WHAT S DIFFERENT ABOUT HEALTH?
HUMAN REPAIR AND CAR REPAIR IN NATIONAL ACCOUNTS AND IN NATIONAL HEALTH ACCOUNTS

JACK E. TRIPLETT

The Brookings Institution
1775 Massachusetts Ave. NW
Washington, D.C. 20036
email: JTRIPLETT@BROOK.EDU

June 25, 1999
(Corrected July 27)

Revised paper for a conference volume,
edited by David Cutler and Ernst R. Berndt:
NATIONAL BUREAU OF ECONOMIC RESEARCH, INC.
CONFERENCE ON RESEARCH IN INCOME AND WEALTH

Discussant: ZVI GRILICHES, Harvard University and NBER

ABSTRACT

What s Different About Health? Human Repair And Car Repair In National Accounts and in National Health Accounts

Jack E. Triplett, Brookings Institution

Existing national health accounts (NHA) focus on sources and recipients of funding for health care. National health accounts have been very successful for the purposes for which they were designed--analysis of resource flows through the medical care system, and answering questions about who pays and who absorbs medical care resources.

However, NHA have been far less successful in measuring real flows--partitioning trends in medical care spending into price change (medical care inflation) and quantity change. The deflation problem means that information from NHA is inadequate for measuring the output and productivity of the medical care sector, and for dealing with questions of welfare: How much does society benefit from expenditure on medical care?

The usual explanation for this state of affairs is some variation on: Measuring health care output is hard. In this paper, I suggest that health care is not a wholly different paradigm from any other analogous service, such as car repair. Consumer ignorance and moral hazard from insurance are encountered in both markets, for example.

The *human repair model* developed in this paper implies measuring health care output by the health implications of medical interventions. Because interventions are, by their nature, specific and relate to specific diseases, the human repair model takes as its starting point expenditures for treating diseases, not expenditures for, say, hospitals. Though the recipient of expenditures may be a hospital, or a doctor's office, past experience abundantly shows that real output measures cannot be constructed by starting from that level of observation.

Implementation of the human repair model in NHA requires two re-orientations to the accounting framework:

(1) NHA need to be integrated into cost-of-disease accounts, which are constructed in a number of countries, and cost-of-disease accounts must pay more attention to time-series comparability than has been the case in the past.

(2) To use information on medical outcomes in NHA requires an adjustment in thinking, to admit going beyond the market transaction--the observational unit that still dominates the measurement of car repair--to consider what medical resources actually do for health. In this respect, human repair is different from car repair--at least in degree.

The model also requires empirical results that fit its framework. The good news is that a growing amount of empirical work is underway that is consistent with the human repair model. The human repair model can in principle be implemented in both market-purchased and government-provided medical care systems. The paper presents an example, an account for expenditures, inflation, and the real quantity for mental health care.

(Revised: June 25, 1999,
with additional corrections, July 27)

What s Different About Health?

Human Repair And Car Repair In National Accounts

and in National Health Accounts

Jack E. Triplett[1]

> The American patient is likely to. . .regard doctors as technicians who are periodically called on to repair his physical machinery.
>
> Aaron and Schwartz (1983, page 17).

Measuring the output of services industries has long been considered hard. "The conceptual problem arises because in many service sectors it is not exactly clear what is being transacted, what is the output, and what services correspond to the payments made to their providers" (Griliches, 1992, page 7). Among the hard-to-measure services, no task has been perceived as more difficult than measuring the output of the health care sector.

Why is measuring health care output so hard? The medical economics literature contains a long list of intimidating and discouraging difficulties. In this paper, I propose to cut through this mostly defeatist list by posing what at first might seem a narrowly-focussed question: Why is health care different from any other analogous service, such as car repair?

Comparing measurement issues in human repair and car repair is instructive. It is not

---

[1] Visiting Fellow, Brookings Institution. A grant from the Eli Lilly Company to the National Bureau of Economic Research supported part of this research. I am greatly indebted to B. K. Atrostic, Ernst R. Berndt, Richard Frank, John Goss, Zvi Griliches, and Thomas Hodgson for valuable discussions and comments on the substance and the exposition of this paper, and to Helen Kim and Jane Kim for research assistance. The paper has also benefitted from seminar presentations at the National Bureau of Economic Research, the Health Care Financing Administration, the Australian Bureau of Statistics, the Australian Institute of Health and Welfare, the Brookings Institution, and the International Symposium on National Health Accounts, Rotterdam, June, 1999.

merely the straightforward analogies: Replacing a shock absorber and replacing a hip are both repairs to a suspension system, diagnostic activity is a crucial part of both production processes, the frequency of costly diagnostic errors is a concern in both types of repairs, and the outputs of both repair industries are enhanced by new technologies for diagnosis, for installation of the part, and also embodied in the part installed. And, as Vaupel (1998) suggests, the subjects of both repair industries are complicated systems, which is why human and automobile mortality functions look remarkably similar.

More importantly, asking why health is different facilitates asking how health is similar. What can we learn from the way we measure the output of car repair that can be applied to the measurement of human repair, and can simplify the health care measurement problem? Health care *is* "different;" but is it so different that we have to start over with a new paradigm?

I contend that health is not *that* different: The paradigm we use for car repair can be applied, with suitable modification, to health care. Emphasizing the similarities in human repair and car repair paradigms makes it easier to design operational measurement strategies. The similarities may also make it easier for national income accountants and users of economic statistics to understand and accept the sometimes controversial extensions to the paradigm that are necessary because health is indeed, in some respects, different.

I. Background

Although one might expect that measuring health care output would entail in some manner measuring "health," most prior economic measurement in health care has been conducted without explicit reference to medical care outcomes. Because output measures in the national accounts of most countries are typically produced through deflation--that is, by dividing health expenditures by a price index--medical care price index methodology has determined the concepts embodied in medical care output measures (except of course in national accounts for countries where medical care is part of the public sector).

Historically in the U.S., the Consumer Price Index (CPI) component for medical care has been used for deflating medical expenditures. This CPI medical care index was until recently constructed from a sample of medical care transactions: a hospital room rate, the price for administering a frequently-prescribed medicine, or the charge for a visit to a doctor's office (see Berndt et al, 1998). Such transactions, which are effectively medical inputs, are sufficiently standardized that the same transaction can be observed repeatedly, which is required for a monthly price index.

The health outcomes of those CPI transactions were never considered explicitly. It is, of course, true that when a consumer paid for an influenza shot, the consumer wanted to reduce the probability of contracting influenza. And if an influenza shot that was more effective in preventing influenza became available, a "quality adjustment" would in principle be made in the CPI to allow for the value of the improvement.

In practice, however, such quality adjustments were seldom carried out in the medical care price indexes, for lack of the required information. A quality adjustment in the CPI requires more than just a measure of health care "quality," which may itself be difficult to obtain. The CPI quality adjustment requires valuation, an estimate of "willingness to pay"--what would a consumer be willing to pay for the improved influenza shot, relative to the unimproved one? For health care, the willingness-to-pay question was hard to answer.

Thus, for two reasons, health outcome measures were ignored. First, the primary focus in constructing the price index was on collecting information on transactions, not on medical outcomes. A collection system that focusses on transactions prices for medical inputs does not routinely yield medical outcomes. Second, when improved medical outcomes did come into the picture (in the form of a CPI quality adjustment), it was not the outcome itself but the consumer's willingness to pay that was relevant.

It was widely noted, even 35 years ago, that the CPI methodology did not adequately account for improvements in medical care. As the influenza shot example suggests, an improvement in medical procedures that raised the cost of treatment but also improved efficacy frequently showed up as an increase in the CPI. When this CPI was used as a deflator, the improved medical care procedure was thereby inappropriately deflated out of the medical output measure.

Two alternatives to CPI methodology surfaced in the 1960's. The first was the idea of pricing the "cost of a cure," estimating the cost of a medical procedure (the treatment of appendicitis, for example). This contrasted with the CPI's focus on hospital billing elements for a medical procedure, such as the hospital room rate and the administration of a pain medication.[2]

Scitovsky (1964, 1967) estimated cost trends for treating selected medical conditions, including appendicitis and otitis media. She reported that the cost of treating illnesses increased faster than the CPI, a result that most economists found puzzling (because the CPI error that it implied went in the opposite direction from what was expected). Scitovsky suggested that the CPI had understated the rate of medical inflation in the 1950's and 1960's because actual charges had advanced relative to the "customary" charges that presumably went into the CPI.[3]

Scitovsky raised some problems with the cost of illness approach that had not previously

---

[2] George Stigler, in testimony on the "Stigler Committee Report" remarked: "...we were impressed by some of the preliminary work that has been done...on problems such as the changing cost of the treatment of a specific medical ailment.... We think it would be possible...to take account of things such as the much more rapid recovery and the much shorter hospital stay...." United States Congress, Joint Economic Committee (1961), page 533.

[3] In recent years, it has been asserted that the error from "list" prices goes the other way; see Newhouse (1989).

been considered: What should be done about potential adverse side effects of a new treatment that was better in some respects (or for some care recipients), but worse in others (or for other recipients)? Her example was a new drug treatment for appendicitis that lowered average hospital stay, reduced recovery time, and was far less painful, but increased the chance of a ruptured appendix, with potentially fatal consequences. Though it was not recognized at the time, the Scitovsky study showed that all the outcomes of a medical procedure must be considered, not just a single one, or just the principal or primary outcome measure. Put another way, it said that looking only at the cost of a unidimensional "cure" (appendicitis treatment), without considering the multidimensional attributes or characteristics of a medical procedure, could produce its own bias. Though this problem was intractable with the analytic tools that were available in the 1960's, it has been addressed in the cost-effectiveness research of the past 10-15 years (see the discussion below).

It is a bit perplexing that, in intervening years since Scitovsky's work, few other estimates of the cost of treating an illness have been made. Cutler et al (1996), Shapiro and Wilcox (1996), and Frank, Berndt, and Busch (1999) followed Scitovsky by three decades.

As a second alternative to the CPI medical care price index, Reder (1969, page 98) proposed to by-pass the medical pricing problem altogether by pricing medical insurance: "If medical care is that which can be purchased by means of medical care insurance, then its 'price' varies proportionately with the price of such insurance." Barzel (1969) estimated an insurance measure of medical price inflation, using Blue Cross-Blue Shield plans.

The medical insurance alternative has not been without critics. Feldstein (1969, page 141) objected that the cost of insurance approach "is almost certain to be biased upward" because "average premiums will rise through time in reflection of the trend toward more comprehensive coverage" and because the insurance plans will purchase "more services or services of higher quality." Moreover, if an epidemic occurred which raised the cost of insurance, it would inappropriately show up as an increase in the cost of medical care, and therefore not an increase in its quantity, unless the medical premium were calculated net of utilization rates. Thus, implementing the insurance alternative requires solving two quality adjustment problems--adjusting for changes in the quality of medical care and of the quality of insurance plans. Additionally, measuring the output of insurance is also hard conceptually (see Sherwood, 1999).

Little empirical work on medical insurance has followed Barzel in the intervening thirty years. Pauly (1999) has recently revived the proposal. He argues that improved methods for measuring willingness to pay make the medical insurance alternative a more attractive option now than it was in the past. In principle, Pauly contends, one could ask how much one would be willing to pay for an insurance policy that covered an expensive medical innovation, compared with one that did not. Weisbrod (1999) noted that no "constant-technology" health insurance contracts exist, no plans promise to pay for yesterday's technology at today's prices, which in itself suggests that the improved technology was worth the increased cost to insurance buyers. Even if the logic of Pauly's proposal suggests an empirical approach, no empirical work exists,

so its applicability to measuring medical price and output has not been tested.

As these references from the 1960's suggest, the major issues on health care output were joined years ago. Until recently, debate on measuring the output of the medical sector largely repeated those 30 year old arguments. The empirical work and the data, too, had not advanced that much beyond the mid-1960's (Newhouse, 1989).

Several things have changed recently in the United States. First, the Bureau of Labor Statistics, initially in the Producer Price Index and more recently in the CPI, has introduced new medical price indexes that are substantial improvements on what existed before (Catron and Murphy, 1996; Berndt et al., 1998; United States Department of Labor, Bureau of Labor Statistics, 1996). Second, a major new research initiative on health care price indexes, using new approaches and new sources of data, has been created by a research group centered in the National Bureau of Economic Research (these studies are described below). Third, information on health care outcomes has been enhanced greatly by recent research on "cost-effectiveness analysis" within the medical establishment itself (Gold et al, 1996).

A task as yet unexplored is building these new price indexes and health outcome measures into an output measure for the medical care sector. The remainder of this paper will develop an approach (which I call the "human repair model"), contrast it with approaches that are used in other parts of national economic accounts and national health accounts, explore the reasons why health care output requires a modification to the measurement conventions typically used for nonmedical services, such as car repair, and, in the last section, present an empirical example of a health account computed from such information.

II. The Conceptual Framework for the Human Repair Model

How do we measure the output of nonmedical services in national accounts? Taking as an example car repair, most countries do something like the following. First, one gathers from some place the total expenditure on car repairs (expenditures on brake jobs, water pump and fuel pump replacements, engine overhauls, and so forth). Next, a government statistical agency takes a sample of car repairs (brake jobs and water pump replacements, say); it computes the price change for brake jobs and the price change for water pump replacements, and from them constructs a price index for auto repair.[4] When the price index is used as the deflator for automobile repair expenditures, the result is the (real) expenditures on the output of the auto repair industry (see United States Department of Commerce, Bureau of Economic Analysis, 1989).

Thus, we have:

[4] This describes, very generally, the BLS methodology for the "auto repair" component of the CPI. See United States Department of Labor, Bureau of Labor Statistics (1992).

(1) $I_{0t} = \sum_i P_{it} Q_{i0} / \sum_i P_{i0} Q_{i0}$
(2a) $Z_{0t} = \sum_i P_{it} Q_{it} / \sum_i P_{i0} Q_{i0} / I_{0t}$
(2b) $= \sum_i P_{it} Q_{it} / \sum_i P_{it} Q_{i0}$
$=$ real expenditure on car repair

The subscript *I* in these equations refers to individual car repairs (replacing brake pads, for example). Equation (1) is the car repair price index, weighted in principle by the quantities of the different kinds of repairs. The first term on the right hand side of equation (2a) is the change in expenditure on auto repair, and equation (2b) gives the expression for the change in real output or expenditure on auto repair.[5]

Constructing a measure of health care output can proceed in ways that are in some respects similar to methods used for nonmedical services. That is, we can assemble data on expenditures on treating groups of diseases, such as, for example, expenditures on treating mental conditions, or circulatory diseases, or, if more detailed data are available, on treating hart attacks or treating depression. If we can construct price indexes by disease, then these disease-specific measures of medical inflation can be used as deflators to obtain measures of the real quantity of medical services by disease, in a manner that is described exactly by equations (1) - (2b). In the rest of this paper, this approach to obtaining real output of the medical care sector is called the "human repair model."

There are great advantages to proceeding by the human repair model. However, there are also some necessary differences between human repair and car repair. The following sections highlight some of those differences.

A. What is the output of the health care sector?

When a human repair expenditure is incurred, it must in some sense add to the stock of

[5] Note that (1) is a Laspeyres price index number, and (2b) is a Paasche quantity index, which is not the usual national accounts index number system. However, at the lowest level of aggregation in the accounts, the price indexes used for deflation come from price statistics agencies in Laspeyres form in most countries. At the detailed level, the resulting deflated output series is therefore Paasche (or worse, a chained series of changes in Paasche quantity indexes). In the U.S., the Bureau of Economic Analysis now uses a Fisher index number system for aggregating over components of GDP, and also for aggregating output in gross product originating by industry data (see: Landefeld and Parker, 1997; Lum and Yuskavage, 1997). BLS has announced that most CPI components were converted to geometric mean indexes in January, 1999 (but not medical services, which remain Laspeyres). No similar announcement has been made so far for the PPI. Currently, PPI medical care price indexes are used for deflation in the medical care components of the NIPA and in the U.S. NHA. At the detailed level, therefore, equation (2b) describes the calculation that is presently in the real medical care components of the U.S. NIPA and NHA.

health, just as car repair adds to the stock of functioning cars.[6] But how should we think about that increment?

There is little disagreement that health is produced by many factors, and not solely by the activities of the medical sector. Diet, lifestyles, environmental factors, genetic endowments, and other influences determine an individual's, or a society's, level of health. It might even be true, as sometimes asserted, that nonmedical influences on health are more important than the medical ones (McKeown, 1976; Mokyr, 1997).

Medical and nonmedical influences on the "production" of health can be represented in a very general way as:

(3) health = H(medical, diet, lifestyle, environmental, genetic, etc.)

"Health" is thus the ultimate output of a "production process" in which medical interventions are one of a number of contributing inputs.

Using equation (3), it is natural to measure the contribution of the health care sector to the production of health by the incremental contribution to health caused by medical interventions. That is:

(4) effectiveness of the health sector = ∂(health) / ∂(medical), other influences constant,

where ∂(health) is the change in health that is attributable to ∂(medical), the incremental resources put into medical care interventions. Equation (4) describes a relation between medical procedures and health, *all other influences on health constant*.

To do this right, ∂(medical) should include the increments of all the resources required by a medical intervention, which may include direct and indirect costs (unpaid caregiving by the patient's family, for example). And ∂(health) should be a comprehensive measure that incorporates all of the effects on health of a medical intervention, including unwanted side effects, if any. Equation (4) implies that the *health outcomes associated with medical interventions define the output of the health care sector*. Let us call this the "medical interventions perspective" on health care output.

The medical interventions perspective on health output requires scientific information on the relation between medical interventions and health status. The information that economists need for measuring health care output is the same as the information needed to determine whether a medical intervention is an effective treatment. The nature of this medical data is

[6] Many medical procedures or expenditures are preventive in nature, they are not strictly speaking human repairs, nor are they disease-related. But car repair expenditures also include preventive maintenance.

discussed more fully in a subsequent section on cost-effectiveness studies.

Notice that equation (4) does not imply that a society's *level* of health is determined by its health expenditures or by the level of medical interventions it supports. Neither does it imply that a society with a higher level of health expenditures necessarily has a higher level of health than another society with lower health expenditures. One often reads or hears statements such as the following: U.S. spending on health care, which amounts to around 14 percent of GDP, must not be productive (says the speaker), because life expectancy in the U.S. is lower than it is in some other countries that spend a smaller amount on health care. This "total health" view of the output of the medical sector is widely expressed. An example is the following:

> Available estimates generally indicate that medical care has been accountable for only about 10% to 15% of the declines in premature deaths that have occurred in this century--the remainder attributable to factors that have helped prevent illness and injury from occurring. This suggests that the promise implicit in many technological inventions may exceed their ability to deliver genuine health gains, at least on a population-wide basis. However, they certainly consume resources. McGinnis (1996)

The total health view implies that one can judge a health care system's effectiveness by comparing a society's level of health with the health sector resources that presumably produce it. I believe this is not a useful way to look at the matter. The "other factors" in equation (1) are not necessarily constant in international comparisons of health and health expenditures, or in comparisons over time.

Distinguishing between the total health and medical interventions perspectives (between a society's level of health and the health implications of its medical interventions) is particularly important where a medical intervention is undertaken to correct the health consequences of unhealthy lifestyles. A car repair analogy may be helpful. Suppose a car owner with a taste for stop light drag races. Severe acceleration has "unhealthy" consequences for the life expectancies of the clutch, transmission, and tires of his car. One would not assess the output of the car repair industry by the life expectancy of clutches on cars used for stop light drag races, nor deduct from the output of the car repair industry an allowance for the low life expectancy of clutches on cars so used. The car mechanic repairs the consequences of the owner's lifestyle. The medical care sector also repairs, to an extent, the consequences of owners' lifestyles, and repairs as well the consequences of other sources of health problems.

Stop light drag races, in the car repairing example, and fatty diets, smoking, sedentary lifestyles, and so forth in the human repair example, are utility-generating activities--people like them, by and large, even though they may recognize quite well that they are harmful to health, or to cars. Although individuals get utility from better health, they also get utility from consumption activities that may have adverse health consequences. The way we want to model the output of health care is not independent of the demand for health care, and the demand for

health care (or the demand for "health") is one of a set of demands for different commodities, some of which have positive and some negative implications for health. These demands, moreover, are complicated by intertemporal considerations, both in the production process for health and in consumers' decision making.

The future level of health is a consequence, at least in part, of actions today--of expenditures for health care and of diet, environmental, and lifestyle influences. Thus, we might modify equation (3) into the intertemporal production process [7]:

(3a) health(t+n) = H(medical(t), diet(t), lifestyle(t), environmental(t), genetic(t), etc.)

Some consumption goods that yield current utility (smoking and fatty diets can serve as examples) have adverse consequences for health in subsequent periods. That is, there are some components of diet where $\partial(\text{health}(t+n)) / \partial(\text{diet}(t)) < 0$, and similarly for some components of lifestyles and of environmental influences.

On the demand side, however, the current level of utility depends on current health (which depends, in part, on lagged values of the right hand side variables in equation 3a) and on the current level of consumption of normal consumption goods, including lifestyle components, such as restful leisure pursuits. Thus:

(5) utility(t) = U(health(t), diet(t), other consumption goods and services(t), lifestyle(t), environmental(t), etc.),

where health(t) is determined by the lagged values in equation (3b).

For some of the goods in equation (5)--goods that I henceforth designate w--$\partial(h(t+n))/\partial(w(t))<0$, but $\partial(U(t))/\partial(w(t))>0$. These are goods whose consumption makes a positive contribution to present utility, but which have an adverse effect on future health. Grossman (1972) emphasized that abstaining from consumption of such goods is like an investment, in the sense that current consumption (utility) is reduced in order to have greater consumption in the future. The future periods may be a long way off, so the adverse consequences of current unhealthy behavior will be discounted by a rational consumer. The future health consequences are normally changes in probabilities, rather than deterministic. Discount rates, assessments of probability changes, and--because of genetic factors, for example--the actual risks of adverse effects may differ greatly across individuals. Thus, their willingness to undertake "investments" in future health--to reduce current unhealthy, but utility-generating,

[7] This specification is not intended to deny that current levels of health care expenditure and current diet or lifestyle affect current utility, but rather to emphasize the time paths of the effects and the fact that individuals' decisions are intertemporal and have intertemporal effects.

consumption activities--may differ greatly.[8] Indeed, Garber and Phelps (1992) remark that a drastic reduction in fatty diets will only increase the (discounted value) of life expectancy by 4 days for men and 2 days for women.

As incomes rise and as consumers as a group become more wealthy, consumption of, say, rich diets and more sedentary lifestyles may increase because these are luxury goods.[9] Because expensive medical procedures are also more readily available in a more wealthy society, income affects health in two ways: It may encourage less healthy behavior, leading to lower health (Grossman, 1972, presents empirical evidence of this). But income also permits more resources to be devoted to medical care, which increases health.

Thus, the effects of fatty diets, sedentary behavior, and smoking on heart disease might merely be offset by the development of expensive treatments, such as heart bypass surgery. If so, the overall death rate from heart disease might be the same as the rate in a society with healthier living and a smaller amount of expensive surgery. But equality of the expected incidence of heart disease in the two cases tells us nothing about the value of the output of the medical sector.[10]

The empirical question that needs exploring is not whether more medical expenditure gives "more" health, in the sense that a society's level of health is positively correlated with its level of medical expenditures. In the specification of equation (3), the levels might not be closely correlated if other influences on health changed adversely. The task is, rather, to compute the marginal value of a medical intervention on health, holding constant or abstracting from nonmedical influences on health. To measure the output of the health sector we need to model the health consequences of medical interventions, and not to compare the aggregate level of health with the resources employed in the health care sector.

On the other hand, lifestyle and other unhealthy behaviors will complicate severely the empirical work necessary to estimate health sector output. It might not be clear whether the clutch failed because the owner continued to indulge his taste for stop light drag racing or because the mechanic installed it improperly. If heart bypass recipients change their lifestyles in more healthful directions, it will lengthen the apparent effect of the medical intervention.

---

[8] There is a remark attributed to the late Mickey Mantle (a famous American athlete) : "If I'd known I would live so long, I'd have taken better care of myself."

[9] Smoking apparently has a low income elasticity, but automobile transportation has a high income elasticity almost everywhere, leading to the observation that automobiles kill more people through reduced exercise than they do in accidents.

[10] It might tell us a great deal about the allocation of public expenditures between, say, medical expenditures and education expenditures that are intended to make individuals more aware of the trade-offs between lifestyles and disease, but that is a different matter.

Conversely, if they revert to unhealthy lifestyles, it will shorten the apparent effect on life expectancy of the medical intervention.

In summary, in this subsection I considered the appropriate conceptual way to think about health care output. I conclude that we should measure it by the health implications of medical interventions, not by the society's level of health.

The medical interventions approach also implies: To find the incremental impact of interventions on health, one cannot proceed by trying to estimate some aggregate of medical interventions.[11] Interventions are, by their nature, specific, and they relate to specific diseases. Measuring the health implications of medical interventions inevitably implies a strategy of examining these interventions on an intervention-by-intervention basis, that is, on a human repair-by-human repair basis.

B. Cost-effectiveness studies and medical outcomes

In the previous section, I proposed that the output of the health care sector be measured, conceptually, by the health impacts of medical interventions. In the cost-effectiveness literature, such an impact is called a "health outcome." Gold et al (1996, page 83) define a health outcome as the end result of a medical intervention, the change in health status associated with the intervention over some evaluation period, or over the patient's lifetime.

A typical cost-effectiveness study compares alternative health care procedures for a particular disease or condition. The numerator of the cost-effectiveness ratio is the total cost difference between two alternatives, including all direct costs and indirect costs such as family provided care during convalescence. The denominator is the difference in health outcomes for the same two alternatives.

United Kingdom Department of Health (1994) provides a tabulated review of cost-effectiveness studies that existed at that time. Garber and Phelps (1992) provide a theoretical framework for cost-effectiveness studies and show that medical cost-effectiveness studies can be interpreted as willingness to pay for medical interventions. Gold et al (1996) provide a common protocol for carrying out such studies.

Health outcomes my be specific to a disease. Gold et al (1996) present examples (in their table 4.1, pages 85-87) of health outcome measures that have appeared in the cost-effectiveness literature. For critical diseases (a heart attack, for example, or cancer), survival probabilities or changes in life expectancy may be used as the health outcome that measures the

[11] An example of what I have in mind here are studies that regress international expenditures on pharmaceuticals on measures of health or longevity. The argument of this section suggests that such regressions are not useful as indicators of the effectiveness of pharmaceutical interventions.

effect of an intervention (by-pass surgery, for example).

Yet, survival is an inadequate measure, because other aspects of health also matter in treatment of life-threatening diseases. For this reason, Gold et al (1996) recommend as the denominator of the cost-effectiveness ratio a relatively new health outcome measure called the QALY--the quality-adjusted life year--a health outcome measure that combines morbidity and mortality into a single measure of health outcome.

QALY is not without controversy. Gold et al (1996) discuss some of its shortcomings, the assumptions required to implement the measure, and the substantial data that it requires. Others have amplified on the shortcomings, arriving at less favorable assessments, at least with respect to its present level of development. Triplett (1999) discusses the relation between cost-effectiveness studies of health care and price index studies, and explains how medical outcome measures such as QALY can be used as adjustments for improvements in medical technology for measuring medical inflation and the real output of medical care services.

C. The national health accounts production boundary, health care output and car repair

Market transactions have traditionally provided the "production boundary" that defines price and output measurement in national accounts and national health accounts. Putting a value on health outcomes crosses this traditional production boundary. Crossing the production boundary has been, and remains, controversial in national accounts and in national health accounts.

Gilbert (1961, page 290) asserted that "...the production boundary must be fixed at the point at which transactions take place between buyer and seller because that is the only point at which value, output and price are settled for things that are bought and paid for. Recovering from an illness is not a unit of output nor its cost a price...." In this view, improvements in mortality or in morbidity are not relevant to measuring the output of the medical care sector because they are not "charged for" explicitly; a measure that combines the two, such as QALY, is doubly condemned. The view expressed by Gilbert is still very much a part of the intellectual heritage of national accounts, and of national health accounts.

In this respect, the health output proposal is not strictly analogous to the way car repair output is measured in national accounts. One can think of car repair as a production process that combines a broken car and a repair to produce a functioning automobile. Yet, no national statistical agency computes in national accounts the increment that car repair makes to the stock of functioning cars, nor calculates explicitly the benefit of the repair to the car owner. In national accounts, the output of the car repair industry equals the quantity or number of, e.g., (constant quality) brake jobs and other repairs--output is measured by deflating car repair expenditures by a price index for brake jobs and so forth. No one tries to assess the output of the car repair industry by some measure of the quality or operational effectiveness of the functioning stock of cars.

Why not just measure the number of health care procedures, as we do for car repair? Doing so preserves the transaction as the unit of observation, which has practical advantages. Alternatively, doing so in a government-provided health care system preserves government expenditures as the relevant resource measure, which has comparable advantages.

One part of the answer is: What we do for car repair is not all that satisfactory if there are significant improvements in the quality of car repair procedures, because the price indexes may not allow for those quality changes very well. Quality change may bias the price and output measures of the car repair industry. Some may think that quality changes in car repair are not a measurement problem (though they probably are).[12] Nearly everyone agrees, however, that improvement in medical procedures is substantial and that quality improvement in medical care is a major part of what we want to include in an output measure for health care. Thus, though both car repair and human repair pose similar price and output measurement problems, the "quality change problem" looms larger in measuring health care output, which justifies, or at least suggests, more radical solutions.

The other part of the answer involves two aspects in which health care differs from car repair or most other services: In car repair, we are willing to assume that the more expensive repair procedure must be better if the consumer chooses it. The consumer could, after all, sell the repaired car (or the unrepaired one). Accordingly, the very fact that the car repair occurred means that it meets a "willingness-to-pay" test.[13]

Economists, and the medical profession, are less convinced of the equivalent assumption in the case of human repair--there are serious doubts that the price of a more expensive medical procedure necessarily measures its greater contribution to health. The consumer has inadequate basis for making informed choices among medical care providers and among options for treatment.

That consumer ignorance makes health care special is frequently asserted, but one can make too much of it. Charging for unnecessary repairs, or for the wrong repair, is also notorious in car repair. A very large proportion of brake pad replacements are coupled with replacing brake disks as well, which should not be the case. Those Cambridge authorities, Click and Clack,[14] recently reported the reason: It is easier to overcharge for the brake repair than to explain

---

[12] Zvi Griliches notes in his comment that the statistics on car repair productivity look peculiar, and suggests that car repair may not be measured very well. See also Levy (1999).

[13] Generally, a representative consumer's willingness to pay guides the determination of how quality improvements should be treated in the CPI. Fisher and Shell (1972) and Triplett (1983) provide theoretical rationales.

[14] Click and Clack are hosts of a popular American radio program on car repair. They were the 1999 commencement speakers at MIT.

the harmless initial noise that normally accompanies replacement only of brake pads. In car repair, as in human repair, the choice of treatments is largely in the hands of professionals, rather than the consumer, and agent problems potentially interfere with the welfare maximizing outcome in both cases.

Additionally, medical economists often emphasize that insurance drives a wedge between payment and valuation. A standard result in medical economics is that insurance causes more demand for medical care than would otherwise be the case. "For many people...[medical care is] paid for through health insurance, and the existence of moral hazard combined with reasonably generous health insurance policies can call into strong question the validity of the simple proposition that prices represent consumers' marginal willingness to pay for the relevant products" (Keeler, 1996, page 189).

However, many car repairs are also paid by insurance, and it is a commonplace observation that car insurance also causes more car repair than would otherwise occur. It is not so clear that insurance makes a fundamental difference between human repair and car repair, although it might be true that the magnitude of its effects are larger in human repair. More likely, insurance gets more attention in the medical industry case because human repair is more important than car repair, both as a share of the economy and in consumer welfare.

Thus, neither consumer ignorance nor insurance creates a fundamental difference between human repair and car repair. The most important difference between human repair and car repair is the fact that the owner can sell the car.

For a car repair, a consumer routinely asks: Considering what the car is worth, should I repair it? Could I get auto transportation services more cheaply by selling the unrepaired car and buying another? Or even: Should I do without a car? If we were to collect the values of the unrepaired and repaired car, we presume that we would find that the repaired car's change in value justified the cost of the repair. But we don't do that, largely because it is not necessary: Because the car could always have been sold, we assume that the car repair meets the willingness-to-pay test.

Suppose for example that for a brake job the car owner had to choose between two different types of brake pads, one which claimed 20 thousand miles life, the other 30 thousand. In principle, one could evaluate the owner's choice by obtaining "outcome" data (did the more expensive pads actually give longer life or more stopping effectiveness?). This would be analogous to measures of medical effectiveness used in cost-effectiveness studies (see the definition at the beginning of the previous section, or Gold et al, 1996). We could then ask, additionally, whether the improved outcome was worth it: We could calculate (value, repaired car) - (value, unrepaired car), and ask whether this difference exceeded the cost of the repair, again in parallel with medical cost-effectiveness studies.

We do not consider carrying out these calculations for valuing the output of car repair.

We assume that the car repair was undertaken because it was economically appropriate for the owner, and for this reason the calculation of cost-effectiveness ratios is unnecessary.[15]

Obviously, in the case of health care expenditure the consumer's decision is different. If the consumer were paying the full cost, the medical expenditure might meet the willingness-to-pay test, in some sense. But because the analogy to selling the unrepaired car is not normally among the consumer's options, ability to pay for medical care influences the result in a way that is not the case in car repair. And even if individuals' willingness-to-pay did dominate medical decision-making, this is generally abhorred for ethical reasons. Additionally, in a government-provided health care system, the consumer's decision on payment is not the element that matters in deciding whether the human repair is "worth it," and provided at public expense.

Thus, in the case of medical care output, it is necessary to estimate societal willingness to pay or something that looks like it. We cannot assume (as in the case of car repair) that because someone undertakes or approves a medical procedure it meets the consumer willingness-to-pay test. In health care, we need data that show that more resource-intensive medical procedures "work," in a sense that we do not need to show that more resource-intensive car repairs are effective (or cost-effective). We need these data in the medical care case mainly because most health care is provided by third party payers, because we do not tolerate social systems where individuals have to make a decision that is analogous to scrapping the car because it was not worth its repair cost--even though someone must eventually make that equivalent decision in allocating scarce resources to health care.

III. Existing Accounting Systems for Health Care Expenditures

Accounting for health care expenditures occurs in three major places in U.S. statistics--national accounts, national health accounts, and cost-of-disease accounts. Several other countries have a comparable three part health accounting statistical system.

Past efforts to create real output or real expenditure measures for heath care have proceeded within one or the other of the first two accounting systems--national accounts and national health accounts. Real output measures have never been developed for cost-of-disease accounts. In this paper, I propose to reorient work on real output of the health care sector toward cost-of-disease accounts.

Developing the proposals of this paper requires, accordingly, an extended overview of the three existing U.S. health care accounting systems. All major countries share one or more of these health expenditure accounting systems, so the discussion and proposals apply to countries

[15] We might also ask, but we don't, whether the car owner really "needed" the better brake pads (possibly because the rest of the car would only last 15 thousand miles). Such a calculation would parallel cost-effectiveness analysis for human repair, where it is common to ask, of a medical intervention that is effective: Are the benefits worth the cost?.

other than the U.S. (I consider explicitly in a separate paper the task of constructing real output measures for countries that have public health care systems, for which prices are not available, and where price indexes therefore are not relevant).

A. National accounts

Expenditures on health care are part of the U.S. National Income and Product Accounts (NIPA), whose best known statistic is Gross Domestic Product (GDP). The statistical agencies of most countries follow, to a greater or lesser degree, the international standard for national accounting, the System of National Accounts, or SNA (Commission of the European Communities et al, 1993). There are no fundamental differences in the treatment of the health sector in the SNA and the NIPA, though the groupings may not be identical across countries, and in countries that have government medical systems, estimating procedures differ substantially from those of the U.S.

In the NIPA, personal health care expenditures are located primarily in Personal Consumption Expenditures (PCE), which means that they are mostly classified as final products. The PCE includes not just consumer out of pocket health spending, but also other payments for health care, such as by employer-provided health insurance. In 1995, medical care expenditures in PCE amounted to $872 billion, about 18 per cent of Personal Consumption Expenditures and about 12 percent of GDP (table 1).

The product and service categories in PCE medical care include drugs, and some other medical goods, but also institutional providers of services (hospitals and nursing homes, for example). Medical *goods* that are inputs to medical care are classified by a product classification system, like other parts of the PCE, but medical services are classified by type of *provider.* Another way to put it is to say that medical services are grouped by an industry classification system, rather than by a product classification system. Thus, a particular pharmaceutical will be counted in the same place whether it is sold by a grocery store or a pharmacy; but if a medical procedure shifts from in-hospital to a doctor's office or clinic, expenditures on it will show up in a different grouping in the PCE.

The distinction between goods and services classifications in the NIPA is driven largely by data availability (or at least I have never seen a conceptual argument supporting the distinction).

The inconsistency between goods and services classifications is perhaps subsidiary to another point about the NIPA classification system: Nowhere in the categories used for medical expenditures in the NIPA does the NIPA distinguish what medical spending is for--the system does not record what is purchased when medical spending takes place. Expenditures for cosmetic surgery and heart surgery are both (if both are done in a hospital) grouped together in hospital expenditures, and pharmaceuticals for acne and for angina are combined in the medical goods components. If hospital expenditures are growing, there is little in the national accounts

(or in the national health accounts) that will tell us very much about the hospital medical procedures that are fueling overall growth, or about the diseases that are being treated.

Moreover, the NIPA classification naturally orients national accounts producers and users to a particular specification of the deflation problem: With the NIPA classification system, it seems natural to look for deflators for "physicians" and "hospitals" (or even "nonprofit hospitals" and "propriety hospitals"), for those are the expenditure categories that require deflation. With the NIPA classification system it seems less natural to ask: "What is the price index, e.g., for coronary disease, or for heart attacks, or for depression, or for eye surgery?" And with the NIPA system for classifying health care, it is not clear what one would do with price measures for treating diseases, even if they became available. Because price indexes for treating diseases or groups of diseases are in fact becoming available, the fact that the NIPA system has no natural place for them is a severe deficiency.

B. National health accounts

A second U.S. accounting for health care expenditure is National Health Expenditures (NHE), often referred to as the National Health Accounts (NHA). Where the NIPA treat health care as one among many products and services purchased or consumed by households, the emphasis in the NHE is on assembling comprehensive data on total national expenditures on health, and on the sources and recipients of those funds. As noted earlier, total U.S. health spending in the NHA equaled 14 percent of GDP in 1995; total *personal* health expenditures were $869 billion in the NHA in the same year, very close to the 12 percent of GDP total in PCE (table 1). The remainder of NHA health expenditures includes health education, investment, and certain other components, which appear in other parts of the NIPA (such as the accounts for government).

The U.S. national health accounts have been produced since 1964 (Rice, Cooper, and Gibson, 1982; Lazenby et al, 1992). Rice, Cooper, and Gibson (1982) refer to a compatible series for private health expenditure that extends back to 1948, and note even earlier estimates of total U.S. health care spending.[16] Health accounts are also constructed for demographic groups, the aged, for example (Waldo et al, 1989).

The national health accounts are organized in the form of a matrix. Table 2 presents a condensed form of the accounts as they are now published (see Lazenby et al, 1992; Levit et al, 1996).

The columns of the matrix arrange health care expenditures by major source of funding (for example, households, private health insurance, government). As table 2 shows, 54 per cent of U.S. health care expenditures ($536 billion) came from private funding, and 46 per cent ($455

[16] According to one of these early studies, health care accounted for 4 percent of U.S. GDP in 1929.

billion) from government funding; private insurance and the federal government are the biggest individual funding sources for total national health expenditures and for expenditures on personal health care. In these proportions, the U.S., of course, differs from most other industrialized countries. More detail is routinely available in the NHA on Federal and on state and local funding sources, and more detailed estimates are periodically provided for business and households (Levit and Cowan, 1991).

The rows of the NHA matrix show the uses of the funds, in the sense that they detail the sectors or economic units that receive the expenditures on health care. The categories are similar to those in the NIPA (see table 1). However, the close agreement between NHA personal health care expenditures and PCE medical care expenditures at the aggregate level does not extend to the components of medical care. Hospital expenditures, for example, differ in the two accounts (see tables 1 and 2), as does "other professional services." Several categories appear in one system but not as a separate entry in the other (home health care is the largest such category). A NIPA-NHA reconciliation is contained in Rice, Cooper, and Gibson (1982); this appears to be the latest published reconciliation, though a new one is forthcoming.

In the case of health care services, the national health accounts distinguish, again in parallel with the NIPA, the organizational unit that receives the funds, rather than, strictly speaking, the type of service. For example, the same type of service for treating a disease might be performed in a doctor's office or in a hospital; the national health accounts would distinguish whether the expenditure was received by a hospital or by a doctor's office, but would not distinguish the expenditure by the type of service performed, or by the disease category for which treatment was rendered. The classification of individual units receiving payments for medical services is based on the U.S. Standard Industrial Classification system (Executive Office of the President, Office of Management and Budget, 1987).

For drugs, eyeglasses, and other durable and non-durable "therapeutic goods," the national health accounts distinguish, as do the NIPA, the type of goods, using product code classifications from the U.S. Bureau of the Census. Expenditures on therapeutic goods count only those goods that are purchased from retail outlets. Any therapeutic goods that are received by patients in hospitals, for example, will be recorded in the expenditures on hospital care.

Thus, the NHA expenditure classification does not strictly speaking correspond to a "goods-services" distinction, nor does it group expenditures by commodities in the usual sense. It is instead a classification based on the institutional structure of the recipient of the funding. In fact, the category "drugs and other therapeutic goods" is really a classification that groups medical expenditures that are received by the retail trade sector. This classification has implications not only for the interpretation of the published components, but also for other aspects of the NHA: For example, the proper deflator for the pharmaceutical portions of NHA will exclude drugs sold to hospitals because they are not included in the drugs that are counted

separately in the NHA "goods" classification scheme.[17]

Because a national health accounts matrix is prepared for each year, it is useful to think of the national health accounts as a 3-dimensional matrix. There are the two dimensions shown in table 2. This is like one page in a book. Then, because there is an equivalent to table 2 for each year, there are a series of pages in the book. One can follow any of the columns, or any of the rows, or any combination of cells from the matrix, through time to construct a time series. The constructors of national health accounts, therefore, pay a great deal of attention to time series comparability (which is not the case for cost-of-disease accounts, considered below).

Like the NIPA accounts, the structure of the NHA also orients producers and users toward deflation for institutional units, such as "hospitals." Severe problems with past deflation at this level, and with available U.S. price indexes (Berndt et al, 1998) has led to increased use in the NHA of a broader list of hospital inputs as a proxy for output price measures (Freeland et al, 1991). It is widely recognized in the price index literature that measuring the prices of inputs usually provides a poor proxy for the movement of output prices, unless there is no productivity in the industry, which is surely not the case for medical care.

The U.S. NHA are well known and are widely used for analyzing the economics of the health care sector. Similar health accounts are produced in other countries.

For example, France has a system of health accounts, Comptes Nationaux de la Sante (Ministere du Travail et des Affaires Sociales, 1996) that is similar to the U.S. NHA, in that it provides information on the source of funds in France for health care and health spending and on the institutions receiving the funding. The French system dates from 1976, and is available, as is the U.S. system, in quarterly and annual time series.

The Comptes de la Sante are referred to as a "satellite account," a term that is not generally applied to the U.S. NHA.[18] Despite this, the groupings of data in the Comptes de la Sante are similar to those in the U.S. NHA, with some exceptions, such as the inclusion of expenditures for spas ("cures thermales") in France. The Organization for Economic Cooperation and Development (1997, page 121) lists ten OECD countries where satellite

---

[17] Pharmaceutical price indexes in the PPI are based on all sales by manufacturers, and thus do not provide appropriate deflators for the NHA, as the NHA are now constructed. Ellison and Hellerstein (1999) found that for one pharmaceutical product (cephalosporins) prices for drugs sold to hospitals moved very differently from those sold to retailers, and so presumably the prices charged to consumers by pharmacies.

[18] On satellite accounts, see the description in the SNA (Commission of the European Communities et al, 1993), though the reader is warned that this chapter is not particularly clear. For the development of the concept of the satellite account, see Vanoli (1975, 1986), Teillet (1988), and Pommier (1981).

accounts for health have been produced or are "under study;" the U.S. NHA is not included in the OECD list of satellite health accounts, presumably because the U.S. NHA are imperfectly articulated into the U.S. NIPA.

Three properties of national health accounts deserve emphasis. First, total national health expenditures, and other NHA aggregates, are built up from the bottom. For the most part, these are not estimates where one starts with a total and distributes the total among the different categories. Rather, one adds up the categories to get the total. Cost-of-disease accounts (described in the next section) distribute totals to categories, and hence cannot be estimated independently of NHA-type accounts.

Second, the fact that the NHA accounts are arranged in a matrix means that there are cross checks. All the row and column totals must add up. But because estimates for different cells of the matrix come from different data sources, which may be compiled by different methodologies and may not be consistent across different sources, adjustments may have to be made to source data to insure that all row and column totals in the matrix balance. Though this assures consistency in the matrix and corresponds to good economic accounting principles, it can mean that the entry in a particular cell of the matrix does not agree with the best independent estimate of the value for that cell.

Third, as already noted, NHA accounts preserve time series comparability. They are explicitly designed for use in analyzing time trends in health care expenditures.

Unlike the case of national accounts, where the SNA provides a standard for producing internationally-comparable data, no international standard for health accounts exists at present. The World Bank has set out informal guidelines for NHA development for borrowing countries (McGreevey, 1996). However, the OECD, with funding from the US agency that produces the US NHA, has released a proposal for an international standard (OECD, 1997). No price index or method for producing real output measures is discussed in the OECD report.

C. Cost-of-disease accounts

1. Overview.

In some respects, the concepts and structure of national health accounts resemble "flow of funds" accounts, in that they focus on financial flows of health expenditures, and on sources and recipients of funds. The two dimensions of the national health accounts matrix--sources of funds and recipients of expenditures--have been useful for many of the analytic tasks for which health expenditure data are required. However, these two dimensions are not the only useful way in which one might array health expenditure data.

Consider the subtotal "personal health expenditures." In the national health accounts this category has the following definition: "Personal health care comprises therapeutic goods or

services rendered to treat or prevent a specific disease or condition in a specific person" (Lazenby et al, 1992). As this definition suggests, one can envision disaggregating personal health care expenditures by expenditures on specific diseases. Such a disaggregation is most commonly performed as part of a "cost of disease" or "burden of disease" study.

For present purposes, I define a *cost-of-disease study* as one that estimates expenditures for treating disease, sometimes referred to as the direct costs. A *burden of disease study* would also include indirect costs of disease--unpaid care provided by family members, and loss or reduction of earnings, and it would put a value on the losses from premature mortality and from the disutility of disease itself. Examples of burden of disease studies are Rice (1966) and Murray and Lopez (1996).

A burden of disease study considers all the social and economic costs of disease, and not just--as in a cost-of-disease study--the direct costs, or direct monetary expenditures. Put another way, a cost-of-disease study estimates the cost of treating diseases that are treated; a burden of disease study would include additionally the economic and social costs of diseases that are not treated, or for which treatment is ineffective.

Burden of disease studies correspond to a broader economic accounting that goes beyond the traditional market boundary adhered to in national accounts and in national health accounts (see the section on the production boundary, above). Mainly for reasons of space, I will not pursue any of the implications of this broader accounting in the present paper. The present inquiry, then, will concern only the direct costs of treating illness, not because those other costs are without relevance, but because understanding the implications of direct resources that are put into the health care sector is a step toward any broader accounting. At present, the goal is creating real measures of the output of medical treatments for disease, and not, or at least not at present, of the real cost of diseases that are not treated.

The first systematic U.S. disaggregation of health expenditures by disease appears to be Rice (1966), although she cites predecessors. Subsequent updates include Cooper and Rice (1976), Hodgson and Kopstein (1984), and Hodgson and Cohen (1998). These accounts are summarized in table 3.

2. Classification Matters

There are of course thousands of diseases, conditions and diagnoses. Some grouping of conditions must be carried out. Classifications systems provide the building blocks for much of economic statistics--though their properties are often ignored by the economists and health care analysts who use them.

The most widely used disease classification system is the International Classification of Diseases (or ICD), which has gone through a number of revisions since its inception. The International Classification of Diseases, Injuries and Causes of Death, 9th Revision (ICD-9) was

developed by the World Health Organization (WHO) and issued in 1977. The classification system is intended to produce comparable cross-country health statistics, particularly on causes of death. A later revision (ICD-10) is not yet in general use for U.S. statistics.

Experiments with using the ICD system for producing U.S. hospital statistics began in the 1950's. Experience led to a U.S. modification of ICD-9, known as ICD-9-CM (for Clinical Modification), that is now in extensive use for coding hospital records. The main differences from the international system are: more detail (that is, more specific and precise codes for medical conditions), elimination of some ambiguities in pregnancy and childbirth conditions and in some other areas, and changes in the presentation to make the system easier to use by data coders. Despite claims on the latter point, private publishers have introduced their own, more user-friendly editions of the classification manual. The growth of such a publication market underscores the increasing practical utilization of ICD-9.

The ICD-9 classification system conforms consistently neither to an anatomical nor to an etiological or causal principle. It does, however, conform generally to the way diseases are treated and to the way medical specialties are demarcated. This makes the ICD an advantageous and natural system to use to generate economic data on the treatment of disease. It is now used, not just for classifying incidences of diseases and causes of death, but also for classifying a wide range of economic data, compiled mainly for administrative purposes.

A second classification system is relevant to this paper--the “Diagnosis Related Groups” system (DRG) of the U.S. Health Care Financing Administration (HCFA). This system is used to classify hospital and medical procedures for the purpose of making government reimbursements under Medicare and other government health programs. The third revision of the DRG system was introduced for Medicare hospital services in 1986 and is still in effect. The Australian DRG system is a modification of the U.S. DRG system.

The structure of the DRG system is designed to organize hospital admissions by the resources that would be expected to be spent in the treatment of a particular admission. It is thus a classification system that yields, by its design, economic data on the costs of illness. The coding of principal diagnoses under the DRG system conforms, with certain exceptions, to ICD-9 chapters, or more precisely, to the ICD-9-CM, which is a modification used for administrative purposes in the U.S.

The DRG system has, nevertheless, two deficiencies for the purposes of this paper. First, it is not an international system, even though the U.S. and Australia share similar DRG systems. Second, it is not well suited to recording the incidences and prevalence of diseases because, below the first level or chapter groupings, each DRG is a grouping of diseases which might be rather different, but which have similar expected treatment costs.

Cost-of-disease studies use ICD-9 chapters. At the chapter level, ICD-9 and DRG systems are similar.

Two of the U.S. studies in table 3 use the ICD-8 system and the other two use the ICD-9. A certain amount of noncomparability is thereby introduced. Another problem is inconsistency in coding practices: special problems here are the coding of diseases of infancy, of old age, of certain respiratory diseases, and of diseases, like diabetes, that typically have extensive co-morbidities.[19] Coding inconsistency is a long-standing problem with medical data. For example, McKeown (1976) notes the long-term decline in "old age" as a cause of death--it has gradually been replaced with more precise coding of a medical condition, which reflects not only increased medical knowledge, but also changes in attitudes and social mores.[20]

3. Estimating methods.

As noted above, similar cost-of-disease disaggregations have been produced recently for several countries, including the U.S. (Hodgson and Cohen, 1998), Canada (Moore et al, 1997), the United Kingdom (actually, England and Wales--National Health Service, 1996), and Australia (Mathers et al, 1998). This section describes their methodologies.

Cost-of-disease studies typically distribute totals for health care expenditures among disease categories. That is, they are not "bottom-up" estimates, as are the NHA, and in fact they typically start from NHA aggregate health care expenditures.

The methodology can best be understood by an example, for which I use the allocation of hospital expenditures. The other components of medical expenditures are calculated in similar ways, but of course the data vary according to the component, and to an extent, according to country. Details for the U.S. are contained in Hodgson (1997), Chapter 5.

Total expenditures for inpatient hospital care are computed and published in the NHA (see table 2). This expenditure is allocated to ICD-9 chapters by the following steps.

The National Hospital Discharge Survey (NHDS) gives the total number of inpatient hospital days and subtotals by ICD-9 Chapter. The average charge per inpatient hospital day, grouped by ICD-9 chapter, is found in the National Medical Expenditure Survey (NMES) for the year 1987; this charge is updated to 1995 by the CPI Hospital Room Price Index. For each ICD-9 chapter, the number of hospital days (NHDS) is multiplied by the average charge per day corresponding to diseases in that chapter (NMES); when each of these products is divided by the total for all ICD-9 chapters, the result is the share of expenditures allocated to each chapter. This share is multiplied by total NHA hospital expenditures (which are, of course, determined

---

[19] Hodgson (1997) contains a special chapter on the problems of estimating costs for diabetes.

[20] Aaron and Schwartz (1983) quote a British physician who remarked that the body gets "a bit crumbly" after age 55. For a very different view of aging, see Vaupel (1998) and Manton and Vaupel (1995).

independently of the two data sources used to calculate the share) to determine the hospital inpatient expenditures for each ICD-9 chapter.[21]

Hodgson and Cohen (1998) were able to allocate 88 percent of NHA personal healthcare expenditures to a medical diagnosis, and 98 percent of major categories such as hospitals, physicians and dentists services, and nursing homes (see table 4). Although all expenditures for prescription drugs were allocated to an ICD-9 chapter, only 35 percent of nonprescription drugs and related goods could be allocated. It is not entirely clear, for example, for what medical condition aspirin will be used. Information on amounts allocated is contained in table 4.

The basic methodology for cost-of-disease studies was developed by Rice (1966). The methodology used in Canada (Moore et al, 1997) is very similar to that used in the United States. Estimates in the UK (United Kingdom Department of Health, National Health Service, 1996) apply a single average cost of hospitalization across all medical conditions, and thus lack the refinement of the Canadian and U.S. studies, which differentiate cost per day across different classes of illnesses. The Australian study contains unique aspects that reflect that country's health care system.

Though the basic estimating methodologies are similar, that does not mean cost of illness studies are exactly comparable over time, or across countries at one point in time. Estimating methods, classifications, data sources, and to some extent medical practice, the diseases themselves and how they are diagnosed, classified and treated, as well as other considerations have changed over the nearly 35 years that are covered by cost-of-disease estimates in the United States. For all of these reasons, intertemporal comparability may be compromised.

Similarly, data sources, national practices, and estimating methods create noncomparabilities in international comparisons. For example, in the U.S. estimates drugs administered in hospitals are included in hospital expenditures; in Canada, drugs administered in hospitals are removed from hospital expenditures and placed in pharmaceutical expenditures.

IV. Comparisons and Trends, Cost-of-Disease Accounts

Exactly comparable or not, it is very useful to examine the trends of expenditures by disease category, and to make international comparisons of them. Doing so is problematic: To date, cost-of-disease studies have not been produced with an eye toward time series comparability or toward international comparability. Though a more comprehensive study would

---

[21] "In effect, HCFA's estimates of inpatient hospital expenditures are distributed by sex, age, and diagnosis, according to the distribution of days of hospital care weighted by the average charge per day." Hodgson, 1997, page 6. In the above, I have ignored the demographic parts of the calculation. Catron and Murphy (1996) present a similar disaggregation of U.S. hospital revenue for 1987. In their data, circulatory diseases rank first in hospital revenues, and digestive system diseases second.

first make adjustments for time series and international comparability, this exceeds the scope of the present study. According, I content myself with a contribution to the demand for future international and time series comparability in cost-of-disease studies (demand for statistics tends to create its own supply).

A. International comparisons of expenditures by disease

Table 5 compares partitions by chapters of the ICD-9 for total health expenditures in the U.S., Canada, England, and Australia. Table 6 shows a similar partition for hospital expenditures.[22]

The proportions of health spending by disease differ from country to country, yet there are also broad similarities. For example, circulatory system diseases are the largest expenditure category in U.S. overall health care spending (nearly 15 per cent of the total) and in U.S. hospital spending (19 per cent); they are also the largest expenditure category in Canada, and the second largest expenditure category in the U.K. (12 per cent) and in Australia (12 percent). Circulatory diseases are only the fourth largest category of spending in Australian hospitals (however, our preliminary concordance for the classification system for Australian hospitals may have noncomparabilities in it). In the U.K., mental disorders are the largest spending category; in the U.S., they are the second largest category of hospital spending, and third in overall expenditure. Endocrine, metabolic and immunity disorders account for a larger proportion of U.S. spending than in the U.K., Canada and Australia, presumably because of AIDS. Digestive system diseases are relatively more important in the U.S. and in Australia than in the U.K., with Canada in an intermediate position--closer to the U.K., overall, but close to the U.S. in hospital spending. Other differences exist.

What accounts for international differences in the composition of healthcare spending? Several potential causes are topics for future research.

First, there are international differences in the incidence of diseases. For example, Australia reportedly has one of the highest rates of skin cancer in the world; that would push up Australia's relative spending on ICD-9 chapter 2 (cancers), which at six percent is somewhat higher than the proportion in the U.S. and England (though lower than in Canada).

Second, there are also case-mix effects. For example, skin cancer is a relatively low cost form of cancer, and it is frequently treated outside of hospitals; possibly for this case-mix reason, the proportion of Australian hospital spending on cancer is substantially lower than in the other three countries.

Third, there may be international differences in cost per case, even aside from case-mix considerations. Costs per case may differ because some countries employ less effective

[22] Sources are given in the footnotes to the tables.

treatments (see Aaron and Schwartz, 1983, for some examples), or because some countries adopt more cost-effective procedures, or because of international differences in medical industry efficiency. These matters are not pursued here.

International comparisons of costs requires information on prices. Cross-country prices for medical care appear in the "purchasing power parity" statistics published by the Organization for Economic Cooperation and Development (OECD). However, the adequacy of PPP indexes has been questioned recently (Castles, 1997). Additionally, international comparisons of the costs of diseases requires price or cost differences by ICD-9 Chapter, and not just for an overall "medical care" aggregate. This point parallels the argument developed in this paper for time series comparisons.

Finally, as already noted, noncomparabilities exist in these data, and these will affect the percentages presented in tables 5 and 6. The totals do not correspond to exactly the same definitions. For example, Australian hospital data exclude certain hospitals. Though the ICD-9 provides an international standard for classifications, it is sometimes not applied consistently. For example, it appears from tables 5 and 6 that complications of pregnancy and childbirth account for a considerably smaller proportion of U.S. health care spending than in the other three countries (which is puzzling because birth rates are similar in the U.S. and Australia, for example). However, an examination of incidence rates suggests that data for the other three countries include normal pregnancies, which are not treated as a disease in U.S. data. An adjustment for pregnancy and childbirth costs can be made to U.S. data, but there are doubtless other noncomparabilities that have not yet been explored.

B. Trends in U.S. health expenditure by disease, 1963-95

Tables 5 and 6 show cost-of-disease accounts for the United States for the four years for which these accounts have been compiled (1963, 1972, 1980, and 1995). To correspond to the “production boundary” of NHA, I tabulate only the direct costs, though the sources also present indirect costs of disease. Table 7 shows average annual rates of increase computed between each of the years for which U.S. cost-of-disease accounts exist.

The first caveat to be expressed about these tables is that researchers who have assembled cost-of-disease accounts warn that they are not comparable over time. For one thing, the classification systems have changed. The first two U.S. studies use ICD-8, the second two, ICD-9. Unlike conventional national accounts, where "bridge" tables would have been constructed to permit moving more or less consistently across changes in classification structures, no such adjustments exist for cost-of-disease accounts.

Another noncomparability arises because the proportion of expenditures that can be allocated to disease changes over time. The effect of this at the aggregate level can be seen from comparing the first two lines of Table 3: When the proportion of unallocated expenditures falls (true between 1972 and 1980), the rate of growth of allocated expenditures will exceed that of

total expenditures. Conversely, when the proportion of unallocated expenditures grows, the rate of growth of allocated expenditures will fall short of the growth rate of total expenditures (as is true for the 1980-1995 comparison).

Changes in unallocated expenditure may affect rates of growth for ICD-9 chapters as well. For example, if data become available to allocate ambulance expenses by disease category, the new allocation would probably affect, disproportionately, ICD-9 chapters "Injury and Poisoning" and "Diseases of the Circulatory System," compared with, say, "Skin Diseases" and "Congenital Anomalies."

Finally, there are other differences in estimating methods, data availability and presentation conventions that also limit time series comparability. Some of these changes can be discerned and adjustments made by users; but for most of them, only the compilers have sufficient knowledge of the data and estimating methods to construct appropriate bridge tables. Again, it is common practice in national accounting and in national health accounting to link out, so far as possible, the effects of changes in data availability, so as to construct a more nearly comparable time series. Little concern for their time series properties is evident in compilation of cost-of-disease accounts.

Judging from the uses and demands for other economic statistics, the lack of time series comparability for cost-of-disease accounts is puzzling. There is, clearly, some value in knowing that circulatory diseases account for nearly 15 per cent of U.S. health care spending in 1995 (table 5), and nearly 19 per cent of U.S. hospital spending in the same year (table 6). But there is also, surely, great interest in the rate of growth of U.S. spending on circulatory diseases, or on mental diseases. Little direct data on rates of growth for expenditure by disease exists. In the future, meeting time series uses for cost-of-disease accounts should be added to the tasks of their producers.

Leaving aside the noncomparabilities and time series inadequacies of the basic data, and taking the data only for what they present, table 7 shows the following. Though diseases of the circulatory system are the largest expenditure category in the U.S., growth rates are only marginally above average for recent years (9.9 compared with the average of 9.3 percent between 1980 and 1995). Diseases of the digestive system, once the largest category of U.S. expenditure (table 3), show a growth rate that is well below average in recent years (7.3, compared with the 9.3 per cent growth of total allocated expenditures--see table 7). The growth rate for expenditures on mental disorders (considered at greater length in a subsequent section) is about average, for the 1980-1995 interval.

V. Implementing the Human Repair Model

For the health sector, national accounts, national health accounts, and health satellite accounts all share an unresolved problem: How does one construct adequate real output measures for medical care? How does one measure the real growth in medical care services?

In the present section, I develop a health accounting structure that is derived from the human repair model, which will yield a real output measure for medical services, using existing and prospective data. The accounting structure is implementable now, in principle. It also facilitates, as I will show, use of new data that are being generated from a variety of sources, including price index studies and cost-effectiveness studies. These new data are difficult to integrate into the existing structure of national health accounts.

However, economic measurement of medical care would be much improved with new data on medical outcomes, prices, and quantities of services. The accounting structure is also intended as a framework that suggests the directions in which we can push data development to improve the measurement of the health care sector's output.

A. Data for estimating the output model

The starting point for estimating real output of medical care using the human repair model is expenditure on diseases. The major existing data that are organized by disease are in cost-of-disease accounts, as discussed in a previous section. Cost-of-disease accounts disaggregate medical expenditures by ICD-9 chapters.

Considering the number of diseases identified in the ICD-9 classification system, disaggregating by ICD-9 chapters is a beginning, but it does not go far as one might like for empirical work on the human repair model. For example, ICD-9 chapter 7 (Circulatory Diseases) covers ICD-9 codes 390-459; of these, codes 393-429 are heart disease codes, of which codes 410-414 are ischemic heart disease, among which code 410 is acute myocardial infarction, or heart attack.

Cutler, McClellan, Newhouse, and Remler (1998) estimate a price index for heart attacks. This is the level at which practical research on price indexes and cost-effectiveness (discussed below) must be carried out. For the circulatory disease chapter of ICD-9, some additional disaggregation of expenditures is available. Hodgson, for example, estimates that coronary (ischemic) heart disease (ICD-9 codes 410-414) accounts for roughly half of total expenditure for all heart disease, and additional detailed estimates may be available in the future.

Although additional disaggregation beyond the ICD-9 chapter is essential, at some point more expenditure detail will be both impossible to obtain and perhaps inappropriate: The greater the detail at which expenditures are disaggregated, the more likely that expenditures on a particular episode of illness encompass multiple individual ICD-9 codes.

Section II developed the idea that the output model for health care must build in data on the outcomes of health care procedures. Two recent bodies of research make use of or generate health outcomes.

1. Cost-effectiveness studies.

Cost-effectiveness studies were described in section II.B. An increasing number of cost-effectiveness studies are being carried out within the health care industry itself.

The effectiveness part of a cost-effectiveness study requires a measure of health outcome. The denominator of the cost-effectiveness ratio is the difference in health outcomes for two or more alternative treatments for the same disease. The increasing employment of cost-effectiveness studies in medical decision-making means that an increasing number of health outcome measures for different diseases are being generated and also that increased research attention is being given to improving measures of health outcomes. The potential value of this research for measuring the output of the medical care sector is tremendous, even if, as Pauley (1999) and others have suggested, substantial problems with existing measures of health outcomes remain to be resolved.

2. Price index research.

A number of recent studies have been undertaken by a group of researchers at the National Bureau of Economic Research that have the explicit objective of measuring a price index for some part of the health care sector. Examples are Cutler et al (1998) on heart attacks, Frank, Berndt, and Busch (1999) and Berndt, Cockburn, and Griliches (1996) on the treatment of depression and depression pharmaceuticals, and Shapiro and Wilcox (1996) on cataract surgery.

In the heart attack study, the medical outcome measure was the increase in life expectancy associated with more resource intensive heart attack treatments. In the depression studies, the outcome measure was the elimination of the symptoms associated with a diagnosis of severe depression, without holding constant methods of treatment (or, to put it another way, without necessarily holding constant the characteristics of the transaction, as with traditional price and output measurements). Other, similar studies are underway.

A price index study such as Cutler et al (1998) is similar to a cost-effectiveness study, differing mainly in the following ways (an extended discussion of the relation between cost-effectiveness studies and price index studies is contained in Triplett, 1999). First, the health outcome measure in Cutler et al (1998) was life expectancy, not QALY (cost-effectiveness studies have also employed life expectancy in the past: see United Kingdom Department of Health, 1994; Gold et al, 1996). If heart disease treatments had no implications for quality of life (for example, the ability to exercise or conduct daily living without chest pain), then an increase in life expectancy is an increase in QALY. Use of QALY would extend and enhance the measures in Cutler et al (1998).

Second, Cutler et al (1998) value the change in life expectancy, that is, they put a dollar value on the medical outcome. Medical cost-effectiveness studies do not do this (see the discussion of this point in Gold et al, 1996). Valuing medical outcomes for price indexes is discussed in Triplett (1999), and in Triplett and Berndt (1999).

VI. An Example: Implementing the Model on Mental Health Care Expenditures

Treatments for mental disorders account for over eight per cent of total U.S. health care expenditures and about a tenth (9.5%) of all allocatable U.S. personal health care expenditures (tables 5 and 8). It is well known that the U.S. spends about one-seventh of its Gross Domestic Product (GDP) on medical care (the largest proportion in the world), so mental health care expenditures make up just over one per cent of gross domestic product (GDP).[23]

By international standards, the U.S. mental health care expenditure share is not particularly high. The mental health share of U.S. hospital expenditures is about 12%, and for hospitals and nursing homes combined, 13.7%. This is about the same as the Canadian share of hospital expenditures going to treat mental illness (13.9%), and considerably lower than the comparable share in the U.K (over 17% of hospital expenditures). Even though the Australian hospital share is smaller than in the U.S. (under 5 percent), the share of mental health in total Australian health expenditures is about the same as in the U.S. (see tables 5 and 6).

To help understand trends in such a significant portion of health care expenditures, I split U.S. mental health expenditure trends into mental health care inflation and quantity of mental health care services. I then adjust the inflation and real medical services trends to take account of new data and recent research. The estimates show how new information on inflation and quantity of medical services can improve National Health Accounts (NHA).

A. Trends in U.S. mental health expenditures

Complete cost-of-disease account have been constructed for the U.S. for only four years, as noted above. However, U.S. mental health expenditures have been estimated much more frequently. For the interval 1954-1996, more than twenty different single year estimates of expenditures on mental health treatment exist. A dozen of them were reviewed in Rice, Kelman, and Miller (1991). A list of studies appears in Appendix table A-1, which also presents each study's estimate of mental health care expenditures.

Most of these estimates originate with two groups of researchers. Dorothy Rice and her collaborators at the National Center for Health Statistics have produced a series of estimates of expenditures on mental health, the earliest covering the year 1963 (Rice, 1966), and the last 1995 (Hodgson and Cohen, 1998). Mental health expenditure estimates of Rice and her collaborators are generally consistent with the cost-of-disease accounts also initiated by Rice (1966). They are consistent as well as with the NHA, because cost-of-disease accounts disaggregate NHA totals for direct expenditures on the treatment of disease.

---

[23] These percentages are based on the important new work of Hodgson and Cohen (1998). As noted above, about 12 percent of personal health care expenditures in 1995 cannot be allocated by disease.

Another group of studies originate with Levine and Levine (1975), and proceed through a group of researchers at the RTI (Research Triangle Institute). Other estimates include Fein (1954, actually the first such study, treated here as an antecedent of studies by the Rice group), Frank and Kamlet (1985), Parsons et al (1980), and Mark et al (1998, condensed as McKusick et al, 1998).

The cost estimates in these studies often include indirect and social costs of mental illness, because a major part of the cost of having mental illness falls on the patient in the form of lost work time and so forth, and on others (family members, for example, or the victims of violence committed by the mentally ill) who experience the effects of mental illness in friends or strangers. However, I address only the direct treatment costs of mental illness in this section because I want to integrate cost of mental illness data with the National Health Accounts, which in principle include only the direct costs of treating an illness, and not the costs of having the illness. This does not imply that I think that the NHA could not or should not be extended to encompass indirect costs of illness, only that such an extension is beyond the scope of the present study.

Not surprisingly, methodologies for estimating the cost of treating mental illness have evolved through the years. For example, in the 1963 estimate (Rice, 1966) pharmaceutical expenditures were not allocated across diseases; when prescription drug allocations first became available (for 1972), they were around 7% of mental health expenditures. From this point on, drug expenditures are included in estimates of mental health care expenditures. As a second example, the earliest estimates exclude from mental health costs the costs of treating alcohol and drug abuse; when data for these mental conditions became available, they were first reported separately, and in the later estimates are folded into the mental health total, without a separate allocation (alcohol and drug related conditions are included in the mental conditions chapter of ICD-9).

Of the major two groups of studies, the RTI studies generally obtain higher *levels* of expenditures for treatment of mental conditions than do the studies of Rice and her colleagues. For example, Harwood et al (1984) report mental health care expenditures (including substance abuse) of $35.5 billion for 1980, when Rice et al (1985) report $19.8 billion for the same year. Similarly, Levine and Wilner (1976) estimate 1974 expenditures at $17.0 billion, when Paringer and Berk (1977) got only $9.4 billion for the following year. When the two sets of authors reference each other (which is seldom), they simply note that different methodologies are used.[24] Yet, the published descriptions of methodologies are remarkably similar.

Rather than attempting to reconcile the levels, I convert mental health treatment cost

[24] Levine and Levine (1975) present information on the quantitative effects of differences in estimating methodologies that were in use at that time. Mark et al (1998, chapters 1 and 8) contains a useful summary of methodological differences between their own study and the methods used by the Rice group and the RTI group.

estimates for various years into annual rates of change by matching published studies on groups of authors (Rice and collaborators, and the RTI group--details are in Appendix tables A-1 and A-1a). As noted earlier, successive studies from the same group of authors incorporate methodological and data improvements, so they are not strictly speaking comparable over time. Yet, research methodology is not chosen randomly, and there is a great amount of commonality among studies by the same group of authors. Matching on groups of authors minimizes the methodological noncomparabilities between groups of authors that result in differences in levels, noted above.[25]

This matching method necessarily omits studies (like Parsons et al 1980) that were "one time" estimates, and do not fit into either of the major groups of studies.

Even with matching, it is not straightforward to convert these data into a time series. When changes to estimating procedures were made, or when the authors decided to employ a different tabular presentation or classification, they seldom linked the methodology used in one paper with the one they employed for estimates two or three years previously. For example, early estimates apply average costs for all hospital stays to mental health; later ones use an explicit estimate for daily costs for mental health treatment (which are usually lower). But, undoubtedly because time series comparability was not a research priority, authors of studies generally do not provide an estimate of the effect of such changes in methodology on the estimates. In producing growth rates for mental health care expenditures, I have linked out major changes in estimating methods so that increased scope of the estimates is not inappropriately treated as increased expenditures on treating mental health conditions. These links are noted in Appendix table A-1a, but they do not exhaust changes that have been made between adjacent estimates by the same group of authors.[26]

When cost estimates are grouped by authors, rates of increase among the various estimates look fairly consistent and plausible, in the sense that trends in expenditures are more nearly similar between the two groups of studies than are estimates of levels of expenditures. For example, Harwood et al (1984) estimate the 1980-1981 expenditure growth to be 15.1%; combining Rice et al (1990) with Rice, Hodgson and Kopstein (1985) suggests an average annual rate of growth between 1980 and 1985 of 19.1% (see Appendix table A-1a, lines (i) and (l)).

---

[25] For 1980, I use the estimate of Rice, Hodgson, and Kopstein (1985), who report that total mental health expenditures for 1980 were $19.8 billion. This estimate is derived from Hodgson and Kopstein (1984), with correction of an error (the correction was published in Health Care Financing Review, Winter, 1984, pp. 128-30--information supplied by Thomas Hodgson).

[26] The 1968-71 match probably overstates the rate of increase, because it appears that Conley and Conwell (1970), the source for the 1968 estimate, excluded substance abuse expenditures and that Cooper and Rice (1976) included them. I could not find data to remove the effect of the change in coverage from the estimates.

However, the RTI estimate for 1977-80, at 19.1 percent annually, exceeds the five-year Rice group estimate (1975-80, 16.1 percent annually) by almost the same amount, suggesting that the differences may in part reflect the difference in year spans of the two studies, and the great spurt of inflation between 1979 and 1982. Where multiple estimates of expenditure growth rates are available for a given year, as in this case, I take the simple average: Thus, the 1980-1981 growth rate obtained by averaging these two studies is 17.1%. For the post-1990 period, the two available estimates of expenditure growth differ more than is the case for most earlier periods. Additional discussion of the post-1990 period is presented below.

Table 9 summarizes rates of growth of expenditures on treatment of mental conditions over different intervals (which depend on data availability) from 1954-1995. Expenditures for the treatment of mental conditions accelerated from their 1950s and 1960s rates, reaching a peak of 17% annually between 1980 and 1985. After 1990, mental health care expenditure grows at the lowest rate in three decades.

The expenditure growth rate for the decade of the 1990's calls for further comment. Two estimates exist.

Mark et al. (1998) is, to my knowledge, the only study of mental health expenditures that actually computes a growth rate--7.2 per cent per year, for the ten year interval 1986-96. Their estimate is higher than the growth rate calculated from other studies. Additionally, information from the Rice group suggests that the rate of growth of mental health care expenditures declined over this interval. Incorporating the Mark et al (1998) study dampens the slowdown in the rate of growth of expenditures that is evident in estimates from the Rice group, possibly because the Mark et al. (1998) definition of mental health excludes some ICD-9 codes included in the definition used by others (Alzheimer's disease, for example).

The domain of the estimate by Mark et al. (1998) differs from other studies, and from the national health accounts, in two respects. First, Mark et al. (1998) include mental health expenditures that go to organizations that are classified in the 1987 U.S. SIC system as social services industries, rather than as part of the medical industries. Because the NHA focus on institutional recipients of expenditures, expenditures received by non-medical "industries" are not included in the NHA definition. Mental health services provided by institutions outside the NHA definition, and included in the Mark et al. (1998) study, amount to about 4 per cent of the total. De-institutionalization of mental health patients has caused a shift in care toward these out-patient facilities, and their exclusion from other estimates (and from NHA) biases downward rates of growth. Surely we want data on expenditures to treat disease, not solely on expenditures received by medical care industries, so the domain chosen by Mark et al. (1998) is the appropriate one.

Second, Mark et al. (1998) adopt a definition of mental health care that encompasses a narrower set of ICD-9 codes than the definition of, e.g., Hodgson and Cohen (1998). For example, Mark et al. (1998) exclude Alzheimer's disease and other dementias from their

definition of mental health care. No estimate of the size of this exclusion appears in the study. Different definitions may be appropriate for different purposes, but this definition is not well motivated by the authors. It has the effect of reducing both the level and the rate of growth of mental health expenditures recorded in nursing homes, for example, because the number of patients admitted for non-dementia forms of mental conditions has declined, while the number of (and presumably the proportion of expenditures on) dementias has risen.

Why have expenditures grown? Has the U.S. increased the levels of services to the mentally ill? Or is it inflation in mental health care? The following sections construct data and explore the aggregate evidence.

B. Deflation methodology for mental health expenditures

Growth in medical care expenditures has three components.

First, there may be an increase in the number of patient-treatments. Because there are many types of patient profiles and of mental disorders, one can think of the patient-treatments component as a change in a constant patient-mix measure, which we would probably want to weight according to the relative costs of care for different conditions.

Second, there may be an increase in the cost or the price of treatments. In parallel to the first component, we could measure the increase in price as the cost change of a constant patient mix, holding constant also the level of treatment efficacy. This is medical care inflation.

Third, there may be changes in the patient mix and in the efficacy of treatment which imply changes in average treatment costs. For example, a shift in the patient mix toward more severe mental disorders implies increasing average treatment costs, even if the cost of no single treatment changes. One would not want an adverse (or a favorable) change in the mix of medical illnesses to influence the measure of the *price* of medical treatments. Additionally, improvements in the efficacy of treatment must be measured and allowed for in some way, for the same reason. One does not want a shift toward more expensive, but more efficacious, care to be confused with medical care inflation.

The standard way to obtain output (quantity) measures in economic statistics is to deflate the change in expenditures on a product (haircuts, say) by a price index for that product. It has not been customary to apply exactly the same deflation procedure to health care. No existing health expenditures account, to my knowledge, calculates the quantity of mental health care by dividing expenditures on mental health care by a price index for mental health care. There are, however, great advantages to doing so.

In a subsequent section, I present a new price index, or deflator, for expenditures on mental health treatment. Because the first step is to assemble existing published government price indexes, and to match them to mental health treatment expenditure information, it is

worthwhile to consider the results of this intermediate price index calculation here.

Details of the match between price indexes and mental health care expenditures are in a data appendix available from the author; excerpts are reproduced as Appendix table A-2c. As an example, hospital and nursing home expenditures on mental health treatment (80% of total expenditures on treatment of mental conditions in 1990) are matched to the Consumer Price Index (CPI) hospital price index before 1992. After 1992, two new Producer Price Index (PPI) hospital indexes became available: An index for the treatment of mental conditions in general hospitals and an index for psychiatric hospitals. The match weights changes in these two PPI indexes equally. A nursing home PPI index begins in 1994, and it is brought into the match at that point. These hospital and nursing home indexes are combined with price indexes for other components of mental health care (mental health care professionals and pharmaceuticals), using weights based on various mental health care expenditures estimates by the Rice group, for consistency with NHA. These weights are documented in Appendix table A-2b. I calculate ordinary Laspeyres indexes, for more nearly straightforward comparison with existing BLS medical care price indexes.[27]

Average annual growth rates in prices of mental health services, computed in this manner from available published government price indexes, are presented in the right-hand column of table 9. As the table shows, published price indexes that provide the closest match for mental health expenditures suggest that most of the increase in mental health expenditures has been caused by medical care inflation. Since 1985, mental health care inflation, computed this way, amounts to about 7 ½ percent per year. These data suggest that price increases for mental health services have actually outstripped both the overall CPI medical care index (up about 6 ½ per cent per year over this period) and mental health care expenditures, which grew 6 per cent per year. This implies that the quantity of mental health services being provided fell after 1985, and fell even more sharply after 1990. Note that this calculation is a total, economy-wide one, not adjusted for population--or patient population--growth, it is not mental health services per capita (which would fall even more).

In subsequent sections, I consider possible biases in these government price indexes for mental health, and estimate new mental health price indexes, which show a very different picture of medical care inflation from those in table 9. I then use the new price indexes to compute a

---

[27] Mark et al. (1998) suggest very different weights, with hospitals and nursing homes receiving only 40 % of 1996 mental health care expenditures. As noted above, their exclusion of dementias from their definition of mental conditions biases downward their estimate of the nursing home share of expenditures, but the NHA-SIC definition of the medical industry biases upward the hospital-nursing home share in estimates of the Rice group, because a portion of total expenditures on mental health care (around 4 per cent, in the Mark, et al., 1998, data) are excluded from the total. But even so, it is hard to see why the proportions could differ so much. This is a case where a careful reconciliation of the two estimates--which as a practical matter is best done by the authors--would produce more confidence in the results.

mental health account that shows trends in expenditure, inflation, and real quantity of mental health services.

1. Problems in estimating a mental health care expenditures price index

It has become commonplace that medical care inflation outstrips the overall inflation rate. For example, between 1985 and 1995, the medical care component of the CPI rose 6.5% per year, when the overall CPI rose only 3.6 per cent.

Until fairly recently, price trends for all types of medical care have been inferred almost exclusively from the medical care component of the CPI. Inflation information that is specific to the treatment of mental conditions and the measurement problems are discussed in this section.

The earliest U.S. price index specific to mental health care is the "psychiatrist office visits" component of the CPI, a price index that was published between 1964 and 1977. As table 10 shows, over the 1964-1977 period the price of a visit to a psychiatrist's office rose less rapidly (about 5.2% per year for the whole period) than did the fees of other medical professionals (a bit over 7% per year), or CPI medical care as a whole (about 7½ per cent per year). In fact, psychiatrists' fees were rising slightly slower than the overall CPI. [28]

The CPI psychiatrist office visit index obviously excludes hospital costs for medical treatment of mentally ill patients. In 1963, 88 per cent of total expenditures for mental health went to hospitals and nursing homes; in 1995, the proportion was still 81 per cent.[29] Non-hospital mental health professionals account for only a small proportion of the cost of treating mental conditions (in 1995, 11 per cent).

A CPI hospital index also existed in the 1960's. It measured hospital costs with a price

---

[28] The price index for psychiatrist office visits ends in 1977 because it was merged into the physician's fee CPI component, which, though it includes psychiatrists on a probability basis, contains no published detail on medical specialities.

[29] The data for 1963 exclude prescription drug expenditures, which are included in the 1995 total. Drugs accounted for 6 per cent of mental conditions treatment costs in 1972, the first year for which drug expenditure estimates by disease are available (and 8½ percent in 1995). Making an allowance for the probable size of drug expenditures in 1963, the proportion of hospital and nursing home expenditures in the mental health total has been remarkably stable through ostensibly major changes in treatment regimens over the last 35 years: The hospital expenditure share has fallen (from 86 per cent to 58 per cent), but the share of nursing home expenditures has increased from 1 per cent of the total to 23 per cent. The de-institutionalization of mentally ill in the 1970's reduced expenditures in mental hospitals, but this has been matched, almost exactly, by increased expenditures on forms of mental illness, primarily among the elderly, that are treated by institutionalization in nursing homes.

index for hospital room rates (for example, the cost of a semi-private room in a hospital). Hospital room rates have risen rapidly throughout the postwar years. Mental health care, however, was not distinguished separately in the CPI, so we do not know whether the cost for a day in the hospital for a mental health patient was advancing more or less rapidly than for other patients.

Most importantly, the cost of a day in the hospital or the cost of a psychiatrist's office visit, is an inadequate measure of the cost of treating mental illness, unless there is some way to adjust the costs for changes in medical efficacy. Do we want to know the price of an office visit, the cost per patient visit, or the cost per incident of depression? For most analytical purposes, we want to know the cost of treating depression and not (or at least not primarily) the cost of one input in treating depression. A CPI that tells us that the cost per psychiatrist's visit has advanced 10% may be accurate in what it tells us, but it is highly misleading in terms of what we want to know.

Ideally, one wants to "adjust" or correct the price index in some fashion for improvements in medical efficacy, and to obtain a price index for the treatment of a disease. Because medical economists generally believe that progress has been made in medical technology--better prognoses, less time spent in the hospital for any given condition, less painful and onerous conditions during treatment, and so forth--they believe that inadequate adjustment for these changes in medical technology creates upward biases in price indexes for medical care.

Merely to state the problem this way underscores the difficulties that statistical agencies face in producing price indexes for medical care. Calculating the change in costs for treating an episode of an illness requires not only the traditional statistical agency skills in gathering prices, but also a great deal of medical knowledge about changes in the efficacy of medical treatments (knowledge which, in many cases, is scientifically uncertain, or in contention). It also requires knowledge about patient valuations of changes in treatments, particularly when treatments change in dimensions that involve the patient's time, tolerance for pain, and valuation of the disutility of side effects, or of the onerous implications of treatments (such as, for example, a frequent treatment regimen for a pharmaceutical).

2. Improved BLS price indexes for mental health care expenditures

Subsequent to 1977, BLS made few changes in its CPI methodology for pricing medical care, until major improvements were initiated in the 1990's. The focus was still on a visit to the doctor's office, the cost of a hospital room, or the administration of a simple medication or a shot, and not on estimating the cost of treating an illness. Moreover, it was, and to an extent still is, a focus on the institution (the hospital, the nursing home, the doctor's office), rather than on the disease (mental illnesses, respiratory or circulatory illnesses, and so forth). And the inherent problems discussed in the previous section were not overcome.

Much criticism of the CPI medical price indexes was voiced after 1977 because many

economists believed that they overstated inflation in medical care. A milestone in that criticism was Newhouse (1989). For the period 1977-1992 we know nothing about trends in medical care prices for mental health treatments, because mental health costs are buried in CPI prices for medical care, which were generally believed to be upward biased.

In 1992, BLS introduced new price indexes for health care in its Producer Price Index (PPI) program. Although the new health care price indexes were still oriented toward the institution or the "industry" (the hospital, the nursing home, the doctor's office or clinic), they introduced a new methodology for measuring the price of medical care.

Rather than pricing the cost of a day in the hospital, the BLS now draws a probability sample of treatments for medical conditions. For example, for the PPI price index for mental health care treatment in a hospital the probability selection might be "major depression." The BLS then collects the monthly change in costs for treating that identical medical condition (see Berndt et al, 1998, and Catron and Murphy, 1996, for more information on BLS procedures).

Overall, the new PPI indexes present a picture of lower medical care inflation, compared to CPI measures, for the period where the two overlap (Catron and Murphy, 1996). The BLS subsequently introduced similar methodology into the CPI (Cardenas, 1996).

PPI hospital price indexes (though not those for physicians and clinics) include detail that is approximately by ICD-9 chapter. For the first time, PPI indexes estimate the cost change for treating mental disorders, and they permit comparison with other medical care costs.

For mental health costs, the new methodology makes a striking difference. For the period 1993-1998, the CPI "hospital and other related services" index rose 24.0 per cent (table 11). The PPI index for "mental diseases and disorders" treated in general medical and surgical hospitals rose about the same amount (21.5 per cent), but the PPI index for psychiatric hospitals rose only 4.6 per cent over this same period (table 11).[30] Taking the simple average of the two PPI indexes for hospital treatment of mental conditions suggests an increase of 13 per cent, which is just over half the rise in the CPI hospital index.

Similarly, the PPI index for "offices of clinics of doctors of psychiatry" rose only 3.9 per cent between 1994 and 1998 (this PPI index only began in 1994--see table 11). The CPI index for "physicians' services" rose 14.9 per cent for the same period.

3. Adjusted price indexes for mental health care

---

[30] It is not clear why these two indexes for hospital treatment of mental conditions should differ so greatly. Perhaps different diseases are treated in the two types of facilities, perhaps they have different cost structures, or perhaps they provide different treatments. Alternatively, the PPI procedures may produce substantial variance in the estimated price index.

The new medical care PPI indexes introduced in 1992 are great improvements on the previously-available CPI medical price information (see the assessment in Berndt et al, 1998). Evidence from the post-1993 period, when both PPI and CPI medical indexes were available suggests substantial upward bias in the CPI. The BLS does not compute historical price indexes when it introduces improvements. There is great need for better historical measures of medical care prices than the CPI has given.

One approach is to "backcast" estimates of the improvement that the PPI indexes represent. For the period following 1992, I match PPI and CPI components. For example, I combine the mental health subindex in the PPI index for general hospitals with the PPI index for psychiatric hospitals (see the discussion above) and match the result to the CPI "hospitals and related services" index. In this case, the ratio of the PPI indexes to the CPI index, from 1993 to 1998, was 0.912.[31] This ratio between the PPI and CPI hospital indexes for the period for which both were published provides a correction factor for the historical CPI hospital index before 1993 to make it more appropriate for measuring cost change in mental health treatment. I adjust the CPI hospital components in this manner for 1972 to 1993.

Similar adjustments are made for doctors offices--see the CPI physicians and PPI psychiatrists indexes in table 11. A PPI index for psychotherapeutics begins in 1982. This index is compared with the CPI prescription pharmaceuticals index for the period 1982-1992 (table 11), and used to backcast a correction to the CPI for the years 1972-1982. Additional details of the match and the resulting adjustment factors are available in the data appendix.[32] The resulting adjusted indexes for components of mental health care are combined with a logarithmic aggregator, rather than the arithmetic one (Laspeyres index) used by the BLS, and for the unadjusted indexes. This logarithmic index is discussed in section C.2, below.

The resulting price indexes are labeled "adjustment 1" in table 12. It is important to emphasize what adjustment 1 corrects. It corrects the historical CPI for the following: The old CPI index collected mostly list prices, the new PPI indexes are more nearly transactions prices. The old CPI pertained only to consumer out-of-pocket payments, so prices paid by health insurance and for procedures that are not normally paid by consumers were not adequately

---

[31] That is, from table 11: ½ (PPI mental conditions, general hospitals + PPI psychiatric hospitals) / CPI hospitals index = ½ (1.215 + 1.046) / 1.240 = 0.912.

[32] There is strong evidence of upward bias in the PPI pharmaceutical indexes in recent years (Berndt, Cockburn, and Griliches, 1996). I have not adjusted the historical PPI for the results of this research, because the PPI errors appear to be uniquely associated with particular events and years, but an improvement to the adjustment 1 index computed here would incorporate the findings of this and other research on pharmaceutical price indexes.

represented;[33] the PPI indexes cover both these lacunas. The PPI indexes are specific to mental health; the old CPI was much broader, and therefore could not represent price movements in mental health treatments that differed from those for treating other medical conditions. The PPI moved substantially toward pricing the cost of treating an illness, rather than the cost of, e.g., a visit to a psychiatrist's office, or hospital room charges. PPI methodology should at least partly pick up changes in medical technology that reduce the cost of treatment, compared with the cost of an office visit or hospital room. Finally, the adjustment 1 indexes contain an approximation for fixed-weight index number bias.

It is also important to emphasize limitations. The backcast will be valid if the joint error from all the above factors is the same for the period of the backcast (1972-93) as it was for the overlap period (1993 or 1994 to 1998). It is unlikely that the backcast is exactly valid, but even given its limitations it is likely to be far better than the historical CPI. The work of Scitovsky (1964) suggests that CPI measurement error might have gone in the opposite direction in the 1950's and early 1960's, which argues against backcasting this far.

Even though it is a great improvement, the new PPI methodology still omits some aspects of the cost of treating disease. It has been difficult for BLS to find data to adjust for changes in the efficacy of treatment (see the discussion in a previous section).

Additionally, some changes in medical treatment cause shifts in expenditures among PPI index categories; the PPI methodology contains no obvious way to take these cost savings into account. As an example, consider increased use of drugs that permit treatment of mental conditions on an outpatient basis, rather than in a mental hospital. Substitution of drugs (and clinical visits) for hospital care will reduce the cost of treatment, but this cost reduction will be reflected inadequately in the PPI because the PPI holds the weights for the various expenditure categories (hospitals, doctors offices, pharmaceuticals, and so forth) constant.

The major research on these problems are the two studies by Frank, Berndt and Busch (1998, 1999)--hereafter, FBB. FBB studied the cost of treating depression by American Psychiatric Association guidelines. Rather than creating a price index for each alternative guideline treatment (for example, treatment by psychotherapy alone), they considered all guideline treatments that had equal clinical outcomes as equivalent.[34] For example, partial substitution of drugs for some time spent in psychotherapy generally reduces the cost of treatment. This cost saving is incorporated into FBB's price indexes.

FBB reported that their price index for treating depression fell over the 1991-95

---

[33] Some of these charges made their way into the CPI health insurance index, but they do not influence directly the CPI price indexes for hospitals and physicians.

[34] They also considered alternative assumptions.

interval.[35] The new PPI indexes for mental health treatment--hospital and non-hospital care--were fully in place only for 1994-95, when, perhaps fortuitously, their rate of increase matches FBB almost exactly (3 per cent). The back-casted index (adjustment 1 index in table 13) rose more rapidly than did the FBB index in the other two years. Over the entire interval, the ratio of the FBB index to the adjustment 1 index is 0.978, a difference in growth rates of about a 0.2 percentage point per year.

However, the costs of treating depression may not be a good proxy for the costs of treating other forms of mental illness. Depression is more frequently treated outside hospitals than is the case for some other mental diseases, and FBB estimate an index for outpatient care.[36] Their results may not well represent the cost experience for disorders that more frequently require institutionalization. Additionally, FBB explore pharmaceutical innovations in mental health care for depression. Even though pharmaceutical advances in medical practice for other mental disorders have taken place (for example, schizophrenia), the technological innovations in the treatments for other disorders may not have the same implications for treatment costs as the experience with treating depression.

On the other hand, BLS price indexes do not, generally, pick up changes in either the efficacy (or quality) of treatments, or changes--such as those documented by FBB--that reduce the cost of treating an illness by changing the treatment itself. The BLS method standardizes on the treatment, and collects costs for a given treatment for a specified condition.

There is thus merit in using the FBB study to adjust a mental health care price index for improvements in medical practice. The only issue is how to use it.

Two options present themselves. One could apply the FBB study as a correction to the entire BLS mental health care price index, on the grounds that FBB is the only study available. For the reasons given above, the FBB study may overstate the improvement in medical care for other mental health conditions, and understate the rate of inflation in other mental disorders.

An alternative is to apply FBB only to the depression portion of the mental health care price index. This alternative implies that *no* comparable gains in medical practice occurred in treating other forms of mental illness. Thus, if applying FBB to the entire mental health

---

[35] For purposes of this paper, I use results in the second FBB study (1998), which differ to an extent from the first (1999) in that their depression price index falls less in the second study.

[36] I have not located data on hospital expenditures for treating depression. DuPont et al (1998) report that about 59 per cent of expenditures on a group of anxiety disorders were spent in hospitals and nursing homes (almost entirely the latter). These anxiety disorders did not include depression, but all are located next to depression in the coding structure of ICD-9, and comorbidities between these disorders and depression suggests that data on the distribution of expenditures on treating depression might be similar.

expenditures creates error (because of the two assumptions noted above), applying the study's results only to the depression part of the index creates error in the other direction. In any case, unless deterioration of medical practice for other forms of mental illness occurred, both errors are bounded by adjusted index 1.

Compilers of national accounts and of national health accounts are typically (but not always) conservative about applying results of one piece of research to another set of data, so I suspect that future compilers of accounts for diseases might prefer the second alternative. I do not know the proportion of mental health care costs made up by the treatment of depression. As an exercise, I assume it is 25 per cent.[37]

The next question is the use of FBB to backcast, in order to correct the historical indexes for improvements in medical practice in treating mental disease. Applying an adjustment from the FBB study implies that the *rate* of improvement in medical practice for treating depression (that is, in health care outcomes) was the same in earlier years. This is debatable. Nevertheless, I use the FBB results to make an additional "backcasting" correction to the historical price series.

Table 13 shows that the difference in trend between the price index for depression in FBB and weighted PPI mental health indexes, for the period in which both are available is about 0.2 percentage point. I weight *one-quarter* of the mental health adjustment 1 price index with this additional adjustment. The result appears as "adjustment 2" price index in the third column of table 12.

For the reasons given above, the FBB adjustment (adjustment 2) may overstate the improvement in medical care in other periods, and understate their rates of inflation. The alternative, as I noted above, is to assume no improvement in medical practice (other than what is incorporated into the new PPI indexes). The "no improvement" assumption yields the adjustment 1 index, which can then be taken as a bound on the error in the adjustment 2 index.

C. An expenditure account for mental health

We are now in a position to form a U.S. health expenditure account for mental health treatment that shows trends in expenditures, prices, and the quantity of mental health services, 1972-1995. Years before 1972 are omitted here, but could be added for a longer historical series. The summary of this account is in table 12.

1. Price index bias.

Backcasting improved price measures shows that the old CPI medical care price indexes overstated inflation in medical care costs for mental health, a result that is consistent with

---

[37] Affective disorders (not including depression) account for 28.7% of mental health care costs (Rice and Miller, 1998). Considerable comorbidity exists between affective disorders and depression.

medical economists' presumptions. For mental health, CPI indexes weighted to reflect mental health expenditures (the unadjusted index in table 12) show double-digit inflation in the 1970's and early 1980's. An adjustment factor based on the improved post-1992 PPI lowers substantially the medical inflation estimates of the 1970's and early 1980's, so they are no longer double-digit. Indeed, the adjustment 1 price index for mental health care is essentially flat since 1985.

Of course the adjustment 1 index relies on the validity of the backcasting exercise and its assumption that corrections taken from the period of the 1990's apply to the decades of the 1970's and 1980's. Medical economists have a strong presumption that the unadjusted indexes are biased upward, and available empirical work is consistent with that presumption of upward bias; the adjusted indexes, however, may have their own biases.

Adjusting price change for the FBB research results has a smaller, thought not negligible effect (Table 12). The adjustment 2 index shows slightly negative inflation in mental health care in the 1990's, because FBB found a declining price for depression in their study, and even more strongly negative mental health care inflation rate for the late 1980's. These estimates compare to the 6 ½ percent and nearly 9 percent rates given by the unadjusted price indexes for these two periods.

Possibly the costs of treating depression have fallen relative to the costs of treating other mental conditions. If not, the adjustments applied to obtain the adjustment 2 index are too small, and the adjusted price indexes for mental health care should fall even more rapidly than the adjustment 2 index of table 12. As noted, if the PPI indexes are correct for the remainder of mental health treatments, then adjustment 1 index bounds the correct index.

I noted earlier that it is commonplace that medical care inflation outstrips the overall CPI inflation rate. The price index numbers in table 12 call this generalization into doubt. At least since the mid-1980s, inflation in mental health care was substantially lower than the general CPI inflation rate (3.6% per year, 1985-1995), and it may have been negative through the whole period. For mental health care, "runaway medical inflation" is a wholly invalid characterization.

2. Growth in mental health services.

I next use the adjusted price indexes to estimate the growth in the quantity of mental health care services (or real expenditure growth).

For this stage of the research, I take a short cut. I begin from the index number system in equations (1) and (2b), above, which is the inverse of the index number system that is normally used in national health accounts. Equation (2b) specifies a Paasche index number for real expenditures. The conventional NHA system employs a Laspeyres index for real expenditures--that is, the real output index (not the price index) is weighted by base-period expenditures. Producing the conventional system implies an additional computational step that has not been

undertaken.

One reason is that the conventional national accounts-national health accounts index number system is subject to substitution bias. A better system for computing price indexes and indexes of real expenditures uses a superlative index number system. The NIPA now uses a superlative index number system, specifically, Fisher index numbers.[38] In principle, the Fisher system provides a better measurement system for NHA also.

However, I have not computed a Fisher system for the data on mental health for technical reasons that could readily be surmounted, but that introduce extraneous issues that are better discussed elsewhere.[39] An alternative that is computationally (and procedurally) simpler is to compute base-weighted logarithmic indexes, which are approximately Cobb-Douglas indexes with shares based on the initial period's expenditures (Moulton, 1996, explains the procedure). Table 12, then, presents real expenditure trend information produced through deflation by a base-weighted logarithmic price index.[40]

Not surprisingly the largest growth rate estimate comes from using both the adjustments--the adjustment 2 index that incorporates adjustments based on the new PPI indexes based and also on the FBB study (the last column of table 12). With this adjustment, growth in mental health care services from 1985 to 1990 was over 8 percent, from 1990 to 1995 about 5 percent. This contrasts with the impression of negative growth in aggregate mental health care services that arises from looking at the unadjusted price indexes (table 12). But even without the FBB adjustment, the adjustment 1 indexes also show real growth in mental health care services, not the negative growth suggested by the available historical data, without adjustment. Even so, the rate of growth of mental health services, which peaked in the early 1980's at over 15 percent per

---

[38] The Fisher price index is: $I(P)_{0t} = [(\sum_i P_{it} Q_{i0} / \sum_i P_{i0} Q_{i0}) (\sum_i P_{it} Q_{it} / \sum_i P_{i0} Q_{it})]^{½}$, and the Fisher quantity index, $I(Q)_{0t}$ is obtained by reversing the P and Q terms in the price index formula. A convenient property of the Fisher system is that the product of the Fisher price index and the Fisher quantity index--that is, $I(P)_{0t} \times I(Q)_{0t}$-- is the change in total expenditure, the first term on the right-hand side of equation (2a).

[39] The essence of the Fisher price index or Fisher quantity index is the application of weights for initial and ending periods to the *same* components. Partly because the available weighting information for mental health expenditures come from years that are fairly far apart, but also because the detailed information on price indexes, as well as the expenditure components that are available, changes from period to period, calculating a Fisher system from existing data on mental health expenditures and price indexes implies applying different weighting structures to different components, not to the same components. There is no reason why this cannot be done, it in fact occurs in the application of Fisher indexes to the NIPA.

[40] Note, however, that the adjustment 2 index incorporates FBB's chain Fisher index. The arithmetic (Laspeyres) form of the adjustment 1 index is presented in Appendix table A-2.

year, has slowed substantially in the 1990's, to under 5 percent annually. Note again that these are aggregate numbers, the type of estimates that typically appear in NHA; they do not estimate mental health care services per capita.

These are, perhaps, just numbers. What is their importance?

A great deal of effort has been put into medical care cost containment in the U.S. The data in table 12 suggest that medical care inflation is not the driving force behind the run-up in medical care costs, at least in mental health care. In the case of mental health care, the aggregate level of services has improved, judging from the best picture that can be assembled from U.S. aggregate statistics, but the rate of growth of the real quantity of mental health care services has slowed in the post-1990 period of cost containment. If these numbers are anywhere near correct, they suggest that health care cost containment may have social costs--curtailment of health care that has real impacts on health--that are more severe than are generally recognized.

And if the numbers are not correct, or if they need refinement before they can be used to inform public debate (the need for refinement in these estimates is hardly debatable), it is also the case that decisions on health policy are being based on statistical trend estimates that are at least as defective, and probably far more misleading, than the ones developed here.

The need is strong for aggregate U.S. data on health care that matches the price and output information that is routinely available for other portions of the U.S. economy.

VII. A Note on Measuring Output in Publicly-Provided Health Care Systems

The examples in this paper reflect the institutional structure of the U.S. health care system, which is atypical for industrialized countries. Three of the countries for which cost-of-disease accounts are shown in tables 5 and 6 have government-funded health care systems. However, the general "human repair" accounting framework can be applied to governmentally-provided health care systems, as well as market-provided ones, and perhaps even better.

For the governmentally-provided health care systems, integration of cost-of-disease accounts into NHA in each country is a first step, just as it is in the U.S. After that the steps are different. Rather than developing deflators, as in the U.S. case, the accounts require information on case costs and on case quantities, which is information that a national health system ought to collect (but sometimes does not) for its own purposes. The emphasis is on developing direct quantity measures, with associated costs, rather than on producing the quantity measures indirectly through deflation. Information on medical outcomes performs an analogous role in such an implementation. The details are left for another paper.

VIII. Where Does This Take Us? Conclusions

Medical outcome measures are disease specific. So are research price indexes of the type

discussed in this paper (the heart attack price index of Cutler et al, 1996, or the depression price index of Frank, Berndt, and Busch, 1999). There are obviously many human repairs to be considered, even in one ICD-9 chapter, let alone across all of them. Moving through the ICD-9 chapters on a disease-by-disease basis is clearly a very big job.

Yet, there is no reasonable alternative but to take samples of disease treatments and compute the value of medical interventions on the health of the recipients. Trying to deflate expenditures on "hospitals," for example, without considering the treatment of individual diseases, has a long history of failure. Indeed, the new PPI price indexes for hospital care *begin* from measuring the price change for treatment of individual diseases. The failure of the conventional focus on the institutional setting for treatment is greater the greater is the possibility that treatment moves from a hospital setting to outpatient care, or that new pharmaceuticals substitute for hospital or clinician resources. The new PPI indexes are not immune to deficiencies from this source because they are still "industry" price indexes, though industry price indexes with useful "product" detail on the costs of treating diseases.

To put the magnitude of this task in context, however, personal health care accounts for about 12 per cent of U.S. GDP in 1995. This is a large share of GDP, much larger than, say, the Producers' Durable Equipment (PDE) portion of investment, which was roughly 7 ½ percent of GDP in the same year. The unpublished detail from which U.S. PDE is calculated runs to something on the order of 800 lines. Not all of these lines have their own deflators, but the deflation detail in PDE incorporates some 400 lines, for which in many cases both domestic and imported products are distinguished and deflated separately. Additionally, the Bureau of Economic Analysis computes a capital flow matrix that distributes these investment components to the more than one thousand industries identified in the U.S. industry classification system, though not to every one of them at the finest level of industry detail.

Against this product detail in measuring PDE, the expenditure detail available on "products" in the NHA is minimal. This lack of expenditure detail in NHA (product detail that is actually already present in cost-of-disease accounts) is the great limitation on the potential for creating real output measures from NHA data. Even though creating it will be a great deal of work, it is work well worth doing.

Measuring health is not a smaller job than measuring PDE, nor a less important one. The difference is, rather, that the investment accounts have many years' head start.

Table 1: Comparison of NHA and NIPA Medical Care Expenditure Categories, 1995

NHA		NIPA	
Category	$ billion	$ billion	Category
Personal Health Care	869.0	871.6	Medical Care
Hospital Care	346.7	310.6	Hospitals
Physician Services	196.4	191.4	Physicians
Dental Services	44.7	47.6	Dentists
Other Professional Services	54.3	104.4	Other Professional Services
Home Health Care	28.4	--	
Drugs and Other Medical Nondurables	84.9	85.7	Drug preparations and sundries
Vision products and Other Medical Durables	13.1	13.1	Opthalmic products and orthopedic appliances
Nursing Home Care	75.2	65.2	Nursing Homes
Other Personal Health Care	25.3	--	(Other Categories)
Program Administration and Net Cost of Private Health Insurance	60.1	40.7	Medical Care and Hospitalization
(Other Categories)	--	12.9	Income Loss and Workers' Compensation

source: HCFA website, http://www.hcfa.gov/stats/NHE-OAct/tables/t11.htm; and *Survey of Current Business*, August 1997, Table 2.4.

Table 2: National Health Expenditures, by Source of Funds and Type of Expenditure: 1995 (Amount in Billions)

		Private					Government		
			Consumer						
	Total	All Private Funds	Total	Out of Pocket	Private Insurance	Other	Total	Federal	State and Local
National Health Expenditures	991.4	536.2	493.6	166.7	326.9	42.6	455.2	328.7	126.5
Health Services and Supplies	960.7	525.3	493.6	166.7	326.9	31.7	435.4	314.7	120.6
Personal Health Care	869.0	480.4	449.4	166.7	282.6	31.1	388.5	301.7	86.8
Hospital Care	346.7	136.2	121.2	9.6	111.6	15.0	210.5	172.3	38.2
Physician Services	196.4	133.1	128.9	29.0	99.9	4.1	63.3	50.7	12.6
Dental Services	44.7	42.7	42.5	21.0	21.5	0.2	2.0	1.1	0.9
Other Professional Services	54.3	41.9	38.1	20.4	17.7	3.8	12.4	9.5	2.9
Home Health Care	28.4	12.4	9.1	5.9	3.2	3.3	16.0	14.1	1.9
Drugs and Other Medical Nondurables	84.9	73.1	73.1	48.6	24.5	--	11.7	6.3	5.5
Vision products and Other Medical Durables	13.1	7.7	7.7	7.1	0.6	--	5.4	5.3	0.1
Nursing Home Care	75.2	30.2	28.8	25.1	3.7	1.4	45.1	29.5	15.6
Other Personal Health Care	25.3	3.3	--	--	--	3.3	22.0	12.9	9.1
Program Administration and Net Cost of Private Health Insurance	60.1	44.8	44.2	--	44.2	0.6	15.3	9.2	6.1
Government Public Health Activities	31.5	--	--	--	--	--	31.5	3.8	27.7
Research and Construction	30.7	10.9	--	--	--	10.9	19.8	14.0	5.8
Research	16.7	1.3	--	--	--	1.3	15.3	12.9	2.4
Construction	14.0	9.6	--	--	--	9.6	4.5	1.1	3.4

note: 0.0 denotes less than $50 million. Research and development expenditures of drug companies and other manufacturers and providers of medical equipment and suplies are excluded from “research expenditures,” but are included in the expenditure class in which the product falls. Numbers may not add to totals because of rounding.

source: Levit et al (1997), Table 11: “ National Health Expenditures, by Source of Funds and Type of Expenditure: Selected Calendar Years 1991-96.”

Table 3: US health expenditures by ICD chapter (amount in millions)

	Estimated direct expenditures[1]	Estimated direct costs[2]	Estimated amounts of direct costs[3]	Estimated amount of personal health care expenditures[4]
ICD Chapter Headings	1963	1972	1980	1995
Total expenditures	29,394[a]	78,537[d]	219,443[g]	897,510[j]
All conditions (total allocated expenditures)	22,530	75,231	206,878	787,510
Infectious and parasitic diseases 001-139	502	1,412	4,300	17,656
Neoplasms 140-239	1,279	3,872	13,049	42,917
Endocrine, nutritional and metabolic diseases, and immunity disorders 240-279	903	3,436	7,329	33,825
Diseases of the blood and blood-forming organs 280-289	156	491	1,155	4,890
Mental disorders 290-319	2,402	6,985	19,824	74,707
Diseases of the nervous system and sense organs 320-389	1,416	5,947	17,132	65,847
Diseases of the circulatory system 390-459	2,267	10,919	32,488	133,196
Diseases of the respiratory system 460-519	1,581	5,931	16,661	61,481
Diseases of the digestive system 520-579	4,159	11,100	30,974	89,656
Diseases of the genitourinary system 580-629	1,210	4,471	12,313	37,462
Complications of pregnancy, childbirth, and the puerperium 630-676	1,391	2,607	--[h]	3,555
Diseases of the skin and subcutaneous tissue 680-709	248	1,525	5,940	18,824
Diseases of the musculoskeletal system and connective tissue 710-739	1,430	3,636	13,124	50,309
Congenital anomalies 740-759	113	381	1,345	5,046
Certain conditions originating in the perinatal period 760-779	30	--[e]	--[h]	3,349
Symptoms, signs, and ill-defined conditions 780-799	624	--[e]	3,815	23,487
Injury and poisoning 800-999	1,703	5,121	18,684	71,806
Supplementary classifications V01-V82	966	--		49,494
other	150[b]	7,398[e]	8,746[h]	
unallocated	6,864[c]	3,306[f]	12,656[i]	110,000

1 Source: Rice (1966) tables 1 and 31, pages 3 and 109. Note that the disease classification used at this time period was the ICD-8 or

ICDA.

2 Source: Cooper and Rice (1976) table 1, page 23.

3 Source: Hodgson and Kopstein (1984), Rice Hodgson, and Kopstein (1985) table 5, page 69.

4 $110 billion could not be allocated by diagnosis, and this constitutes 12% of personal health care expenditures. Source: Hodgson (1997).

a From Rice (1966) table 1: Differs from NHA personal health care expenditures by $0.1 billion.

b The "Other" category for this year was actually labeled "Miscellaneous" from the publication, and includes special conditions and examinations.

c Not given explicitly; calculated as "Total expenditures" less "All conditions (total allocated expenditures)."

d From Cooper and Rice (1976) table, page 22: the entry for "other health services" is understood to be the equivalent of the "unallocated" category in this table. This number is added to the "All conditions" number to give the total expenditure.

e The "Other" category for this year includes certain causes of perinatal morbidity and mortality, symptoms of ill-defined conditions, and special conditions without sickness and symptoms.

f From Cooper and Rice (1976) table, page 22: "other health services" is the equivalent of "unallocated."

g Differs from NHA by 2x.

h The "Other" category includes complications of pregnancy, childbirth, and puerperium and certain conditions originating during the perinatal period.

i Includes $8.3 billion of personal health care expenditures that could not be allocated by age and sex, and $4.265 billion that could not be allocated for other reasons.

j Differs from NHA. Original total was taken from a projection to 1995; the initial NHA actual estimate was about 2% below the projection and the revised estimate is lower than the projection by 3.2%. See Hodgson (1997).

Table 4: Personal health care expenditures, by allocation status and type of health service: US, 1995 estimates[1]

Type of health service	NHA Total	Hodgson's Allocated Total	Unallocated Amount	Unallocated Percent
	Amount in billions			
All personal health care	897.7	787.5	110.2	12.3
Total hospital care	364.5	360.3	4.2	1.2
Total physician services	198.0	185.3	12.7	6.4
Dental services	42.9	42.9	0.0	0.0
Total other professional services	62.9	21.7	41.2	65.5
Home health care	27.9	27.9	0.0	0.0
Drugs and other medical non-durables	84.7	55.2	29.5	64.8
Total vision products and other medical durables	13.9	13.9	0.0	0.0
Nursing home care	80.2	80.2	0.0	0.0
Other personal health care	22.7	0.0	22.7	100.0

[1] Estimates are based on Hodgson (1997) and additional information supplied by Thomas Hodgson. NHA total refers to preliminary estimates. 1995 numbers are updated in Levit et al (1997): 191.

Table 5: Total Health Expenditures, United States, England, Canada, and Australia, Disaggregated by ICD-9 Chapters

Diagnosis and ICD-9-CM chapters and codes	United States[1] 1995 All personal health care		England, net public expenditure[2] 1992/93 NHS and PSS expenditure		Canada 1993 Total Direct Costs		Australia 1993-94 Total Health System Costs	
	millions of US dollars	% of total	millions of UK pounds	% of total	millions of Can. dollars	% of total	millions of Aus. dollar	% of total
All conditions	$787,510	100.0%	£31,060	99.9%	$44,130	100.0%	$31,397	100.0%
1) Infectious and parasitic diseases 001-139	17,656	2.0%	311	1.0%	787	1.8%	849	2.7%
2) Neoplasms 140-239	42,917	4.8%	1,273	4.1%	3,222	7.3%	1,905	6.1%
3) Endocrine, nutritional and metabolic diseases, and immunity disorders 240-279	33,825	3.8%	497	1.6%	1,334	3.0%	966	3.1%
4) Diseases of the blood and blood-forming organs 280-289	4,890	0.5%	155	0.5%	274	0.6%	192	0.6%
5) Mental disorders 290-319	74,707	8.3%	5,156	16.6%	5,051	11.4%	2,634	8.4%
6) Diseases of the nervous system and sense organs 320-389	65,847	7.3%	2,609	8.4%	2,252	5.1%	2,333	7.4%
7) Diseases of the circulatory system 390-459	133,196	14.8%	3,758	12.1%	7,354	16.7%	3,672	11.7%
8) Diseases of the respiratory system 460-519	61,481	6.9%	1,926	6.2%	3,787	8.6%	2,510	8.0%
9) Diseases of the digestive system 520-579	89,656	10.0%	2,578	8.3%	3,326	7.5%	3,712	11.8%
10) Diseases of the genitourinary system 580-629	37,462	4.2%	1,118	3.6%	2,248	5.1%	1,658	5.3%
11) Complications of pregnancy, childbirth, and the puerperium 630-676	3,555	0.4%	1,025	3.3%	2,025	4.6%	1,051	3.3%
12) Diseases of the skin and subcutaneous tissue 680-709	18,824	2.1%	528	1.7%	892	2.0%	955	3.0%
13) Diseases of the musculoskeletal system and connective tissue 710-739	50,309	5.6%	2,423	7.8%	2,460	5.6%	2,971	9.5%
14) Congenital anomalies 740-759	5,046	0.6%	124	0.4%	305	0.7%	191	0.6%
15) Certain conditions originating in the perinatal period 760-779	3,349	0.4%	217	0.7%	551	1.2%	247	0.8%
16) Symptoms, signs, and ill-defined conditions 780-799	23,487	2.6%	1,273	4.1%	1,851	4.2%	1,336	4.3%
17) Injury and poisoning 800-999	71,806	8.0%	1,180	3.8%	3,122	7.1%	2,607	8.3%
Supplementary classifications V01-V82	49,494	5.5%	1,553	5.0%				
Supplementary (health status)			93	0.3%				
Well-patient					2,741	6.2%		
other					549	1.2%	1,607	5.1%
unallocated	110,000	12.3%	3,230	10.4%				

[1]Total allocated expenditures (from Table 3).

[2]The definition of "net public expenditure" is assumed to undertake the explanation provided in Annex A, part A.4.1.2 that describes expenditure data: "The analysis includes the majority of health andsocial services expenditure, around 85%. The major exclusions comprise NHS headquarters administration, ambulance and accident and emergency services, day hospital care, services classified in the programme budget as "other hospital", and social services for children. Department of Health administration costs and centrally financed services (such as, for example, Departmental grants to voluntary organisations) are also excluded. Income support expenditure for residents in independent residential care is included...in view of the community care reforms."
NHS is the National Health Service, and PSS is the Personal Social Service; total percent for all conditions do not add up to 100.0% due to rounding.

Sources: Hodgson (1997); National Health Service (1996), table 6.1; Moore et al (1997); Mathers et al (1998), table C.2, page 34.

Table 6: Hospital Expenditures, United States, England, Canada and Australia, Disaggregated by ICD-9 Chapters

Diagnosis and ICD-9-CM codes	United States, 1995		England, 1992/93		Canada, 1993		Australia, 1994-95	
	Hospital care[1]		NHS hospital expenditure[2]		Direct costs, hospitals		Private acute hospitals[3]	
	millions of US dollars	% of total	millions of UK pounds	% of total	millions of Can. dollars	% of total	millions of Aus. dollars	% of total
All conditions	$360,341	100%	£16,200	99.9%	$26,096	100.0%	$2,399	100.0%
1) Infectious and parasitic diseases 001-139	9,426	2.6%	162	1.0%	345	1.3%	14	0.6%
2) Neoplasms 140-239	28,104	7.8%	1,021	6.3%	2,467	9.5%	79	3.3%
3) Endocrine, nutritional and metabolic diseases, and immunity disorders 240-279	14,643	4.1%	194	1.2%	527	2.0%	20	0.8%
4) Diseases of the blood and blood-forming organs 280-289	2,641	0.7%	113	0.7%	157	0.6%	3	0.1%
5) Mental disorders 290-319	43,172	12.0%	2,770	17.1%	3,632	13.9%	113	4.7%
6) Diseases of the nervous system and sense organs 320-389	13,247	3.7%	810	5.0%	793	3.0%	210	8.7%
7) Diseases of the circulatory system 390-459	67,604	18.8%	1,847	11.4%	4,862	18.6%	225	9.4%
8) Diseases of the respiratory system 460-519	31,039	8.6%	940	5.8%	1,788	6.9%	97	4.0%
9) Diseases of the digestive system 520-579	28,688	8.0%	826	5.1%	2,093	8.0%	326	13.6%
10) Diseases of the genitourinary system 580-629	18	4.9%	778	4.8%	1,076	4.1%	260	10.8%
11) Complications of pregnancy, childbirth, and the puerperium 630-676	2,121	0.6%	875	5.4%	1,650	6.3%	217	9.1%
12) Diseases of the skin and subcutaneous tissue 680-709	6,411	1.8%	324	2.0%	223	0.9%	96	4.0%
13) Diseases of the musculoskeletal system and connective tissue 710-739	20,512	5.7%	923	5.7%	1,286	4.9%	486	20.2%
14) Congenital anomalies 740-759	2,728	0.8%	113	0.7%	232	0.9%	--	--
15) Certain conditions originating in the perinatal period 760-779	2,535	0.7%	211	1.3%	518	2.0%	35	1.5%
16) Symptoms, signs, and ill-defined conditions 780-799	7,682	2.1%	761	4.7%	845	3.2%	10	0.4%
17) Injury and poisoning 800-999	40,433	11.2%	940	5.8%	2,253	8.6%	22	0.9%
Supplementary classifications V01-V82	21,572	6.0%	810	5.0%			63	2.6%
Well-patient care					1,349	5.2%		
unallocated	--	--	1,761	10.9%			125	5.2%

[1]Total allocated expenditures. Expenditures include: services provided in short-term community hospitals, $9 billion of expenditures by the Department of Defense, patients seen in hospital outpatient departments and emergency rooms, and may include those for hospice services. From the data by Thomas Hodgson, listed below in Sources.

[2]NHS is the National Health Service; total percent for all conditions do not add up to 100.0% due to rounding.

[3]Australian figures were converted from an alternative classification system by diagnoses; as such, congenital anomalies were not assigned figures, and additional information on the classification will be built into a subsequent revision.

Sources: Hodgson (1997); National Health Service (1996), table 6.1; Moore et al (1997); Mathers et al (1998).

Table 7: Direct Cost or Public Health Expenditure Average Annual Rate of Increase by Disease Category, Various Years[1]

ICD-9-CM / ICDA chapter heading	1963-1972[2]	1972-1980[3]	1980-1995[4]
Total expenditures	11.54	13.71	9.85
All conditions (total allocated expenditures)	14.35	13.48	9.32
Infectious and parasitic diseases	12.18	14.94	9.87
Neoplasms	13.10	16.40	8.26
Endocrine, nutritional and metabolic diseases, and immunity disorders[5]	16.01	9.93	10.73
Diseases of the blood and blood-forming organs	13.60	11.29	10.10
Mental disorders[6]	12.59	13.93	9.25
Diseases of the nervous system and sense organs	17.29	14.14	9.39
Diseases of the circulatory system	19.08	14.60	9.86
Diseases of the respiratory system	15.82	13.78	9.09
Diseases of the digestive system	11.53	13.69	7.34
Diseases of the genitourinary system	15.63	13.50	7.70
Complications of pregnancy, childbirth, and the puerperium[7]	7.23	--	--
Diseases of the skin and subcutaneous tissue[8]	22.36	18.53	7.99
Diseases of the musculoskeletal system and connective tissue[9]	10.93	17.40	9.37
Congenital anomalies[10]	14.46	17.08	9.21
Certain conditions originating in the perinatal period[11]	--	--	--
Symptoms, signs, and ill-defined conditions[12]	--	--	12.88
Injury and poisoning[13]	13.01	17.56	9.39
Supplementary classification (V or E codes)[14]	--	--	--
Other	--	--	--

[1] All calculations originate from sources giving direct cost figures, except for the 1995 data set, where only public health expenditures were available.

[2] Source: 1963 data: Rice (1966) table 21, page 109.
1972 data: Cooper and Rice (1976) table 1, page 23.

[3] Source: 1972 data: Cooper and Rice (1976).
1980 data: Rice, Hodgson, and Kopstein (1985) table 1, page 62.

[4] Source: 1980 data: Rice, Hodgson, and Kopstein (1985).
1995 data: Hodgson (1997).
Note: for the "All Conditions" category, the original amount used from the 1995 data source is the amount originally cited ($787.5 billion) added to the amount not originally included due to the uncertainty of allocation by diagnosis ($110 billion).

[5 - 14] These chapter headings, listed as found in the ICD-9-CM, is listed differently in the ICDA (ICD-8), which affects the 1963 and 1972 data sources, and which may or may not affect comparisons.

Note: Calculations of rates were made using $x_1(1+r)^{1/n-1}=x_n$.

Table 8: Estimated Amount of Personal Health Care Expenditures, Total and for Mental Disorders, by Type of Provider: United States, 1995

	All conditions ($ millions)	Mental disorders ($ millions)[a]	Mental disorders, as proportion of all conditions
All personal health care	787,510[b]	74,707	9.5%
Hospital care	360,341	43,172	12.0%
Physician services	185,329	7,761	4.2%
Prescription drugs	55,224	6,057	11.0%
Nursing home care	80,200	16,968	21.2%
Hospital care and nursing home care, combined	440,541	60,140	13.7%

[a] ICD-9-CM codes 290-319.
[b] Excludes $110 billion, 12 percent of personal health care expenditures, that cannot be allocated by diagnosis.
Source: Hodgson and Cohen (1998) table 2.

Table 9: Mental Health Treatments: Average Annual Increases in Expenditure and Price (percent change)

	Average annual expenditure growth rates	Average annual price index growth, unadjusted BLS price indexes[1]
1954-63	3.76	---
1963-72	12.57	9.96
1972-80	13.11	10.76
1980-85	17.09	10.81
1985-90	7.06	8.78
1990-95	4.94	6.47
1985-95	6.00	7.63

[1] BLS price indexes, matched to expenditure categories for mental health care, and weighted with mental health care treatment costs (see explanation in the text). 1954-63 not calculated.

Table 10: Consumer Price Indexes, All Items and Selected Medical Care Services, Average Annual Rates of Change, 1964-1977

	Average annual growth rates (percent) 1964-77
CPI, all items	5.29
Medical care, total	6.68
Medical care services	7.51
Physicians' services	7.03
Psychiatrist, office visits	5.19
Hospital Daily Service Charges	10.79
Semiprivate rooms	11.60

Source: United States Department of Labor, Bureau of Labor Statistics (1999).

Table 11: CPI and PPI Medical Price Indexes, Annual Average Growth Rates

Hospitals (1993-1998)	Total change	Average annual change
CPI: hospital and other related services	23.98	4.39
PPI: general medical and surgical hospitals	21.52	3.98
PPI: psychiatric hospitals	4.63	0.91
Physicians (1994-1998)		
CPI: physicians' services	14.86	3.53
PPI: offices and clinics of doctors of medicine, psychiatry	3.89	0.96
Drugs (1982-1992)		
CPI: prescription drugs and medical supplies	138.03	9.06
PPI: psychotherapeutics	264.38	13.80
Drugs (1992-1998)		
CPI: prescription drugs and medical supplies	20.45	3.15
PPI: psychotherapeutics	254.61	23.49

Source: United States Department of Labor, Bureau of Labor Statistics (1999).

Table 12: Growth Rates, Expenditures and Prices, Mental Health Treatments, 1972-95

	Annual expenditure growth rates	Price indexes (percent increase)			Real expenditure growth		
		Unadjusted	Adjustment 1[a]	Adjustment 2[b]	Unadjusted	Adjustment 1[a]	Adjustment 2[b]
1972-80	13.11	10.76	4.22	3.65	2.25	8.65	9.25
1980-85	17.09	10.81	2.32	1.76	5.76	14.61	15.24
1985-90	7.06	8.78	-0.57	-1.11	-1.54	7.72	8.31
1990-95	4.94	6.47	0.51	-0.05	-1.37	4.45	5.02

[a] adjustment based on ratio of PPI/CPI detailed indexes, 1993-98
[b] adjustment based on Frank, Berndt, and Busch (1998, Table 7, Chained Weights, Fisher Ideal)

Table 13: Alternative Estimates of Price Change for Mental Health Treatments

	1992-93	1993-94	1994-95	Average annual change 1992-95
Unadjusted[1]	1.076	1.038	1.029	1.047
Adjustment 1 Index[1]	0.983	1.021	1.029	1.011
FBB[2]	0.963	0.974	1.030	0.989

note: ratio of FBB/ Adjustment 1 Index = 0.989/1.011 = 0.978

[1] source: Appendix Table A-2
[2] source: Frank, Berndt, and Busch (1998, Table 2); see also Appendix Table A-2

References

Aaron, Henry J. and William B. Schwartz. 1983. *The Painful Prescription: Rationing Hospital Care.* Washington, DC: The Brookings Institution Press.

Averill, Richard F., Robert L. Mullin, Barbara A. Steinbeck, Norbert Goldfield, and Enes D. Elia. 1997. *Diagnosis Related Groups, Definitions Manual*, version 15.0. 3M Health Information Systems.

Barzel, Yoram. 1969. "Productivity and the Price of Medical Services." *Journal of Political Economy* 77(6): 1014-1027, November/December.

Berndt, Ernst R., Iain M. Cockburn, and Zvi Griliches. 1996. "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs." *Brookings Papers on Economic Activity* 0(0): 133-199, Microeconomics.

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 Costs," *National Bureau of Economic Research Working Paper* 6817 (November).

Cardenas, Elaine M. 1996. "Revision of the CPI Hospital Services Component." *Monthly Labor Review* 119(12): 40-48, December.

Castles, Ian. 1997. "Review of the OECD-Eurostat PPP Program," presented at the Meeting on the Eurostat-OECD Purchasing Power Parity Programme, Chateau de la Muette, Paris, November 5-6.

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

Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, and World Bank. 1993. *System of National Accounts 1993.* Office for Official Publications of the European Communities Catalogue number CA-81-93-002-EN-C, International Monetary Fund Publication Stock No. SNA-EA, Organisation for Economic Co-operation and Development OECD Code 30 94 01 1, United Nations publication Sales No. E.94.XVII.4, World Bank Stock Number 31512.

Cooper, Barbara S., and Dorothy P. Rice. 1976. "The Economic Cost of Mental Illness Revisited." *Social Security Bulletin* 39(2): 21-36, February.

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.

DuPont, Robert L, Dorothy P. Rice, Leonard S. Miller, Sarah S. Shiraki, Clyton R. Rowland, and

Henrick J. Harwood. 1996. “Economic Costs of Anxiety Disorders.” *Anxiety* 2: 167-172.

Ellison, Sara Fisher and Judith K. Hellerstein. 1999. “The Economics of Antibiotics.” In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatments*. Washington, DC: The Brookings Institution Press.

Executive Office of the President, Office of Management and Budget. 1987. *Standard Industrial Classification Manual*. Springfield, VA: National Technical Information Service.

Fein, Rashi. 1958. *Economics of Mental Illness.* Monograph Series Number 2. New York: Basic Books Incorporated.

Feldstein, Martin S. 1969. “Discussion.” In Victor R. Fuchs, ed., *Production and Productivity in the Service Industries*, as a response to “Some Problems in the Measurement of Productivity in the Medical Care Industry,” by M. W. Reder. National Bureau of Economic Research, Studies in Income and Wealth, vol. 34. New York: Columbia University Press.

Fisher, Franklin M., and Karl Shell. 1972. *The Economic Theory of Price Indices: Two Essays on the Effects of Taste, Quality, and Technological Change*. New York: Academic Press.

Frank, Richard G. and Mark S. Kamlet. 1985. “Direct Costs and Expenditures for Mental Health Care in the United States in 1980.” *Hospital and Community Psychiatry* 36(2): 165-8, February.

Frank, Richard G., Ernst Berndt, and Susan H. Busch. 1998. “Price Indexes for Acute Phase Treatment of Depression.” *National Bureau of Economic Research Working Paper* 6799 (November).

Frank, Richard G., Ernst R. Berndt, and Susan Busch. 1999. “Price Indexes for the Treatment of Depression.” In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatments*. Washington, DC: The Brookings Institution Press.

Freeland, Mark S., George S. Chulis, Aaron P. Brown, David Skellan, Brenda T. Maple, Naphtale Singer, Jeffrey Lemieux, and Ross H. Arnett, III. 1991. “Measuring Hospital Input Price Increases: The Rebased Hospital Market Basket.” *Health Care Financing Review* 12(3): 1-14, Spring.

Garber, Alan M. and Charles E. Phelps. 1992. “Economic Foundations of Cost-Effective Analysis.” Working Paper No. 4164. Cambridge, Mass.: National Bureau of Economic Research (September).

Gilbert, Milton. 1961. “Quality Changes and Index Numbers.” *Economic Development and Cultural Change* 9(3): 287-294, April.

Gold, Marthe R., Joanna E. Siegel, Louise B. Russell, and Milton C. Weinstein. 1996. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press.

Griliches, Zvi. 1992. “Introduction.” In Zvi Griliches, ed., *Output Measurement in the Service Sectors*. National Bureau of Economic Research, Studies in Income and Wealth, vol. 56. Chicago: University of Chicago Press.

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.

Harwood, Henrick J., Diane M. Napolitano, Patricia L. Kristiansen, and James J. Collins. 1984. *Economic Costs to Society of Alcohol and Drug Abuse and Mental Illness, 1980*. Research Triangle Park, NC: Research Triangle Institute.

Health Care Financing Administration. 1998. http:www.hcfa.gov/stats/NHE-OAct/tables/t11.htm.

Hodgson, Thomas A. and Andrea N. Kopstein. 1984. “Health Care Expenditures for Major Diseases in 1980.” *Health Care Financing Review* 5(4): 1-12, Summer.

Hodgson, Thomas A. 1997. “Medical Care Expenditures for Major Diseases, 1995,” study carried out at the National Center for Health Statistics, Centers for Disease Control and Prevention. Unpublished paper.

Hodgson, Thomas A. and Alan J. Cohen. 1998. “Medical Care Expenditures for Major Diseases, 1995,” National Center for Health Statistics, Centers for Disease Control and Prevention, unpublished paper.

Keeler, Theodore E. 1996. Comments to “Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs,” by Ernst R. Berndt, Iain M. Cockburn, and Zvi Griliches. *Brookings Papers on Economic Activity*: 189-99, *Microeconomics*.

Lazenby, Helen C., Katherine R. Levit, Daniel R. Waldo, Gerald S., Adler, Suzanne W., Letsch, and Cathy A. Cowan. 1992. “National Health Accounts: Lessons from the U.S. Experience.” *Health Care Financing Review* 13(4): 89-103, Summer.

Landefeld, J. Steven and Robert P. Parker. 1997. “BEA's Chain Indexes, Time Series, and Measures of Long-Term Economic Growth.” *Survey of Current Business* 77(5): 58-68, May.

Levine, Daniel and Dianne Levine. 1975. *The Cost of Mental Illness, 1971.* DHEW Publication No. (ADM) 76-265 Rockville, MD: National Institute of Mental Health.

Levit, Katherine R. and Cathy A. Cowan. 1991. “Business, Households, and Governments: Health Care Costs, 1990.” *Health Care Financing Review* 13(2): 83-93, Winter.

Levit, Katherine R., Helen C. Lazenby, Bradley R. Braden, Cathy A. Cowan, Patricia A. McDonnell, Lekha Sivarajan, Jean M. Stiller, Darleen K. Won, Carolyn S. Donham, Anna M.

Long, and Madie W. Stewart. 1996. "National Health Expenditures, 1995." *Health Care Financing Review* 18(1): 175-214, Fall.

Levit, Katherine R., Helen C. Lazenby, Bradley R. Braden, Cathy A. Cowan, Arthur L. Sensenig, Patricia A. McDonnell, Jean M. Stiller, Darleen K. Won, Anne B. Martin, Lekha Sivarajan, Carolyn S. Donham, Anna M. Long, and Madie W. Stewart. 1997. "National Health Expenditures, 1996." *Health Care Financing Review* 19(1): 161-200, Fall.

Levy, Frank. 1999. "Some Initial Results on Productivity in Car Dealerships and Auto Repair." Unpublished paper, presented to the Brooking Workshop on Measuring the Output of Business Services, Washington DC, May 14.

Lum, Sherlene K.S. and Robert E. Yuskavage. 1997. "Gross Product by Industry, 1947-96." *Survey of Current Business* 77(11): 20-34, November.

McGinnis, J. Michael. 1996. Preface to *Cost-Effectiveness in Health and Medicine by* Marthe R. Gold, Joanna E. Siegel, Louise B. Russell, and Milton C. Weinstein. New York: Oxford University Press.

Manton, Kenneth G. and James W. Vaupel. 1995. "Survival after the Age of 80 in the United States, Sweden, France, England, and Japan." *New England Journal of Medicine* 333(18): 1232-5, November 2.

Mark, Tami, David McKusick, Edward King, Henrick Harwood, and Jim Genuardi. 1998. *National Expenditures for Mental Health, Alcohol and Other Drug Abuse Treatment, 1996*. Rockville, MD: U.S. Department of Health and Human Services. DHHS Publication No. SMA 98-3255.

Mathers, Colin, Ruth Penm, Rob Carter, and Chris Stevenson. 1998. *Health System Costs of Diseases and Injury in Australia 1993-94: An Analysis of Costs, Service Use and Mortality for Major Disease and Injury Groups*. Canberra: Australian Institute of Health and Welfare.

McGreevey, William. 1996. NHA at the World Bank, Past and Prospect. Paper presented at the Workshop on National Health Accounts: Developing Internationally Comparable Data on Health Care Expenditure and Financing (November 15). Washington, DC: World Bank.

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

McKusick, David, Tami L. Mark, Edward King, Rick Harwood, Jeffrey A. Buck, Joan Dilonardo, and James S. Genuardi. "Spending for Mental Health and Substance Abuse Treatment, 1996." *Health Affairs* 17(5): 147-57, September/October.

Ministere du Travail et des Affaires Sociales. 1996. *Comptes Nationaux de la Sante 1993-1994-1995*. Paris: Ministere du Travail et des Affaires Sociales, Service des Statistiques, des Etudes et

des Systemes d'Information (July).

Mokyr, Joel. 1997. "Valuation of New Goods Under Perfect and Imperfect Competition." In Timothy F. Bresnahan and Robert J. Gordon, eds., *Economics of New Goods.* National Bureau of Economic Research, Studies in Income and Wealth, vol. 58. Chicago: University of Chicago Press.

Moore, Rachel, Yang Mao, Jun Zhang, and Kathy Clarke. 1997. *Economic Burden of Illness in Canada, 1993.* Catalogue No H21-136/1993E. Canada: Minister of Public Works and Government Services Canada.

Moulton, Brent R. 1996. "Constant Elasticity Cost-of-Living Index in Share-Relative Form." Bureau of Labor Statistics, unpublished paper, December.

Murray, Christopher J.L. and Alan D. Lopez, eds. 1996. *Global Burden of Disease: a Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020*. 2 volumes. Cambridge, Massachusetts: Harvard University Press.

National Health Service. 1996. *Burdens of Disease*. London: Health Authority Chief Executives, District Directors of Public Health, Directors of Finance, Directors of Purchasing and Regional Directors of Research and Development, October.

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

Organization for Economic Cooperation and Development. 1997. *Principles of Health Accounting for International Data Collections: Working Party on Social Policy Ad Hoc Meeting of Experts in Health Statistics*. Paris, France: Organization for Economic Cooperation and Development.

Parsons, P. Ellen, Richard Lichtenstein, S.E. Berki, Hillary A. Murt, James M. Lepkowski, Sharon A. Stehouwer, and J. Richard Landis. 1986. *Costs of Illness United States, 1980.* National Medical Care Utilization and Expenditure Survey, Survey C, Analytical Report No. 3, DHHS Publication No. 86-20403, National Center for Health Statistics, Public Health Service, Washington, DC, April.

Pauly, Mark. 1999. "Medical Care Costs, Benefits, and Effects: Conceptual Issues for Measuring Price Changes." In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatments*. Washington, DC: The Brookings Institution Press.

Pommier, Philippe. 1981. "Social Expenditure: Socialization of Expenditure? The French Experiment with Satellite Accounts." *Review of Income and Wealth* 27(4): 373-86, December.

Reder, M. W. 1969. "Some Problems in the Measurement of Productivity in the Medical Care Industry." In Victor R. Fuchs, ed., *Production and Productivity in the Service Industries*.

National Bureau of Economic Research Studies in Income and Wealth, vol. 34. New York: Columbia University Press.

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., Barbara S. Cooper, and Gibson. 1982. "Accounting for Health: An International Survey." In *Le Santé Fait ses Comptes* edited by Emile Levy. Paris: Economica.

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.

Rice, Dorothy, Sarah Dunmeyer, Sander Kelman, and Leonard S. Miller. 1990. *The Economic Costs of Alcohol and Drug Abuse and Mental Illness: 1985.* DHHS Publication No. (ADM) 90-1694 University of California, San Francisco: Institute for Health and Aging. Alcohol, Drug Abuse, and Rockville, MD: Mental Health Administration.

Rice, Dorothy P. and Leonard S. Miller. 1998. "Health Economics and Cost Implications of Anxiety and Other Mental Disorders in the United States." *The British Journal of Psychiatry* 173(34): 4-9, July.

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, Matthew P. and David W. Wilcox. 1996. "Mismeasurement in the Consumer Price Index: An Evaluation." *NBER Macroeconomics Annual* 11: 93-142.

Sherwood, Mark K. 1999. "Output of the property and casualty insurance industry." *Canadian Journal of Economics* 32(2): 518-46, April.

Survey of Current Business August 1997, Table 2.4.

Teillet, Pierre. 1988. "A Concept of Satellite Account in the Revised SNA." *Review of Income and Wealth* 34(4): 411-439, December.

Triplett, Jack E. 1983. "Concepts of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource-Cost Debate." In Murray F. Foss, ed., *The U.S. National Income and Product Accounts: Selected Topics*. National Bureau of Economic Research, Studies in Income and Wealth, vol. 47. Chicago: The University of Chicago Press.

Triplett, Jack E. And Ernst Berndt. 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. Washington D.C.: The Brookings Institution Press.

Triplett, Jack E. and Ernst Berndt. 1999. "New Developments in Measuring Medical Care." In *Measuring the Prices of Medical Treatments* edited by Jack E. Triplett: 1-33. Washington D.C.: The Brookings Institution Press.

United Kingdom Department of Health. 1994. *Register of Cost-Effectiveness Studies*. London, England: Department of Health, Economics and Operational Research Division (August).

United Kingdom Department of Health, National Health Service. 1996. *Burdens of Disease*. London, England: Department of Health, Economics and Operational Research Division (October), catalog number 96CC0036.

United States Congress, Joint Economic Committee. 1961. *Government Price Statistics: Hearings before the Joint Economic Committee, Congress of the United States*. Eighty-seventh Congress, first session pursuant to Section 5(a) of Public Law 304 (709th Congress). Part 2. Washington, D.C.: United States Congress (May 1, 2, 3, 4 and 5).

United States Department of Commerce, Bureau of Economic Analysis. 1989. *Government Transactions*. Methodology Paper Series MP-6. Washington, D.C.: Government Printing Office.

United States Department of Health and Human Services. 1989. *The International Classification of Diseases, 9th Revision, Clinical Modification: ICD-9-CM*, third edition, vols. 1-3. Washington, D.C.: United States Department of Health and Human Services, Public Health Service, Health Care Financing Administration.

United States Department of Labor, Bureau of Labor Statistics. 1992. *Handbook of Methods* Bulletin 2285. Washington D.C.: United States Government Printing Office (April).

United States Department of Labor, Bureau of Labor Statistics. 1996. "Changing the Hospital and Related Services Component of the Consumer Price Index." *CPI Detailed Report*: 7-8, June.

United States Department of Labor, Bureau of Labor Statistics. 1999. "Consumer Price Indexes - All Urban Consumers (current series)," http://146.142.4.24/cgi-bin/dsrv?cu, "Producer Price Index Revision - Current Series," http://146.142.4.24/cgi-bin/dsrv?pc, Kay Ford, Dan Ginsberg.

Vanoli, A. 1975. *Comment Structurer le Système Statistique d un Domaine et l articuler avec les Autres: L exemple de la Santé.* International Statistical Institute, Warsaw.

Vanoli, A. 1986. "Sur la Structure Générale du SCN à Partir de l'expérience du Système Élargi de Comptabilité Nationale Française." *Review of Income and Wealth* 32(2), June.

Vaupel, James W. 1998. "Demographic Analysis of Aging and Longevity." *American Economic*

Review 88(2): 242-47, May.

Waldo, Daniel R., Sally T. Sonnefeld, David R. McKusick, and Ross H. Arnett, III. 1989. "Health Expenditures by Age Group, 1977 and 1987." *Health Care Financing Review* 10(4): 111-20, Summer.

Weisbrod, Burton A. 1999. "Measuring Health Care Prices." In *Measuring the Prices of Medical Treatments* edited by Jack E. Triplett: 251-5. Washington D.C.: The Brookings Institution Press.