NEW DEVELOPMENTS IN MEASURING MEDICAL CARE

Jack E. Triplett and Ernst R. Berndt*

How serious is "medical care" inflation in the United States? For many years, price indexes for medical care have outstripped the overall rate of inflation. For example, between 1986 and 1996, the medical care component of the Consumer Price Index (CPI) rose 6.5 per cent per year, which far exceeds the 3.6 percent annual increase in the overall CPI during this period.

However, constructing accurate price indexes for medical markets is especially difficult, and many economists believe that economic statistics on medical care do not accurately measure medical care price changes. Some very recent research, reported in this volume, suggests that—contrary to the usual presumption of runaway medical inflation—prices for at least some medical care interventions are not rising rapidly and may even be falling.

Understanding medical care inflation is important for topics such as policy toward medical care cost containment. If medical care inflation is not the driving force behind the run-up in medical care costs, as the studies included in this volume suggest, then medical care costs are being driven by real changes in the quantity of care. This finding, if confirmed by studies on other medical procedures, has major implications for health care cost containment policy, because it suggests 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.

Medical care price indexes are not only indicators of the rate of medical care inflation, they also affect other economic statistics on medical care. For example, measures of the output of the

medical care sector and of the consumption of medical care in national accounts use medical care price indexes as deflators, as explained in a subsequent section. Errors in the deflators create parallel errors in the measures of medical care output, and of consumption. Thus, our understanding of the most important economic trends in the medical care sector is vitally dependent on accurate medical care price measures.

This volume brings together state-of-the-art methodological and empirical work on the measurement of medical outcomes and prices. The papers in it were originally presented at a conference that was jointly sponsored by the Brookings Institution and the American Enterprise Institute for Public Policy. This introductory essay sets the stage for the conference papers, and provides a summary of the papers and the conference discussion.

I. Background

Information on aggregate U.S. medical care expenditures appears in the National Income and Product Accounts (NIPA) of the Bureau of Economic Analysis (BEA) and in the National Health Accounts (NHA) of the Health Care Financing Administration (HCFA). The latter provide the frequently cited statistic that the United States spends one-seventh of gross domestic product on medical care, the highest percentage in the world.

Medical care expenditures, of course, are the product of price and quantity. In national accounting language, the quantity of medical care is called "real" medical care expenditures. We can interpret real medical care expenditures as the output of the medical care sector and as the consumption of medical care by the U.S. population.

In national accounts, one normally estimates real medical care expenditures by removing estimated medical inflation from the increase in actual expenditures on medical care. This process is generally known as "deflation." Effectively, the part of the increase in medical care expenditures that is *not* inflation is an increase in the quantity of health care, where the quantity, in this way of looking at it, includes both changes in the number of treatments and changes in the quality or effectiveness of treatments.

With the deflation methodology, any error in the price indexes that measure medical inflation creates an equal error of opposite sign in the real expenditure, or quantity, data. Thus, accurate price indexes for medical care are important both because they measure medical care inflation and because they are used to measure medical care real expenditures.

Historically, the major source of information on U.S. medical care inflation was the medical care component of the Consumer Price Index (CPI), published by the Bureau of Labor Statistics (BLS). Health economists have long suspected that the CPI medical care price indexes have errors, and most have believed that they rise too rapidly. If measured medical price inflation has been overstated, then the NIPA and the NHA have understated improvement in the quantity of medical care. That problem has motivated research on medical care price indexes.

The CPI is designed to answer the question: How much *out-of-pocket* expenditures would be necessary in the current year to buy the basket of goods and services that were bought in the base year?[1] Medical care goods and services are included in the CPI basket.

However, on the CPI's definition, only those medical care commodities that are purchased directly by households, or the proportion of them that are financed by out-of-pocket expenditures, are within the scope of the index. For this reason, the items chosen for CPI pricing were

weighted toward those that are more frequently purchased from out-of-pocket expenditures.

For many purposes, economists and health policy analysts wanted to know—not the increase in prices for the consumer out-of-pocket portion of medical expenses, which corresponded to the CPI concept—but the price change for the full medical care sector. When they used the CPI as an indicator of price change for the whole medical care sector, they found it had a persistent shortcoming: Some expenditures, such payments by employer-provided medical insurance, were outside the scope of the CPI definition of consumption. Additionally, the CPI did not collect discounts that third-party payers, such as insurance companies, obtained from medical care providers.[2]

Another problem arose because in the old CPI (before 1997) the BLS chose a sample of medical care charges—a visit to a doctor's office, administration of a standardized medicine (such as an influenza shot), the daily room rate in a hospital, and so forth. When the CPI medical care component was constructed from the price of a hospital room rate, one expects a substantial bias relative to the cost of treating a disease because the average number of days in the hospital for a specified disease is declining. For example, the average length of stay for heart attack patients has fallen from fifteen days in 1975 to eight days in 1995 (Heidenreich and McClellan 1998), and the average hospital stay for cataract surgery has fallen from seven days in 1952 to zero currently because the surgery has become an outpatient procedure (Shapiro and Wilcox 1996). Similarly, collecting the price for a doctor's office visit would, one expects, yield a substantially biased price index when new technology greatly expands what the doctor can do during that visit.

As Dennis Fixler (this volume) pointed out, the BLS has recently addressed both these problems. Beginning in 1992, a new series of industry medical care price indexes was introduced

into the Producer Price Index (PPI). The PPI price indexes for hospitals and for physician's offices are based on charges to all payers, not just consumer out-of-pocket expenditures.

Additionally, the BLS adopted a new, improved methodology for pricing medical care for the PPI indexes. In the new methodology the BLS obtains prices for a sample of specified treatments for particular diseases, rather than obtaining the price for a day in the hospital. The improved pricing methodology was extended to the CPI in 1997. Economists agree that this new methodology is a great improvement, compared with the old one.

In its new price indexes the BLS has not, however, fully solved the difficult problem of allowing for quality change in medical treatments--changing medical technology that results in improved medical outcomes. This problem motivates most of the research summarized in this volume.

An additional part of the health measurement problem concerns price indexes for prescription pharmaceuticals. Here, again, research suggests that price indexes have in the past overstated the increase in medical price inflation.

Researchers, primarily those at the National Bureau of Economic Research (NBER), undertook several detailed evaluations and critiques of the BLS's CPI and PPI prescription drug components. Zvi Griliches and Iain Cockburn (1994) pointed out that the BLS treatment of generic drugs—as new goods, not directly comparable to their originator branded predecessors—resulted in overstated price inflation, because no price decreases were recorded in the indexes when new, cheaper generics became available. Ernst Berndt, Zvi Griliches, and Joshua Rosett (1993) also presented evidence suggesting that the PPI program oversampled middle- and older-aged drugs and undersampled newer pharmaceutical products. Because price increases were

larger for older than for younger drugs, those results implied that growth in producer price indexes for pharmaceuticals overstated true price growth. Studies by IMS, a data firm specializing in pharmaceuticals (Ristow 1996), and by BLS researchers (Kanoza 1996; Kelly 1997) corroborated the empirical findings of Berndt, Griliches, and Rosett.

Partly in response to that set of research findings, the BLS implemented changes in both its CPI and PPI prescription drug components. The CPI now treats a branded drug and its generic equivalent as the same product, rather than as different products—the previous price index treatment. With the new methodology, the introduction of a generic causes the price index to fall, as it should. U.S. Department of Labor, Bureau of Labor Statistics (undated) provides a more detailed description of the new BLS procedures involving the CPI for prescription drugs.

The BLS's prescription drug PPI program has implemented two sets of changes—one involving generics and the other supplementary sampling to bring more quickly into the sample newly introduced pharmaceuticals. Kanoza (1996) and Kelly (1997) outlined those changes. Their impact has been substantial. As Kelly has discussed, the combined impact of changing the treatment of generics and supplementary sampling of newer drugs has reduced measured inflation in the prescription drug PPI from 4.1 percent (under older procedures) to 3.3 percent (using new procedures) between January 1996 and March 1997.

II. What Do We Want in Economic Statistics on Medical Care? And Why Do We Want It?

The conference organizers and the researchers who conducted the studies included in this volume implicitly took it as obvious that economists, public policy analysts, and health industry professionals need to separate spending increases on health care into price increases and quantity increases. But in some of the conference discussion, some policy analysts present questioned the premise: Why do we want measures of price change for medical care? And why do we need to adjust the price indexes for the value of improved medical technology? Charles Phelps (this volume) and Mark Pauly (this volume) addressed the question, but perhaps it deserves a lengthier and more explicit discussion.

Phelps and Pauly gave compatible answers. Pauly contended:

> A price index should tell us how much better or worse off price changes make a particular set of consumers. Slightly more formally, if the expenditure on a particular good or basket of goods changes, a price index measures the amount of additional income needed to be provided to the consumer (for price increases) or taken away from the consumer (for price decreases) to leave him or her as well off as before a change.

In addition, both scholars gave the same reason for wanting to treat quality improvements in health care as if they were increases in the quantity of care. According to Pauly:

> It is obvious that when quality changes, the measure of the change in total expenditure needs to be adjusted to account for the quality change. . . . One reason

> why an improved price index for medical services . . . is an issue is precisely because we have strong reasons to believe that the quality of those services has been changing over time.

The Phelps-Pauly way of looking at the problem of measuring medical care prices is not so different from the way one measures economic activity in the rest of the economy. In the national health accounts (Levit et al. 1996), as in the national income and product accounts (U.S. Department of Commerce 1985), one objective is to produce measures of "real" spending, so that increased quantities are separated from price movements (inflation). Increased quantities are associated with increased consumption. Other things equal, increasing per capita consumption translates into increased economic welfare of consumers.

Yet one suspects that the answer that "we do it for the other parts of the economy" may not convince everyone. One stereotype about health has increased medical spending driven by medical care inflation. Much of the discussion surrounding the Clinton administration's health care reform initiative repeated that idea: Health care prices were rising too rapidly, many health care policy analysts contended, and the policy problem was to control those "out of control" price increases. Indeed, as Phelps remarked, rapid health care inflation has been "the fundamental premise of health care policy over the past decades."

But medical care has another stereotype: The quantities produced and consumed are excessive. Observers often allege that the United States has too many high-tech medical procedures, overprescription of pharmaceuticals, excessive care of various types, an excess of "defensive" medicine, and so forth. One suspects that the policy analysts who are unconvinced about the need for measuring the quality of medical care would say, "So you have shown me that

some of what I thought was medical price inflation is actually increased medical quantities; but I don't care, because the quantities are too great anyway." They are contending, in effect, that medical care is different: Increased consumption of apples and oranges may represent increased welfare for consumers, but increased consumption of medical care does not necessarily increase welfare.

Indeed, Pauly asserted:

> Virtually the entire American health care policy analysis in the past thirty years has been based on the view that the medical market did not function like ordinary markets. . . . Whether the cause was health insurance, consumer ignorance, and consequent demand inducement or a postulated "technological imperative," the U.S. medical care sector fosters some changes in technology whose value fell far short of their cost.

Researchers have also confronted that challenge. One does not want an accurate price index for medical care solely to determine that part of medical expenditure increases were driven by increased quantities of medical procedures. Indeed, we knew that, even without a price index.

The underlying issue, as several of the studies stated explicitly, is to determine whether the increased use of medical procedures improves patient outcomes. If new medical technologies have a value that falls short of their cost, as sometimes asserted, then price indexes that accurately account for the value of quality change in health care will rise; additionally, real health care expenditures (or quantity) will grow less rapidly than actual expenditures.

Conversely, if the value of new technology exceeds its cost, the price indexes will fall, even if medical expenditures are rising; and measures of real medical expenditures will rise even

more rapidly. For example, David Cutler, Mark McClellan, and Joseph Newhouse (this volume) stated:

> We define the change in the cost-of-living index as the increased spending on medical care over time less the additional value of that care. If medical care increases in cost without much improvement in health, that would be an increase in the cost of living. If medical care increases in cost, but the value of that care rises over time, the cost-of-living index would be falling.

To estimate such "quality-adjusted" price indexes, one needs to be able to connect measures of medical outcomes with expenditures on medical treatments. The crucial measurement issue, then, is the valuation of medical care improvements. Obtaining accurate information on the relations among price change, quantity change, changes in medical outcomes, and expenditure change is the real objective of improved economic statistics on medical care treatments.

Quantitative measures of medical outcomes thus play an essential role in improving economic measurement of the medical sector. Cutler, McClellan, and Newhouse (this volume), Richard Frank, Ernst Berndt, and Susan Busch (this volume), Pauly (this volume), and also Hay and Yu (this volume), who call attention to the lack of a sufficient quantity of medical outcome information, emphasize the importance of that role. Jack Triplett (this volume) points out that the absence of medical outcome measures in the past has left the nation's economic statistics on medical care inadequate and frequently criticized.[3] John Eisenberg (this volume) pointed the way to government-sponsored improvements in information on medical outcomes that could in principle be used to make economic statistics on health care better reflect the contribution of

medical care to improvements in health.

The policy analysts' objection has some validity as a general statement of a problem. Medical care differs in some ways from many other consumption goods and services (though not necessarily from all of them), in the sense that a great amount of concern exists within the medical profession and within the policymaking community about the relative efficacy and the cost-effectiveness of some forms of medical treatment. But economists have not ignored that problem. Far from it. One of the recurring themes in the studies reviewed here was the need—for economic measurement as well as for other purposes—of exactly the measures of medical outcome that would tell us whether increased quantities of medical care really do improve the welfare of medical care recipients.

A panel discussion, involving Willard Manning, David Meltzer and Burton Weisbrod (this volume), explored some deeper issues that are closer to the frontiers of economic understanding than were the policy analysts' questions, though superficially the questions look similar. Except for very narrowly specified purposes, we do not ultimately want to know about the price change for a particular treatment. Pricing treatments is, as Willard Manning (this volume) put it, a move in the right direction because it moves us closer to what is ultimately wanted. Ideally, economists want to measure the cost of producing increased survival and increased quality of life. We ultimately want to know, as David Meltzer (this volume) put it, "whether expenditures on health care are worth it." That requires us to put a value on health care and a value on improving health.

Manning and Meltzer, as well as Weisbrod, also emphasized that a price index is closely related to a measure of economic welfare, though that is sometimes not recognized in the pragmatic use of the statistics that government agencies produce. The problem—and this is what

makes the measurement of health care really hard—is that knowing how to measure health in its broad dimension requires knowing a great deal about individuals' preferences for different health states and for different alternatives that might be produced for the same expenditure of medical resources. For example, measures of medical outcomes such as quality-adjusted life years[4] explicitly require information about preferences among different health outcomes.

Manning (this volume) and Weisbrod (this volume) also addressed another serious omission in existing measures of the cost of health care. BLS price indexes, the national accounts as produced by the BEA, the national health accounts as produced by HCFA, and most estimates of cost-of-living indexes (such as Cutler, McClellan and Newhouse, 1998), omit all nonmarket costs of illness and of receiving health care. Such nonmarket costs are outside the traditional "production boundary" of the national accounts. In Manning's example, if someone withdraws from the labor force to care for a family member (instead of continuing to work and placing the family member in a nursing home), the national accounts assume that the time devoted to such care is worth nothing. Of course, that is true of the national accounts (and CPI) treatment of all economic activities so it is not unique to the national accounts treatment of health care. But as these economists pointed out, the issue looms larger in the case of health care, because time costs are associated with the care of others and because time costs are associated with obtaining care for oneself. Omitting those costs seems likely to bias significantly our understanding of changes in the cost, or price, of health care.

III. The Papers for the Conference

The remaining sections of this essay discuss the papers included in this volume and the issues they raise.

The Costs and Benefits of Intensive Treatment for Cardiovascular Disease

Cutler, McClellan, and Newhouse (this volume) summarize a continuing research program on the treatment of heart attacks that they are conducting in collaboration with other researchers. In this research, they examined changes in the technology for treating heart attacks, estimated the impact of changes in technology and in pharmaceuticals on survival probabilities, determined the sources of increased costs for treating heart attacks, and finally proposed an answer to the question, Was the increased expenditure worth it?

Cutler and his colleagues demonstrate what economic research on medical treatments can do and provide a greater understanding of what remains to be done. Coupled with the quite different work on a different disease—mental depression—by Frank, Berndt, and Busch (this volume), the study provides a model that researchers could well follow to analyze other medical conditions.

The setting for the study by Cutler, McClellan, and Newhouse is the remarkable decline in the mortality implications of heart attacks. According to the National Center for Health Statistics (NCHS), the number of recorded U.S. heart attack deaths declined by about a third between 1975 and 1995 (table 1, line 1a). Over that twenty-year period, however, the population grew, and the age structure of the population shifted toward the older groups that are more vulnerable to heart

attacks, so the improvement in mortality experience was greater than the impression conveyed by the unadjusted numbers. The NCHS calculated that the heart attack death rate declined by 45 percent between 1975 and 1995 and that the age-adjusted rate declined by nearly 60 percent (table 1, lines 1b and 1c).[5]

Why has the heart attack death rate declined? A heart attack is not the only outcome of heart disease, nor is it the only cause of death from heart disease. Heart attacks account for a little less than half of recorded deaths from ischemic heart disease.[6] Changes in heart disease and heart attack incidences (and possibly severity) might be attributable to direct medical intervention—primary prevention and secondary prevention—or to nonmedical sources, such as changes in diet or in exercise habits or less smoking.

Some of the changes that we might observe in the population, had we sufficient data and research, undoubtedly offset each other. New pharmaceuticals that reduce blood vessel clogging or reduce blood pressure, for example, might counteract the effects of more sedentary behavior and richer diets, which would produce offsetting influences on aggregate incidence rates. But the risk factors for heart disease are themselves not independent of medical knowledge and medical practice.

We can next ask, What happens after a victim suffers a heart attack? A heart attack, when detected, requires treatment in a hospital. The first step to treatment is therefore to get the victim to a hospital, and some die before receiving any hospital treatment—indeed, most heart attack deaths occur outside hospitals (compare lines 1a and 2 in table 1).

Table 1 implies a substantial drop in out-of-hospital deaths from heart attacks. Although there is some ambiguity in the aggregate data, the total number of heart attack deaths has declined

substantially between 1975 and 1995—107 thousand fewer deaths, a reduction of about one-third (line 1a). Over the same interval, the number of heart attack deaths occurring in hospitals was effectively unchanged (line 2). Those two numbers suggest that most of the total 1975–1995 decline in heart attack deaths occurred because of a sharp reduction in the number of heart attack victims who die before reaching a hospital and, perhaps, after discharge from a hospital—the decline in out-of-hospital heart attack deaths is over 40 percent.[7]

Both medical and nonmedical factors influence prehospitalization death rates. In a review of the medical literature on heart attack treatments, Paul Heidenreich and Mark McClellan (1998) reported that existing studies of initial treatment changes (such as advanced cardiac life-support facilities in emergency vehicles) fail to explain much of the out-of-hospital death rate decline: "The actual improvements in prehospital technologies appear to account for only a modest increase in the number of [heart attack] patients reaching the hospital alive." Heidenreich and McClellan also noted that changes in coding practices may account for some of the recorded decline in nonhospital heart attack deaths (because those cases are sometimes difficult to classify in the absence of an autopsy) and that reporting errors appear to be correlated with reported incidence rates. Yet the aggregate numbers indicate that nonhospital heart attack deaths have dropped substantially, and it is difficult to believe that all of that drop could result from measurement error.

Of course, a nonhospital heart attack death averted might only result in an in-hospital death. Indeed, one expects that the most severe heart attack incidences, on average, involve victims who are more likely to die before hospitalization. If a higher proportion of heart attack victims did reach the hospital in 1995 (as noted, it is not certain that such is the case), then those

cases increase at the margin the severity of the cases the hospital must treat. Heidenreich and McClellan (1998) cited fragmentary evidence in favor of that "marginal" patient hypothesis.

The actual story seems to be that the *average* severity of heart attack hospitalizations has declined, even if the marginal cases have increased severity. Cutler, McClellan, and Newhouse reviewed several measures of heart attack severity among hospital patients and concluded that "average [heart attack] case severity decreased slightly, accounting for 10 percent to 20 percent of the decline in the average . . . mortality rate." Heidenreich and McClellan (1998) concluded that "the bulk of the evidence suggests that changes in the nature of [heart attacks] in hospitalized patients accounts for a significant part of the observed improvements in outcomes," particularly before 1985.

Thus, something has reduced the number of heart attacks that cause death before hospitalization and presumably has also reduced the severity of cases that are hospitalized. We cannot determine whether the causal factors are improvements in general health of the population,[8] changes in primary prevention methods, or something else. Aggregate U.S. data on the incidence and severity of heart attacks would be useful for exploring this question.[9]

What has happened to heart attack victims once they reach the hospital? According to NCHS data, the actual number of in-hospital heart attack deaths has hardly changed (table 1, line 2). Because the population has grown, a more relevant statistic is the hospital death rate, or case fatality rate. But the aggregate NCHS data on heart attack hospitalizations, and thus also the hospital death rate, are ambiguous.

Two NCHS classifications for heart attack hospitalizations exist. "All listed" includes any hospital discharge where heart attack is mentioned among the diagnoses. "First listed" means that

a heart attack was the principal condition "established after study to be chiefly responsible for occasioning the admission" (U.S. Department of Health and Human Services, 1995). Constructing a consistent time series is complicated by a coding change in 1982 that transferred some two hundred thousand cases into the first-listed category of heart attacks.[10]

We have adjusted the published NCHS first-listed hospital discharge time series by estimating the effect of the 1982 coding change.[11] The result, shown on line 3a of table 1, is an increase of 216 thousand heart attack hospitalizations between 1975 and 1995, a little less than 40 percent.

However, Heidenreich and McClellan (1998), working with National Hospital Discharge Survey (NHDS) microdata and other longitudinal data sources, rather than published NCHS tabulations, have noted substantial increases in transfers and readmissions. They adjust heart attack discharges additionally for these effects and report (table 1, line 3b) that adjusted new heart attack discharges actually fell. There are probably inconsistencies between the number of heart attack hospitalizations computed by researchers from NCHS microdata and the total that is estimated from the same data by NCHS, and in interpreting the data in table 1, those should be kept in mind. Nevertheless, hospital fatality rates, computed with either adjusted series (3a or 3b), fell substantially between 1975 and 1995--at least by a quarter (line 4a), and perhaps by nearly 40 percent (line 4b). The latter estimate contains an additional adjustment for age-sex composition of the heart attack population (which is much older than the US population) and for the fact that some reported heart attack deaths occurred in cases that had been readmitted, rather than truly new cases.

Heidenreich and McClellan (1998) contend that an even better indicator of the efficacy of

treatment is the age-adjusted 30-day fatality rate, which is influenced by treatment in the hospital and in the post-hospitalization period. That rate has also fallen by 36 per cent (table 1, line 4c).

In summary, the age-adjusted population death rate from heart attacks declined by nearly 60 percent in the twenty-year period 1975-1995 (table 1, line 1c). The age-sex adjusted thirty-day fatality rate for hospitalized cases fell by a little less than 40 percent (line 4c). Although the two rates are not exactly comparable, it is clear that the improvement in heart attack mortality experience applies to both nonhospitalized and hospitalized cases, with perhaps a somewhat greater aggregate improvement in the cases where deaths occur outside hospitals.

For the phase they call "acute management," which encompasses the thirty-day interval after the onset of a heart attack, Cutler, McClellan, and Newhouse estimated that changes in the hospital treatments of heart attacks accounted for about 55 percent of the decline in acute-phase mortality. New or more widespread use of pharmaceutical treatments accounting for 50 percentage points of the total, by far the largest effects on heart attack mortality. The increased use of more invasive cardiac procedures (angioplasty and heart bypasses) accounted for a limited share of the reduction—only about 5 percent, in their preferred estimate. Results from the medical literature cannot explain the other 45 percent of reduction in the in-hospital death rate.[12]

Moreover, table 1 tabulates only mortality. Heart disease affects the quality of life as well as the quantity of it, and the available evidence (see Cutler, McClellan, and Newhouse, this volume, their table 2) suggests improvements as well in the quality of life for heart disease sufferers.

Heart Attack Price Indexes. Although cardiovascular health has improved, expenditures for

treating cardiovascular disease have soared. Nearly one-seventh of all medical spending is for cardiovascular disease (Hodgson and Cohen 1998), and that share has grown over time. Medicare hospital spending per heart attack is nearly $15,000. With about 230,000 new heart attack cases annually, the Medicare program currently spends more than $3 billion on hospital care for heart attacks alone. Reductions in mortality have been gratifying; but were the dramatic spending increases worth it? Are we better off devoting money to cardiovascular disease, or would we be gain by reallocating resources elsewhere?

To answer these questions, Cutler, McClellan, and Newhouse first examine why Medicare spending on heart attacks has increased over time—from $11,000 per case in 1984 to $15,000 in 1991 (they adjusted both numbers for the overall rate of inflation). Using detailed claims records for the Medicare population, the researchers found that essentially all of the increased spending resulted from more intensive treatment of heart attacks.

For each type of heart attack treatment, the change in Medicare reimbursement was virtually identical to the overall inflation rate. Medicare spending rose more rapidly than the inflation rate because more patients received more intensive services, for which Medicare pays more, not because the rates for each of the services themselves rose faster than the general rate of inflation.

What were the health outcome effects of increased treatment intensity? Were the increased expenditures worth it? To help answer those questions, Cutler, McClellan, and Newhouse computed price indexes for heart attack treatments that allowed for the value of the more intensive treatments that heart attack patients have received. They compared the results with the CPI medical care component.

The authors appropriately emphasize a number of qualifications to the CPI comparisons they made. The CPI medical care component has a coverage that is broader than just heart attacks, and a price index for heart attacks is not necessarily representative of price indexes for all medical treatments. Additionally, because of data availability, Cutler, McClellan, and Newhouse computed price indexes from Medicare data for elderly heart attack victims and from heart attack cases in a large teaching hospital, whose name must be kept confidential. A price index for one hospital may not be representative of the entire country, and the cost experience of Medicare may not correspond very well to the costs of non-Medicare, nonelderly medical care. Nevertheless, when the authors mimicked CPI procedures on data from the major hospital, the two trend rates of change differed by only one-tenth (0.1 percentage point) over the 1984-1991 interval.

Although we should keep all those qualifications in mind, the Cutler, McClellan, and Newhouse results are provocative. In table 2, we have rearranged information from their table 10 to present their information in a slightly different form.

As we noted earlier, the CPI is in concept a fixed-weight approximation to a cost-of-living index, and the CPI concept includes monetary expenditures only. Before 1997, CPI medical components were constructed from a sample of standardized charges. Cutler, McClellan, and Newhouse referred to that as a "service price index," though one might also call it a "price index for medical care inputs."

The BLS adopted an improved methodology for the CPI in 1997 and earlier for the PPI program: the BLS now obtains prices for specified treatments for a particular disease. Cutler, McClellan, and Newhouse refer to that as a "treatment regimen price index." One expects that the treatment regime price index will rise less than the medical input price index if, for example, the

number of hospital days required to treat a medical condition declines with improved medical methods.

Another implementation question concerns the weights for the price index. As the Bureau of Labor Statistics implemented the "old" CPI method (pricing medical inputs), the weights in the index were held constant for a relatively long interval; the BLS used the average expenditure proportions in 1982–1984 until early 1998. The price index literature is dominated by concern that holding the weights unchanged might give an inflation measure that rises more than an index number that either adjusts the weighting structure more rapidly (usually referred to as a chain index) or one that uses a more complex index number formula that accommodates changes in consumers' market baskets of goods and services.

Panel A of table 2, which we base on Cutler, McClellan, and Newhouse's "major teaching hospital" price indexes, shows that both aspects of CPI implementation—pricing the inputs to medical care and holding the weights fixed—bias upward the price indexes for medical care. Holding the weights for medical inputs fixed at their 1984 proportions (which is roughly coincident with the CPI's former 1982–1984 weighting period) creates an upward bias of 1.7 percentage points per year, compared with a chain price index that updates the weights annually. A price index that is based on medical inputs creates an additional 0.3 percentage point of upward bias per year, compared with the treatment regimen price index, when one computes both on a chain basis.

Thus, the constant-weight aspect of the old BLS method had an effect (1.7 points) that was much larger than the agency's former decision to price inputs rather than treatments (0.3 points). Economists will probably find that result surprising. As we noted earlier, one expects a

substantial bias when the price index is formed from the hospital room rate or on the price for a doctor's office visit, because medical technology has reduced the length of hospital stays and because the doctor can now do more during that visit. The vision that pricing medical care inputs must bias the CPI upward has long been part of the health economist's tool bag. The Cutler, McClellan, and Newhouse research indicates that the vision is correct, but that it only accounts for a relatively small (0.3 percentage point) effect. Cutler, McClellan, and Newhouse's results suggest that the CPI could have been converted into a fairly effective cost-of-treatment price index simply—and cheaply—by frequently updating the weights attached to the medical inputs that it was pricing all along.

Of course, Cutler, McClellan, and Newhouse's regimen price index for heart attack treatments might not move closely with a treatment regimen price index for other diseases, and perhaps there was something atypical about the price indexes constructed for their major teaching hospital. But their finding that price index weights have a greater effect on medical price indexes than does technical change that reduces medical inputs per medical treatment is provocative and suggestive for future research on sources of measurement error in medical price indexes.

BLS materials emphasize that the CPI is an approximation to a cost-of-living index—though, as noted above, an approximation that is circumscribed in specific ways. One can also look at the price of medical treatments within the cost-of-living index framework. When we are dealing with health, the way the cost-of-living index is defined matters.

For a cost-of-living index for goods—apples, oranges, carrots, and cars—the consumer's monetary resources, monetary income or monetary wealth, is what matters, mostly. It is true that nonmonetary costs also exist for nonmedical consumption goods—for example, time costs for

shopping for apples and oranges, food preparation time for the carrots, and driving time for use of the car. Economists have recognized, at least since the work of Becker (1960), that the omission of the consumer's time in studies of consumption may neglect an important element that is necessary to analyze consumption behavior. If time cost is necessary to analyze consumption behavior, it is also necessary for a full accounting for a cost-of-living index. In the case of most consumption goods, neglecting time and other nonmonetary costs has a long tradition, and staying within the traditional market boundary that encloses money income and money expenditures exclusively is probably adequate for most purposes. For that reason, the CPI concept has no room in it for the time required for any consumption activity, or for improvements in any broader concept of income or of consumption that the household does not receive in monetary terms.

In the case of health, and in some other services as well, it becomes essential to go beyond the usual market or production boundary. One seeks medical care for a heart attack to prolong life. Few individuals ask explicitly: What is the value of the life for which I am purchasing medical treatment? Yet, it is hard to understand the outcome of a medical treatment for heart attacks without considering the expected change in mortality that the patient wants the treatment to bring about.

Additionally, if mortality change is a medical outcome measure that needs to be brought into the price index calculation in some manner, it is hard to see immediately how to do that without being willing to put a value on the life so extended. Additional expenditures on apples and oranges give more consumption of apples and oranges (even if what is really wanted is in some sense nutrition). But no normal person buys more medical care to consume more medical care; what one expects to consume is more life or more health.

Accordingly, Cutler, McClellan, and Newhouse specified that a cost-of-living index should measure "the increased spending on medical care over time less the additional value of that care." To compute the value of improvements in heart attack care, they valued increased years of life expectancy for heart attack victims by an estimate of the value of an additional life year, taken from economic estimates that were computed for other purposes. The researchers chose a value of $25,000 for an additional life year, which is on the conservative side of current practice among economists.

Even with that low value, however, the researchers found that the benefits of increased survival (eight months at $25,000 per year, or about $12,000 in present value terms) are greater than the additional spending (about $4,000). Using a lower estimate of the change in life expectancy from improved medical treatments (four months) gave a lower estimate of the value, but one that was still well above the increment to medical costs. The lower estimate assumes that the entire Medicare population (heart attack victims and those who did not suffer a heart attack) enjoyed increased health from nonmedical sources, so that nonmedical factors caused part of the improvement in life expectancy for heart attack victims. The larger estimate assumes that improved medical care, for heart attacks and for other conditions, accounted for all the improvement in life expectancy.

As panel B of table 2 shows, allowing for the value of increased life expectancy for heart attack victims lowered the cost-of-living index by 1.1 to 1.7 percentage points per year. The range corresponds to the two alternative assumptions, noted above, about sources of increased life expectancy in the Medicare population.

Cutler, McClellan and Newhouse's cost-of-living index estimate has been controversial,

mainly because many people are reluctant to put a value on additional years of life. Though the ethical reasons for that reluctance are well known, Triplett (this volume) contended that existing statistical methods for measuring medical price indexes implicitly contain the same assumption. The very nature of the economic problem--measuring the cost of medical treatments and the value of medical outcomes--forces such valuation. For a similar view, see Gold et al. (1996).

We noted earlier that third-party discounts have been an issue in measuring medical care prices. The discount effect is important to analyses that formerly used the CPI in the absence of a PPI index for the medical care industry. Estimates of discounts show the difference between the CPI's out-of-pocket measure for consumers and what insurance companies and others were actually paying for medical services.[13]

Cutler, McClellan, and Newhouse used their hospital data to calculate the effect of discounts, within a framework that was otherwise comparable to the CPI. The results appear in panel C of table 2. As many medical economists suspected, discounts were rising substantially over the 1984–1991 interval. The effect on the price index was also substantial: 1.1 percentage points annually.

Whether discounts belong in the CPI is a somewhat complicated question, which we need not discuss comprehensively here. The BLS includes consumer-purchased medical insurance in the CPI, and for that the information on discounts to third-party payers is relevant. Thus, for the medical *insurance* part of the CPI, the total bias estimated by Cutler, McClellan, and Newhouse approximated 3.1 percentage points annually (2.0 percentage points on the out-of-pocket concept, plus the 1.1 percentage points from industry discounts—see table 2). But for the out-of-pocket medical expenditures portion of the CPI itself, discounts to third-party payers are not

relevant.[14]

In summary, an important implication of the Cutler, McClellan and Newhouse (this volume) research is that the cost of living for a person with a heart attack—the cost of restoring health—has risen less rapidly over time than the general rate of inflation. The NBER research team estimated the decline in the cost of living, relative to the overall inflation rate, at about 1 percent annually. That is in sharp contrast to a conventional price index for heart attacks, which Cutler, McClellan, and Newhouse also computed from their data for the teaching hospital. This conventional price index increased more rapidly than the overall rate of inflation (about three percentage points more, annually). Accounting for the benefits (improved outcomes) of changing medical treatments thus has a fundamental impact on our view of the magnitude of price increases for medical care.

An Exploratory Analysis of Price Indexes for the Treatment of Acute Depression

What has happened to the cost of treating an episode of care for a common mental illness such as acute depression? Richard G. Frank, Ernst R. Berndt, and Susan M. Busch (this volume) found that the cost of treating acute depression to guideline standards of care has fallen during the 1990s. The recent increase in mental health expenditures therefore reflects very substantial growth in the quantity of treatments for depression and not in the price of those treatments.

Treatments for major depression have advanced rapidly over the past twenty years. Innovative techniques in psychotherapy include interpersonal therapy, behavior therapy, family therapy, and cognitive behavior therapy. Each of those treatments has been shown to reduce depressive symptoms for less severe forms of major depression at comparable efficacy and with

similar outcomes. Innovative advances in antidepressant pharmaceuticals have been dramatic as well, especially since the introduction of the selective serotonin reuptake inhibitors (SSRIs) in 1988. The SSRIs are associated with significantly fewer side effects and are easier to take than the older generation of tricyclic antidepressants (TCAs). In treating major depression, doctors frequently combine psychotherapeutic interventions with antidepressant medications.

Using guidelines published by the Agency for Health Care Policy Research and the American Psychiatric Association, Frank, Berndt, and Busch identified nine major sets of "treatment bundles" for acute depression that employ various mixes of psychotherapy, antidepressant drugs, and medical management. They identified those bundles by using criteria based on data from clinical trials. Thus, the episodes of care the authors considered corresponded directly to treatments tested in the clinical trial literature. The researchers constructed medical price indexes based on the cost of treating an entire episode of acute depression in a manner that met published treatment guidelines.

A notable initial finding was that only 15 percent to 25 percent of episodes of depression treatment met guideline standards. The authors compute prices only for those treatment episodes meeting guideline standards. They noted that in subsequent research they will relax the "pure" definition of an episode of care, thereby pricing an episode of care that more realistically captures actual treatments.[15]

The researchers' review of the clinical trial and medical literature revealed the following conclusions regarding expected outcomes from those various treatment bundles:

- Compared with no treatment, psychotherapies of all kinds result in superior outcomes.
- Psychotherapies alone, TCAs with medical management, TCAs alone, and SSRIs alone appear

to produce comparable outcomes for less severe forms of major depression in terms of short-term symptom reduction.

- For more severe forms of major depression, TCAs alone, SSRIs alone, and combination of drugs and psychotherapy have comparable levels of efficacy.
- Compared with TCAs, SSRI use is associated with a higher rate of achieving recommended lengths of treatment.

To develop prices for treatment bundles, Frank, Berndt, and Busch relied on outpatient and pharmaceutical claims data from the Medicaid Statistical Reporting and Analysis System that were drawn from four large self-insured employers covering about a half-million people, annually from 1991 through 1995. Those claims data are associated with alternative treatment bundles meeting published guideline standards of care. The researchers found that prices received by providers differed markedly across treatment bundles. Even for therapeutically similar treatment bundles, for example, in 1993 prices ranged from $254 for short-term SSRI treatment alone, to $924 for short-term psychotherapy alone.

Alternative Price Indexes for the Treatment of Acute Depression. Prices of alternative treatment bundles for each year over the 1991–1995 time period provided the basis for constructing alternative price indexes for treating depression. Price measures for the treatment of acute major depression were computed in several different ways. One price measure, analogous to the PPI, is based on the total revenues received by the provider—the sum of the insurers' payments plus those of the patient. Another, analogous to the CPI, is based only on the direct payments made by the patient for treatment—copayments plus deductibles.

Frank, Berndt, and Busch computed price indexes under a variety of assumptions. In one case, they computed the price index based on the assumption that the treatment bundles could not be substituted for one another. That index conforms to the Laspeyres formula used by the BLS. In another, they considered various treatment bundles to be perfectly substitutable. Their final price index used an index number formula (the Tornqvist) endorsed by the CPI Commission (Boskin, et al, 1996); that index incorporates the changing value shares of the various bundles over time and made no a priori assumption concerning the substitutability of the various treatment bundles, so the index represents a "compromise" of the two extreme assumptions about substitutability.

No matter how they constructed the index, Frank, Berndt, and Busch found that the price of treating acute major depression fell by close to 30 percent over the 1991–1995 time period (price declines of roughly 7 percent per year). In the PPI-like index, the 1991–1995 price decline ranged from 22 percent to 30 percent.

Price declines of this magnitude imply that the recent increases in expenditures involve a substantial growth in the quantity of treatments and not in their prices. Price declines reflect the price concessions won by the rapidly growing managed-care sector, as well as the changing mix of treatment bundles—increasing the prescription pharmaceutical component and SSRI use and reducing the intensity of psychotherapy, which is by far the most expensive form of treatment.

Although no exact match exists among published PPI and CPI price indexes, the most closely related components of those indexes all rose over the period. For example, the inpatient treatment PPI rose about 10 percent and the medical care services CPI more than 25 percent over the same interval. The authors concluded:

Our results [suggest] that the [medical] CPI and PPIs may be particularly prone to

> distortion . . . where managed care has potentially large impacts on both input prices and the composition of treatment and important technical change in treatment methods has not occurred.

An Exploratory Study of Price Indexes for Cephalosporins

Sara Ellison and Judith Hellerstein (this volume) have produced price indexes for the cephalosporin class of antibiotics—a large and important subclass of antibiotics indicated for a wide range of infections—that are used in different clinical settings. They computed price indexes for branded and generic cephalosporins, distinguished separately cephalosporins sold across different channels of distribution, and compared those new price indexes with the cephalosporins component of the PPI. The researchers reported that their new price indexes rose less rapidly than the PPI indexes. Measured by the new price indexes, the average annual increase in cephalosporin prices was modest—only .76 percent, compared with a 4.54 percent average annual increase in the cephalosporins PPI.

The Ellison-Hellerstein study adds to the existing body of NBER research suggesting that a significant upward bias existed in the prescription drug PPI under the old BLS methodology. As we noted earlier, however, the BLS has recently made significant changes in both the PPI and CPI programs to correct for bias caused by not linking generics and by oversampling older products and undersampling newer prescription drugs (Kelly 1997). An important topic for future research will be the extent to which the new procedures employed by the BLS have eliminated the upward bias existing under the old methodology.

Although the slow growth in those cephalosporins price indexes indicates that antibiotics

have remained inexpensive, Ellison and Hellerstein noted that those price data mask one other very significant cost-saving feature of antibiotics. In the past, certain bacterial infections led to routine hospitalizations for some segments of the population (for example, very young children). To the extent that antibiotics now keep people out of the hospital when they contract a bacterial infection, the price of treating bacterial infections has become much lower, both because the direct cost is lower and because the indirect cost—the emotional and developmental cost of hospitalization of young children—is reduced. That extra cost saving from substituting outpatient treatment with antibiotics for hospital care is not reflected in the pharmaceutical price indexes computed by the authors. In a sense, pharmaceuticals are inputs to the production of medical care. Although it is extraordinarily important to get the inputs measured correctly, in the end we want to know the costs of what the pharmaceuticals do to health conditions and the contribution they make to improving health and to reducing the cost of health care.

The Economics of Antibiotics

and

Drug Patents and Prices

Two studies, in addition to addressing price measurement, considered the economic incentives for developing new pharmaceuticals and for generating adequate information about medical outcomes for new and for existing pharmaceuticals.

Sara Ellison and Judith Hellerstein (this volume) developed an economic model of the incentives for research on new antibiotics. Seventy years ago, bacterial infections such as pneumonia, tuberculosis, and typhoid fever were leading causes of death worldwide. Although

McKeown (1976) and his collaborators demonstrated that in industrialized countries the decline in the death rates from those diseases began long before the development of effective antibacterial pharmaceuticals, treatment with readily available and inexpensive antibiotics has now essentially eradicated typhoid fever in the developed world and reduced most cases of pneumonia and tuberculosis to readily curable conditions.

Ellison and Hellerstein pointed out that when an antibiotic is used successfully to combat an infectious disease, the benefit accrues not only to the person who takes the antibiotic but to other persons who, if the antibiotic were not used, might have contracted the contagious disease. Thus, individuals who do not have the disease have an interest in promoting more use of antibiotics to reduce their probabilities of contracting it. However, their interests are not factored into the decision to take the antibiotic. In the language of economics, this is an externality that results in too small a value being placed on the antibiotic and too little of it being used.

On the other hand, more use of an existing antibiotic increases the probability of bacterial resistance to it, which makes the drug less effective in the future. In that case, Ellison and Hellerstein contended that private decisionmakers will fail to make sufficient use of a diversity of antibiotics, which also reduces the demand for *new* antibiotics. Their study does not estimate the magnitudes of those two effects.

Thomas Croghan, Patricia Danzon, and Henry Grabowski found the model somewhat too rudimentary, at its current stage, to be fully convincing. Each had somewhat different reasons. Croghan (this volume) pointed out that the model was not realistic for medical reasons. Danzon (this volume) noted that the model did not consider adequately at its current stage the distinction between broad-spectrum antibiotics and narrower-spectrum ones; incentives to produce drugs

with excessively broad spectrums might be too great, so the model might have been misleading. Grabowski (this volume) added that the Ellison-Hellerstein model did not consider alternatives, such as vaccines to prevent disease. Incentives to develop preventive drugs (vaccines) may be too low for a variety of reasons he listed, including federal government action, and one cannot determine whether the incentives to produce antibiotics are nonoptimal without considering as well alternative medical courses of action against infectious diseases.

Thus, Crogan, Danzon, and Grabowski contended that the Ellison-Hellerstein model's results were very sensitive to special assumptions that the authors made. Though those criticisms were telling, formal economic models are always dependent on their assumptions, and the early stage of modeling on most economic problems frequently leaves substantial room for subsequent improvement. The Ellison-Hellerstein model was no different from others in that respect, and it is best viewed as an interesting start on a difficult economic problem.

Joel Hay and Winnie Yu (this volume) emphasized that a large proportion of our information about the effectiveness of pharmaceuticals is based on clinical trials sponsored by manufacturers of patented pharmaceuticals. That, they argued, creates a bias in incentives, because some kinds of pharmaceutical knowledge can be patented, whereas for others a patent cannot be obtained.

For example, if a new use is discovered for an old drug (an example might be the discovery that aspirin was an effective drug for treating heart attacks), one cannot obtain a patent for that discovery because the drug itself has long since passed the limits on patent protection. Thus, Hay and Yu contended, pharmaceutical firms have no incentive to develop new information about the usefulness of "old" drugs, which implies that private research resources are biased

toward the development of new (and patentable) drugs, even when society might be better off putting resources into the exploration of the properties of known chemical compounds. Hay and Yu presented quite a number of anecdotes to make their point. They proposed an expansion of use patents to change the present incentive structure.

While conceding the bias to private research incentives that Hay and Yu pointed out, Crogan, Danzon, and Grabowski raised an impressive number of objections to the desirability and feasibility of the Hay-Yu use-patents proposal. For example, both Croghan and Grabowski pointed out that use patents already exist, but that enforcing a patent on the use of a drug, rather than on the right to manufacturer it, requires producers to sue infringing physicians, HMOs, and patients. Grabowski wryly noted "the disinclination on the part of firms to sue their customers." Crogan, Danzon, and Grabowski also cited some undesirable incentive effects that the Hay-Yu use-patent proposal would create.

Additionally, defects in private-sector incentives to produce new knowledge on older drugs might be remedied by publicly funded research, rather than by trying to improve private-sector incentives to carry out research that is not profit-making. Hay and Yu, however, dismissed the latter possibility because, they contended, the National Institutes of Health are "limited by competing demands for medical research initiatives, the micro- and macropolitics of bioscientific R&D resource allocation decisions across diseases and therapies, the dogma of prevailing scientific paradigms, the prejudices and political correctness of academic experts, and the red tape of government bureaucracy."

The use-patent proposal is not, in itself, directly relevant to the topic of measuring the prices of medical treatments.

Cost, Effects, Outcomes and Utility

and

Integrating Price Index and Cost-Effectiveness Research

Two studies dealt with the broad view—how can one combine research on measuring medical care into an overall economic picture of medical inflation, the quantity and quality of medical treatments, and the consumption of medical care?

Mark Pauly (this volume) proposed an alternative to the approach taken in the empirical studies in this volume. Rather than measure the price of medical treatments with appropriate quality adjustments, he proposed to measure the change in willingness to pay for insured medical services:

> My suggestion is that the object of valuation be defined as the difference between the willingness to pay premiums for a managed care plan covering the technology available in the preceding period and the willingness to pay for the same plan covering the technology available in the current period.

As Pauly noted, pricing medical care by pricing insurance plans was proposed at least thirty years ago (Reder 1969). What is new, Pauly contended, is the development of willingness-to-pay techniques in economics. Now, one could ask respondents to put a valuation on an insurance policy that covered some new technique, or a bundle of new techniques, compared with an insurance policy that did not cover those techniques.

One advantage of pricing insurance policies is that such an approach captures behavior toward risk in a way that is perhaps neglected in studies that only address the cost of treating a disease. If one has a disease, and if there is no insurance, the cost of treating the disease matters.

If one does not have the disease, then insuring against the risk of a costly medical bill if the disease is contracted is important.

Pauly pointed out that in the United States insurance covers some 80 percent of medical expenses. For the insured population, the cost of treating a disease is not quite what matters, in the sense of computing, say, a cost-of-living index for medical care. Instead, Pauly contended, what matters in a cost-of-living index is the cost of medical insurance.

Pauly buttressed his case for pricing insurance by discussing many of the problems that arise in constructing quality-adjusted life years [QALY], which is an essential component of the alternative approach proposed by Jack Triplett (this volume). He concluded, "The advantage of this method [QALY] then is not its credibility."

There is little question that Pauly's list of problems with measuring QALY are legitimate concerns. Indeed, one can find similar discussions in the QALY literature itself. For example, the report of Gold et al. (1996), in which QALY was proposed as the standard medical outcome measure for cost-effectiveness studies, contains an extensive discussion of the primitive stage of development of the QALY measure and the problems that have yet to be resolved.

In addition, Pauly made a new point about the use of QALY in price indexes. For most cost-effectiveness studies—the purpose for which QALYs are usually computed—finding a precise value for QALY is not required. All that is necessary, in most cases, is to determine whether cost per QALY is below some threshold number or to compare cost per QALY for two or more alternative treatments. For those purposes, a precise measure of QALY is not necessary, because there are few "close calls" for which real precision matters.

In the case of price indexes, however, using QALY as a measure of medical outcome to

make an adjustment for changes in medical technology requires, in principle, a precise measure. Even when more QALY estimates become available, they may not be measured precisely enough to meet the needs of economic statistics.

> Pauly concluded that a mix of pricing insurance and use of QALY may be appropriate:
>
> For analysis of expenditure change for populations heavily subsidized by public programs for equity reasons—Medicaid and the part of Medicare spending that goes to low income elderly—it may be best to use the monetized QALYs approach, precisely because the monetization, as well as the entire program, is really an object of collective choice. There is little point, say, in adjusting Medicaid spending for the value Medicaid beneficiaries place on improved quality, even if we knew it.

Jack Triplett's study (Triplett, this volume) focuses on an augmented national health account that would better accommodate research on medical prices and medical outcomes. He proposes integrating the present national health accounts constructed by HCFA (Levit et al. 1996) with the "cost of disease" accounts originally constructed by Rice (1966) and recently implemented in expanded form by Hodgson and Cohen (1998). Cost-of-disease accounts already disaggregate total expenditures on medical care by disease classifications (the International Classification of Diseases, 9th revision, or ICD-9). For example, ICD-9, chapter 7 covers circulatory diseases, including heart attacks—the focus of the Cutler, McClellan, and Newhouse study. Triplett remarked that organizing health accounts by disease classifications (rather than, as now, by source and recipient of funds) is analogous to the "product" side of conventional national accounts, because diseases are what funds are spent on.

Estimating real output of medical care in Triplett's augmented national health account system could proceed by using the new PPI medical care indexes that are already arranged by Diagnostic Related Groups (DRG's). DRG's are consistent, for the most part, with ICD-9 classifications. Triplett also contended that it would be natural and straightforward to build into the deflation process new research on price indexes by disease, as it became available. For example, the price indexes for heart attacks and for mental depression contained in this volume, and the Shapiro and Wilcox (1996) study of cataract surgery, could be substituted for PPI indexes, or they could be used to supplement them. New research could be integrated piece by piece as new studies become available.

Additionally, the burgeoning growth of cost-effectiveness studies provides information about medical outcomes that could be also used in a systematic way, if the national health accounts were organized by ICD-9 classifications, rather than, as now, by sources of funding and recipients of funds. Triplett noted that the approach outlined in his study could be implemented in countries where no private markets for medical care or for medical insurance exist, which is a great advantage for future international comparisons of medical costs and real consumption of medical care.

To execute such an expanded health account system, one needs to understand how price indexes for medical care and cost-effectiveness studies fit together. Triplett's study explains the relationships between those two bodies of research. This technical demonstration only matters for practical purposes when and if an accounting for health care is established that can make use of both new price index studies and new cost-effectiveness studies as they become available.

Neither Pauly nor Triplett proposed one approach to the exclusion of the other. Indeed,

the real problem is that both approaches have difficulties, and too little information is available for implementing either approach.

On the Pauly insurance proposal, Weisbrod (this volume) noted:

> There does not seem to be available today an insurance contract that provides yesterday's technology. . . . One of the interesting questions is why there is no such market, and what does that tell us about people's preferences and notion of quality?

With respect to the use of information from cost-effectiveness analysis, Hay and Yu (this volume) maintained that not only is the information currently available insufficient, but too few incentives exist to produce more of it. Pauly, and also Manning (this volume) and Meltzer (this volume), pointed out inherent problems with QALY and some of the difficulties in using QALY to construct better price indexes.

Probably the best way to look at this debate is to say that there are two glasses of water, both of which are nearly empty. If either one were close to full (with empirical studies), it could be used effectively to improve measures of medical care. Pauly found conceptual reasons for preferring that the glass marked "measuring insurance policies" be the one into which the refreshing water of empirical research be poured. Triplett pointed out that the springs of new research are now filling the other glass. Even if there is now the sediment that others claim to see in the glass (because of problems with QALY and so forth), he expected the filters to be improved as the glass is filled (methods will be improved as more research is done). He therefore opts to build a conceptual framework that will make use of the studies that are accumulating, rather than to wait for researchers to implement the price-of-insurance approach.

IV. Conclusion

An encouraging number of new developments in measuring medical care have occurred very recently. Most of them are not yet well known. Together, they are changing our perceptions of medical care inflation, and of the quantity of medical care services that are provided by U.S. expenditure on medical care.

Statistical agencies have greatly improved government price indexes for medical care-- for medical care providers, such as hospitals and physicians, and also for pharmaceuticals. Some of these PPI and CPI improvements are the direct result of previous research into new methods for medical care price index measurement, particularly by a group of NBER researchers.

The papers in this volume also present promising new developments in measuring medical care. New methodologies and new empirical estimates for selected medical care disease categories suggest that improvements in medical care are greater than the usual statistics suggest, and that medical care price inflation is lower. The papers and discussion also offer insights into future directions for research.

It is also true that the contributions in this volume suggest how much is yet to be done. Only a small number of medical procedures have been studied intensively. One cannot infer that price indexes for heart attacks, for depression, and for antibiotics are representative of what researchers will find when studies are completed for other medical procedures. Even though much that is promising has occurred very recently, it is still too soon to see how the full economic picture of medical care will look when more of the pieces have been studied.

But the good news is: Progress is being made. That is the story of this volume.

References

Becker, Gary S. 1960. An economic analysis of fertility. In *Demographic and Economic Change in Developed Countries*. New York: National Bureau of Economic Research.

Berndt, Ernst R., Zvi Griliches, and Joshua G. Rosett. 1993. Auditing the Producer Price Index: Micro evidence from prescription pharmaceutical preparations. In *Journal of Business and Economic Statistics* 11(3)(July): 251-264.

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgensen. December 4, 1996. Toward a more accurate measure of the cost of living: Final report to the Senate Finance Committee from the Advisory Commission to study the consumer price index.

Croghan, Thomas W. Forthcoming. Improving incentives for pharmaceutical innovation. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Cutler, David, Mark McClellan, and Joseph Newhouse. Forthcoming. The costs and benefits of intensive treatment for cardiovascular disease. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Danzon, Patricia M. Forthcoming. Comments on Ellison and Hellerstein and Hay and Yu. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Eisenberg, John M. Forthcoming. Types and uses of outcome research: The view from a federal agency. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Ellison, Sara Fisher, and Judith Hellerstein. Forthcoming. The economics of antibiotics: An exploratory study. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Fixler, Dennis. Forthcoming. Discussion of "The Cost and Benefits of Intensive Treatment for Cardiovascular Disease" by Cutler et. al. and "Price Indexes for the Treatment of Depression" by Frank et. al. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Frank, Richard G., Ernst R. Berndt, and Susan Busch. Forthcoming. Price indexes for the treatment of depression. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

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

Grabowski, Henry. Forthcoming. Improving the incentives for pharmaceutical R&D: Comments on the Ellison-Hellerstein and Hay-Yu papers. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Griliches, Zvi and Iain Cockburn. 1994. Generics and new goods in pharmaceutical price indexes. In *American Economic Review* 84(5) (December): 1213-1232.

Hay, Joel W. and Winnie Yu. Forthcoming. Drug patents and prices: Can we achieve better outcomes? In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Heidenreich, Paul and Mark McClellan. 1998. Trends in heart attack treatment and outcomes, 1975-95: Literature review and synthesis. Manuscript.

Hodgson, Thomas A. and Alan J. Cohen. 1998. Medical care expenditures for major diseases. Manuscript. Hyattsville, M.D.: National Center for Health Statistics.

Kanoza, Douglas. 1996. Supplemental sampling in the PPI pharmaceuticals index. In *Producer Price Indexes Detailed Price Report* (January): 8-10.

Kelly, Gregory G. 1997. Improving the PPI samples for prescription pharmaceuticals. In *Monthly Labor Review* 120(10) (October): 10-17.

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.

Manning, Willard G., Jr. Forthcoming. Discussion. U.S. Department of Health and Human Services, National Center for Health Statistics. 1995. *Vital and Health Statistics: National Hospital Discharge Survey: Annual Summary, 1993* Ser. 13(121) (August): 62.

Meltzer, David. Forthcoming. Discussion. U.S. Department of Health and Human Services, National Center for Health Statistics. 1995. *Vital and Health Statistics: National Hospital Discharge Survey: Annual Summary, 1993* Ser. 13(121) (August): 62.

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

Newhouse, Joseph P. 1989. Measuring medical prices and understanding their effects--the Baxter Prize address. *Journal of Health Administration Education* 7(1): 19-26.

Pauly, Mark V. Forthcoming. Costs, effects, outcomes and utility: Concepts and usefulness in medical care price indexes. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Phelps, Charles E. Forthcoming. Comments on price index papers. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

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*. Washington, DC: US Government Printing Office.

Ristow, William. 1996. IMS presentation to the BLS. Mimeo. IMS America, Plymouth Meeting, PA (July).

Shapiro, Matthew P., and David W. Wilcox. 1996. Mismeasurement in the consumer price index: An evaluation. In *NBER Macroeconomics Annual* 11: 93-142.

Triplett, Jack E. Forthcoming. Accounting for health care: Integrating price index and cost-effectiveness research. In Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

U.S. Department of Commerce, Bureau of Economic Analysis. 1985. *Introduction to National*

Economic Accounting. Methodology Paper Series MP-1. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Commerce, Bureau of Economic Analysis. 1990. Improving the quality of economic statistics. *Survey of Current Business* 70(2) (February): 2.

U.S. 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.

U.S. Department of Health and Human Services, National Center for Health Statistics. 1995. *Vital and Health Statistics: National Hospital Discharge Survey: Annual Summary, 1993* Ser. 13(121) (August): 62.

U.S. Department of Health and Human Services, National Center for Health Statistics. 1997. *Monthly Vital Statistics Report* 45(11)(Supp. 2)(June 12).

U.S. Department of Labor, Bureau of Labor Statistics. Not dated. *Improvements to CPI Procedures: Prescription Drugs.* Washington, D.C.

Weisbrod, Burton A. Forthcoming. Measuring health care prices: What research is needed? In

Jack E. Triplett, ed., *Measuring the Prices of Medical Treatment*. Washington, D.C.: Brookings.

Table 1:

Changes in US Heart Attach Mortality, 1975-1995

	1975	1995	change	% change
1a) Total heart attack deaths (thousands)	325	218	-107	-32.9
1b) Heart attack deaths per 100,000 population	152	83	-69	-45.5
1c) Heart attack deaths per 100,000 population (age-adjusted)	107	44	-64	-59.2
2) In-hospital heart attack deaths, by first-listed diagnosis (thousands)	75	76	+1	+1.3
3a) Hospital discharges, heart attacks, first-listed, adjusted for coding change (thousands)	555	771	+216	+38.9
3b) Hospital discharges, heart attacks, adjusted for transfers, etc. (thousands)	542	540	-2	-0.5
4a) In-hospital fatality rate (% of first-listed heart attacks, adjusted for coding change)	13.5%	9.9%	-3.6	-26.7
4b) In-hospital case fatality rate, age-sex adjusted	23%	14%	-9	-39
4c) 30-day fatality rate	27%	17%	-10	-36

Notes:

1a) The 1975 number was calculated from the rate of heart attacks (.1524%), and the population (213,032,000). Source: U.S. Department of Health Education and Welfare, National Center for Health Statistics, 1979, *Vital Statistics of the United States, 1975: Volume II--Mortality, Part A*, Hyattsville, Maryland, p. 1-7, Table 1-7 and p. 6-23, Table 6-2.
(Table 1-7: Death Rates for 69 Selected Causes: United States, 1968-75; Table 6-2: Estimates of total resident population of the United States, by age, race, and sex: July 1, 1975).
Note: According to the table, the population number was retrieved from the Bureau of the Census in *Current Population Reports*, Series P-25, No. 643. This number does not correspond to updated corrections later reported by the Census Bureau.

The 1995 number comes directly from U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Center for Health Statistics. 1997. *Monthly Vital Statistics Report* 46(1) (Sept. 11): 24, Table 11: Deaths, death rates, and age-adjusted death rates for 72 selected causes, Human immunodeficiency virus infection, Alzheimer's disease, Injury by firearms, Drug-induced deaths, Alcohol-induced deaths, and Injury at work: United States, final 1995 and preliminary 1996.

1b) 1975 data source: U.S. Department of Health, Education and Welfare, National Center for Health Statistics. 1979. *Vital Statistics of the United States, 1975: Volume II--Mortality, Part A.* Hyattsville, MD: National Center for Health Statistics.

1995 data source: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Center for Health Statistics. 1997. *Monthly Vital Statistics Report* 45(11) (Supp. 2) (June 12): 20, table 6.

1c) 1975 data was age-adjusted by using the method used by the National Center for Health Statistics (U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Center for Health Statistics, *Technical Appendix from Vital Statistics of the United States: 1994, Mortality*. Hyattsville, MD: National Center for Health Statistics, 1994, http://www.cdc.gov/nchswww/data/techap94.pdf, pp. 28, 32.),

$$R'' = \Sigma w_i R_i$$

where R" is the age-adjusted rate, w_i is the population weight (provided in the *Technical Appendix*), and R_i is the age-specific rate for the i^{th} age group. The weights are based on the age structure of the U.S. population in 1940. The NCHS advised that the comparison between data before 1979 and after is not completely compatible, since the ICD-9-CM replaced the ICDA/ICD-8 in 1979, and coding changes may have occurred.

1995 data source: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Center for Health Statistics. 1997. *Monthly Vital Statistics Report* 45(11) (Supp. 2) (June 12): 20, table 6.

2) The number of in-hospital AMI deaths recorded by NCHS.

1975 source: U.S. Department of Health and Human Services, National Center for Health Statistics, 1978, *Vital and Health Statistics: Inpatient Utilization of Short-Stay Hospitals by Diagnosis: United States, 1975*, Ser. 13 (35): Table R: Number of deaths in short-stay hospitals, and hospital fatality rates, excluding newborn infants, by age and first-listed diagnostic conditions: United States, 1975.

1995 source: U.S. Department of Health and Human Services, National Center for Health Statistics, 1998, *Vital and Health Statistics: National Hospital Discharge Survey: Annual Summary, 1995*, Ser. 13(133): 13, Table L: Number of deaths and fatality rate for discharges from short-stay hospitals, by age and selected first-listed diagnoses: United States, 1995.

3a) In 1982, a coding change affecting first-listed heart attack diagnoses became effective. The 1981 number of first-listed heart attacks is corrected for this coding change, as follows (data come from various issues of NCHS *Vital and Health Statistic* Series 13):

First, the average annual percentage changes of first-listed diagnoses (2.97 percent) and of all-listed diagnoses (1.57 percent) were calculated for 1975 through 1981. There is a steady upward drift, of first-listed relative to all-listed heart attack discharges, of a little over one percentage point per year. On the assumption that the same drift would be present in consistently coded series for 1981 and 1982, the difference in these two rates (1.38 percentage points) was added to the 1981-82 percent change for *all-listed* diagnoses; the resulting percentage change became the *corrected* percent change in *first-listed* diagnoses between 1981 and 1982. The adjusted time series of first-listed diagnoses was calculated for each preceding year (1975-81) by

applying to the adjusted 1981 number the percentage changes in first-listed diagnoses from the published NCHS data.

1975 source used for calculation: U.S. Department of Health and Human Services, 1978 (op. cit.), pp. 26, 50; tables 1 and 5.

1995 source: U.S. Department of Health and Human Services, 1998 (op. cit.), p. 5, table B.

3b) Adjusted for transfers and readmissions.

Source: Heidenreich, Paul, and Mark McClellan, 1998, "Trends in Technology Use for Acute Myocardial Infarction," manuscript, table 1.

4a) The 1975 and 1995 data are calculated by dividing line 2) by line 3a).

4b) Adjusted for age and gender (1975 levels), in Heidenreich and McClellan 1998, table 1.

4c) Based on an exponential decline in daily mortality from day 7 to day 30 in, Heidenreich and McClellan 1998, table 1.

Table 2:

Sources of CPI Bias,

Based on Estimates for Heart Attack Treatment

Sources	Percentage points (annual rate of change)
A. CPI Concept: Out-of-Pocket, Money Expenditure Only	
1) Didn't change weights (1984 weights index less annual chain weights index)[1]	1.7
2) Didn't allow for quality change (treatment price index less annual chain weights index)[2]	0.3
3) Total bias, on own concept[3]	2.0
B. COL Concept	
1) Didn't allow for valuation of increased life expectancy	1.7
C. "Industry" Concept	
1) Didn't get discounts to third party pagers	1.1

Source Notes (all entries based on Cutler, McClellan and Newhouse, this volume, Table 10):

Line A1 Synthetic CPI for MTH-Costs, less Annually Rebased Price Index

Line A2 Annually Rebased Price Index, less Treatment Regimen Price Index (MTH data)

Line A3 Sum, lines A(1) and A(2)

Line B1 Cost of Living Index, less Treatment Regimen Price Index (Medicare data)

Line C1 Synthetic CPI for MTH-Costs, less Synthetic CPI for MTH - changes

Endnotes

[1] The price index is formed from a ratio that takes as the numerator the hypothetical expenditure level that is the answer to this question, and as the denominator the actual expenditures in the base year. The price index is usually expressed with the base year set equal to 100. In concept, the CPI is a fixed weight approximation to a cost of living index, where the cost of living index measures the cost in the current period of the standard of living achieved in the base period.

[2] Employer-provided medical insurance was outside the scope of the CPI definition of consumption.

[3] For an example of that criticism, see Newhouse (1989).

[4] "The QALY (quality-adjusted life years) is a measure of health outcome which assigns to each period of time a weight, ranging from 0 to 1, corresponding to the quality of life during that period, where a weight of 1 corresponds to perfect health and a weight of 0 corresponds to a health state judged equivalent to death. The number of quality-adjusted life years, then, represents the number of healthy years of life that are valued equivalently to the actual health outcome." Gold, et. al. (1996). Page 29. QALY estimates are being employed in cost-effectiveness studies on medical interventions.

[5] The published NCHS rates are given in table 1. Regrettably, the NCHS age-adjusted rates use as weights the age structure of the U.S. in 1940. In 1940, ages 65 and over accounted for only about 7 percent of the population, so the weights used by NHCS are not close to either those of 1975 or 1995. The magnitudes of the declines reported in table 1 are so large that reweighting is unlikely to affect the major conclusion.

[6] NCHS *Monthly Vital Statistics Report*, June 12, 1997, table 8. In 1995, the numbers were 481 thousand and 218 thousand, respectively, for ischemic heart disease and heart attack deaths, giving a ratio of 45 percent. Ischemic heart diseases are those that are caused by or which cause obstruction of the blood supply to the heart. Deaths from old heart attacks contribute a portion deaths classified among non-heart attack, ischemic heart disease deaths.

[7] The unadjusted numbers, calculated from the data in table 1, are: 250 thousand in 1975 and 142 thousand in 1995. Out-of-hospital deaths from heart attacks are not published separately on an age-adjusted basis.

[8] Heidenreich and McClellan (1998) review evidence on changes in population risk factors.

[9] Though it seems odd, there are apparently no reliable national data on the number of heart attacks that are reported to physicians and hospitals. Some heart attacks are not noticed by either victims or physicians. A recent and highly publicized case was the prominent manager of the Los Angeles Dodgers baseball team, whose multiple heart attacks were discovered well after the events. Heidenreich and McClellan (1998) review evidence from longitudinal panels that have been constructed to analyze heart attacks and other diseases, but the statistical design for a longitudinal survey is not necessarily intended to yield good population estimates.

[10] The coding change, which raised the proportion of heart attack discharges that were coded as first listed, was intended to correct improper coding (when a heart attack victim is hospitalized, the heart attack is seldom a secondary consideration). No consistent time series of heart attack hospitalizations is published by NCHS. From piecing together annual reports, it is clear that the entire change in the relation between first-listed and all-listed series took place in 1982, in which year the first listed series increased by 230 thousand (fifty percent), the all-listed series by only 13 thousand.

[11] The derivation of the adjusted estimate is explained in the footnote to line 3a of table 1.

[12] Hunink et al. (1997) reached similar conclusions for the decline in the death rate from coronary heart disease (which is, of course, broader than heart attacks).

[13] Employer-provided medical insurance has always been outside the scope of the CPI definition of consumption.

[14] One could also argue that the full cost of living index, to which the CPI is an approximation, ought to include in its consumption measure the medical services purchased by employer-provided medical insurance. This debate takes us too far into the details of price index construction.

[15] Subsequent research, not reported here, determined that the price indexes were in fact sensitive to whether of not the patient received full guideline level of care.