost), direct comparison still understates the true cost decline because an allowance or adjustment should also be made for the value of the improvement.

The best way to do it is to adjust the prices or quantities by a measure of the medical outcome, even though outcomes do not, strictly speaking, concord with the theory that underlies the measurement of output. Considering two treatments used for the same diagnosis, and letting m_a and m_b be the value of medical outcome measures corresponding to treatments a (the old one, used in period 1) and b, while p_2 and p_1 are the prices of treatments in periods 2 and 1, respectively, then the quality-adjusted price measure (η_{21}) is:

(1.9) $\eta_{21} = p_2 / (p_1 (m_b / m_a))$

Expenditures on this treatment are $p_1 q_1$ and $p_2 q_2$ in the two periods, so that the deflated quantity change in treatments (λ_{21}) is:

(1.10) $\lambda_{21} = (q_2 (m_b / m_a)) / q_1$

---

[11] For example, Schreyer (2009) says that the health measure in the UK national accounts is "a cost weighted activity index." However, these UK "activities" are not completed treatments, they are a mix of mostly intermediate functions (see the description in Dawson et al, 2005). The UK measure is not an adequate output index.

Notice that the adjustment makes the improvement in outcomes part of the quantity change in medical care. Thus, if the new treatment is 20 percent better by the medical care outcome measure, it should be recorded (whether the adjustment is made in the price index used for deflation or in the direct quantity measure) as approximately 20 percent more output.[12]

Where the direct quantity index method is used to measure the output of medical care, the same result would be obtained through a parallel adjustment to the quantities, equivalent to equation (1.10). In this case, no price index adjustment is needed. The implicit, quality-adjusted change in the price of the treatment is obtained by dividing the expenditure change by equation (1.10)—refer back to equation (1.8).

Whether deflation or direct quantity index method is used, the great difficulty is the lack of sufficient medical care outcome measures to facilitate an adjustment.[13] The medical outcome measure should, of course, refer to the specific change in treatment and only to that change. Other changes (e.g., hospital meals and amenities) should be held constant, to avoid the danger of double-counting. An outcome measure obtained from a medical trial is an example where other influences are held constant; an outcome measure from a hospital patient survey may not do so.

b) ***An equivalent treatment (judged by a medical care outcome measure) becomes available at lower cost in a different facility***. The archetypal example is cataract surgery (Shapiro, Shapiro, and Wilcox, 2001), once done in a hospital with a lengthy recovery period and now performed as an outpatient procedure, often in a non-hospital facility. If deflation is employed, any price reduction will be missed because prices are gathered by repeat visits to the same facility; in consequence, output increase is understated (refer to equation 1.8), and so is productivity change. If the direct quantity index is computed, the quantities from both providers may be recorded, but the change in the cost weights will be missed or misinterpreted, and again, output increase and productivity change are understated.[14] The solution to this problem is again to obtain a medical outcome measure, which will tell, when the treatment moves to a new, lower-cost facility, whether or not it is an equivalent treatment.

The major work on these two problems has been done in the deflation context, and includes: Cutler, et al. (1996) on heart attacks, Berndt, Busch and Frank (1999) on depression, Shapiro, Shapiro, and Wilcox (2001) on cataract surgery, and several other similar studies summarized in Cutler (2004). They show that true inflation in medical care in the United States was substantially lower than had been thought. Consequently, the output growth of medical services was greater and productivity growth was also understated in the available government data (consequently, the medical care productivity estimates in Triplett and Bosworth, 2004, are biased downward). The studies also hint that policy measures to slow the growth of medical costs may impact medical care services, rather than (as most policy makers have hoped) simply slowing the growth of medical care inflation.

It has sometimes been thought that one can circumvent the twin change-in-treatment problems noted above by collecting data from insurance company claims, rather than from medical care providers. One advantage of claims data is their large sample size. However, even though in principle claims data could be

---

[12] This is an oversimplification because m_a and m_b are valuations, and the value of improvements in medical outcome measures will usually not be linear; as with everything else, diminishing marginal utility sets in. A 20 percent improvement in the medical outcome measure itself would normally be valued at less than 20 percent.

[13] On medical outcome measures, see chapter xx in this Handbook. The quality change problem in economic statistics generally is discussed in Triplett (2005); for the link between medical outcome measures and traditional index number quality adjustments, see Triplett (1999). See also the references in section III.C.

[14] The first facility has a decline in quantities (say), with relatively high cost weights; the second facility we suppose has an increase in quantities of treatments, with relatively low cost weights. Then, the shift between the two will be recorded as a decline in a (weighted) quantity index, when the quantities should not decline. Instead, a decline in medical inflation (the reduction in cost) should be recorded.

fitted into either side of equation (1.8)—that is, they could be used to calculate a weighted price index or a weighted quantity index—in practice, researchers have calculated a unit value index:

(1.11) unit value index = $(\sum P_1Q_1 / n_1) / (\sum P_0Q_0 / n_0)$

where n_0 and n_1 are the numbers of treatments in periods 0 and 1, respectively, and of course, P and Q are prices and quantites of treatments. Deflation by the unit value index gives[15]:

(1.12) output change = $\{(\sum P_1Q_1 / \sum P_0Q_0)\} / \{(\sum P_1Q_1 / n_1) / (\sum P_0Q_0 / n_0)\} = n_1 / n_0$

Thus, the unit value index implies a quantity index of medical care services that is just a count of the number of treatments. Even though a researcher can subdivide the data to achieve more homogeneous groupings (to calculate separate indexes for types of heart attacks, for example), it is hard to accept the idea that any grouping of medical treatments can be regarded as homogeneous. The extreme heterogeneity in treatments, therefore, is simply assumed not to matter, but the assumption is tenuous, at best.

All the research cited above has focused entirely on estimating improved price indexes, to get better measures of inflation in the provision of medical care. This reflects the special character of the US system, and even in the US the use of deflation for estimating the growth in medical care services is problematic, as indicated earlier. Too little work on the change-in-treatment problems has been done from the direct quantity measurement side, that is, by estimating outcome-adjusted direct quality indexes of medical care (direct estimation of the right-hand side of equation 1.8 and of equation 1.10). Such future work will presumably originate from researchers in countries outside the United States (and its unique deflation environment for the measurement of medical care output). Some of the data required for the research, especially costs and quantities of treatments, should be (but may not be) in the files of government health care programs. However, the detailed information on health outcomes that is required is scarce. Schreyer (2009) presents information on current research and development activities on medical care output measurement within OECD statistical agencies.

C. Estimating Productivity: Inputs to Medical Care

Of the inputs to medical care listed in equation (1.1), **capital** poses the most troublesome problems. The contribution of capital to current output is the flow of capital services provided by the industry's stock of capital goods—physical capital such as buildings and equipment, and also intangible capital.[16] The flow of capital services is derived from the stock. The price of capital services, in concept, is the charge for the use of a capital good for a unit of time. Estimating the price and quantity of capital services is complex, because estimation requires (except when the capital good is rented or leased, when the lease payment provides the price measure) determining the capital stock, estimating deflators for the stock, and measuring depreciation, all of which pose major empirical difficulties. However, medical capital goods pose no problems that are not already familiar in work on measuring capital services in other sectors of the economy. The methods, accordingly, will not be reviewed here (for a summary, see Schreyer, 2001).

[15] See also Balk (1999; 2008), who presents the same result.

[16] A long history of debate over concepts of capital in production analysis spilled over into national accounts measurement. The debate is now settled, on the lines suggested in Jorgenson, Gollop and Fraumeni (1987). See, in particular, the OECD handbook on productivity measurement by Paul Schreyer (2001), which records the consensus and makes recommendations for the measurement of capital stocks and capital services in national accounts that are wholly consistent with the methods now generally accepted in production analysis. See also Hulten (2001). From the *conceptual* view, medical care presents no unique problems in the measurement of capital.

However, few countries publish much data on the detailed types of capital equipment used by the medical care sector. Where information is available, the fundamental measurement problem is quality change.

For many years, economists have known that quality change poses serious price measurement problems. Among the many relevant references that could be cited are Stigler et al (1961), Griliches (1994) and Boskin, et al (1996). Quality change measurement errors are likely when goods and services experience rapid technological change. Medical care industries purchase a range of very highly technological equipment, which is importantly linked with technical change in treatments, and it is hard to measure accurately.

For example, a scanner is essentially an imaging device coupled to a computer. We know that imaging has made vast strides in recent years, and computers have declined in price at the rate of 20-30 percent *per year* for more than 50 years (Triplett, 2005). Moreover, the only research study of scanner prices that exists (Trajtenberg, 1990) found that CT scanner prices fell at a rate not too dissimilar from those of computers. Few countries even have a price index for scanners in their investment goods price indexes, and the U.S. Producer Price Index for diagnostic equipment (which includes scanners) shows only a relatively modest (by computer standards) decline. If scanner prices are declining more rapidly than government price indexes indicate, then investment in scanners (the quantity of scanners) is under-estimated (refer to equation 1.8), and accordingly so is the capital input in equation (1.1). More accurate measurement of high-tech medical equipment is an urgent need. Methods that have been used for computers and other technological goods, such as hedonic indexes (Triplett, 2005), are also appropriate to scanners.

A similar point applies to some **materials** inputs, including pharmaceuticals and medical devices such as stents. Much recent technological change in medicine has proceeded through materials inputs that are unique to the medical care sector. It is not clear how well these intermediates are represented in the national statistics of most countries—the data are suspect because of problems of coverage, of detail, of the quality of information on inter-industry flows, and again, quality change errors in the measured deflators are paramount. Some good work on pharmaceuticals can be cited (see the summary in Berndt et al, 2000), but it is not nearly comprehensive. No similar research on medical devices has appeared, despite their importance in some recent improvements in medical treatments.

Of course, some intermediate materials inputs to the medical sector present no unique problems. Hospitals consume paper products and cleaning supplies, like any other business.

In medical care industries, **services** include purchased medical services from other medical establishments (which present the same measurement concerns already covered in section III.B), and other services, such as accounting, business services, communications and so forth. Triplett and Bosworth (2004) suggest substantial measurement issues for these and similar services; the discussion applies as well to their roles as inputs to the production of medical care.

Most economists probably believe that measuring the **labor** input is much easier than measuring capital, but problems persist. Though the basic unit of measurement is labor hours, human capital is a major contributor to output, in all industries, but especially in medical care. Jorgenson, Gollop, and Fraumeni (1987) provide methods for incorporating labor quality, or human capital, in industry accounts and productivity analysis. Dawson, et al. (2005) present information on medical human capital in the U.K. However, the traditional human capital measures distinguish mainly years of schooling or levels of educational attainment; they need extending to take account of the unique sets of skills and training in modern medicine.

Unpaid volunteer labor in hospitals, hospices and so forth, and time spent caring for ill relatives and friends, is an important input into producing medical care. For productivity measurement of medical care, there is no reason to conform to all the conventions of national accounts (where unpaid labor is excluded), even if the measure of medical care output retains the market output boundary of the national accounts. Estimates of unpaid labor belong in a study of medical care productivity.

Implications: Most, though probably not all, measurement errors in the inputs to the medical care sector imply that measured input growth is too low. For example, if the technical change in scanners is imperfectly captured in the data, then the growth of capital inputs to medical care is understated, and similarly for the growth of medical devices such as stents and for pharmaceuticals, where there is in fact evidence of understatement. Moreover, the growth of human capital in medicine is greater than the growth of labor hours, so to the extent that human capital is not accounted for in medical care industries, growth in the labor input is also understated.

Since productivity growth estimates have the inputs in the denominator, understatement of input growth has the effect of *overstating* productivity growth. We have already noted in the Introduction, however, that most empirical estimates of productivity growth in medical care have reported negative growth. Putting those two facts together suggests the size of the likely bias to output measures: The measurement error in output growth therefore must be large enough to eliminate negative productivity growth indicated by the current statistics (provided we are right in our priors that medical care productivity growth should not be negative) and also additionally large enough to nullify the effects of understatement of input growth. That suggests a very large error indeed in measured output growth in the medical care sector.

D. A Conceptual Error

A relatively minor slip in the Atkinson (2005) report opened a hole that the U.K. health and education bureaucracies initially exploited for their own ends. Atkinson noted that services such as health and education are valued more highly by a society with a higher income. This is quite correct. If one is doing a cost-benefit analysis, for example, the increased valuation that is associated with income growth should be factored in, and this increased valuation, relative to cost, would lead to production of more of the services.

In the U.K., the Atkinson suggestion induced government agencies to add to their productivity numbers “wealth” adjustments (some quite substantial) to incorporate a presumed increased willingness to pay for their outputs. These adjustments have in several cases converted negative agency productivity growth estimates into putative positive ones.

For productivity measurement, willingness to pay adjustments are quite wrong. The agencies have muddled the P’s and the Q’s.

Productivity is the ratio of outputs to inputs, the quantities. When a society is willing to pay more for a service, it is an increase in its price. The quantity of the service is whatever the quantity is, the quantity is not increased by a presumed increase in its value.

One way to see this is to consider productivity change from the standpoint of the dual. Duality theory tells us that one can compute productivity in two alternative and equivalent ways: as the ratio of the quantity of outputs to the quantity of inputs or as the ratio of output prices to input prices.[17] In a competitive economy, those two calculations yield the same result (one can hardly presume that this applies to a

[17] Calculating MFP from the price side is not done very often, but it has a long history. One of the first ever MFP calculations compared growth in output price indexes and input price indexes—Copeland and Martin (1938).

government agency, of course, but leave that aside for present purposes). Productivity increases as the quantity ratio *increases* and as the price ratio *decreases* (greater efficiency makes it possible to lower the price of output relative to the cost of inputs). Thus, an increase in the price of output relative to inputs is a sign of decreasing productivity. Making a positive adjustment to government productivity numbers when their output valuation rises relative to their input costs turns the signal of prices on its head.

The U.K. Office for National Statistics has now rescinded the adjustment for presumed increased willingness to pay for government services (ONS, 2009). This willingness-to-pay adjustment should not be mimicked elsewhere.

IV. Health, Its Determinants, and the Productivity of Producing It.

Grossman's (1972) approach to the analysis of health has become the standard for economists' thinking about the subject (see Bolin, chapter XX). It is a bit surprising, then, that attempts to implement the Grossman model empirically are few. The difficulties in doing so are evident from equation (1.4): Analysis of the production of health requires measuring health (which is multi-dimensional), identifying and measuring the determinants of health and then finding an appropriate way to summarize them, analogous to the index number formula for inputs in equation (1.2). Representing the h (.) function in equation (1.4) may be complicated because of the heterogeneity in the units in which the variables in it are expressed.

Abraham and Mackie, eds. (2005, Chapter 6) suggest constructing a "health account." The health account would provide a welfare-oriented measure as a counterpart to the market-oriented measures of the System of National Accounts, or SNA (Commission of the European Communities, et al 1993), and would link the change in health to the inputs that determine health, of which medical care is but one. Thus, it would be structured by analogy to the familiar national accounts that record economic activity, but would be built around the functional relation and the variables in equation (1.4).

Output in such an account requires measuring health. QALY is one empirical possibility, so the health account would benefit from the substantial ongoing research to improve this and similar measures (these are reviewed in chapter xx).

As inputs for the health account, the Abraham-Macke report lists: medical care (the output of the medical care account), time invested in one's own health (in principle, based on time-use surveys, which are conducted in a number of countries, though the report lists several conceptual problems in using these data in a health account), "other consumption items" (discussed below), research and development (R&D data are also compiled in many industrialized countries), and environmental and "disease state" variables, with pollution mentioned specifically.

Presumably, environmental effects would encompass such things as improved public health and sanitation measures and changes in working conditions, a greatly neglected contributor to historical health improvements. The Sheffield Industrial Museums Trust, South Yorkshire, U.K., has a display of the appalling working conditions that prevailed in U.K. steel production even as late as the 1930s. Leaving aside the distressing number of industrial accidents that must have accompanied the lack of even rudimentary safety precautions, working long hours in the searing heat (workers tied wet rags around their faces to protect their lungs) undoubtedly exacted a frightful toll on life expectancy and on the morbidity of workers who survived to old age. Employees in some industries were literally worked to death.[18]

---

[18] The same was true in the United States. Costa (2000) found that at least a quarter of the decline in respiratory ailments among U.S. elderly men over the 20th century can be attributed to employment shifts out of occupations associated with the highest risks for such diseases.

The list of inputs into equation (1.4) is no doubt long. In many instances, a virtuous circle appeared: For example, part of the increase in income over time was taken in improved working conditions, which led to improved health, which in turn (since the link between improved health and greater earnings is well established) led to further income increases which fed back to greater improvements in working conditions. For some purposes, taking account of the feedbacks matters, even though it greatly complicates the modeling because it suggests that the single equation representation in equation (1.4) is too simple.[19]

Cutler and Rosen, in an ongoing project, are estimating disease models and combining them with economic data (see their proposal in Cutler and Rosen, 2007). Their research uses the perspective of the Grossman model (and the related epidemiological perspective), so one can think of it as estimating equation (1.4), on a disease-by-disease basis. They have selected disease categories that account for a large portion of national health care expenditures and are working toward determining the factors—including of course medical care—that affect changes in mortality and morbidity. At present, they have eschewed the more complicated task of aggregating their disease models into an overall accounting for health. A second effort along similar lines, but for cancer and circulatory diseases in the U.K., is reported in Morton, Rice and Smith (2008). Their work is directed toward making comparisons of the determinants of health costs in different countries.

More research on the determinants of health, especially the non-medical determinants, is vital. Statistical agencies should be encouraged to develop data to enable such research.

Does putting the data into an SNA-type economic *accounting* structure further the objective? Abraham and Mackie and their co-authors start from "holes" or "gaps" in the existing national accounts that arise because the accounts do not measure welfare, and because their measures of outputs (and inputs) are incomplete—only market inputs and outputs are included. For example, uncompensated labor expended on the care of the ill is excluded from the SNA, by convention. If the task is filling SNA holes or gaps then it seems reasonable to work out an expanded accounting system that meshes with, or is patterned after, the traditional national accounts structure. Many of the details have yet to be worked out.

Data requirements for an economic welfare account for health go beyond the database for research on health determinants. For example, Abraham and Mackie list diet among health determinants, certainly an important consideration. The report of the World Cancer Research Fund/American Institute for Cancer Research (2007) summarizes evidence connecting dietary factors to different types of cancer—consumption of red and processed meats raises the risk of colorectal cancers, and excessive consumption of salt for stomach cancer, while on the other hand, consumption of fresh fruits and vegetables reduces risks of a number of digestive system cancers. A research model for cancers might be designed in which dietary data are employed in conjunction with medical care data to determine the relative impacts of diet and medical care on cancer death and incidence rates. If the objective is to estimate determinants of health, then information on consumption of the foods of interest (data which are readily available) is the main requirement.

Fitting dietary (and lifestyle) influences on health into an SNA-type accounting structure is more difficult. In the Grossman (1972) model, and from a welfare perspective, it is necessary to compute the *net* gain—the value of the increment to health, less the loss in utility from abstaining from steak and ham, or from eating vegetables when the individual does not like them.[20] One can conceive of a research project to

[19] Actually, equation (1.4) oversimplifies Grossman's (1972) model. He implies time subscripts on some variables to incorporate delayed effects of current life style and consumption on future health. The time subscripts have been suppressed for present purposes.

[20] "[Nonmarket accounts] should address the major conceptual issues by measuring income and output in ways that best correspond to net economic welfare" (Nordhaus, 2006, page 143).

measure such utility losses, and, though more difficult, perhaps putting them into an aggregate welfare estimation. It is much less clear, however, how utility losses should be fitted into an *accounting* structure.

One possible parallel is with environmental accounting. Some production processes (electricity generation, for example) produce both goods and "bads" (the "bad" in this case is pollution). In an environmental account, one subtracts the values of the bads from the goods to get a net welfare measure. In the health case, the bads are utility losses from pursuing more healthy life styles,[21] so equation (1.4) should be rewritten as:

(1.13) (λ (health, utility loss) = h (medical care, time, consumption, R&D, environment, etc)[22]

In the structure implied by equation (1.13), the net output (the value of welfare gains from improved health less the value of utility losses from dietary and behavioral changes) is the relevant measure for welfare, and therefore for the SNA-analog health account. The net measure, however, is not the measure that is useful for research on health. For that task, only the gross measure, health, is wanted.

The difference between gross and net is reduced the larger is the proportionate contribution of medical care to the change in health. Though it is clear from the historical record that medical care was not the main determinant of past health improvements (McKeown), Cutler (2004) contends that medicine has been the main factor over the last half century, particularly if one adds in the contribution of medical research to the information that has led to changed life styles. Even so, he suggests that the utility losses from life style changes subtract approximately half the utility gain from improved health (Cutler, 2004, pp. 139-40). Thus, the net output measure is still likely to differ substantially from the gross.

In programs to generate data choices must be made, so priorities must be established. It is not premature to recommend collection of more information about the determinants of health. It may be premature to recommend that statistical agencies organize health data to accommodate a welfare-oriented health account: The effort to estimate the utility losses from behavorial changes will likely steal resources away from the generation of data on the determinants of health. For most purposes, it seems clear that developing data on the determinants of health should receive priority.

As was true of the development of national accounts in the 1930s, health accounts have not evolved very far. The situation will change with more work on their conceptual underpinnings and practical needs.

V. Conclusion.

The main limitation on the analysis of health system productivity lies in deficiencies in data. For the medical care sector, the major needs are not only better input data but also expenditures data that are arrayed by disease or illness classification and better methods for separating expenditure changes on health care into increases in price or cost and increases in the quantity of medical services.

Though they are needed for productivity research, these data are needed primarily for other, more important purposes. Nearly all industrialized countries are concerned with advancing medical care expenditures. Yet, they have very little information on what their medical expenditures buy. The National Health Accounts produced in many countries have become nearly irrelevant to the current policy debates on medical care costs because they show only who provides the money for health care (consumers,

---

[21] This is not the definition of "bads" used in Ch. ZZZ, where bads are things that contribute to current utility, but have negative long-term influences on health. Here, bads are things that subtract from current utility.

[22] Or, to be more general: $0 = \eta$ (health, utility losses, medical care, time, consumption, R&D, environment, etc)

governments, insurance companies) and who gets it (hospitals, doctors, pharmaceutical companies). They do not show what is bought for health care expenditures—treatments for disease.

It is naïve to think that expenditures across all types of disease are growing at the same rate: Are expenditures growing more rapidly for treatments of circulatory diseases, or for mental health care, or for cancer? How many countries have this information? How can one make sensible policy decisions on medical expenditures without knowing where the increases originate?

And of course, just knowing which disease categories have advancing expenditures is not enough: We also need to know whether the more rapidly increasing sectors have growing expenditures because of rapidly increasing treatment costs or rapidly increasing numbers of treatments. Again, there is no reason to believe that treatment costs are advancing at the same rates across medical conditions. In the U.S. Producer Price Indexes for hospitals (which are published by ICD categories), the past inflation rate for blood and blood forming organ diseases is *two and a half times* the inflation rate for treating infectious diseases. These are hospital costs, not total costs of treating the disease, but they suggest strongly that the costs of treating different diseases are not advancing at the same rate. How can anyone design intelligent policy on advancing medical care expenditures without knowing where costs are rising? And why? It is commonly alleged, for example, that costs are rising because of technological advance; but do the diseases that have the most innovations in new treatments have the greatest cost increases? Little data exist to answer that question. Information on price or cost growth by disease is sadly not available in most countries that have had debates about medical care expenditure containment.

If expenditure growth is contained, is the effect to reduce medical care inflation? Or is it, instead, to reduce the growth of medical care services? We don't really know. Only after determining this can we move to the next stage and ask whether the growth in the number of medical care services is worth it, a question that has too often been "answered" with a minimal amount of relevant data.

Improving the database for the analysis of medical care productivity deserves high priority. It deserves high priority, not so much because productivity analysis is necessarily that important, but because vital questions of health care policy demand exactly the same data.

References

Abraham, Katharine and Christopher Mackie, eds. 2005. *Beyond the Market: Designing Nonmarket Accounts for the United States*. Washington, D.C.: The National Academies Press

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

Balk, Bert M. 1999. On the Use of Unit Value Indices as Consumer Price Subindices.. In Walter Lane, ed., *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*. Washington, D.C.: U.S. Department of Labor, pp. 112-20.

Balk, Bert M. 2008. *Price and Quantity Index Numbers: Models for Measuring Aggregtate Change and Difference*. Cambridge: Cambridge University Press.

Berndt, Ernst R, Susan Busch, and Richard G. Frank. 1999. Price Indexes for the Treatment of Depression. In *Measuring the Prices of Medical Treatments*, Jack E. Triplett, ed.: 71-102. Washington, D.C., Brookings Institution Press.

Berndt, Ernst R., David M. Cutler, Richard Frank, Zvi Griliches, Joseph Newhouse, and Jack E. Triplett. 2000. "Medical Care Prices and Output." In *Handbook of Health Economics v. 1A*, eds. Anthony J. Cutler and Joseph P. Newhouse. Amsterdam: Elsevier: 119-18

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson. 1996. Toward a More Accurate Measure of the Cost of Living. Advisory Commission to Study the Consumer Price Index. Interim Report to the Senate Finance Committee, December.

Busse, Reinhard, Jonas Schreyogg and Peter C. Smith. 2008. Variability in Healthcare Treatment Costs Amongst Nine EU Countries—Results from the Healthbasket Project. *Health Economics*, Vol 17, Issue S1, pp S-1-S8.

Castelli, Adriana, Diane Dawson, Hugh Gravelle, and Andrew Street. 2007. Improving the Measurement of Health System Output Growth. *Health Economics*, Vol. 16, pp. 1091-1107.

Catron, Brian, and Bonnie Murphy. 1996. "Hospital Price Inflation: What Does the New PPI Tell Us?" *Monthly Labor Review* 120(7): 24-31, July.

Caves, Douglas W., Laurits R. Christensen and W. Erwin Diewert. 1982. The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica*, Vol 50, Issue 6, pp. 1393-1414.

Commission of the European Communities—Eurostat, International Monetary Fund, Organisation for Economic Cooperation and Development, United Nations and World Bank. 1993. *System of National Accounts, 1993*.

Copeland, Morris A. and Edwin M. Martin. 1938. The Correction of Wealth and Income Estimates for Price Changes. Conference on Research in Income and Wealth, *Studies in Income and Wealth* 2: 85-119 and 131-35. New York: National Bureau of Economic Research.

Costa, Dora L. 2000. Understanding the Twentieth Century Decline in Chronic Conditions Among Older Men. *Demography*, Vol 37, No 1 (February), pp.53-72.

Cutler, David M. 2004. *Your Money or Your Life: Strong Medicine for Americ's Health Care System*. Oxford University Press.

Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler. 1996. "Are Medical Prices Declining? Evidence from Heart Attach Treatments." *Quarterly Journal of Economics* 113(4): 991-1024, November.

Cutler, David and Ernst Berndt, eds. 2001. *Medical Care Output and Productivity*. Studies in Income and Wealth, Vol. 62. The University of Chicago Press for the National Bureau of Economic Research.

Cutler, David M. and Allison B. Rosen. 2007. Measuring Medical Care Productivity: A Proposal for U.S. National Health Accounts. *Survey of Current Business*, June, pp. 54-58.

Dawson, Diane, Hugh Gravelle, Mary O'Mahoney, Andrew Street, Martin Weale, Adriana Castelli, Rowena Jacobs, Paul Kind, Pete Loveridge, Stephen Martin, Philip Stevens, and Lucy Stokes. 2005. Developing New Approaches to Measuring NHS Outputs and Productivity: Final Report. York and London: Centre for Health Economics, University of York and the National Institute for Social and Economic Research.

Diewert, W. Erwin. 1976. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4(2): 115-45, May.

Eurostat. 2001. "Handbook on Price and Volume Measures in National Accounts." March, 2001.

Fare, Rolf, S. Grosskopf, B. Lindgren, and P. Roos. 1994. Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach. *Data Envelopment Analysis: Theory, Methodology and Applications*, A. Charnes, W.W. Cooper, A. Lewin, L. Seiford (eds.) Kluwer Academic Publishers (1994) 253-272.

Fisher, Franklin M. 1993. *Aggregation: Aggregate Production Relations and Related Topics*. Cambridge Massachusetts: MIT Press.

Gilbert, Milton. 1961. Quality Changes and Index Numbers. *Economic Development and Cultural Change* 9 (April): 287-94.

Griliches, Zvi. 1988. *Technology, education, and productivity*. Oxford: Basil Blackwell.

Griliches, Zvi, ed. 1992. *Output Measurement in the Service Sectors*. National Bureau of Economic Research. Studies in Income and Wealth. Volume 56. Chicago and London: University of Chicago Press.

Griliches, Zvi. 1994. "Productivity, R&D, and the Data Constraint." *American Economic Review*, 84(1) (March), pp. 1-23.

Grossman, Michael. 1972. *The Demand for Health: A Theoretical and Empirical Investigation*. National Bureau of Economic Research, Occasional Paper 119. New York: Columbia University Press.

Harper, Michael J., Bhavani Khandrika, Randal Kinoshita, and Steven Rosenthal. 2008. Multifactor Productivity Contributions of U.S. Non-Manufacturing Industry Groups: 1987-2005. Paper presented at the World Congress on National Accounts and Economic Performance Measures for Nations, Arlington, Virginia, USA, May 13-18.

Heijink, J., M. Noethen, T. Renaud, M. Koopmanschap, and J.J. Polder. 2008. Cost of Illness: An International Comparison of Australia, Canada, France, Germany and The Netherlands. Health Policy, Vol 88, Number 1, pp 49-61. [**need to check this ref**]

Hodgson, Thomas A. and Alan J. Cohen. 1999. Medical Care Expenditures for Major Diseases, 1995. Health Care Financing Review 21 (Winter), 119-64.

Hulten, Charles R. 2001. Total Factor Productivity: A Short Biography. In Charles R. Hulten, Edwin R. Dean and Michael J. Harper, eds., *New Developments in Productivity Analysis*. Studies in Income and Wealth, Vol. 63. University of Chicago Press for the National Bureau of Economic Research.

IHAT (listed as author of the document), OECD-Eurostat-WHO. 2009. Notes on Improving the Quality of the Estimation of Capital Formation in the SHA. Document code SHA-REV-A1001. Paris, OECD, October 7.

Jorgenson, Dale W., Frank M. Gollop, and Barbara M. Fraumeni. 1987. Productivity and U.S. Economic Growth. Cambridge, Massachusetts: Harvard University Press.

Jorgenson, Dale W., J. Steven Landefeld, and William D. Nordhaus. 2006. A New Architecture for the U.S. National Accounts. Studies in Income and Wealth Volume 66. Chicago: University of Chicago Press.

Martin, S., N. Rice, and P. Smith. 2008. Does Health Care Spending Improve Health Outcomes? Evidence from English Programme Budgeting Data. *Journal of Health Economics* 27, 826-42.

McKeown, Thomas. 1976. *The Role of Medicine: Dream, Mirage, or Nemesis?* London: Nuffield Provincial Hospitals Trust.

Meltzer, David. 2001. Theoretical Foundations of Medical Cost Effectiveness Analysis: Implications for the Measurement of Benefits and Costs of Medical Interventions. In David M. Cutler and Ernst R. Berndt, eds. *Medical Care Output and Productivity*. Studies in Income and Wealth Vol. 62. The University of Chicago Press for the National Bureau of Economic Reseach.

Murray, Richard. 1992. “Measuring Public-Sector Output: The Swedish Report.” In *Output Measurement in the Service Sector*, Zvi Griliches, ed. National Bureau of Economic Research *Studies in Income and Wealth* Vol. 56, pp. 517-42. Chicago: University of Chicago Press.

National Research Council. 1979. *Measurement and Interpretation of Productivity*. Report of the Panel to Review Productivity Statistics. Washington, D.C., National Academy of Sciences.

Newhouse, Joseph P. 1989. “Measuring Medical Prices and Understanding Their Effects--the Baxter Prize Address.” *Journal of Health Administration Education* 7(1): 19-26.

Newhouse, Joseph P. 1992. Medical Care Costs: How Much Welfare Loss?” Journal of Economic Perspectives 6(3), pp.3-21.

Nordhaus, William D. 2006. Principles of National Accounting for Nonmarket Accounts. In Jorgenson, Landfeld and Nordhaus (2006, 143-160).

Organization for Economic Cooperation and Development. 2001. *Measuring Productivity, OECD Manual: Measurement of Aggregate and Industry-Level Productivity Growth*. Paris: OECD.

Organization for Economic Cooperation and Development. 2003. *A Disease-Based Comparison of Health Systems: What is Best and at What Cost?* Paris: OECD.

Philipson, Tomas and Richard Posner. 2008. Is the Obesity Epidemic a Public Health Problem? A Decade of Research on the Economics of Obesity. NBER Working paper no. 14010. Cambridge, Massachusetts: National Bureau of Economic Research. May.

Polder, J.J., and P.W. Achtenberg. 2004. Cost of Illness in the Netherlands. Bilthoven, The Netherlands: Centre for Public Health Forecasting, National Institute for Public Health and the Environment

Price Statistics Review Committee. 1961. *The Price Statistics of the Federal Government*. U.S. Congress, Joint Economic Committee. Government Price Statistics, Hearings, Part 1. 87th Congress, 1st Session, 1961; also published as: National Bureau of Economic Research, General Series, Number 73.

Rice, Dorothy P. 1966. Estimating the Cost of Illness. Health Economic Series No. 6. DHEW Publication No. (PHS) 947-6 (US Department of Health, Education, and Welfare, Rockville, MD).

Rice, Dorothy P., Thomas A. Hodgson, and Andrea N. Kopstein. 1985. "The Economic Cost of Illness: A Replication and Update." Health Care Financing Review 7(1): 61-80, Fall.

Roehrig, Charles, George Miller, Craig Lake and Jenny Bryant. 2009. National Health Spending by Medical Condition. *Health Affairs*, 28 no. 2: 358-367.

Schreyer, Paul. 2001. Published as OECD (2001).

Schreyer, Paul. 2009. Towards Measuring the Volume of Health and Education Services: OECD Draft Handbook. Paris: Organisation for Economic Co-operation and Development. Paper STD/CSTAT/WPNA(2008)12

Scitovsky, Anne A. 1964. "An Index of the Cost of Medical Care--A Proposed New Approach." In Solomon J. Axelrod, ed., *The Economics of Health and Medical Care*. Ann Arbor: Bureau of Public Health Economics, University of Michigan.

Scitovsky, Anne A. 1967. Changes in the costs of treatment of selected illnesses, 1951-65. *American Economic Review* 57(5): 1182-1195.

Shapiro, Irving, Matthew D. Shapiro and David W. Wilcox. 2001. Measuring the Value of Cataract Surgery. In Cutler and Berndt (2001): 411-437.

Stigler et al. 1961. (See Price Statistics Review Committee.)

Trajtenberg, M. 1990. *Economic Analysis of Product Innovation—The Case of CT Scanners*. Cambridge, MA: Harvard University Press.

Triplett, Jack E. 1999. Accounting for Health Care: Integrating Price Index and Cost-Effectiveness Research. In *Measuring the Prices of Medical Treatments* edited by Jack E. Triplett: 220-50. Brookings Institution Press.

Triplett, Jack E. 2001. "What's Different About Health? Human Repair and Car Repair in National Accounts and in National Health Accounts." In Cutler and Berndt, eds. (2001): 15-94.

Triplett, Jack E. 2004. Performance Measures for Computers. In Dale W. Jorgenson and Charles Wessner, eds. *Deconstructing the Computer*. Washington, D.C. National Academies Press. pp. 99-139.

Triplett, Jack E. 2005. *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*. Paris: OECD Publishing.

Triplett, Jack E. and Barry P. Bosworth. 2004. *Productivity in the U.S. Services Sector: New Sources of Economic Growth*. Brookings Institution Press.

Triplett, Jack E. and Barry P. Bosworth. 2007. The Early 21st Century U.S. Productivity Expansion is *Still* in Services. *International Productivity Monitor*. Number 14, Spring 2007, pp. 3-19.

World Cancer Research Fund/American Institute for Cancer Research. 2007. *Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective.* Washington, D.C.: AICR

World Health Organisation. 1977. *Manual of the International Classification of Diseases, Injuries and Causes of Death (ICD-9).* Geneva: World Health Organisation.