

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Medical Care Output and Productivity

Volume Author/Editor: David M. Cutler and Ernst R. Berndt, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-13226-9

Volume URL: <http://www.nber.org/books/cutl01-1>

Publication Date: January 2001

Chapter Title: Introduction to "Medical Care Output and Productivity"

Chapter Author: David M. Cutler, Ernst R. Berndt

Chapter URL: <http://www.nber.org/chapters/c7623>

Chapter pages in book: (p. 1 - 12)

---

# Introduction

David M. Cutler and Ernst R. Berndt

---

Measuring the output of the medical sector has been a long-standing policy concern. With the United States and other developed countries spending so much on medical care (7 to 14 percent of GDP), health care analysts invariably ask what one obtains for all these expenditures. Answering this question requires measuring the output of the medical care sector. National income accountants face the same problem. National income accounts divide nominal spending growth into changes in prices and changes in quantities. But neither prices nor quantities can be estimated without an accurate measure of the medical care industry output.

At the conceptual level, an appropriate measure of medical care output is clear—the health gains resulting from medical care. If increased medical spending leads to health improvements worth more than their cost, then medical care productivity is increasing. If spending increases lead to less valuable increases in health, medical care productivity is falling.

The difficulty rests in implementing this framework. Measuring health status is difficult, and attributing changes in health status to a particular factor such as medical care is even more challenging. Thus, productivity estimation for medical care has necessarily been infrequent and tentative.

Substantial research has recently been directed at this issue, however. The “outcomes movement” in health economics is an attempt to measure the health impact of medical care. Research conducted at and supported by public agencies such as the Bureau of Labor Statistics (BLS), the Bu-

David M. Cutler is professor of economics at Harvard University and a research associate of the National Bureau of Economic Research. Ernst R. Berndt is professor of applied economics at the Sloan School of Management of the Massachusetts Institute of Technology and director of the NBER's Program on Technological Progress and Productivity Measurement.

reau of Economic Analysis (BEA), and the National Institute on Aging, as well as by private pharmaceutical firms such as Eli Lilly and Company, have investigated the pricing of medical services. In 1996, the National Bureau of Economic Research, drawing together many of its own research efforts in this field, organized a study team on the Economics of Medical Treatments.

After several years of research, and in consultation with the Bureau of Labor Statistics, the Bureau of Economic Analysis, the National Institute on Aging, and Eli Lilly and Company, it was decided to take stock of where we are. This volume is the result of that assessment.

The chapters in this volume cover a very wide range, from theoretical discussions of how medical care productivity ought to be measured, to empirical analyses of prices and productivity for different illnesses and medical conditions, to nonmedical factors that affect health status and mortality. The chapters in this volume make a substantial contribution to the measurement of medical care prices, output, and productivity. Together, they illustrate where we are now, what can be done in the near future, and where price index measurement in this vital field can move feasibly and fruitfully in years to come.

This volume is organized around four thematic areas, which we now briefly summarize.

### **Conceptual Issues in Medical Care Prices and Productivity**

The first topic is the methodology for measuring productivity in medical care. Medical care is not the only sector of the economy where productivity measurement is difficult, of course. Productivity measurement is difficult in many of the service sectors. The first section of the volume compares and contrasts the theoretical issues in productivity measurement in medical care with those in other sectors.

In his chapter "What's Different about Health? Human Repair and Car Repair in National Accounts and in National Health Accounts," Jack E. Triplett argues that health care is not in fact wholly different from other services, such as car repair. Both markets, for example, have characteristics of asymmetric information and moral hazard.

Triplett takes as a starting point in the human repair framework that the focal point of expenditures is that for treating diseases and conditions, and not, for example, expenditures for hospitals. This implies that national health accounts must be organized and integrated into cost-of-disease accounts. Moreover, instead of looking just at market transactions, for human repair one must consider what medical resources actually do for health. This means assessing the "production function" for health status outcomes. The integration of economic and outcomes research distinguishes, at least in degree, human repair from car repair accounting.

In his chapter “Theoretical Foundations of Medical Cost-Effectiveness Analysis: Implications for the Measurement of Benefits and Costs of Medical Interventions,” David Meltzer explores the theoretical foundations underlying measures of medical care outcomes. The difficulty in measuring the costs and benefits of medical treatments is one of the most salient differences between medical care and other industries. Meltzer notes two key dimensions along which the literature has had substantial debate: how to measure the quality of life for individuals with particular health states, and how cost-effectiveness analysis should deal with costs occurring in the future, after the initial medical treatment has been provided.

On the first point, Meltzer argues that ideally quality of life should be measured with revealed preference analysis—using the choices that individuals actually make between different treatment options to infer their value of different health states. An example is treatment for diabetes. In comparing therapies with different levels of intensity and long-run quality of life gains, Meltzer argues that the choices that diabetic patients make are a reasonable guide to the underlying utility they receive. He shows how this insight can be turned into an estimate of quality of life useful for outcome valuation purposes.

On the second point, Meltzer argues that productivity measurement in medical care must consider the future costs resulting from medical treatments provided today. If medical intervention lengthens an individual’s life by ten years, for example, Meltzer shows that the costs of sustaining that person for the ten additional years must be attributed to the intervention, just as the benefits of the ten years of additional life are. Meltzer shows that treating future costs and benefits consistently can have a large effect on our perception of the relative ranking of different interventions.

Another way in which health care differs from many (but not all) other service industries is that most production takes place in nonprofit organizations. Yet relatively little is known about the economic forces affecting productivity in this industry. In their chapter “Medical Care Output and Productivity in the Nonprofit Sector,” Tomas Philipson and Darius Lakdawalla specify medical service provider objective functions that allow for productivity differences between for-profit and nonprofit producers to be interpreted as evidence for differences in preferences, rather than in costs. Their model predicts that when both coexist, nonprofits are larger and less efficient in input use (“input preferrers”), but nevertheless become more numerous than for-profit firms, even though the nonprofits exhibit higher average and marginal costs. The nonprofits thrive because of their dependence on donations and tax advantages. The authors interpret empirical evidence from the nursing home industry as supportive of their framework, although discussant Richard G. Frank interprets this evidence as being more ambiguous.

This thematic section on conceptual issues ends with a chapter by Ernst

Berndt, David Cutler, Richard Frank, the late Zvi Griliches, Joseph Newhouse, and Jack Triplett, "Price Indexes for Medical Care Goods and Services: An Overview of Measurement Issues," summarizing the NBER methodology to price and productivity measurement. The Berndt et al. chapter has two goals. First, it discusses the conceptual and measurement issues that underlie construction of medical care price indexes in the United States, particularly the medical care consumer price indexes (MCPIs) and medical-related producer price indexes (MPPIs). The problems in price measurement are multiple: Insurance makes the cost of medical care to consumers lower than its cost to society; physicians may not do exactly what well-informed patients would want; technological progress makes the basket of goods different over time; and organizational changes such as managed care change the basis of pricing in the market. Because of these various factors, market outcomes are not likely to be efficient, and therefore market transactions data cannot be relied upon to reveal marginal valuations. This implies that productivity measurement must go beyond the method of hedonic analysis that is now standard in other markets or industries.

The second part of the Berndt et al. chapter describes procedures currently used by the BLS in constructing MCPIs and MPPIs, including recent revisions, and then considers alternative notions of medical care output. Whereas historical BLS methods relied on pricing all of the detailed services that a patient used, the authors argue instead that the focus should be on more aggregated pricing of an episode of treatment. The authors outline features of a proposed new experimental price index—a medical care expenditure price index—that is more suitable for evaluation and analyses of medical care cost changes than are the current MCPIs and MPPIs. They propose that, in addition to the MCPIs and MPPIs, some federal statistical agency publish an experimental medical care expenditure price index.

### **Current State of Measurement**

In the second section of the volume, the current state of official government measurement regarding price and productivity measurement in medical care is summarized and reviewed. The first two chapters discuss the current BLS methodology for the consumer and producer price indexes in medical care, respectively.

In "Medical Care in the Consumer Price Index," Ina Kay Ford and Daniel H. Ginsburg provide details and discuss implications of changes introduced into the January 1998 major revision of the Consumer Price Index (CPI). Item weights are of course changed in this revision. Because the scope of the CPI is limited to consumers' direct out-of-pocket payments plus the employee-paid share of employer-provided health insur-

ance, the expenditure share of medical care in the CPI (7.4 percent in 1995) is much less than its share in the personal consumption expenditure component of GDP (17.9 percent). The Laspeyres fixed quantity basis of the CPI also resulted in the 1997 relative importance weight for medical care of about 7.5 percent being considerably larger than the new weight of 5.6 percent embodied in the 1998 revision.

Other CPI revisions introduced in 1998 include rotating the sample of categories of commodities rather than subsets of the pricing areas, which permits more frequent updating of selected item strata. To capture possible inpatient-outpatient substitution, the new hospital CPI incorporates hospital room expenses, charges for other inpatient services, and the cost of outpatient services all under one umbrella, rather than as distinct item strata. The authors note continuing conceptual and implementation challenges in pricing medical insurance, and conclude with a discussion of input versus outcomes approaches to the choice of what it is that is being priced by the medical CPI.

For quite some time now, the BLS has expanded the scope of its producer price indexes (PPIs) to include service sector industries, not just goods producing sectors. In their chapter, "Health Care Output and Prices in the Producer Price Index," BLS economists Dennis Fixler and Mitchell Ginsburg provide a very detailed and thorough description of how PPIs are now being constructed for the main industries composing the health care sector of the economy. By linking price quotes to patient diagnoses and procedural codes, these health care PPIs now attempt to price much more closely episodes of typical treatments, rather than just pricing discrete medical inputs such as hospital days. These developments in the PPI are very significant. A major continuing difficulty, however, is that the PPI program has as its organizing structure the Standard Industrial Classification (SIC) system, and many medical procedures combine outputs from several SIC industries. For example, the treatment of outpatient mental health combines inputs from psychiatrists and pharmaceuticals, the latter being classified within the SIC manufacturing sector.

The third paper in this section examines one of the long-standing issues in medical care measurement—the differences between different federal health care accounts in the United States. Both the Health Care Financing Administration and the Bureau of Economic Analysis produce and publish a set of national health accounts, where they track medical spending in the economy as a whole. But the two accounts differ by a considerable amount—3 to 4 percent in 1996, for example. With multiple measures of medical care spending, it has been difficult for users to obtain an accurate handle on the scope of the medical sector. Arthur Sensenig and Ernest Wilcox report on an attempted reconciliation between these two accounts in their chapter "National Health Accounts/National Income and Product Accounts Reconciliation." Sensenig and Wilcox highlight both conceptual

and data differences between the two accounts. Conceptually, the two accounts differ in their use of revenues versus expenditures. The data used also differ, depending on whether they are benchmarked to other industry surveys. In a changing industry such as medical care, where revenues and costs may differ considerably (as they do, for example, in not-for-profit hospitals), these differences can be quite significant. Tracking these differences might therefore have a substantial impact on output and productivity estimates in the industry.

### **Recent Developments**

The third section of this volume presents new empirical analyses of price, output, and productivity measurement in medical care. The authors of these papers generally focus on treatment of a particular condition, such as heart attacks or acute phase depression. The reason for this narrow focus was detailed in the earlier conceptual section of the volume: Because productivity analysis requires direct measurement of health outcomes, it needs to be done at the level of a specific disease or illness.

In their chapter “Pricing Heart Attack Treatments,” David Cutler, Mark McClellan, Joseph Newhouse, and Dahlia Remler estimate price indexes for heart attack treatments, demonstrating the techniques that are currently used in official price indexes and presenting several alternatives. Cutler et al. consider two types of price indexes: a service price index, which prices specific treatments provided, and a cost-of-living index, which prices the health outcomes of patients. Both indexes are complicated by price measurement issues. For example, list prices and transactions prices are fundamentally different in the medical care marketplace. The development of new or modified medical treatments further complicates the comparison of “like” goods over time. Furthermore, the cost-of-living index is hampered by the need to determine how much of health improvement results from medical treatments in comparison to other factors, such as lifestyle changes.

Cutler et al. describe methods to address each of these obstacles. They employ national data on treatments and outcomes for Medicare beneficiaries who have had a heart attack to measure the cost of living for medical care. Cutler et al. conclude that, while traditional price indexes when applied to heart attack treatments are rising at roughly 3 percent per year above general inflation, a corrected service price index is rising at perhaps 1 to 2 percent per year above general inflation, and the cost-of-living index is falling by 1 to 2 percent per year relative to general inflation.

An alternative method for measuring the impact of medical care involves use of clinical trial evidence on medical treatments. Meta-analyses based on clinical trials can be used to infer the impact of medical treatments on outcomes. The more medically oriented chapter by Paul Heiden-

reich and Mark McClellan, “Trends in Heart Attack Treatment and Outcomes, 1975–1995,” demonstrates this methodology. Heidenreich and McClellan summarize the voluminous medical literature on the efficacy of treatment for myocardial infarction (heart attacks). They conclude that over the 1975–1995 period, medical treatments played a large role in improving heart attack survival rates. Specifically, they conclude that about two thirds of increased survival resulted from medical advances, most notably the diffusion of pharmaceutical therapies such as aspirin, beta blockers, and thrombolytics. Invasive technologies such as primary angioplasty also contributed significantly to increased survival. Heidenreich and McClellan’s analysis thus complements the findings of Cutler et al. on the efficacy of medical treatments for myocardial infarction.

The next chapter, “Measuring the Value of Cataract Surgery,” provides a dramatic example of the impact of changing medical technologies on reducing the burden of an ophthalmic condition, namely cataracts. The authors, Irving Shapiro, Matthew D. Shapiro, and David W. Wilcox begin by documenting changes in the last half-century for surgically extracting cataracts. Incisions and suturing have improved, intraocular lens implants have eliminated the need for cataract eyeglasses or contact lenses, and surgery is now done largely on an outpatient basis, resulting in greatly improved quality of outcomes. Patients have much faster ambulation, face lower rates of complications, and have better postoperative visual outcomes. Because of these cost reductions and quality improvements, patients with less severe disease are now increasingly having cataract surgery, earlier in the course of the disease, with more long-lived benefits. While precise measurement is still challenging, it is clear that although monetary costs have been relatively flat over time, in any reasonable quality-adjusted world the real price of cataract extraction has fallen sharply with time. The authors conclude with a discussion about the generalizability of their findings, and urge that future research focus on taking into account how the population of patients receiving treatment changes endogenously in response to changes in cost.

In their chapter “Hedonic Analysis of Arthritis Drugs,” Iain Cockburn and Aslam Anis focus on measuring price indexes for pharmaceuticals used in the treatment of rheumatoid arthritis. Although this is a widespread and debilitating disease with very substantial impacts on the health of patients and on the economy, currently the available drugs have limited efficacy and serious side effects. Clinical research conducted since these products were approved has resulted in substantial revisions to the body of scientific information available to physicians. The relative quality of these drugs (as represented by efficacy and toxicity measurements reported in peer-reviewed clinical trials) has changed markedly over the past fifteen years. Cockburn and Anis examine how prices relate to quality. Somewhat surprisingly, they find that in this therapeutic class, prices are only weakly

related to quality. They do, however, find a relationship between changes in reported efficacy and toxicity, and the evolution of quantity shares within this therapeutic class. Thus the Cockburn-Anis research reminds us that generalization across diverse medical conditions is hazardous, and that disease-by-disease analysis is instead necessary.

The chapter by Ernst Berndt, Susan Busch, and Richard Frank, "Treatment Price Indexes for Acute Phase Major Depression," focuses on alternative price indexes for acute phase unipolar (major) depression. Making use of results from the published clinical literature and from official federal government treatment guideline standards, Berndt et al. begin by identifying therapeutically similar treatment bundles. These bundles can then be linked and weighted to construct price indexes for specific forms of major depression. In doing so, Berndt et al. construct CPI- and PPI-like medical price indexes that deal with prices of treatment episodes rather than prices of discrete inputs, that are based on transaction rather than list prices, that take quality changes and expected outcomes into account, and that employ current, time-varying expenditure weights in the aggregation computations. Berndt et al. find that, regardless of which index number procedure is employed, the treatment price index for the acute phase of major depression has hardly changed, remaining at 1.00 or falling slightly to around 0.97. This index grows considerably less rapidly than the various official PPIs. Thus, relative to overall inflation, the price index for the treatment of the acute phase of major depression has fallen over the period 1991–95.

Berndt et al. further find that a hedonic approach to price index measurement yields broadly similar results. These results imply that given a budget for treatment of depression, more could be accomplished in 1995 in terms of outcomes than in 1991. The results suggest that at least in the case of acute phase major depression, aggregate spending increases are due primarily to a larger number of effective treatments being provided, rather than being the result of price increases.

### **Extensions of the Frontier**

The final section of the volume tackles more exploratory issues in output measurement—previously unstudied issues or extensions to other methodologies. Sherry Glied focuses on nonmedical sources of health improvements for children and adolescents. In her chapter, "The Value of Reductions in Child Injury Mortality in the United States," Glied notes that childhood mortality rates have declined steadily over time and across causes of death. She investigates alternative explanations for this decline. Glied focuses on several potential factors in improved health: changes in children's living circumstances, changes in the professional child injury knowledge base, changes in the information imparted to parents, and

changes in the regulations surrounding childhood behavior. Using data from the National Mortality Detail Files on the number of child deaths by age, cause, and state, combined with information from the Current Population Survey on the characteristics of children and their families by state, Glied finds that changes in children's living circumstances can explain little, if any, of the change in child health. There is limited evidence that regulatory interventions intended to change behavior have been important. Most important is evidence suggesting that changes in the knowledge available to parents about child health have become increasingly important. Over time, parents' time has become less important in producing health. These results provide a first effort in understanding the dramatic reduction in child injury mortality and also illustrate how the development of scientific information, a public good, is translated into private outcomes. However, this scientific information can generate growing inequality in those outcomes.

In their chapter, "Patient Welfare and Patient Compliance: An Empirical Framework for Measuring the Benefits from Pharmaceutical Innovation," Paul Ellickson, Scott Stern, and Manuel Trajtenberg develop an empirical framework for evaluating the patient welfare benefits arising from pharmaceutical innovations. Ellickson et al. extend previous studies of the welfare benefits from innovation, unpacking the separate choices made by physicians and patients in pharmaceutical decision making. They develop an estimable econometric model which reflects these choices. The proposed estimator for patient welfare depends on whether patients comply with the prescriptions they receive from physicians, and the motives of physicians in their prescription behavior. By focusing on compliance behavior, the proposed welfare measure reflects a specific economic choice made by patients, and thereby addresses to some extent the principal-agent relationship that confounds analysis of medical care.

Ellickson et al. review evidence that the rate of noncompliance ranges up to 70 percent, suggesting an important gulf between physician prescription behavior and realized patient welfare. Because physicians act as imperfect but interested agents for their patients, the welfare analysis based on compliance must account for the nonrandom selection of patients into drugs by their physicians. The paper integrates the choices made by both physicians and patients into a unified theoretical framework and suggests how the parameters of such a model could be estimated from health claims data.

The final chapter is by Frank Lichtenberg, entitled "The Allocation of Publicly Funded Biomedical Research." Lichtenberg develops a simple theoretical model of the allocation of public biomedical research expenditure and presents selected empirical evidence about the determinants of this allocation. Lichtenberg notes that the composition of expenditures should depend on the relative costs as well as the relative benefits of

different kinds of research. Analysts of technical change typically have data on neither of these, but Lichtenberg argues that the burden of illness is indicative of the potential benefit of achieving advances against different diseases, allowing him to infer how closely disease costs and benefits are aligned.

In his empirical work, Lichtenberg calculates distributions of government-funded biomedical research expenditure, by disease, from records of all research projects supported by the U.S. Public Health Service. To obtain a reasonably complete accounting of disease burden, he utilizes data on both the dying (from the Vital Statistics–Mortality Detail file) and the living (from the National Health Interview Survey). Lichtenberg finds a very strong positive relationship across diseases between total life years lost before age sixty-five and public R&D expenditures. He also finds that the amount of publicly funded research on a disease decreases with the share of life years before age sixty-five lost to the diseases by nonwhites. This could reflect the fact that lack of scientific knowledge is a less important cause of premature mortality among nonwhites than it is among whites. The number of research grants mentioning a chronic condition is uncorrelated with the number of people with the condition, but is very strongly positively related to the number of people whose activities are limited by that condition. Finally, Lichtenberg finds that there tends to be more research about chronic conditions that are prevalent among people living in low-income households, and that are prevalent among the youngest (under age eighteen) and the oldest (above age seventy-five).

### **Concluding Observations**

Although the chapters in this volume are diverse, two themes predominate—one expected, and the other a surprise to many. First, accounting for changes in medical outcomes is difficult but essential—particularly as outcomes improvements increasingly involve the quality rather than the length or quantity of life. While current measurement methods are not entirely satisfactory, progress is being made and additional enhancements to measuring medical care productivity are very likely in the near future. Second, the conventional wisdom that technological advances in medicine are a driving force of increasing health care costs is much too simplistic and deserves much more careful empirical scrutiny.

Some technological developments, such as those involving treatment for cataracts and for acute phase depression, enable a larger proportion of the affected population to tolerate and benefit from treatment. In turn, the greater treatment effectiveness creates incentives for more intensive and more frequent diagnoses. While the number of patients receiving effective treatment may increase as a result of the technological developments, and total treatment expenditures may increase, in many cases the treatment

cost per patient episode has fallen. Treatment quantity rather than treatment price may well be the largest driver of expenditure escalation.

Finally, the chapters in this volume were written several years ago, and one might ask, why publish them now? There are several compelling reasons. First, the issues addressed in this volume—the reliable measurement of medical care sector output and productivity—continue to persist and frustrate private and public sector analysts, in the United States and elsewhere. As the population age structure becomes older in the coming decades in much of the world, benefits from being able to measure more reliably the efficiency and productivity of a growing medical care sector will be increasingly valuable. In brief, issues involving medical care output measurement are persistent and increasingly important. Second, the literature on measuring medical care output and productivity is still unsettled and in flux, and it is important to understand the diverse and wide-ranging natures of the alternative approaches. This allows us to take stock of where we are. It also allows statistical agencies to evaluate changes in price and productivity measurement for medical care. Third, to facilitate future research, it is particularly valuable and useful to assemble and put into one volume a diverse collection of conceptual and empirical analyses, authored by leading researchers. That goal, we believe, is achieved in this volume.

This Page Intentionally Left Blank