

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: Price Indexes for Medical Care Goods and Services -- An Overview of Measurement Issues

Chapter Author: Ernst R. Berndt, David M. Cutler, Richard Frank, Zvi Griliches, Joseph P. Newhouse, Jack E. Triplett

Chapter URL: <http://www.nber.org/chapters/c7628>

Chapter pages in book: (p. 141 - 200)

Price Indexes for Medical Care Goods and Services

An Overview of Measurement Issues

Ernst R. Berndt, David M. Cutler, Richard G. Frank,
Zvi Griliches, Joseph P. Newhouse, and Jack E. Triplett

Statistics on medical prices should be improved; indexes of medical productivity should be developed; and the search for an understanding of the determinants of medical price and cost behavior should be developed.

—U.S. Department of Health, Education, and Welfare (1967)

4.1 Introduction

The measurement of the output of the medical care system is necessary to assess the productivity levels and growth of a country's economy and of course of its medical care system. This is true in countries with universal health care coverage or incomplete coverage, and regardless of the mix of public and private provision of medical care.

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. David M. Cutler is professor of economics at Harvard University and a research associate of the National Bureau of Economic Research. Richard G. Frank is professor of health economics at Harvard Medical School and a research associate of the National Bureau of Economic Research. The late Zvi Griliches was the Paul M. Warburg Professor of Economics at Harvard University and director of the Productivity and Technical Change Program at the National Bureau of Economic Research. He was past president of the Econometric Society and of the American Economic Association and a member of the National Academy of Sciences. Joseph P. Newhouse is the John D. MacArthur Professor of Health Policy and Management and is on the faculties of the Kennedy School of Government, the Harvard Medical School, the Harvard School of Public Health, and the Faculty of Arts and Sciences at Harvard University, and a research associate of the National Bureau of Economic Research. Jack E. Triplett is a visiting fellow at the Brookings Institution.

The authors gratefully acknowledge research support from the U.S. Bureau of Economic Analysis (Berndt, Cutler, Griliches), the U.S. Bureau of Labor Statistics (Berndt, Cutler, Griliches), Eli Lilly and Co. (Berndt, Frank, Griliches, Triplett), the National Institute of Mental Health (Frank), the National Institute on Aging (Cutler), the National Science Foundation (Berndt, Griliches), Hans Sigrist Stiftung (Newhouse), and the Alfred P. Sloan Foundation (Cutler, Frank, Newhouse). The authors have benefited from conversations with numerous colleagues and friends, and in particular they thank Dennis Fixler of the Bureau of Labor Statistics, Brent Moulton of the Bureau of Economic Analysis, and Richard Zeckhauser of Harvard's Kennedy School of Government for comments on earlier drafts. A substantial portion of this chapter also appears in Berndt et al. (2000), and is reprinted with permission from Elsevier Science.

For most industries in most countries, real output measurement is accomplished by dividing data on revenues or sales by a price index to obtain a measure of real output. Reliable output measurement for an industry therefore requires correspondingly reliable revenue data and a price index. A number of conceptual difficulties and institutional characteristics of medical care markets, however, make reliable price measurement of medical goods and services particularly difficult and challenging.

For countries where medical care goods and services are provided by the government without direct charge, or with only nominal direct charges, data on revenue or receipts for medical care may not be available or may not be relevant. For these countries, the problem of measuring the output of medical care goes well beyond the inherent difficulty of measuring medical care prices. In such cases the difficult problem of measuring prices and output of medical care is combined with the equally formidable problem of measuring the output of the government sector.¹

Medical price indexes have uses other than those involving output and productivity measurement. In the United States, both within the health sector and more generally, contracts occasionally contain provisions that depend on growth of the medical Consumer Price Index (CPI).² Medical CPIs and medical Producer Price Indexes (PPIs) are also employed in updating fee schedules for certain administered pricing schemes and payments to some health plans. Medical CPIs and PPIs are also employed by public policy analysts in projecting the impacts of changes in public policy.

Although medical CPIs and PPIs play prominent roles in private and public sector transactions and analyses, both the U.S. Bureau of Labor Statistics (BLS) and its critics have acknowledged that current BLS practices for tracking price changes in the medical care industries, industries that are characterized by dynamic technological and organizational changes, are likely to be inaccurate and in need of substantial improvement and overhaul.³

Several aspects of the medical care industry make the BLS's task of constructing accurate and readily interpretable medical CPIs and PPIs particularly difficult. Output measurement of the health care system is inherently difficult when mortality is but one possible outcome from treatment. Mortality is particularly inappropriate as an output measure for treatments of a variety of acute conditions that are not life threatening, and for many increasingly prevalent chronic illnesses. Additional attributes such

1. For a discussion of measurement issues in public sector output, see Kendrick (1991) and Griliches (1992, 18–19). Murray (1992) and the Swedish Ministry of Finance (1997) contain empirical analyses of publicly provided health sector output and productivity growth.

2. For general discussion of CPI use in escalation clauses, see Triplett (1983).

3. See, for example, U.S. Senate Finance Committee (1996); U.S. Department of Labor, Bureau of Labor Statistics (1997b); and Abraham, Greenlees, and Moulton (1998).

as morbidity, pain and suffering, functional and emotional impairment, and quality of life are each highly valued aspects of treatment response.

Another output measurement challenge that is rather unique to the medical care sector arises from the moral hazard, caused by health insurance, that causes marginal private and social costs to diverge. As emphasized by, among others, Newhouse and the Insurance Experiment Group (1993), the existence of demand-side moral hazard or administratively set prices makes it inappropriate to attribute the usual normative properties to medical CPIs that are commonly associated with other such price indexes. The provision of medical care services also involves a principal-agent relationship: In choosing treatment, patients typically rely considerably on the advice and counsel of their physician, whose incentives and financial interests may or may not align well with those of the patient. Any misalignment of interests may result in inefficient outcomes.

A third dimension of medical care that poses significant price measurement challenges relates to technological progress. While not unique to medical care, technological progress is nevertheless of great significance in this sector of the economy. New treatment technologies are continually emerging and being introduced into common clinical practice. This creates many of the problems of new goods, with which economists interested in index number and productivity measurement have struggled for many years.⁴

Finally, organizational changes have been dramatic in the medical care sector. The manner in which medical technologies are rationed, delivered, and even priced has evolved rapidly during the last decade. Managed care arrangements have resulted in changes in the locus of care, the organization of medical practice, contractual relations between buyers and sellers, and the manner in which inputs are combined to create treatment. Thus the way in which typical treatment for an illness such as depression is organized and provided has been remarkably altered in just a few years. Even given a known set of treatment technologies, important qualitative differences have emerged in the supply of treatment and in the way care is experienced by patients.

In this chapter we review in considerable detail the measurement issues that underlie construction of medical care price indexes; we describe procedures employed by the BLS in the construction of its medical CPIs and PPIs (including recent revisions and changes); we discuss alternative notions of medical care output that involve the price of an episode of treatment rather than the prices of fixed bundles of inputs; we outline salient features of a new medical care expenditure price index; and we suggest future research and measurement initiatives that are likely to be most fruitful. We begin with a description of the market environment underlying medical care CPIs and PPIs in the United States.

4. See, for example, the chapters and references in Bresnahan and Gordon (1997).

4.2 The Market Environment Underlying Medical Care CPIs and PPIs

Viewed by an economic statistician, the medical care sector is large and intimidating. As with an elephant, one can employ several approaches in cautiously observing, walking around, and measuring it. We begin by describing the principal actors, characteristics, and incentive structures that must be taken into account in providing a foundation for the measurement of medical care prices.

4.2.1 Distinguishing Features of the U.S. Medical Care Marketplace

Economists generally presume some form of consumer optimization and efficiency in the purchase of goods and services. As in other markets, consumers of medical services are envisaged as maximizing some notion of utility, buying goods and services that generate direct utility, and using some of these goods and services as intermediate goods to produce utility. In the medical care marketplace, however, this optimization and efficiency is exceedingly complex; it involves behavior based on the use of asymmetric information and personnel who act as imperfect agents for consumers, under rationing constraints that are not nearly as pervasive as in other consumer markets.

The medical care industry provides goods and services in a number of specific subsectors: hospitals (including hotel and cafeteria services), physician practices, laboratories, pharmaceuticals, clinics, medical devices, nursing homes, home health agencies, and so on. These services are provided to consumers, but consumers typically do not value these services *per se*. Rather, they value the health outcomes resulting from medical interventions provided by the medical care industry.⁵ These impacts on health are conceptually the composite good that we want to price, but the nature of transactions in this industry is exceedingly complex.

As in any industry, market structure affects the industry's price level, and perhaps the rate of price growth, particularly if production efficiency is affected over time. Licensing, reputation, the regulatory environment, and intellectual property rights provide suppliers of medical services with varying amounts of market power, particularly because some medical service suppliers such as hospitals and physicians face limited competition outside rather narrow geographical market boundaries. In many cases fixed costs are high, to a great extent consumers arrive at random times, and price is greater than short-run marginal cost.

Buyers also have market power. Although the federal government has long been a major purchaser of medical services (providing funds for about 39 percent of personal health care expenditures in the United States in 1996),⁶ within the last decade there has been much consolidation of

5. For further discussion, see Triplett (1999, chap. 1 in this volume).

6. Levit et al. (1998), exh. 3, p. 39 and exh. 4, p. 43.

buying power among health maintenance and managed care organizations. Thus on both the supply and demand sides of the medical care marketplace, market power is present. Moreover, since most medical care services are not resellable, price dispersion is not easily eliminated by arbitrage, price discrimination is prevalent, and thus the law of one price typically does not hold.

There are several other features of the market structure of the medical sector that, while present to some extent in other sectors, are particularly pervasive in medical goods and services. First, the vast majority of medical care payments are not made directly by consumers. Indeed, in the United States in 1996, out-of-pocket payments by consumers accounted for only about 19 percent of total personal health care service expenditures.⁷ The remainder of medical care is largely paid for by insurers.⁸ Insurance programs may be run publicly, as with Medicare, Medicaid, and other federal, state, and local funds, which together accounted for 53 percent of personal health care service expenditures in 1996; or the sources of funds may be private, which in 1996 made up 37 percent of personal health care service expenditures, primarily for the nonelderly. Ultimately, the insurance payments not paid directly by individuals are passed back to individuals, in the form of higher taxes or reductions in other forms of government spending when the insurance payments are by the public sector, or in the form of an adjusted employee compensation package when insurance is provided by employers.⁹

The predominance of the indirect nature of payments creates several difficulties for constructing and interpreting consumer price indexes. The most significant of these is moral hazard. If consumers pay for only, say, 20 percent of medical care at the margin, they will seek to consume medical care until its marginal value is only about twenty cents per dollar of spending. This is true even though people *on average* must pay for the full dollar of medical care. Individuals will therefore tend to overconsume medical resources—resources will be consumed that cost society \$1 (less if there are rents) but are worth less than that at the margin.

The second important feature of the medical care market is that consumers do not always know what services they want. Patients tend to rely on physicians both to provide them services and to recommend the services they need. As a result, there is a principal-agent problem: Patients would like physicians to act in the patients' best interests, but physicians might not always have an interest in doing so.

In traditional U.S. health insurance arrangements, physicians and patients both had incentives for excessive medical care. Patients were well-insured at the margin and physicians were paid on a fee-for-service basis—

7. Ibid.

8. In the United States, however, about 4 percent derives from other philanthropic sources.

9. For a discussion of the incidence of employer-provided health insurance, see Gruber (1994, 2000) and Pauly (1997).

earning more when they did more because fees were generally above marginal cost. The result was an incentive structure on both the demand and supply side that induced excessive care. Today's environment in the United States is much changed, and increasingly involves more complicated rationing. Health plans now often operate under fixed budgets, whereas before they typically passed costs through to the employer or government. Thus they have begun to employ administrative mechanisms and financial incentives to control health care spending. The result is that, with increasing frequency, patient demand incentives are at odds with those of their health plans or physicians.

The implication of both of these pervasive features of the medical care industry is that revealed consumer purchases are *not* a reliable guide to the marginal value of medical care. This is in contrast to other markets, such as that for, say, compact discs, where consumers' marginal valuations are likely to be well reflected in prices and expenditures. Consumers may receive too much medical care, as they likely did under traditional insurance arrangements, or too little care, as some allege they do under managed care or capitated insurance (one price per patient per year, independent of the amount of services the patient actually receives). In many markets, it is eminently reasonable to relate relative prices to marginal rates of substitution in consumption, but in medical care this assumption is simply not tenable. As a practical matter, this inability to employ the assumptions underlying traditional revealed preference theory severely hampers the ability of economic statisticians to construct accurate and readily interpretable price indexes for medical care.

The extent to which medical care services differ from other services can be illustrated by considering a hypothetical transaction in a restaurant. Suppose an individual places an order for a particular set of items on the menu, and then leaves. Another person enters the restaurant, sits down at a table, eats the meal that was ordered, and then leaves. Finally, a third person comes in and pays for the meal. In medical care, these three persons are the physician, the patient, and the insurer. Whose valuation shall one measure?

As with many other services such as ATM banking services, the production function for medical care involves interdependent efforts of suppliers and consumers. This interdependent aspect of medical care production makes it more difficult to distinguish between producer and consumer price indexes. Moreover, for consumers, medical care, health, and utility are quite different. This occurs in part because the production function for health has a number of arguments, other than medical care. One formulation of the production function is as follows:

health =

$H(\text{medical care, knowledge, time, lifestyle, environment, etc.}).$

It is useful to consider these other inputs into the production of health.¹⁰ Many of the inputs have been shown to contribute more to health than medical care (Fuchs 1974, 1983). Knowledge, for example, mediates between medical care and health. Medical treatments must not only be produced, but they must be used as well, and knowledge about how to use them changes over time. To the extent that knowledge is nonrivalrous in nature and has public good properties, its existence together with the interdependent nature of production makes it difficult to assess uniquely the impacts of changes in knowledge on prices for suppliers versus those facing consumers.

As an example of the importance of medical knowledge, suppose that medical research discovers that a particular pharmaceutical agent is just as effective when taken in half a dosage strength as when taken at full strength; something like this occurred for contraceptives several decades ago. Has the price of medical care changed? From the consumers' perspective, the answer is likely yes, for the cost of achieving a particular health state has fallen.¹¹ From the pharmaceutical manufacturers' perspective, however, the answer may be no, for the marginal cost of producing a milligram of the medication may not have changed. From the vantage of the family practitioner physician, whether the price has fallen depends in part on how one views physician services. The advice provided by the physician to the patient may still take the same amount of billable time, but if the physician is part of a staff model health maintenance organization with total pharmaceutical coverage, the price of providing family planning services may have fallen.

Knowledge, of course, is just one form of technological change. The reason we distinguish knowledge from other forms of technological progress is that knowledge is often envisaged as *disembodied* technological progress, while most new technologies are *embodied* in a particular service or product. Although the absorption of knowledge is not without cost, to some extent knowledge has public good properties and is nonrivalrous, quite unlike, say, a piece of medical diagnostic equipment.

These two types of technological change overlap, yet it is useful to distinguish between them. Significant quality changes are often embodied in new medical care-related goods and services, but use of the new good may require additional knowledge. For example, a new noninvasive operation is typically performed by physicians using novel inputs (endoscopic instru-

10. For a more complete discussion, see Grossman (1972a, 1972b).

11. Even here, matters are complex. For a number of brand name pharmaceuticals in the United States (but not in Europe), the price per tablet is the same regardless of strength. Moreover, in the example here it is implicitly assumed that the consumers' cost for a contraceptive medication is not fully covered by insurance. Until recently in the United States, unlike the case for most medications for which the patient makes a copayment, for contraceptives consumers have generally borne the entire direct cost of the prescription.

ments), and knowledge about such new treatments can be usefully employed by patients and clinicians alike. In any case, it is appropriate to envisage knowledge as an input into the production of health, distinct from medical treatments.¹²

Another input into the production of health is time. Producing better health requires time inputs from households as well as providers: Time is spent in seeking and receiving treatment, in recovery, and in assisting others. Some medical innovations (for example, new anesthetics) have shortened the recovery time for patients.¹³ This too will reduce consumers' cost of better health, but may well leave producers' costs unchanged (say, if the new anesthetic cost the hospital as much as the old).

An individual's lifestyle is another input into the production of health. Eating habits, drinking patterns, exercise regimens, and the pursuit of risky behavior all affect an individual's health. In some cases greater use of medical care and unhealthy behavior occur simultaneously to maintain a given health state. For example, with the introduction of over-the-counter H₂-antagonists such as Pepcid AC, individuals can preemptively take a heartburn prevention medication and then eat a high-calorie, highly spiced meal.

The environment is yet one more input that affects health. Environmental changes may improve or retard health. For example, new diseases like AIDS may develop and be discovered while other diseases like smallpox may be eradicated. Changes in air and water pollution, in climate and weather, as well as in rates of criminal activity will also have important impacts on health. It is important that such environmental changes be envisaged as primarily affecting the quantity or quality of medical services provided, not their price.

The age distribution of the population might also be envisaged as an environmental input affecting health care spending of populations. As people age, typically more inputs of medical care are required to maintain health or to mitigate deterioration. Increased medical care expenditures, or increased health insurance premiums that reflect impacts of an aging population, are appropriately viewed as quantity rather than price increases; in such cases medical expenditures rise due to increased quantity consumption, not price changes.

Finally, it is important to emphasize that while the marginal utility of health is positive, health is not the only argument in an individual's utility function. Thus a utility function might be envisaged as follows:

$$\text{utility} = U(\text{health, lifestyle, leisure time, other consumption goods and services, environment, etc.}).$$

12. For an exchange of views on this, see Gilbert (1961, 1962) and Griliches (1962).

13. Because the largest cost of receiving medical care is often the patient's time cost, there are substantial incentives to develop innovations that conserve on time, particularly when the cost of these innovations is covered by insurance.

Note that some factors such as lifestyle and environment not only have a direct impact on utility, but also have an indirect impact via health. One important implication of this, as has been emphasized elsewhere by Triplett (1999, chap. 1 in this volume), is that the output of the medical care industry is not something like the *average* health of the population, but rather is best viewed in *marginal* terms as the health implication of a medical intervention, conditional on lifestyle, environment, and other inputs affecting health.

The production function for health, as well as the utility function, has intertemporal aspects. Some current consumption goods affect future health states as well as current utility, while some medical interventions impact future consumption possibilities and patterns. Here we put these complications aside, but see Grossman (1972a, 1972b) and Meltzer (1997) for further discussion.

4.2.2 Pricing Medical Care Services

With this discussion of salient characteristics of the medical care marketplace as background, we now consider approaches to price measurement. A representative consumer can be envisaged as an individual who is making decisions before knowing what diseases he/she might eventually experience. Extensions to heterogeneous consumers complicate matters, but for our purposes it is sufficient initially to work here with the simpler representative consumer framework.¹⁴ Let this representative consumer have a utility function that depends on consumption of goods and services (other than medical care) and health. For concreteness, assume there is only one disease, which everyone contracts; extending the analysis to multiple diseases and probabilities of having each disease is straightforward. Denote Y as exogenous income, H as the health state, M as the quantity of medical care and P_M its (normalized) price, I and P_I as the quantity and price (premium) of a constant-quality insurance policy, K as medical care knowledge, E as the state of the environment, L as leisure time, and T_M as time allocated to receiving medical treatments. For simplicity, M and P_M include both consumers' direct health expenditures and indirect medical services obtained through health insurance in a competitive, actuarially fair insurance market where changes in costs of medical services to insurers are passed on to consumers via insurance premium changes. In this context, I and P_I are associated only with pure insurance services. The utility function is then written as

$$(1) \quad U = U[Y - P_M M - P_I I, H(M, K, E), L - T_M].$$

The first term is nonmedical care consumption (nonmedical expenditures divided by numeraire price), the second is health, and the third is

14. For an extension to heterogeneous consumers, see Pollak (1980, 1998) and Fisher and Griliches (1995).

nonmedical care time.¹⁵ Although equation (1) embeds a multiyear framework, for simplicity we assume but one time period. Notice also that equation (1) makes no assumption about how medical treatment decisions are made, or how medical prices are set.¹⁶

Over time, medical care and its price may change, or there may be changes in knowledge, the environment, and time devoted to medical care. For concreteness, consider changes between periods 0 and 1. The question posed is, What is the correct price index for changes between periods 0 and 1, assuming that consumers optimize in each time period? We can define the *cost of living* in one of several natural ways—the change in the cost of living between periods 0 and 1 is the additional funds the individual needs in period 1 to be just as well off as he/she was in period 0. This amount may be positive, in which case the cost of living has increased, or it may be negative, in which case the cost of living has fallen.¹⁷ This hypothetical is associated with the Laspeyres base period utility notion of cost of living. An alternative, associated with the Paasche notion, uses the current period utility as the point of reference, and asks, What is the change in funds the individual needs in period 0 to be just as well off as he or she is in period 1? We consider this distinction in further detail below. A third index, the Fisher ideal, is the geometric mean of the Laspeyres and the Paasche.

Consider the amount C of additional money the consumer requires in period 1 to make him/her indifferent between living in periods 0 and 1 (the Laspeyres notion):

$$(2) \quad U[Y - P_{M_1}M_1 - P_{I_1}I_1 + C, H(M_1, K_1, E_1), L - T_{M_1}] \\ = U[Y - P_{M_0}M_0 - P_{I_0}I_0, H(M_0, K_0, E_0), L - T_{M_0}].$$

C is the change in the cost of living—a positive C implies an increase in the cost of living, and a negative C a decrease.

To form a price index, one can scale C by the income required to produce utility in period 0, or Y . The cost-of-living index could therefore be

15. For simplicity, time at work is omitted.

16. The relationship between utility maximization and index numbers relies critically on a number of assumptions. In the present context, such assumptions might well be that the consumer chooses M , I , K , and T_M given Y , P_M , P_I , P_K (which could be zero at the margin if knowledge is nonrivalrous), E , and L , so as to maximize U in each time period. As we point out at various times in this paper, these assumptions are likely to be particularly untenable in the medical care marketplace.

17. The issues under discussion here involving measurement of the cost of living are very different from those raised by the Boskin Commission, who recommended that the BLS move from a Laspeyres price index formula to a superlative index such as the trailing Törnqvist, the latter more closely approximating a much more narrow notion of a cost-of-living index. See U.S. Senate Finance Committee (1996); Boskin et al. (1998); Abraham, Greenlees, and Moulton (1998); and Persky (1998).

$$(3) \quad \text{cost-of-living index} \equiv 1 + C/Y.$$

Using a first-order difference approximation, we differentiate and rearrange equation (2), yielding

$$(4) \quad C \equiv \{d(P_M M + P_I I)/dt - (U_H/U_X) [H_M(dM/dt) + H_K(dK/dt) + H_E(dE/dt)] + (U_L/U_X)(dT_M/dt)\},$$

where U_H is the marginal utility of health; U_X is the marginal utility of nonmedical consumption ($X \equiv Y - P_M M - P_I I$); U_L is the marginal utility of leisure; and H_M , H_K , and H_E are partial derivatives of H with respect to M , K , and E .¹⁸ Several comments are worth noting.

The first term on the right-hand side of equation (4), $d(P_M M + P_I I)/dt$, is additional spending on medical care and insurance services over time. A spending increase may be due to increased quantities of medical services provided (direct or via health insurance), increases in the prices paid for those medical services by consumers/insurers, increases in the carrying cost of insurance, or increases in the quantity of pure insurance services provided. Thus it is clear that an increase in the cost of medical services, *ceteris paribus* (in particular, health outcomes and environment assumed constant), increases the cost-of-living index. Notice that if the medical environment changes, such as if a new disease such as AIDS appears, medical expenditures will likely increase, but this is not properly viewed as a change in the cost of living, for the latter assumes an unchanged environment. As Griliches (1997) has pointed out, price index computations assume an average, unaging, unchanging individual living in a world in which nothing changes except prices. When a country's population becomes more aged, medical expenditure and the quantity of medical resources consumed increase, but as stated earlier, this is properly viewed as an expenditure and quantity increase, not a price increase. Similarly, because outcomes are being held fixed in the Laspeyres-type hypothetical, if bacteria develop drug resistance and low-priced antibiotics are replaced by more expensive drugs, the price index should increase, reflecting the reduced efficacy (quality deterioration) of the older antibiotic.

The second set of terms in the first line of equation (4), $-(U_H/U_X) [H_M(dM/dt) + H_K(dK/dt) + H_E(dE/dt)]$, is the dollar value of change in health over time. Health may change because the quantities of M , K , and/or E change, or because of, say, changed efficacy of a given medical treatment (change in H_M). The $-U_H/U_X$ term is the marginal rate of substitution between health and all other goods. Multiplying the health change by this amount expresses health in dollars. Note that an improvement in

18. We assume here that $dC/dt = U_X$, and that the marginal price of nonrivalrous additional knowledge is zero.

health through any of these three channels, *ceteris paribus*, reduces the cost-of-living index, or in other words, $C < 0$.

The final term in equation (4), $(U_l/U_x)(dT_M/dt)$, is the change in the time cost of receiving medical care. Hours are converted into dollars by multiplying hours by the marginal rate of substitution between leisure and goods (in most cases, equal to the after-tax wage). If more efficient delivery of medical care reduces patient travel and waiting time for medical care, or if recovery time from orthopedic surgery is reduced due to increased use of arthroscopic surgery, *ceteris paribus*, the cost of living falls.

Our discussion to this point on cost of living is that for a representative consumer. There are various ways in which group or aggregate cost-of-living measures and price indexes can be constructed, even when consumers' preferences are diverse and income (or total expenditure) has an unequal distribution. As discussed by, among others, Pollak (1980, 1998) and Fisher and Griliches (1995), a common aggregation procedure is to weight each person's utility in each of the two time periods by his/her dollar share in total expenditures; the share weights can be base period, current period, or some average of the two (the Törnqvist index). The aggregate cost-of-living indexes, analogous to equation (4), then include terms that represent share-weighted averages of various expenditures. Notice also that such aggregate cost-of-living indexes are conditional on the distribution of income and demographic composition of the population.¹⁹

Several other issues merit attention. First, it is useful to consider the $P_M M + P_I I$ term in equation (4) further. One possible price index to compute would involve asking what price the consumer would pay in two adjacent time periods for an actuarially fair medical care insurance policy to keep on the same expected level of utility, *ceteris paribus*? Note that this is not the same as a disability insurance policy, for with that the beneficiary only recovers lost income and medical care costs, and is not compensated, for example, for lost utility due to loss of vision. The consumer may have an expected life pattern in mind, with age-related probabilities of experiencing certain diseases. Thus the price index would be based on the price of contracting for a year of medical costs, given expected disease susceptibility, technology, efficacy, environmental factors, and so forth.

Realizations over the ensuing time period could well change the market price of such an insurance contract, for a variety of reasons, with differing implications for price and quantity. If the consumption of more medical related goods is induced by the expansion of technological opportunities (new artificial hips) or changes in the environment (increased sensitivity to allergens), the change is appropriately viewed as one of quantity or

19. Pollak (1980, 1998) therefore calls these aggregate price indexes "plutocratic," and contrasts them with ones he names "democratic."

quality, not price. The premium paid on repriced insurance policies might increase as a result, but that is because of the changed technology or environment, not because of a price change. Note also that because of moral hazard, when improvements in technology occur, they may be difficult to value properly within the medical marketplace.

Earlier we noted that a Laspeyres type of cost-of-living index uses the base period utility as reference, whereas the Paasche employs the current period utility as the reference point. Suppose that in the time interval between the base and current periods, the individual experiences a deterioration in health state so severe that it no longer is biologically feasible for the individual in period 1 to maintain the period 0 level of utility, for example, if the individual loses eyesight or develops an illness such as AIDS. In such a case, there may not be any feasible answer to the Laspeyres question, but one might still be able to answer the Paasche question.

A still deeper problem occurs when unexpected changes take place, such as those that result in the unanticipated lengthening of life expectancy. An individual might want to alter considerably his/her lifetime optimization plan given a change in information, yet he/she may lack the resources to modify consumption to a new path that has now become optimal. Hence it is possible for the cost of living *per year* to increase, for the cost of living *a lifetime* to increase, all as a result of this unanticipated *benefit* of increased life expectancy. This raises difficult issues, and mixes up changes in cost-of-living indexes with technological progress. Cost-of-living indexes typically refer to the cost of a flow of services over a relatively short time period. Converting from, say, a lifetime stock to an annual flow may be reasonable if the population is assumed to be ageless or has a fixed age composition, if there are no unexpected changes, and if *ex ante* decisions are still correct *ex post*. If these conditions are not met, paradoxes may well emerge.

4.2.3 Forming a Price Index

A fundamental issue is how one estimates the values of the variables in cost-of-living index equations such as those in equations (3) and (4). Suppose we focus attention just on how changes in the medical sector affect the cost-of-living index. Although current procedures used by the BLS in its medical care-related CPIs and PPIs are discussed in detail below, here we briefly consider several alternative procedures.

One approach used in other settings is hedonic price analysis. If one estimates a regression model where the price of a medical service is the dependent variable, and where attributes of the medical procedure, the patient, and the provider are explanatory variables, then one can decompose price changes over time into changes in the value of services to patients and pure changes in price. An example of this type of hedonic price

measurement, for the treatment of acute phase depression, is in Berndt, Busch, and Frank (chap. 12 in this volume).²⁰

There are a number of problems with using hedonic analysis in this market. At the level of individual diseases, hedonic prices are not necessarily equal to consumers' marginal valuations. Because consumers are insured, the price they pay for medical care at the margin is different from the cost of medical care to society. Further, providers have their own incentives in recommending treatment decisions, which may reinforce or contradict consumer preferences. Thus, both because of moral hazard and principal-agent issues, we would not necessarily expect treatment decisions to be made optimally. Estimated parameters in hedonic price equations could therefore be based on data points reflecting socially (and privately) inefficient actions by consumers, physicians, and/or insurers. This raises difficulties in placing any social welfare interpretations on movements in hedonic price indexes over time.

Alternatively, one might perform hedonic analysis at the level of the insurance plan, as was recommended by Reder (1969) and as has been implemented by Jensen and Morrissey (1990). One could estimate a hedonic model for the price of insurance, using the attributes of the insurance policy as regressors, and thus infer the residual price increase. The difficulties here are both theoretical and practical.

At the theoretical level, a theory about how consumers choose health insurance plans is required that incorporates consumers' self-selection, moral hazard (augmented by tax subsidies), and preferences for compensation. Because most private health insurance is provided through employment, this involves a link between workers and employers, and between different workers within a firm. Our knowledge about how insurance decisions are made in firms is very limited.²¹ Hedonic analysis also presumes that consumers are fully aware of the attributes of the good they are buying. But with health insurance, there are often fundamental parts of the insurance contract that consumers do not know—indeed, cannot know—in advance.²²

At the practical level, we probably are unable to control for many of the other factors that influence plan costs. For example, plan premiums will depend on the health status of people who are enrolled in the plan as well

20. For an introductory discussion to the hedonic method, see Griliches (1988), chaps. 7 and 8, and Berndt (1991), chap. 4. Other applications in the medical context include Trajtenberg (1990), Berndt, Cockburn, and Griliches (1996), and Cockburn and Anis (chap. 11 in this volume).

21. See Gruber (2000), Pauly (1997), and Summers (1989).

22. For example, most consumers do not know the details of who they are allowed to see for cancer care in advance of being diagnosed with cancer. Indeed, the specific benefits may depend on the severity of the cancer of the person and may change with new knowledge about cancer treatment.

as the benefits offered. But plan enrollment reflects adverse selection. When premiums change, we need to be able to decompose them into changes in the cost of a given set of benefits, and changes in the sickness of the people enrolled in the plan. Without knowing in detail who is enrolled in each plan and what their expected medical spending would be, we cannot adequately control for the many factors involved in premium variation. Moreover, data to control for these factors are typically unavailable.²³ As with observations on disease-specific treatment costs, the analysis of cross-sectional and/or time series insurance policy data might well be comparing various inefficient equilibria.²⁴ This is particularly likely with noncontractable aspects of health plan rationing under managed care.

Pauly (1999) has recently revived Reder's proposal to use medical insurance prices as the basis of a price index for medical care. Pauly contends that the development of willingness-to-pay techniques in economics has become sufficiently advanced that one could now ask respondents to put evaluations on an insurance policy that covered some new medical technique, or a bundle of new medical techniques, compared with an insurance policy that did not cover those techniques. One advantage of this form of pricing insurance policies would be that it would in theory capture behavior toward risk in a way that is typically neglected in studies that address only the *ex post* cost of treating an illness/condition. If one has a disease, 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. Though this alternative approach may have advantages over attempting to construct price indexes for the treatment of specific diseases, pricing insurance policies also has significant disadvantages, as discussed earlier in this chapter and by Feldstein (1969) many years ago, and willingness-to-pay techniques remain subject to framing reference point and other issues. Moreover, empirical work to implement Pauly's suggestion is not yet available.

Yet another alternative to the hedonic and insurance policy approaches is to make specific assumptions about the way that medical treatment decisions are made. For example, one can assume that consumers have a specified known distribution of preferences for one prescription drug over another, as in Fisher and Griliches (1995) and Griliches and Cockburn (1994), or that consumers are making purchase decisions for goods with a high out-of-pocket share, such as for prescription drugs in Cockburn and Anis (chap. 11 in this volume). One can then combine this model with observed data on treatment and prices to form a component of the cost-

23. This data situation is gradually improving. See Cohen et al. (1996) for a discussion of the Medical Expenditure Panel Survey.

24. For further discussion, see Feldstein (1969) and Prescott (1997).

of-living index. While this approach is reasonable in some applications, it does not work well in markets where consumer information is poor and the share of out-of-pocket costs is low, as occurs in most medical care markets.

A third option involves more direct measurement. Suppose one focuses on a particular disease or condition and estimates empirically the changes in treatment costs and medical outcomes for that disease. If in addition one makes an assumption concerning the dollar value of health improvements, one can calculate the various individual factors in a cost-of-living index. This approach has recently been implemented by Cutler et al. (1998, chap. 8 in this volume). If such an approach were to be followed more generally, it would of course be necessary to undertake such analyses for a representative mix of illnesses, where outcomes could be reliably measured. We return to a discussion of this approach later in this paper.

With this discussion on the difficulties of conceptualizing and implementing price measurement of medical care services as background, we now turn to a review of price measurement procedures currently employed by the BLS in its medical care-related CPIs and PPIs. As we shall see, while changes have recently been implemented at the BLS in its medical care CPI and PPI programs, for the most part the BLS still treats medical care in the same way it treats other industry and consumer prices. The combination of inherent difficulties in measuring service industry prices, distinctive features of the medical care industry, and use of traditional index number procedures for measuring prices makes clear interpretation of the BLS's current medical care CPIs and PPIs very difficult.

4.3 Construction of Medical Care CPIs and PPIs at the BLS

The U.S. Bureau of Labor Statistics (BLS) constructs and publishes CPIs and PPIs for various components and aggregations of medical care goods and services. Hereafter we denote these medical care CPIs and PPIs by the acronyms of MCPIs and MPPIs, respectively. Although the essential structure and conceptual foundations of these price index measurement efforts have been in place for some time, the BLS has recently announced and undertaken a considerable number of changes in its MCPI and MPPI programs. Here we summarize both continuing and recently changing procedures. We begin with a more general overview of the CPI and the PPI, and then we consider issues particularly important to measuring prices and quantities of medical care goods and services in the MPPI and MCPI programs.

4.3.1 A Brief Summary of the CPI

According to the BLS, the CPI is "a measure of the average change in the prices paid by urban consumers for a fixed market basket of goods

and services.”²⁵ It is calculated monthly and is published about two weeks after the end of the month to which it refers.²⁶ From its first regular publication in 1921 until the end of World War II, the CPI was called a cost-of-living index. In March 1944 the chairman of the President’s Committee on the Cost of Living appointed a group of technical experts (Wesley Mitchell, Simon Kuznets, and Margaret Reid) to examine whether the BLS’s cost-of-living index was properly accounting for war-related quality deteriorations in goods and services, as well as for the effects of rationing and shortages. Controversy had emerged in part because in 1942 the “Little Steel Formula” had been adopted, which linked permissible wartime wage increases to the index, but representatives from organized labor argued that the cost-of-living index understated true price inflation.²⁷

Along with a special committee of the American Statistical Association appointed in 1943, the technical experts concluded that “the index understated the wartime price rise to some extent because of a number of factors, of which incomplete account of quality deterioration was only one.”²⁸ To avoid confusion with popular notions of cost of living, the president’s committee, as well as the union critique of the index, also recommended that the name be changed to Consumers’ Price Index, a change which was adopted in August 1945.²⁹

Since 1945 many changes have occurred involving the BLS’s construction of the CPI, but its underlying hierarchical structure has been relatively stable. The identity and number of items sampled, and the weights used in aggregating sampled items into increasingly comprehensive subindexes, constitute a hierarchical structure of market baskets that the BLS changes infrequently. Based on data from its Consumer Expenditure Surveys (CEX), the BLS identifies and defines a fixed “market basket” of goods, employing a classification system known as the item structure. The item structure has been updated approximately every ten years; the most recent update took place in January 1998.

For example, based on data from the 1993–95 CEX, the BLS identified

25. U.S. Department of Labor, Bureau of Labor Statistics (1992), 176.

26. Here we focus primarily on the CPI for All Urban Consumers (CPI-U), introduced in 1978 and representative of the buying habits of about 80 percent of the U.S. noninstitutional population. An alternative index, CPI-W (wage earners and clerical workers only), was introduced much earlier for use in wage negotiations, and represents only 32 percent of the U.S. population. The methodology for producing CPI-U is the same as that for CPI-W.

27. Hoover (1961), 1175. Union criticism of the index was written up in a “Meany Report,” mentioned by Hoover. Also see Persky (1998).

28. Hoover (1961), 1175. Hoover reports that from January 1941 to September 1945, the estimated downward bias was 5 percentage points. Also see Weiss (1955).

29. Hoover (1961), 1175, n. 2; also see Weiss (1955), 23. Incidentally, the Meany report argued that “To most people, ‘cost of living’ means the amount of money a family spends. If it buys more food and finer clothes, or moves to a roomier home, its cost of living goes up. That interpretation is so widespread that we think the Bureau’s index is misnamed” (quoted in Hoover 1961, 1175, n. 2).

eight major product groups of items for representation in the CPI beginning in January 1998: food and beverages, housing, apparel and upkeep, education and communication, transportation, medical care, entertainment, and other goods and services. In turn, these major groups are divided into seventy expenditures classes, which are disaggregated further into 211 item strata. Weights for the 211 item strata are fixed in between major revisions, as are those for the higher level of aggregations of strata into expenditure classes, intermediate aggregates, major groups, and all items indexes. The CPI calculations are done separately for thirty-eight geographic areas.³⁰

CPI calculations are undertaken based on a modified Laspeyres price index. The Laspeyres price index is a weighted sum of price relatives, where the weights are revenue shares of each of the N item strata in the market basket. For month t , the Laspeyres price index is

$$(5) \quad L_t \equiv \sum_{i=1}^N w_{ib} [p_{it}/p_{ib}], \quad w_{ib} \equiv \frac{P_{ib} q_{ib}}{\sum_{i=1}^N P_{ib} q_{ib}},$$

where p_{it} is the price of the i th item in time period t ; $i = 1, \dots, N$; p_{ib} are base period prices; q_{ib} are fixed base period quantities; and w_{ib} is the fixed base period expenditure weight. The term p_{it}/p_{ib} is often called the “price relative” of good i . An attractive feature of the Laspeyres index is that it is consistent in aggregation; in other words, one obtains the same composite Laspeyres index by aggregating over all items simultaneously, or first aggregating items into a set of subindexes, and then constructing a master aggregate from the weighted subindexes.

Because the CPI has 211 item strata, the terms p_{it}/p_{ib} in equation (5) are in fact price indexes, often called “basic components” or “elementary aggregates.” There are thousands, perhaps millions, of items in a modern economy. Within each of the 211 item strata, BLS takes a probability sample of the detailed items (termed entry-level items, or ELIs) that are grouped together into each of the item strata. For example, in the medical care component, there are 13 item strata. Based on a nonlinear programming optimization algorithm, the BLS determines the optimal number of price quotes at the expenditure class level.³¹ For some of the item strata (e.g., baby-sitting, carpooling), it is very difficult to obtain sample price data; thus for 27 of the 211 item strata, the BLS does not sample prices, but instead imputes prices from other goods and services in the same expenditure class. When a detailed item is selected for pricing in the CPI, a

30. Lane (1996), 22. Also see Ford and Ginsburg, (1998, chap. 5 in this volume). Several of these numbers have been revised since publication of these articles. We thank Dennis Fixler for providing final updates.

31. This optimization problem and its implementation are discussed in Leaver et al. (1997).

price for the exact same item is collected at regular intervals, usually monthly or bimonthly. These detailed prices are formed into the basic component price indexes, which are the lowest level for which price index information is published in the CPI.

To accommodate practical issues involving the fact that some products are discontinued and cannot be repriced, that consumers' point of purchasing items changes, and that the CEX provide data on expenditures rather than prices, current BLS practice incorporates a number of modifications. The related issues are discussed in further detail in U.S. Department of Labor, Bureau of Labor Statistics (1997b, 1998) and in Moulton and Stewart (1997).³² Here it is worth emphasizing that while weights may change for elementary items within item strata (due in part to sample rotation), at the item strata level and above, the weights are fixed over time between major revisions, and thus for aggregate price indexes at the level of item strata and higher, use of equation (5) with its fixed weights is essentially what is done by the BLS.

In 1997 the BLS began issuing a monthly experimental measure constructed with use of the geometric mean formula for all index components at levels of aggregation underneath the item strata. The geometric mean index permits limited substitutability among products within the item strata. Provided that commodity substitution is the primary economic behavior that affects these lower level indexes, the difference between geometric and arithmetic mean item strata indexes can be interpreted as a measure of "lower level" substitution bias.³³ This experimental index using geometric means appears to lower the growth of the all-items CPI by approximately 0.25 percent per year.³⁴

Recently the BLS announced that beginning in January 1999, the aggregating formula for constructing most of the elementary aggregates (comprising approximately 61 percent of total consumer spending) will be moved over to a geometric mean. Medical care CPI components are largely exceptions, however; all but prescription drugs, nonprescription drugs, and medical supplies will continue to be constructed by the traditional arithmetic mean calculation.³⁵ Note that for each of the more highly aggregated 211 item strata, fixed quantity weights will still be employed, reflecting the continuing assumption of zero substitutability *between* these strata.

With this overview of the CPI hierarchical structure, weights, and aggregation formulas as background, we now move on to a brief summary of the PPI.

32. Also see Moulton (1996) and Moulton and Moses (1997).

33. See Pollak (1998) for further discussion.

34. U.S. Department of Labor, Bureau of Labor Statistics (1997a).

35. U.S. Department of Labor, Bureau of Labor Statistics (1998), updated in Eldridge (1998).

4.3.2 A Brief Summary of the PPI

The Producer Price Index (PPI) “measures average changes in selling prices received by domestic producers for their output.”³⁶ Before 1978 the BLS named this price series its Wholesale Price Index (WPI). The change in name to Producer Price Index emphasized that its conceptual foundations were based on prices received by producers from whomever makes the first purchase, rather than on prices paid to wholesalers by retailers or others further down in the distribution chain.³⁷ At the same time, the structure of the index was changed substantially. The old WPI corresponded, roughly, to the P in the well-known quantity theory of value expression, $MV \equiv PT$. In this view of an inflation index, all transactions mattered, so the WPI combined into one index the prices of, for example, iron ore, steel, and the automobile in which the steel was an input.³⁸ The resulting substantial double counting in the WPI was regarded as a serious problem.³⁹ In response, the BLS converted the old price index to the concept of an industry output price index.⁴⁰ As a result, the basic measurement unit for the PPI has become an industry—in the case of medical care, hospitals, physicians’ offices and clinics, and nursing homes are each separate industries. The PPI publishes separate price indexes for the outputs of each of these industries.

The PPI is calculated monthly and is usually published in the second or third week following the reference month. The PPI involves pricing the output of domestic producers, while the BLS’s International Price Program publishes price indexes for both imports and exports.

The PPI program at the BLS takes as its definition of an industry the one based on the Standard Industrial Classification (SIC) code.⁴¹ Since its inception in 1902, the PPI has focused heavily on the goods-producing sectors of the U.S. economy, but ever since 1986, in recognition of the growing importance of services in the U.S. economy, the BLS has gradually begun to broaden the PPI’s scope of coverage into the service sectors.

Currently the BLS does not calculate and publish an economy-wide aggregate goods and services PPI, although it plans to do so beginning in January 2002. Rather, PPIs are published by industry (based on the SIC four-digit industry code and higher levels of aggregation), by commodity

36. U.S. Department of Labor, Bureau of Labor Statistics (1992), 140.

37. *Ibid.*, 141.

38. The WPI did not implement this theory completely, however, for it omitted nearly all service prices and also transactions in financial and secondhand assets.

39. See Council on Wage and Price Stability (the “Ruggles Report”) (1977).

40. The implementation of an industry output price index was based on the theoretical model developed by Fisher and Shell (1972), and amplified by Archibald (1977) and Diewert (1983).

41. Issues concerning how industries are defined and aggregated, as well as economic issues underlying the SIC code system, are discussed in Triplett (1990).

classification (by similarity of end use or material composition, regardless of whether these products are classified as primary or secondary in their industry of origin, for fifteen major commodity groupings), and by stage of processing (according to the class of buyer and the amount of physical processing or assembling the products have undergone); separately for finished goods, intermediate materials, and crude materials; and both by commodity and industry classifications. In this chapter we focus primarily on PPIs by industry.

Within each industry, the BLS calculates aggregate PPIs using the Laspeyres price index formulas (see equation [5] above). At the most disaggregated level of PPI price measurement (called the cell index), the BLS defines a price as “the net revenue accruing to a specified producing establishment from a specified kind of buyer for a specified product shipped under specified transactions terms on a specified day of the month.”⁴² Prices are for output currently being provided or shipped, and not for order or futures prices.⁴³ Although in general the BLS seeks transactions rather than list prices for its price quotes, responses by firms are less cumbersome when list rather than transactions prices are reported.⁴⁴ Participation in the PPI by firms is on a voluntary basis. As of December 1992, the overall PPI “productive” response rate was 63 percent (Catron and Murphy 1996, table A.2).

The PPI is also based on a hierarchical system, though as noted above, unlike the case of the CPI, currently there is no economy-wide measure of the PPI. The BLS constructs and publishes aggregate PPIs for the total mining and total manufacturing industries, but apparently because of a lack of sufficient coverage, the BLS does not currently publish an aggregate PPI for total services; the BLS hopes to publish such an aggregate services industry PPI by January 2002.⁴⁵

Price quotes from the most disaggregated cell indexes are aggregated via a Laspeyres weighting scheme, in which fixed weights are based on value of shipments data collected primarily by the Bureau of the Census; industry net output weights are employed to account for intraindustry sales. The net output weights therefore vary with the level of industry aggregation (e.g., four-digit to two-digit); the detailed industry flow data required to distinguish net from gross output are derived for the most part from use of input-output tables compiled by the Bureau of Economic Analysis. Beginning in January 1996, industry price indexes have been calculated pri-

42. U.S. Department of Labor, Bureau of Labor Statistics (1992), 141. Net revenue is net of any discounts as opposed to net of production costs.

43. Problems can emerge for industries in which a great proportion of currently shipped output is covered by long-term price contracts, but for which “spot” prices differ from contracted prices.

44. U.S. Department of Labor, Bureau of Labor Statistics (1992), 141–42.

45. U.S. Department of Labor, Bureau of Labor Statistics, 1996a, 1.

marily with net output weights based on 1987 input-output relationships. The 1992 input-output tables have just recently been released, and the BLS envisages using them by the end of 1998 or early 1999 (Lawson 1997).

With respect to the specific establishments and items sampled by the BLS in its PPI program, the BLS currently draws a sample of items for each industry on average every seven years or so, and then reprices this fixed set of items monthly until an entirely new sample is drawn. Since 1978, the BLS has attempted to employ a sampling procedure that makes the probability of selection proportional to a product's value of shipments. Because it recognized that in some technologically dynamic industries a seven-year time lag between samples could result in a sample of products and services much older and quite unrepresentative of market transactions, in 1996 the BLS announced that for certain industries, including pharmaceuticals and electronics, samples would be supplemented at one- or two-year intervals (Kanoza 1996).

Issues surrounding the reliability and possibility of biases in price index measurement have recently received much less attention for the PPI than for the CPI. Use of the Laspeyres weighting procedure, accounting for unmeasured quality changes, and discontinuation and exit of sampled goods and services raise issues which in many respects are similar for the CPI and PPI. On the other hand, a number of significant differences exist between the CPI and PPI medical care components.

First, the lowest level of aggregation is defined differently in the two indexes: The PPI is defined on four-digit SIC industries, and below that, the item detail is defined specifically to each of the diagnosis-related medical care industries (major groups for the hospital index, medical specialties for the physicians' index, and so forth). The item strata in the CPI are based on groups that, in principle, should correspond to consumer demand categories (Lane 1996).

Second, the frequency and nature of major revisions differ. The CPI has been revised every ten to twelve years, when new weights are assigned based on the consumer expenditure survey. The PPI is normally rebased every five years, with weights drawn from the economic censuses.⁴⁶

For the medical price indexes, another major difference exists between CPI and PPI. In the case of the PPI, revenues and output prices collected from the sampled unit refer to revenues from all sources—government, industry, and final consumers. For the CPI, only consumers' out-of-pocket costs are included. Government expenditures made on behalf of consumers and financed by taxes and health expenditures by insurance companies

46. Another difference involves the length of time between collection of underlying sales revenue census/expenditure survey data and the introduction of new weights into the Laspeyres index. For the PPI, this is about one to two years, but for the CPI it has been about three to four years.

in which employers (but not consumers directly) pay the premiums are out of scope for the present definition of the CPI. We return to discuss this difference in scope in section 4.3.4 below.

With these overview discussions of the CPI and PPI as general background, we now move to consideration of issues of particular importance to medical care goods and services. We begin with the MPPI.

4.3.3 PPIs for Medical-Related Goods and Services

As noted above, the BLS does not construct and publish a PPI for an aggregate of services. Nor does the BLS publish a PPI for an aggregate consisting of medical-related goods and services. Indeed, it is only within the last decade that the BLS, as part of its increased effort to measure prices in the various service industries, has begun publishing price indexes for hospital and physician services.

Among the manufacturing industries associated with health care, the BLS has published PPIs for some time for industries such as pharmaceuticals; hospital beds; medical books; surgical, medical, and dental instruments and supplies; ophthalmic goods; and others. Among the service industries, separate PPIs for numerous health care-related industries are a rather recent development. A PPI for health services was introduced by the BLS effective December 1994, for offices and clinics of doctors of medicine in December 1994, for skilled and intermediate care facilities in December 1994, for hospitals in aggregate and by type in December 1992, and for medical laboratories in June 1994.

If the BLS is ever to construct an aggregate MPPI, as with other industry aggregates, it will need to distinguish net from gross output by industry, using some form of an input-output matrix to measure inter- and intraindustry flows. Given the major changes in the health care sectors over the last decade, including impacts from the growth of managed care, it will of course be necessary to employ input-output matrices that are based on much more recent data than the 1987 input-output matrix currently employed by the BLS for defining net output in other industries. While the BLS plans to begin using 1992 input-output data beginning in late 1998, these data will already be six years out of date, and much organizational and technological change has occurred in the health care industries since 1992.

Output Measurement in the MPPI

A central measurement issue in the construction of MPPIs involves the specification and implementation of a concept of industry output. Although the PPI program utilizes the four-digit SIC classification system to identify and define industries, this SIC structure does not provide information enabling the BLS to define what is the appropriate real output concept

in medical care industries, or on how this output quantity and output price can best be measured.⁴⁷ As we shall see, important problems also emerge when medical treatments from distinct SIC industries are substituted for each other in treating an illness or condition.

In the United States, medical goods and services were traditionally paid for by fee-for-service arrangements. In a fee-for-service context, a reasonable business procedure involves identifying and separately billing for each particular component of medical care from, say, a physician, a hospital, and a pharmacy. The fee-for-service was essentially the price for the inputs to medical care.

In 1983 the Health Care Financing Administration (HCFA) introduced major changes in how general acute care hospitals treating Medicare patients were to be reimbursed. Specifically, beginning in 1983 HCFA implemented a prospective payment system for inpatient hospital care whereby general acute care hospitals received a fixed payment for almost every Medicare patient admission, regardless of the amount or duration of services actually provided to the patient. This prospective payment mechanism represented a sharp departure from the retrospective cost-based accounting framework used for many years.

Medicare prospective payment schedules are based on estimates of (average accounting) costs for the resources utilized in providing services for a typical patient in a given geographical area being treated for a particular medical case. As of 1995 payments were distinguished for treatments of twenty-four major diagnostic categories, which are broken down further into 495 medical and surgical groupings, known as diagnostic related groups (DRGs).⁴⁸ The DRG prospective payment schedules have been updated regularly by Congress utilizing recommendations from the Secretary of Health and Human Services and the Prospective Payment Assessment Commission (now the Medicare Payment Advisory Commission); updates include changes in "medical costs" and case-mix indexing to account in part for secular trends in upcoding, also known as "DRG creep."

DRGs provide one possible output concept, and while DRGs in theory are applicable to all populations, Medicare currently employs DRGs only to reimburse hospitals for inpatient hospital care; many outpatient commodities (e.g., home health care) and services for illnesses of the elderly, and particularly of the nonelderly, are not included in the DRG system.⁴⁹

Classification schemes used for other services include version 4 of Cur-

47. For a discussion of the economic foundations underlying SIC definition, see Triplett (1990).

48. A number of these 495 DRGs are no longer valid. For a recent list, see Prospective Payment Assessment Commission (1995), app. E.

49. DRG weights have been calculated for nonelderly patients for Maryland, New Hampshire, and New York, but no DRG nonelderly weights exist based on national data. However, a limited number of private insurers use DRGs for nonelderly beneficiaries, as do several state Medicaid programs. Also see note 51.

rent Procedural Terminology (CPT4) codes, a list containing thousands of procedures for which physicians and hospitals can bill; these CPT4 codes can be envisaged as inputs into the treatment of an illness or condition.⁵⁰

A systematic structure of diagnostic codes for illnesses and conditions is version 9 (now, version 10, but for most of this time period version 9; thus we refer to version 9) of the International Classification of Diseases (ICD-9).⁵¹ Relationships among ICD-9, CPT4 and DRG codes are multifaceted. A single DRG encompasses treatment of somewhat arbitrary aggregations of distinct ICD-9 diagnoses, alternative combinations of CPT4 codes can be used in the treatment of a particular ICD-9 diagnosis, and a given CPT4 procedure can be used in the treatment of various ICD-9 diagnoses. Other diagnostic-related systems used in setting risk-adjusted capitation rates include the Ambulatory Care Group algorithm (Weiner et al. 1996) and the Diagnostic Cost Group's Hierarchical Coexisting Conditions models (Ellis et al. 1996).

DRGs and their offspring represent the beginning of a structure which could facilitate defining, measuring, and pricing the output of medical care providers. In particular, the output of a particular DRG billing involves the treatments for an episode of hospitalization for a particular condition/diagnosis. Instead of pricing each of the components of a hospitalization, with DRGs the composite bundle of hospital services is given a single ex ante price.⁵²

Along with the development of CPT4 and ICD codes, the notion of an episode of illness or treatment has expanded far beyond the hospitalization realm, suggestive of yet alternative ways of measuring medical care output. Numerous professional medical associations, as well as the Agency for Health Care Research and Quality (AHRQ), an agency of the Public Health Service in the U.S. Department of Health and Human Services, have developed clinical practice guidelines and treatment protocols for various illnesses and conditions. These treatment guidelines, which change over time, define ex ante medically acceptable and often therapeutically similar bundles of treatment involving medical inputs such as laboratory tests, pharmaceuticals, minutes of service from physicians and other medical personnel, and various other inpatient and outpatient procedures. Health insurance plans, hospitals, and pharmaceutical companies have developed programs and protocols for the management of certain diseases. These disease management programs implicitly, and sometimes explicitly, suggest outputs of the medical sector that facilitate the pricing of treat-

50. For a discussion of CPT4, see American Medical Association (1990).

51. ICD-9 codes are discussed and listed in U.S. Department of Health and Human Services (1980). The ICD-9 system with clinical modifications is called ICD-9-CM, and it has recently been updated to version 10.

52. The Medicare payment scheme reserves 5 percent of its payments for outlier or exceptionally expensive cases. At the margin, these are reimbursed on a cost basis.

ment bundles and the accounting framework for assigning payments to providers.⁵³

With this as background, we now turn to a discussion of how the BLS's PPI program has implemented medical care sector output price and quantity measurement, and how it has built on the notion of DRG treatment episodes as output measures in the health care sector. As noted earlier, within the last decade the BLS's PPI program has made major changes in health care related PPIs and introduced many new ones. We begin with a discussion of medical services—physicians and hospitals—and then we discuss selected medical goods, such as pharmaceuticals.

Physicians' Services in the MPPI. The PPI program has initiated procedures for constructing medical service PPIs at two rather aggregate levels, physicians' services and hospital services. Each of these two classes of services in turn encompasses a variety of more detailed physicians' and hospital service industries. In table 4.1 we list the entire set of detailed physician, hospital, and medical laboratory industries in SIC 80 for which the BLS is currently constructing health services subindex PPIs.

With respect to offices and clinics of doctors of medicine ("physicians' services"), the new BLS procedures distinguish Medicare from non-Medicare treatments. Within the non-Medicare treatments, multispecialty group practices are treated separately from one- and two-physician practices and single specialty group practices, with the latter in turn being broken down into nine specialties. For skilled and intermediate care facilities, public payers are distinguished from private.

The second principal subindex within health services is hospital services. As is seen in table 4.1, general medical and surgical hospitals are differentiated from psychiatric hospitals and other specialty hospitals aside from psychiatric. Both inpatient and outpatient treatments are separated into those involving Medicare patients, Medicaid patients, and all other patients; for the non-Medicare and non-Medicaid inpatients, hospital treatments are differentiated involving twenty-three distinct illnesses/diseases/conditions.

Development of the BLS's PPIs for physician services has benefited considerably from the prior implementation and common usage of the DRG, CPT4, and ICD classification systems by insurers, hospitals, physicians, and other providers.⁵⁴ Based on a sampling universe including all physician practices in the United States, the BLS employs probability sampling stratified by size and specialty. The size of a physician practice is based on the

53. See Triplett (1999) for further discussion.

54. It is interesting to note, however, that in 1996 the percentage of preferred provider organizations reimbursing hospitals by DRG-based methods was only 31.7 percent (80.2 percent used per diem methods), and that only 7.7 percent of hospitals were reimbursed by PPOs using DRG-based methods. See Hoechst Marion Roussel (1997), 86.

Table 4.1 Subindexes of the Health Services Producer Price Index

Industry	SIC Code
Health services	80
Offices and clinics of doctors of medicine	8011
Primary services	8011-P
Medicare treatments	8011-1
Non-Medicare treatments	8011-3
One- and two-physician practices and single specialty group practices	8011-31
General/family practice	8011-311
Internal medicine	8011-312
General surgery and other surgical specialties	8011-313
Pediatrics	8011-314
Obstetrics/gynecology	8011-315
Radiology	8011-316
Psychiatry	8011-317
Anesthesiology	8011-318
Other Specialty	8011-319
Multispecialty group practices	8011-331
Skilled and intermediate care facilities	8053
Primary services	8053-P
Public payers	8053-101
Private payers	8053-301
Other receipts	8053-SM
Hospitals	806
General medical and surgical hospitals	8062
Primary services	8062-P
Inpatient treatments	8062-1

(continued)

Table 4.1 (continued)

Industry	SIC Code
Medicare patients	8062-131
All medical diagnosis-related groups	8062-13101
All surgical diagnosis-related groups	8062-13103
Medicaid patients	8062-171
All other patients	8062-171
Diseases and disorders of the nervous system	8062-17101
Diseases and disorders of the eye	8062-17102
Diseases and disorders of the ear, nose, mouth, and throat	8062-17103
Diseases and disorders of the respiratory system	8062-17104
Diseases and disorders of the circulatory system	8062-17105
Diseases and disorders of the digestive system	8062-17106
Diseases and disorders of the hepatobiliary system and pancreas	8062-17107
Diseases of the musculoskeletal system and connective tissue	8062-1708
Diseases and disorders of the skin, subcutaneous tissue, and breast	8062-17109
Endocrine, nutritional, and metabolic diseases and disorders	8062-17111
Diseases and disorders of the kidney and urinary tract	8062-17112
Diseases and disorders of the male reproductive system	8062-17113
Diseases and disorders of the female reproductive system	8062-17114
Pregnancy, childbirth, and puerperium	8062-17115
Newborns and other neonates with conditions originating in the perinatal period	8062-17116
Diseases and disorders of the blood and blood-forming organs and immunological disorders	8062-17117
Myeloproliferative diseases and disorders, and poorly differentiated neoplasms	8062-17118
Infectious and parasitic diseases (systemic or unspecified sites)	8062-17119
Mental diseases and disorders	8062-17121
Alcohol/drug use and alcohol/drug-induced organic mental disorders	8062-17122
Injuries, poisonings, and toxic effect of drugs	8062-17123
Burns	8062-17124
Factors influencing health status and other contacts with health services	8062-17125

Outpatient treatments	8062-3
Medicare patients	8062-311
Medicaid patients	8062-331
All other patients	8062-351
Other receipts	8062-SM
Psychiatric hospitals	8063
Primary services	8063-P
Inpatient treatments	8063-1
Medicare patients	8063-101
Non-Medicare patients	8063-103
State and county hospitals	8063-10301
Private hospitals	8063-10303
Outpatient treatments	8063-2
Other receipts	8063-SM
Specialty hospitals, except psychiatric	8069
Primary services	8069-P
Inpatient treatments	8069-1
Rehabilitation hospitals	8069-101
Children's hospitals	8069-104
Alcoholism and other chemical dependency hospitals	8069-107
Other specialty hospitals except psychiatric	8069-108
Outpatient treatments	8069-3
Other receipts	8069-SM
Medical laboratories	8071
Primary services	8071-P
Pathology and laboratory	8071-1
Urinalysis	8071-102
Chemistry, toxicology, and therapeutic drug monitoring	8071-103
Hematology	8071-104
Pathology	8071-107
Profiles and panels	8071-108
Radiological tests	8071-3

number of physicians in a given practice (not the number of employees, or revenues); the sample is stratified further into nine single specialty categories and one multispecialty category. Initially in 1993–94 it was expected that the total number of physician practices sampled would be about 400 and the number of quotes obtained would be about 1,150,⁵⁵ but by mid-1997 only 158 units remained in sample, yielding 845 quotes (Fixler and Ginsburg, chap. 6 in this volume).

Given the sampling unit, at the price quote initiation point in time, the BLS randomly chooses a bill that measures the net prices paid to a physician practice for the entire set of services or procedures provided during an office visit, distinguished by type of payer (cash, third party insurance, Medicaid, Medicare, etc.).⁵⁶ The physician's output from this visit is represented by the content of the patient's bill, including all the CPT codes associated with that visit. To ensure that the unique combinations of inputs listed on a bill associate with a particular medical condition or surgical procedure, an association which is critical for repricing, the BLS also employs the ICD system, a coding scheme with which physician offices have considerable familiarity.⁵⁷ It is worth noting that the net transactions price by payer type requested by the BLS represents the actual anticipated revenues, including discounts, and not billed charges based on, for example, a "chargemaster."

With this sample bill, the BLS contacts the sampled physician unit each month, and asks it to reprice what the current net transactions prices would be for that particular bundle/payer of services. Thus items on the sample bill remain fixed over time (between major revisions), but item prices could change. Because transaction prices may vary from private payer to private payer, this may present considerable difficulty in practice. Indeed, some payers pay the physician in part or in whole by capitation, thereby making the price for any specified mix of services arbitrary.

PPIs for physicians' services have been published since December 1993. Monthly repricing of physicians' bills presents the BLS with numerous practical difficulties. In some cases, bills are purged from the physicians' accounting systems, and therefore cannot be repriced; this has occurred for about thirty-five (4–5 percent of all) quotes each year. In other cases, the reporter at the sample unit has refused to provide line by line quotes; this has transpired for about twenty-five (3 percent of all) quotes each year (Fixler and Ginsburg, chap. 6 in this volume).

In addition to facing such repricing difficulties at physician practices that continue to cooperate with the PPI, the BLS is operating in an envi-

55. See U.S. Department of Labor, Bureau of Labor Statistics, n.d.-a, 2.

56. According to Fixler and Ginsburg (chap. 6 in this volume), in 1996 12 percent of physician revenues came from Medicaid, 43 percent from private insurance, 18 percent from consumers out-of-pocket, and 27 percent from Medicare.

57. How the pattern of comorbidities is allocated in such cases is not clear.

ronment in which the organization of physician practices has undergone dramatic changes in the last few years as practices have been consolidated and sold to larger provider groups. Thus it is not surprising that sample attrition for physicians' services has been considerable. The impact of this physician practice and bill repricing attrition on the representativeness of the current sample frame is currently unknown.

Finally, in terms of quality change, serious difficulties remain, even with the use of CPT codes. For example, if a new laboratory test becomes available that is more sensitive, reliable, and expensive, yet is used for diagnosis of the same condition and has the same CPT code as its predecessor, it will be considered a price change.⁵⁸ In such a case, quality improvements will not be incorporated. On the other hand, if the laboratory tests are read and examined by less experienced technicians having larger error rates but price is constant, quality declines would be overlooked. Currently the BLS makes no quality adjustments for the physician or laboratory services component of the MPPI.⁵⁹

Hospital Services in the MPPI. We now turn to the PPI for hospital services, which the BLS has published since its December 1992 base period. The hospital services PPI measures anticipated net prices paid to hospitals for the entire bundle of services received during a hospital stay, given the type of payer. The hospital's output is represented by the content of a patient's bill, including all room charges, medical supplies, drugs, and ancillary services provided the patient during a single hospital stay; for an outpatient visit, the hospital output is the anticipated net revenues to be received for medical supplies, drugs, and ancillary charges accruing from a single hospital visit.

As with the PPI for physicians' services, the hospital service PPI attempts to be based on patients' bills that specify the purpose of the hospitalization, as recorded by ICD codes; such an association is important so that repricing is based on a unique combination of inputs listed on the bill with a particular medical condition or surgical procedure. This focus on hospitalization episode for a particular treatment is preferable to pricing based on bed-days, drugs, tests, and other factors irrespective of the patient's illness. To take into account the possibility that price per bed-day is increasing along with a reduction in average length of stay, when repricing the BLS's PPI program now explicitly asks whether there has been a change in average length of stay for the hypothetical price quote. If such a change has occurred, it is treated as a quality change, not simply a price change. As of 1998, the change in average length of stay is the only adjust-

58. If the CPT code changes, either a new bill will be constructed and repriced, or the new and old laboratory test will be linked in.

59. For further discussion of quality adjustments, see Moulton and Moses (1997) and Nordhaus (1998).

ment the hospital PPI makes for quality change (personal communication with Fixler).

In principle, net transactions prices incorporate effects of discounts, and therefore are not “list” or “chargemaster” prices. It is not known what proportion of transactions in hospitals *actually* involve only list prices, list prices less certain adjustments, or capitation, and how this has changed since, say, 1992. Although the BLS clearly seeks to obtain price quotes based on net transactions prices, in a recent General Accounting Office (GAO) report involving the MCPI it was noted that only about 15 percent of the hospital price quotes obtained by the BLS included discounts (U.S. General Accounting Office 1996, 58). In Catron and Murphy (1996), however, it is reported that with the MPPI, 43.4 percent of the sampled inpatient price quotes and 64.6 percent of its outpatient price quotes initially collected in 1992 were based on list prices. As with physicians’ services, capitation for hospitals raises further issues, for it calls into question the whole basis of pricing, because it is based on health plan enrollment rather than use of hospital services by any given patient.

The sampling frame for the hospital services PPI is based on a universe compiled by the American Hospital Association, with the probability of a hospital being sampled being proportional to its revenues.⁶⁰ The sample is stratified on the basis of size (measured by number of beds), public versus private ownership, and type of medical specialty. When initially implemented in 1992, given an expected voluntary response rate of 63 percent (similar to that for other PPI industries), the expected sample size was 558, and the total number of expected monthly price quotes was 2,707. By mid-1997, however, the actual sample size was 42 percent smaller at 322,⁶¹ and the number of quotes was 15 percent smaller at 2,302.⁶²

Once a hospital is identified as a sample unit, at the time of sample initiation the BLS chooses a fixed subset of DRGs, and each hospital is then asked on a monthly basis to report on net transactions prices of a single representative patient bill (typically, the last patient bill on file for that DRG) for each of the randomly assigned DRGs. The DRGs are selected using selection probabilities proportional to expenditures in each DRG based on HCFA and other data from a number of payer sources. Since the identical treatment bundle is not always observed in subsequent

60. As noted by Catron and Murphy (1996), 25, federal hospitals, such as those associated with the military, Veterans Administration, and the National Institutes of Health are excluded from both the CPI and PPI hospital universe, because there are no measurable economic transactions between hospital and patient at these federal hospitals—many services are rendered free to the patient from a budget allocated to a federal entity.

61. Fixler and Ginsburg (chap. 6 in this volume). The breakdown of actual versus expected is 211 versus 358 for general hospitals, 39 versus 75 for psychiatric hospitals, and 72 versus 125 for specialty hospitals.

62. *Ibid.* The breakdown on actual versus expected quotes is 1,602 versus 1,889 for general hospitals, 209 versus 283 for psychiatric hospitals, and 72 versus 125 for other specialty hospitals.

months, BLS reporters construct subsequent hypothetical DRG bundle prices by repricing the identical inputs. BLS notes that when a particular hospital does not perform the targeted DRG service, the hospital can instead provide quotes for several alternative DRGs listed by the BLS on the quote assignment sheet.⁶³ Attrition in the BLS's hospital repricing program is likely to be affected by movement away from DRG billings by hospitals, particularly for non-Medicare patients, and is therefore an important issue worthy of close scrutiny in the very rapidly changing hospital marketplace.

It is also worth noting that in recent years, as hospital length of stay has fallen, the use of post-acute care services such as skilled nursing facilities and rehabilitation units has increased. Often these treatment centers are owned by and even physically located in the hospital. Pricing a hospital stay may present a substantially biased picture of the price of an episode of treatment.

Finally, as noted earlier, the PPI distinguishes as "industries" the "hospital industry" and the "physicians' office industry," largely because the mixes of production processes observed in these two types of establishments are, if not completely disjoint, at least demonstrably not the identical set of production processes. On its own terms, this is clearly reasonable. However, this industry distinction creates a substitution bias with respect to an index for the *purchasers* of health care. Specifically, the problem that arises is that from the purchasers' vantage, the same "product" or service might be "produced" by different industries or by different production processes. For example, with both the physicians' services and hospital services PPI, the nature of the fixed and itemized components for the price quotes requested by the BLS does not permit major input substitution for the treatment of a condition, such as changing the mix of psychotherapy and psychotherapeutic drugs used for the treatment of acute phase depression. When this occurs, even if the industry price indexes are in some sense measured correctly, the PPI measures will miss the purchasers' gain from shifting between different suppliers.

Medical Products in the MPPI: Pharmaceuticals. To this point we have discussed the services component of health care, rather than the goods or commodities components. Although numerous manufacturing products are related to the provision of health care, here we focus on one industry class that has received considerable treatment to date and is perhaps the most significant medical goods industry: prescription pharmaceuticals.⁶⁴

Prescription pharmaceuticals is a relatively research intensive industry

63. U.S. Department of Labor, Bureau of Labor Statistics, n.d.-b.

64. Since sampling and disaggregation procedures for prescription pharmaceuticals are very similar to those in other PPI industries, we do not discuss construction of the pharmaceutical PPI in detail here. See Berndt, Griliches, and Rosett (1993), and the references cited therein, for further discussion.

characterized by a considerable number of new product introductions, and therefore it creates substantial challenges for accurate price measurement. Not surprisingly, the BLS's treatment of prescription pharmaceuticals has long been the subject of controversy. As Dorothy Rice and Loucele A. Horowitz noted thirty years ago (1967, 14) for many years the BLS sample tended to focus excessively on old products: "Until 1960, only three prescribed drugs—penicillin, a narcotic, and a non-narcotic—were included. In that year the list of prescribed drugs was increased to 16 items." Describing the Stigler Commission's Report of 1961, Rice and Horowitz noted (p. 15) that "The Subcommittee urged more prompt introductions of new products—a matter of particular importance in the case of drugs and prescriptions."⁶⁵

More recently, a detailed audit of the BLS's PPI for prescription pharmaceuticals was conducted by Berndt, Griliches, and Rosett (1993), which was updated and extended by Berndt and Greenberg (1995). Although these studies examined transactions at a slightly different point in the distribution chain than does the PPI (transactions from wholesalers to retail drug stores, rather than from manufacturers to their initial customer, typically wholesalers), the Berndt studies raised a number of significant issues. In particular, three important findings from these studies were that (1) the BLS oversampled older goods and undersampled new and middle-aged pharmaceuticals, and (2) prices of older products increased more rapidly than those of products earlier on in their life cycle.⁶⁶ As a result, (3) the BLS overstated prescription drug price inflation, by perhaps as much as 3 percentage points a year over the 1986–91 time period. Corroborating evidence has since been reported by others, including the BLS.⁶⁷

Partly in response to this research, the BLS implemented a new sampling method by which newer products are introduced more rapidly. Specifically, to compensate for the age bias in the BLS prescription pharmaceutical sample, in 1995 the BLS linked in a Supplement I sample of about 49 additional drugs newly approved by the FDA since 1992 (the original 1993 sample had 522 products from ninety-two manufacturers, but attrition to 1995 reduced the 571 to 544), and included these in their sample effective December 1995. As noted by Kelly (1997), the resulting PPI with supplemental sampling rose 2.1 percent in 1996. Had this supplement not been introduced, the PPI would have risen 2.7 percent (based on a BLS

65. The Stigler Commission report is found in U.S. Congress, Joint Economic Activity (1961). Also see U.S. Department of Health, Education, and Welfare (1967), 35.

66. These findings were essentially anticipated almost thirty years earlier by John Gardner, Secretary of Health, Education and Welfare. In his report to the president, Gardner stated "It is difficult to adjust the drug component of the CPI for the rapid changes in the character of the drugs prescribed. By the time a prescription item is incorporated into the index, its price may have fallen to a lower level than in previous years" (U.S. Department of Health, Education, and Welfare 1967, 35).

67. Kanoza (1996), Ristow (1996), and Kelly (1997).

research index); in three of the fourteen months since the introduction of the supplement, price changes in the published index exceeded that of the research index. One year later the BLS constructed and linked in a Supplement II sample, bringing the total number of observations to 561 (after additional attrition). As is noted by Kelly (1997, 17), "In the 14 months since January 1996, the published index has risen 3.3%. Had Supplements I and II not been introduced, the index would have risen 4.1%." Apparently the BLS now plans to add supplements to this industry on an annual basis.

Another area in which the BLS's MPPI has recently made substantial changes involves generic drugs. Until several years ago, the BLS's procedures for its pharmaceutical PPIs treated generic drugs as entirely unrelated to their patented antecedents. Griliches and Cockburn (1994) noted that generic drugs were a special case of the more general "new goods" problem facing statistical agencies such as the BLS. Since the U.S. Food and Drug Administration certifies generics as being therapeutically equivalent to brand name versions of the same chemical entity, conventional problems encountered when valuing new goods are much simpler with generic drugs. Griliches and Cockburn illustrated the empirical significance of linking generic drugs to their patented antecedents (based on an assumed uniform distribution of tastes between brands and generics), and contrasted their preferred price index construction procedure with that employed by the BLS at that time. Based on data for two antibiotic drugs, Griliches and Cockburn showed that with a Paasche approximation to the "true" index, using reservation prices based on the uniform distribution yielded a price index 25 percent lower after two years than a Törnqvist index that introduced generics as quickly as was feasible but treated them as new goods, and was 36 percent lower than an index that mimicked the procedures then employed by the BLS. Several years later, Berndt, Cockburn, and Griliches (1996) extended the Griliches and Cockburn research and showed that for the entire class of antidepressant drugs, the BLS's overstatement of price inflation due to the way it handled generic drugs was more than 4 percentage points per year from 1986 to 1996.⁶⁸

The BLS has announced major changes in how it treats generic drugs in its PPI; these changes are summarized in Kanoza (1996) and Kelly (1997). In particular, effective January 1996, for drugs in the BLS sample losing patent protection and experiencing initial generic competition, the BLS split the fixed weight for that molecule into two parts—64.2 percent for the generic, and 35.8 percent for the brand. Thus the new BLS procedure treated the composite molecule price change as a pure price change.

68. In both the Griliches and Cockburn (1994) and Berndt, Cockburn, and Griliches (1996) studies, transactions were measured at the point of wholesaler to drugstore, and not at the initial point in the distribution chain, which is the focal point for the PPI.

The 64–36 percentages were arrived at as a result of a BLS literature review on typical generic penetrations following the expiration of patent protection. The percentage splits were the same for all molecules and were fixed over time. Beginning with the Supplement II sample introduced in late 1996, however, the BLS brand-generic split was based on actual brand-generic dollar sales, using data purchased by the BLS from IMS America.⁶⁹

There is one other curiosum involving the prescription pharmaceutical PPI. As noted in Berndt, Cockburn, and Griliches (1996), for historical reasons involving preferential federal tax treatment, many U.S. pharmaceutical firms currently manufacture drugs in Puerto Rico; the Puerto Rican value of shipments for prescription pharmaceuticals is roughly 20–25 percent of that on the mainland, and is likely to be higher for newer molecules. For purposes of its PPI calculations, however, the BLS is mandated to treat Puerto Rico as outside the U.S., and thus the PPI excludes all Puerto Rican production.

It turns out that how one deals with Puerto Rican economic accounts differs across government statistical agencies, and even within the BLS. For example, the national income and product accounts of the Bureau of Economic Analysis exclude Puerto Rican production and that of other dependencies, but in the balance of payments accounts, Puerto Rico is treated as domestic. The Census Bureau defines the United States as the U.S. customs territory, which consists of the fifty states, the District of Columbia and Puerto Rico, plus U.S. foreign trade zones and the U.S. Virgin Islands. However, within the BLS's International Price Program (IPP), Puerto Rico is considered part of the United States, and thus currently no IPP price quotes are obtained for Puerto Rican pharmaceutical products shipped to the fifty United States.

The issue of how one treats Puerto Rican production is important to the reliability and interpretation of the prescription drug PPI. If Puerto Rico is to be excluded, as is now the case for the PPI, then to the extent public policy analysts and others seek to track the price growth emanating from U.S. producers (many of whom have chosen to produce significantly in Puerto Rico), it will be necessary to collect and publish “import” price series from Puerto Rico, and then to combine those data with the more narrowly defined “domestic” mainland price series.⁷⁰ Of total pharmaceu-

69. The relative growth rates of the published and research PPIs for the pharmaceutical industry, discussed in several earlier paragraphs, reflect the impacts of incorporating both new generics and new branded products into the sample.

70. One incentive for Puerto Rican production has been section 936 of the Internal Revenue Code, which has provided tax benefits to firms producing in Puerto Rico. It is worth noting that under the omnibus minimum wage bill enacted by the U.S. Congress in 1996, these tax incentives will be phased out over the next decade. Thus it is possible that the empirical significance of this out of scope Puerto Rican production will gradually decline. It is also worth noting that active ingredients of pharmaceuticals could be manufactured in

tical shipments “imported” into the United States from throughout the world, it appears that about 15 percent emanate from Puerto Rico.⁷¹

4.3.4 Medical Care Products and Services in the CPI and MCPI

Medical components of the CPI and PPI programs at the BLS have rather different heritages. It is only within the last decade that the BLS’s PPI has extended coverage to a wide variety of service industries, such as medical care. Thus, construction and design of the recently introduced medical care–related PPIs, such as those for physicians’ and hospital services, have had the opportunity of benefiting from recent thinking and developments on what in fact are the outputs of the service industries, and how one might measure prices in the context of rapidly changing market structure. By contrast, the medical CPI has been published for a very long period of time, regularly since 1935.⁷²

A remarkable fact in the BLS’s medical CPI is summarized in figure 4.1.⁷³ Since 1927, the first year for which MCPI data are available, and for each decade since then, measured medical inflation has generally been greater than that for all goods and services.⁷⁴ Over the entire 1927–96 period, the MCPI has risen at an average annual growth rate (AAGR) of 4.59 percent, almost half again as large as the 3.24 percent for the overall CPI.

Beginning with its January 1998 major revisions, the BLS has regularly published an aggregate medical care Consumer Price Index (MCPI), as well as price indexes for nine of the thirteen item strata in the MCPI. Separate MCPIs are also published for two expenditure groups (medical care commodities and medical care services). The four major subindexes of the MCPI, along with their 1993–95 percentage base period weights within the aggregate MCPI, are prescription drugs (14.9 percent); nonprescription drugs and medical supplies (7.6 percent); professional medical services (also called physicians’ services, although dentists are included; 49.5 percent); hospital and related services (22.9 percent); and health insurance (5.0 percent).⁷⁵ Each of these price indexes is based on consumers’ out-of-pocket (OOP) expenditures including employees’ contributions to employment-based insurance, and thereby excludes all payments by gov-

Puerto Rico, shipped to the domestic United States, and then be encapsulated with inert materials into tablets and capsules in the United States. In such a case, the BLS’s PPI program would consider it as within the scope of the PPI.

71. See table 2 in U.S. Trade with Puerto Rico and U.S. Possessions on the web site <http://www.census.gov/prod/3/98pubs/ft895-97.pdf>. We thank Dennis Fixler of the BLS for providing information on this matter.

72. For historical discussions, see Langford (1957) and Getzen (1992).

73. This table is taken from Berndt et al. (1998).

74. However, for several years within the 1927–46 period, year-to-year changes in the CPI were greater than for the MCPI. See Getzen (1992) for a discussion.

75. Taken from Ford and Ginsburg, chap. 5 in this volume, table 5.1. By December 1997, these relative importance weights were 14.6 percent, 7.2 percent, 50.0 percent, 23.8 percent, and 4.5 percent, respectively.

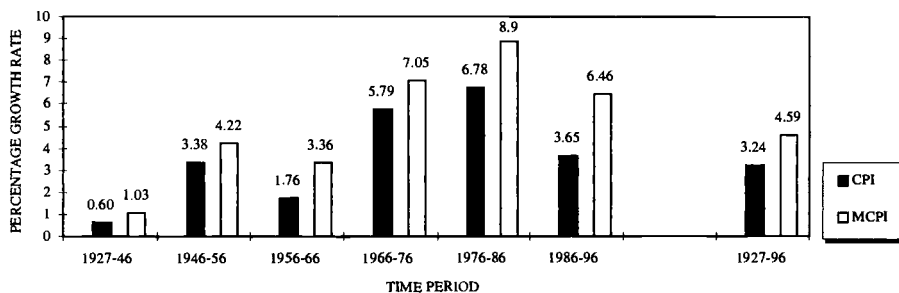


Fig. 4.1 Price inflation in the overall CPI and in the medical CPI, 1927-96

Source: Getzen (1992) and U.S. Bureau of Labor Statistics (<http://stat.bls.gov/blshome.htm>).

ernments and a portion of that from third party insurers. Any health insurance reimbursements for medical services received by a member of the sampled household are netted out to obtain a net out-of-pocket expenditure (Cardenas 1996a). Only that portion of third party insurance paid for out of pocket by consumers (and excluding employers' contributions to employee health insurance) is included within the scope of the MCPI.⁷⁶

However, in constructing weights for the BLS's MCPI, the OOP payments for health insurance are in turn distributed into payments by insurers for medical services, medical commodities, and health insurers' retained earnings.⁷⁷ Analogous to equation (4) earlier, for each MCPI component, OOP expenditures plus the consumer-paid health insurance premium allocation yields a total component weight, which until recently was typically applied to list prices paid by cash-paying customers. Note that over the last decade, actual transaction prices were frequently considerably less than list prices, particularly as discounts to managed care organizations became more common.⁷⁸ To the extent that this occurred, over that time period it is likely that measured MCPIs overstated true price growth. However, particularly more recently, it is possible that discounts have become smaller and less frequent, in which case use of list prices could understate true price growth.

76. We defer additional discussion of OOP expenditure issues to later in this paper.

77. See Fixler (1996), Daugherty (1964), Ford and Sturm (1988), and Getzen (1992). In Ford (1995), for private insurance the allocation is 39.7 percent for hospital services, 28.4 percent for physician services, 5.7 percent for dental services, 0.3 percent for eyeglasses and eye care services, 6.2 percent for services by other medical professionals (including home health care), 6.2 percent for prescription drugs and medical supplies, and 0.6 percent for nursing home care. For Medicare Part B, there is only a four component breakdown: outpatient hospital services, 27.2 percent; physicians' services, 56.8 percent; services by other medical professionals, 9.2 percent; and supplies and durable medical equipment, 6.8 percent. The BLS's treatment of pure health insurance has been criticized by the U.S. Senate Finance Committee (1996).

78. On this see, for example, Dranove, Shanley, and White (1991).

The Item Structure of the MCPI

The basic unit of the hierarchical CPI involves definitions of the item strata. Identifying and defining item strata presents considerable difficulties, particularly when markets are undergoing dramatic change during times within the approximately once-each-decade major CPI revisions.

From January 1987 through January 1997, for example, the CPI hierarchical structure distinguished inpatient hospital services as an item stratum separate from outpatient hospital services. Over this same period of time, cost containment efforts by managed care and other health providers resulted in the transformation of many surgical procedures from inpatient to outpatient hospitalization. By shifting patients from inpatient to outpatient surgeries, hospitals and insurers were frequently able to cut down on total costs. Moreover, the average length of hospital stays declined over the 1987–97 period, as skilled nursing facility days and home health visits were substituted for hospital days.

One consequence of this change in place of service was that the case mix severity in both inpatient and outpatient settings increased, resulting in greater costs for the average case in both settings, even as total inpatient plus outpatient costs decreased. The mean inpatient severity likely increased, because the less complex and critical surgeries were shifted to the outpatient venue, leaving only the more critical and complex surgeries as inpatient. The mean outpatient severity also likely increased over this time, for outpatient surgeries were now being done on a much larger set of more complex patient cases. Total costs of treatment, taking into account the substitution from inpatient to outpatient, were lowered as a result.

It is illuminating to consider price index measurement implications of this cost containment approach employed by managed care. Because the BLS treated inpatient and outpatient hospitalizations as distinct item strata, and because the inpatient to outpatient substitution resulted in greater severity/complexity for both inpatient and outpatient services, price indexes for each item strata grew substantially, and given fixed weights for these item strata, the aggregate hospitalization price index also grew, even as total costs were likely to have decreased. Moreover, the CPI, though not the PPI, priced hospital days. Their average severity also grew, due to shorter stays. Although empirical evidence is not available, we conjecture that over the January 1987–January 1998 period, the BLS's measured CPI inflation for hospitalization considerably overstated true hospital inflation, because it failed to account for substitution from inpatient to outpatient, and also failed to account for treatments involving greater severity of case mix in each component. This overstatement is consistent with the increased spread of the MCPI over the CPI between 1986 and 1996 (fig. 4.1) at a time when increased price competition should have decreased the spread.

The BLS has recognized the problem, and in January 1997, one year before its major CPI revisions introduced in January 1998, it began treating the aggregate of hospital inpatient and outpatient services as a single item stratum. It has also shifted to measuring hospital services by the stay rather than by the day, and it has classified inpatient and outpatient hospital services as substratum indexes (Ford and Ginsburg, chap. 5 in this volume) similar to the ELIs discussed in section 4.3.1. Information is not available, however, on how linking is implemented when, for example, a shift occurs from inpatient to outpatient surgery. Simple redefinition will not fully address the problem of inpatient-outpatient substitution unless a satisfactory linking procedure is developed and implemented as well.

A number of other important changes have recently been introduced into the CPI for hospital services, even before the 1997 and 1998 revisions. Until at least 1990, for example, in most cases procedures for the MCPI involved pricing specific input items at list prices, like chargemaster fees for x-rays, laboratory tests, and physicians' office visits rather than at the average actual charge for treatment of, say, a child's forearm fracture to a managed care organization obtaining a hospital discount.⁷⁹ According to Cardenas (1996b), since 1993, when redrawing outlet and item samples, the BLS has attempted to obtain quotes from hospitals for specific payers, thereby seeking to obtain transaction rather than list prices. Cardenas (1996a, 40) reports that "Employing the sample rotation construct as the vehicle for increasing the number of transaction prices in the CPI, however, has yielded slow progress to date." According to the 1996 GAO study cited earlier, only about 15 percent of the CPI hospital price quotes obtained by the BLS included discounts (U.S. General Accounting Office 1996, 58).

Obtaining transaction rather than list prices is not an easy task, particularly because with price discrimination and alternative pricing methods currently in the medical marketplace, there are frequently many transaction prices. Consider, for example, hospital services. Some insurers pay for medical care on a per diem basis—one price per day to cover all services provided. Other insurers pay on a DRG basis—one price per admission, differentiated only by the severity of the admission. Still other insurers pay on a capitated basis—one price per patient per year, independent of the amount of services the patient actually receives. Because the market has not settled on one basis of price, appropriate price indexes must be able to handle payments using all of these methods. Obtaining the transaction prices for all three methods will be difficult, particularly because transaction prices are frequently considered highly proprietary and confidential

79. For further discussion, see Armknecht and Ginsburg (1992); Cardenas (1996b); Daugherty (1964); Ford and Sturm (1988); and Ginsburg (1978).

by insurers. Moreover, because health plans have different degrees of bargaining power, the same provider may negotiate varying per diem rates with different health plans. These problems are not unique to the MCPI, but are relevant for the MPPI as well. Cooperation and joint efforts by the MCPI and MPPI programs in securing price quotes could be very fruitful.

Other recent changes implemented by the MCPI for hospitals involve item descriptions. At one extreme, one can assume zero substitutability among medical care goods and services for treatment of a condition, and simply take quotes of discrete hospital goods and services. An alternative, discussed above, is to employ DRGs. Although the BLS apparently employed non-Medicare DRG prices in three states beginning in 1990, two of those states have since terminated their state-regulated DRG programs. As of September 1992, approximately 6 percent of the CPI hospital quotes consisted of DRG descriptions (Cardenas 1996b). According to Cardenas (1996a, 40), use of DRGs is problematic, because "a DRG treatment path can be wide-ranging, contingent upon the treating physicians' approach"; for example, coefficients of variation range from 0.30 to over 1.5, thereby indicating considerable variation in the treatment strategies used to treat a case as defined by a DRG (Frank and Lave 1985). Currently the BLS is instead considering use of a "package" treatment, consisting of "highly standardized and tightly defined components and risk factors" for conditions such as appendectomies, tonsillectomies, and cataract surgery. Details on how such treatment packages would be defined and how representativeness would be ensured have not been released, nor have any data concerning the composition and nature of hospital quotes being obtained by the BLS's MCPI since the major revisions of January 1998.

Our MCPI discussion to this point has focused on hospital services. We have not seen comparable literature dealing with MCPIs for physicians' services, although informal conversations with BLS personnel suggest to us that issues of item description, list versus transaction prices, and lack of quality adjustment are similar for physicians' and hospital services.

Like the MPPI, the MCPI has a prescription drug component. Issues discussed earlier in the context of the MPPI concerning the linking of prices of newly entering generic drugs, just after branded drugs lose patent protection, to prices of their pioneer antecedents apply here as well. The MCPI program implemented changes involving generic drugs earlier than the MPPI. Effective January 1995, procedures involving the MCPI prescription drug treatment of generics changed considerably.⁸⁰ For branded drugs in the CPI sample losing patent protection, six months after patent expiration the BLS now follows a procedure whereby branded and generic

80. See Armknecht, Moulton, and Stewart (1994), and U.S. Department of Labor, Bureau of Labor Statistics, "Improvements to CPI Procedures: Prescription Drugs," n.d.-c.

versions of the molecule are randomly selected, in which the probability of selection is proportional to the sales of each version of the drug during the sixth month. If a generic substitute is selected, the entire price difference between the original drug and its generic substitute is treated as a price change. Obviously, if the branded version is selected, repricing will continue as before. Drugs entering the CPI sample after their patent has expired would of course not be affected by this new procedure, because during the CPI sample rotation process generic versions would also have had a chance of being selected. Note that use of a six-month window is somewhat problematic, for in many cases considerable additional diffusion of generics occurs beyond the six months immediately following patent expiration.⁸¹

Finally, regarding sample sizes for the MCPI and its components, as of February 1998 the total number of MCPI current price quotes was 7,676. This was broken down as follows: prescription drugs, 694; over-the-counter drugs, 423; nonprescription medical equipment and supplies, 252; physicians' services, 1,051; dental services, 927; eye care, 290; services by other medical professionals, 365; hospital services, 3,236; and nursing home services, 438 (Ford and Ginsburg, chap. 5 in this volume, table 5.4).

Weighting Issues in the CPI and MCPI

From January 1987 until January 1998, the item strata weights employed by the BLS in its CPI program were those based on the 1982–84 CEX; beginning in January 1998, the new weights are those based on the 1993–95 CEX. Thus weights used just before the most recent CPI revision were about fifteen years out of date, and the newly introduced “current” weights were already almost four years out of date at the time of unveiling. Up-to-date weights are particularly important in the case of medical care, where technological change may result in substantial shifts across weighting categories. For example, Cutler et al. (1998, chap. 8 in this volume) compare the old CPI medical methodology (pricing the hospital room rate and other hospital inputs with weights held fixed over a long time interval) with (1) a price index that priced the inputs but reweighted annually, and (2) a price index that was based on the cost of treating heart attacks. The quantitative impact on the price index from annual reweighting was greater than the impact of moving from pricing medical inputs to pricing the cost of treating heart attacks.

In some goods and services markets characterized by relative tranquility and stability, it is possible that use of old weights in price index construction would not be problematic. In the health care goods and services

81. See, for example, Berndt, Cockburn, and Griliches (1996), table 2, 152.

markets, however, the last fifteen years—indeed, the entire post–World War II era—have been marked by dramatic changes in the number and quality of products offered and consumed, the identity of the payers (cash vs. third-party private or government payer), and in how and by whom the services are provided (e.g., from inpatient to outpatient hospitalization, and from fee-for-service to managed care). The pace of both institutional and technological change has been particularly rapid in the health care sector. Moreover, the role of health care expenditures in the overall consumer budget has changed considerably, in part because the BLS's measured MCPI has increased much more rapidly than that for the all-item CPI (6.46 percent for the MCPI 1986–96, 3.65 percent for the all-item CPI over the same time period).⁸² We now examine some of the implications of these changes for CPI and MCPI measurement.

In the CPI hierarchical system used from January 1987 until January 1998, seven major product categories were represented, and in the 1998 revisions an eighth was added. In column (1) of table 4.2 we present 1982–84 CEX-based weights for the seven major product categories when they were originally introduced into the 1987 revision of the Consumer Price Index. As is seen there, when the 1987 basket was introduced, the Medical Care major product category received a weight of 4.80 percent. Because the BLS's measured price of medical care rose more rapidly than that of the overall CPI, the implicit budget share consistent with fixed 1982–84 base period quantity weights (inflating all base period quantities by CPI-measured price changes) increased over time; as is seen in column (2), by December 1995 the implicit relative importance of medical care increased to 7.36 percent. This raises a number of very important issues.

First, data from other government agencies, such as the Health Care Financing Administration, indicate that national health expenditures as a proportion of GDP are much higher than 7 percent; for example, Levit et al. (1998) report that in 1996, this proportion was 13.6 percent. Why is the CPI weight for medical care so low?

One important reason for this difference is that the MCPI weight reflects only a portion of total medical care outlays. Specifically, the MCPI weight incorporates only direct OOP cash outlays, plus direct household purchases of health insurance (including Medicare Part B), plus employee contributions to health insurance premiums purchased through work. Significantly, the MCPI excludes employer health insurance premium contributions, treating them as a business expense; MCPI also excludes Medicare Part A, 75 percent of Medicare Part B (the fraction paid from general government revenues), and Medicaid outlays. More generally, the MCPI excludes all medical services purchased by the government on behalf of

82. For a discussion of some of these changes, see Berndt et al. (1998).

Table 4.2 Major Product Groups of Items in the CPI: 1982–84 Weights, Implicit Relative Importance, and 1995 Actual Budget Shares

Product Group	1982–84 Weights in 1987 Revision (%) (1)	Implicit Relative Importance 1995:12 (%) (2)	1995 CEX Budget Share (%) (3)	Implicit Relative Importance 1997:12 (%) (4)
Food and beverages	17.84	17.33	15.57	16.31
Housing	42.64	41.35	44.37	39.56
Apparel and upkeep	6.52	5.52	5.57	4.94
Transportation	18.70	16.95	18.47	17.58
Medical care	4.80	7.36	5.21	5.61
Entertainment	4.38	4.37	4.78	n.a.
Recreation	n.a.	n.a.	n.a.	6.14
Education and communication	n.a.	n.a.	n.a.	5.53
Other goods and services	5.13	7.12	5.74	4.32
Total	100.01	100.01	100.00	99.99

Sources: (1) U.S. Department of Labor, Bureau of Labor Statistics (1987), fig. 1, all urban consumers; (2) U.S. Department of Labor, Bureau of Labor Statistics (1996b), all urban consumers; (3) U.S. Department of Labor, Bureau of Labor Statistics, *Consumer Expenditure Survey, 1995*, Table 1300; (4) U.S. Department of Labor, Bureau of Labor Statistics (1997c), table 1 (New Series), CPI-Urban.

Note: n.a. = not available.

its citizens/residents, and weights and prices only those components paid for out of pocket by consumers or from payroll deductions borne by employees.⁸³ Given this conceptual foundation of the MCPI, it is therefore not surprising that the MCPI weight is much smaller than the share of national health expenditures in GDP.⁸⁴

Another issue is whether the implicit relative importance of the medical care component in the CPI (col. [2] of table 4.2) accurately portrayed actual average consumer budget shares in 1995. If the 1982–84 fixed quantity weights provide a poor approximation to actual quantity weights in, say, 1995, then these implicit relative importance percentages could be unreliable and inaccurate as well, thereby compromising the accuracy of the measured CPI and MCPI. Thus it is of interest to compare actual budget shares with implicit relative importance percentages based on fixed weights.

Actual average budget share data based on the 1995 CEX, in which budget shares are weighted averages over the various geographical areas making up the BLS sample, are presented in column (3) of table 4.2. As is seen there, the 1995 average budget share for medical care items is 5.21 percent, which is substantially smaller—2.15 percentage points, about 29 percent—than the implicit relative importance of medical care items (7.36 percent) based on the BLS's fixed 1982–84 quantity weights; alternatively, by 1995 BLS's use of the fixed weight index in its CPI resulted in the implicit relative importance of medical care being about 41 percent larger (7.36 percent vs. 5.21 percent) than was warranted.

The implicit relative importance of the eight major CPI components in the recently revised CPI, based on the 1993–95 CEX and updated to December 1997, are given in the final column of table 4.2. Interestingly, the new relative importance of medical care is 5.61 percent. An implication of this is that because updated data from the 1993–95 CEX replaced outdated data from the 1982–84 CEX, with the January 1998 revisions the weight given medical care fell 1.75 percentage points from 7.36 percent to 5.61 percent, a relative overstatement of 31 percent. This overstatement of the health care relative importance is greater in the 1998 major revision than it was for the major revision eleven years earlier in 1987. Then, as reported by Ford and Sturm (1988, table 1), the corresponding overstatement in December 1986 was 5.74 percent versus 4.66 percent, at 23 percent still substantial but considerably smaller than the 31 percent overstatement in 1998.

There are at least three reasons why the actual budget shares could di-

83. For further discussion, see Armknecht and Ginsburg (1992), particularly 124–42.

84. The appropriateness of this decomposition into employee out-of-pocket versus employers' contributions depends in part on the incidence of the income tax, and the extent to which employees are willing to substitute employers' health insurance contributions for other forms of wage and nonwage compensation. While very important, these issues are beyond the scope of this review. For a recent discussion, see Gruber (2000) and Pauly (1997).

verge so materially from implicit relative importance based on fixed quantity weights. First, the relative quantity weights could have changed over time, reflecting nonzero price substitutability inconsistent with the Laspeyres fixed weight assumption. For example, it is possible that efforts by managed health care organizations to contain medical expenditures have resulted in physicians' and hospitals' performing a smaller number of laboratory tests, scheduling fewer specialist physician visits, and shortening lengths of hospital stay. Hence it is possible that as a result of growth in managed care and other cost containment methods, the relative quantities of medical care items for which consumers made OOP expenditures has fallen since 1982–84.

Second, suppose that demand for health care had a zero price elasticity of demand. In such a case, the divergence would simply reflect overstated medical care price inflation, perhaps from failure to measure transaction prices accurately.

Third, if, however, the demand price elasticity for medical care were greater than unity (say, particularly for those components undergoing dramatic but not fully measured quality change), then the implicit relative importance of these items would be greater than the actual budget share, *ceteris paribus*.

Discovering which of these three reasons, or what weighted combination, contributed to the divergence between the actual 1995 budget shares and implicit relative importance requires additional empirical research. Econometric studies of demand for health care such as those based on the RAND Health Insurance Experiment report modest but price-inelastic demand; it is worth noting that the experimental design of that study in effect controlled for quality variations (Newhouse et al. 1993). Additional research that focused on price measures incorporating quality change, and then evaluated the responsiveness of demand to quality changes, would be useful.⁸⁵

These discrepancies between actual budget shares and implicit relative importance values, resulting from the use of outdated CEX surveys, suggest that more frequent weighting could considerably strengthen the reliability of the MCPI. The frequency of such revisions does not necessarily need to be uniform across the entire CPI, but could involve more frequent updatings in some major product groups (such as medical care) than in others (such as housing). For the rapidly changing medical care sector, decennial updates of weights result in price indexes whose accuracy and reliability can legitimately be called into question.⁸⁶

85. For discussion and references, see Ellison et al. (1997).

86. Suggestions for implementing alternative weighting schemes with time-varying weights have been proposed and evaluated by Shapiro and Wilcox (1997).

4.4 Related Research on Medical Care Price Indexes

The average consumer of medical care is not as interested in the price of a visit or a hospital day as he is in the total cost of an episode of illness.

—U.S. Department of Health, Education, and Welfare (1967)

For quite some time now, health economists and government statisticians have made recommendations concerning directions toward which the pricing of medical care services should move, particularly concerning the definition of the item or product that is to be priced. For example, already in 1964 Anne Scitovsky proposed “an index which would show changes, not in the costs of such items of medical care such as drugs, physicians’ visits, and hospital rooms, but in the average costs of the complete treatment of individual illnesses such as, for example, pneumonia, appendicitis, or measles” (p. 133).⁸⁷ In Scitovsky (1967), this approach was implemented on an illustrative basis for five medical conditions. Notably, in the 1950s and 1960s the BLS price indexes appeared to have *understated* medical price inflation, in large part because physicians’ “customary” pricing in an environment of extensive price discrimination began to change as the proportion of patients covered by Medicare insurance increased.⁸⁸ Hence, the BLS’s alleged upward bias in measuring medical price inflation has not always been the indictment.

Shortcomings in the BLS’s MCPI approach, and preference for the treatment episode–outcomes-adjusted approach to price measurement, have appeared steadily since 1967; see, for example, the chapter “Measuring Changes in the Price of Medical Care” in various editions of a well-known health economics textbook by Paul Feldstein (1979, 1983, 1988) as well as the Baxter Foundation Prize Address by Newhouse (1989).

More recently, price indexes for several specific medical treatments, taking outcomes changes into account, have been constructed, thereby demonstrating again the feasibility and importance of the Scitovsky approach. Using one data set of hospital claims from a major teaching hospital and another very large data set consisting of Medicare claims, Cutler et al. (1998, chap. 8 in this volume) have contrasted input price indexes for the cost of heart attack treatment that rise by 6.7 percent over 1983–94, with an outcomes-adjusted index that takes into account changing treatment regimens and a conservative valuation for the extension of life expectancy attributable to new heart attack treatments; the latter price index increases by only 2.3 percent per year (in real terms, an annual decrease of 1.1 percent), implying a net upward bias of 4.4 percent per year for an MCPI-like index.

87. See Scitovsky (1964) and related discussions in Scitovsky (1967).

88. For further discussion, see Feldstein (1969, 1970).

Similarly, Shapiro and Wilcox (1996) have constructed a price index for cataract surgery, 1969–93, and find that a CPI-like fixed weight input-based price index increases by a factor of about nine; a preferred alternative price index incorporating realized reduced levels of hospital services (input changes), but ignoring any improvements in the quality of medical outcomes, increases by only a factor of three, implying an annual differential of 4.6 percent.

A number of other studies, based on retrospective medical claims data, provide additional evidence that implementation of disease or condition-specific measurement procedures that uses treatment episodes of care as a measure of output, is in fact feasible; see, for example, other papers presented at this conference by Berndt, Busch, and Frank (chap. 12 in this volume), Cockburn and Anis (chap. 11 in this volume), Cutler et al. (chap. 8 in this volume), and Shapiro, Shapiro, and Wilcox (chap. 10 in this volume).

4.5 A New Medical Care Expenditure Price Index Based on Episode Treatment Costs

One could envision an ideal medical care price index as providing accurate and reliable measures for use in at least five very important functions: (1) the measurement of quality of life, (2) the deflation of nominal industry output for the calculation of real output and productivity growth, (3) the indexing of health care benefits as a component of employee compensation, (4) the indexing of payments by health plans to providers of medical care, and (5) the indexing of payments in government transfer programs. Undoubtedly, additional purposes can be envisaged. Unfortunately, these various functions and purposes are very different, and there is no way a single index like the medical CPI (or PPI) can provide an accurate and reliable basis for such diverse needs. The search for a single price index that meets all these purposes is a futile one. But these diverse needs are real and important. We recommend that, rather than trying to change dramatically the conceptual foundations and measurement procedures of the MCPI and MPPI in an attempt to accommodate conflicting needs, the BLS consider constructing and publishing, on an experimental basis, a new price index that we tentatively call a *medical care expenditure price index*.

As we have discussed in considerable detail, the CPI and PPI medical price indexes are very different, they correspond to distinct index number concepts, and thus the appropriate uses to which they are applied must differ as well. The CPI is, in concept, a fixed weight approximation to a cost-of-living index, in which the cost-of-living index is defined as follows: What is the minimum change in expenditure necessary to purchase the set of market goods and services yielding the same standard of living as the set of market-purchased goods and services consumed in the base period?

The manner in which the BLS has implemented this cost-of-living definition in the case of the medical care CPI is to define the scope of the index to apply only to out-of-pocket expenditures. The reasoning is that employer-provided medical insurance is a nonwage part of compensation; BLS does not believe it to be appropriate to add consumption out of nonwage compensation into the consumer expenditures that are defined, implicitly, to be relevant to the wage part of compensation.⁸⁹

Nevertheless, even if the CPI is continued to be defined to include only out-of-pocket expenditures, there are many important purposes for which one needs a price index covering all medical expenditures, no matter who (consumer, employer-provided health insurance, or government) is the nominal payer. This, for example, would be the concept of price change that one would want for most policy analytic purposes, such as containing medical care cost inflation, or examining the impact of new treatment technologies.

The PPI organizes and presents information by medical care *industry*, that is, hospitals, physicians' offices, nursing homes, pharmaceuticals, and so forth. The underlying PPI concept is an industry output price index. This index is useful for a number of purposes, such as comparing hospital price movements with the cost of hospital inputs (though one of the great weaknesses of the U.S. statistical system is its inadequacy of information on industry input quantities and input prices). Moreover, the PPI is a price index for *domestic* industries. It provides, for example, information about price movements for domestically produced pharmaceuticals at the manufacturer's level. But the PPI is not a price index for all pharmaceuticals consumed in the United States. It excludes, for example, imported pharmaceuticals and also, because of a definitional oddity in the U.S. national accounts, pharmaceutical production in Puerto Rico. Additionally, the PPI includes pharmaceuticals and medical devices that are produced in the United States and sold abroad.

Thus, just as the CPI does not provide a comprehensive price index for health care to U.S. purchasers, neither does the PPI provide this information. Even though the CPI and PPI measures are useful on their own terms (and we are not asserting that these measures are not useful or appropriate ones), there is a great gap in medical care price information. The missing part, regrettably, is probably the part that is most vital for medical care policy analysis; namely, the United States needs a comprehensive medical care price index for *expenditures* on medical care. Such a *medical care expenditure price index* would apply to all purchases of medical care, and it would take into account, as the present CPI and PPI do not, substitution by buyers or financiers of medical care across providers or industries that

89. For discussion of the incidence of these employer-subsidized health benefits, see Pauly (1997) and Gruber (2000).

produce medical care. In principle, separate medical care expenditure price indexes could be constructed for public and private sector expenditures, and for the elderly. The medical care expenditure price index would cover all consumption of medical care goods and services, be the providers/producers domestic or foreign. And it would, we believe, be profitably structured around determining the costs of treating an episode of a representative set of illnesses or conditions.

4.6 Where Do We Go from Here?

It is striking that a variety of researchers and government statisticians all agree that price measurement of medical care goods and services is particularly difficult, that current methods of measurement can be improved by measuring episode treatment costs and taking outcomes into account, and that research in this area should receive a high priority. For example, about six years ago BLS economists Paul Armknecht and Daniel Ginsburg (1992) stated, "Although it will be difficult to develop the methods for pricing total treatment, BLS is planning further research in this area" (pp. 141–42). Later, Armknecht (1996) added: "A new dimension needs to be included in the pricing of medical services that includes outcomes, so that if cancer treatment results in improved survival rates, this is reflected in the index. This area appears to have promise and is one that also needs to be pursued as part of the research agenda in quality adjustment" (p. 33). The Boskin Commission (Boskin et al. 1998) made similar recommendations, stating that "This new research . . . opens up the potential for major improvement in our understanding of the economics of medical care. This category should receive a substantial component of the CPI's future research investment, and we strongly endorse a move in the CPI away from the pricing of health care inputs to an attempt to price medical care outcomes. . . . This program should explore measuring the value of time saved by new medical procedures and communication devices, the value of life extended and its associated quality . . ." (pp. 60, 84).

The BLS's response to the Boskin Commission recommendations suggests considerable agreement concerning measurement problems in the medical care component of the CPI: "Although we acknowledge that there have been enormous improvements in medical technology over time, we also note the heterogeneity of the medical services category, which includes services as diverse as dentistry, eyeglasses and eye care, psychological counseling, podiatry, chiropractic, and physical therapy. . . . Some kinds of quality change are difficult to evaluate, involving changes in patient outcomes, such as improved mortality or reduction in pain. The BLS is continuing to support and encourage research on this topic. . . . The BLS recognizes the importance of the health insurance price movements to consumers as well as to policymakers and will continue to search for

ways to overcome the obstacles to accurate adjustment for changes in policy characteristics.”⁹⁰

But where do we go from here? In the previous section we put forward a recommendation that the BLS consider constructing and publishing, on an experimental basis, a medical care expenditure price index. What new research, and what new related measurement initiatives, appear to offer the most promise if such a new price index is to be created?

A number of directions appear promising. First, although most of the public attention in recent years has focused on the CPI, as we have noted several times in this paper, many measurement issues involving medical care goods and services are common to the CPI and PPI. A medical care expenditure price index would have elements from both the CPI and PPI programs. An obvious implication is that the CPI and PPI programs should be encouraged to expand their joint efforts on designing research and measurement initiatives.

Second, as has been emphasized by, among others, Triplett (1999, chap. 1 in this volume), complementary research efforts on health care outcomes by medical researchers involving cost-effectiveness analyses, as well as the public availability of large retrospective health claims databases, now allow the BLS to build on others’ research that defines and identifies episodes of treatment. This research would be particularly important were the BLS to initiate a medical care expenditure price index program. Note that in principle, outcomes research can help somewhat in overcoming the moral hazard problem underlying the failure of revealed preferences as measures of willingness to pay in medical care markets. Together with retrospective claims data, the outcomes studies provide a framework for identifying medical care outputs that incorporate quality change. What Anne Scitovsky proposed in 1962 and illustrated with a small sample of conditions in 1967, and what Health, Education, and Welfare Secretary John Gardner requested more generally in 1967 (U.S. Department of Health, Education, and Welfare 1967), is clearly possible on a much larger scale today.

Third, although the usual source of information for output measurement is based on actual market transactions, use of medical outcomes data to define measures of output implies an adjustment in thinking—to look outside of market transactions to consider what medical resources actually do for health.⁹¹ A medical care expenditure price index program should, to as great an extent as is feasible, combine actual transactions data underlying treatment costs of an episode of an illness, with outcomes data from cost-effectiveness and related medical studies.

Fourth, it is likely that treatment episode price index measurement will

90. U.S. Department of Labor, Bureau of Labor Statistics (1997b), 18, 19, and 31.

91. For further discussion, see Triplett (1999, chap. 1 in this volume).

need to be done at a very disaggregated level of detail for a finite number of representative illnesses or conditions. The extent of medical care progress, as well as the underlying increases in medical scientific knowledge, have varied considerably across illnesses and disorders, with spectacular gains in treating conditions such as cataracts, retinal detachment, schizophrenia, and cystic fibrosis, but with apparently less progress for other conditions such as rheumatoid arthritis, Alzheimer's disease, and the common cold. While the Hicksian aggregation assumption of common proportional price changes over time across a variety of products may be a useful approximation within a number of other industries, for medical care it is not plausible. As suggested already in 1969 by Martin Feldstein, for the BLS to obtain useful measures of medical care output, it would appear to be most useful to obtain a sample of "a representative mix of illnesses" (p. 363). Research that helps identify an appropriate mix of illnesses and their treatments, ones for which outcomes measures and/or published treatment guidelines are available, and ones for which sample sizes in retrospective claims databases are sufficiently large, would seem to be particularly helpful.

Fifth, the pricing of health insurance remains a difficult issue. As noted above, empirical implementation needs theoretical foundations on how employers and employees choose health insurance, and on that issue there has not been much theoretical effort of late. This area seems particularly fruitful for economic theorists, since potentially complementary data are now becoming available. With funding from the AHRQ, the Census Bureau is now in the midst of collecting data for the Medical Expenditure Panel Survey (MEPS).⁹² This is not a one-time survey, but instead involves a panel of 9,500 households, and links the households to the employers, the employers' set of health insurance plans, and the health care providers. If theoretical foundations can be established, it is possible that this database could provide the basis for ongoing hedonic price analyses of health insurance plans.⁹³

However, in terms of measuring the price of health insurance, the growth of managed care makes measurement and the interpretation of hedonic price equations more problematic. In most nonmedical markets, if price increases, it is plausible to assume that those consumers rationed out of the market had a willingness to pay that was less than the now higher price. Although in some cases one might be able to argue that with fee-for-service medical care rationing was reflective of willingness to pay, with managed care consumers have no way by which to reveal their will-

92. Details on the MEPS are given in Cohen et al. (1996).

93. See Jensen and Morrissey (1990) for a hedonic price analysis of group health insurance plans.

ingness to pay at the time services are delivered. The rationing that occurs is instead based on clinical grounds. Thus the growth of managed care makes it difficult to derive economic welfare implications from the estimation of hedonic price equations.⁹⁴ This leads us to suggest that in the near future, research on the construction of a medical care expenditure price index might be fruitfully based on retrospective claims data, although research on the pricing of health insurance policies is also likely to be important.

References

- Abraham, Katharine G., John S. Greenlees, and Brent R. Moulton. 1998. Working to improve the Consumer Price Index. *Journal of Economic Perspectives* 12 (1): 27–36.
- American Medical Association. 1990. *Current procedural terminology (CPT)*. 4th ed. Chicago: American Medical Association.
- Archibald, Robert B. 1977. On the theory of industrial price measurement: Output price indexes. *Annals of Economic and Social Measurement* 6 (1): 57–72.
- Armknrecht, Paul A. 1996. Improving the efficiency of the U.S. CPI in the future. Washington, D.C.: International Monetary Fund. Unpublished manuscript, March.
- Armknrecht, Paul A., and Daniel H. Ginsburg. 1992. Improvements in measuring price changes in consumer services: Past, present and future. In *Output measurement in the service sectors*, ed. Zvi Griliches, 109–56. NBER Studies in Income and Wealth, vol. 56. Chicago: University of Chicago Press.
- Armknrecht, Paul A., Brent R. Moulton, and Kenneth J. Stewart. 1994. Improvements to the food at home, shelter and prescription drug indexes in the U.S. Consumer Price Index. CPI Announcement—Version I, 20 October. Washington, D.C.: U.S. Department of Labor Bureau of Labor Statistics.
- Berndt, Ernst R. 1991. *The practice of econometrics: Classic and contemporary*. Reading, Mass.: Addison-Wesley.
- Berndt, Ernst R., Iain M. Cockburn, Douglas L. Cocks, Arnold Epstein, and Zvi Griliches. 1998. Is price inflation different for the elderly? An empirical analysis of prescription drugs. NBER Working Paper no. 6182. Cambridge, Mass.: National Bureau of Economic Research.
- Berndt, Ernst R., Iain Cockburn, and Zvi Griliches. 1996. Pharmaceutical innovations and market dynamics: Tracking effects on price indexes for antidepressant drugs. *Brookings Papers on Economic Activity: Microeconomics*, 133–88.
- Berndt, Ernst R., David M. Cutler, Richard G. Frank, Zvi Griliches, Joseph P. Newhouse, and Jack E. Triplett. 2000. Medical care prices and output. In *Handbook of health economics*, ed. Joseph P. Newhouse and Anthony J. Culyer. Amsterdam: Elsevier Science.

94. For an example of ambiguous welfare interpretations of increases in a medical CPI (the treatment of depression), see Berndt, Busch, and Frank (chap. 12 in this volume), who argue that increases in patient copayments might also reflect reduced moral hazard, and thus an indeterminate net effect on patient welfare.

- Berndt, Ernst R., and Paul E. Greenberg. 1995. An updated and extended study of the price growth of prescription pharmaceutical preparations. In *Competitive strategies in the pharmaceutical industry*, ed. Robert B. Helms, 35–48. Washington, D.C.: American Enterprise Institute.
- Berndt, Ernst R., Zvi Griliches, and Joshua G. Rosett. 1993. Auditing the Producer Price Index: Micro evidence from prescription pharmaceutical preparations. *Journal of Business and Economic Statistics* 11 (3): 251–64.
- Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson. 1998. Consumer prices, the Consumer Price Index, and the cost of living. *Journal of Economic Perspectives* 12 (1): 3–26.
- Bresnahan, Timothy F., and Robert J. Gordon. 1997. *The economics of new goods*. NBER Studies in Income and Wealth, vol. 58. Chicago: University of Chicago Press.
- Cardenas, Elaine M. 1996a. The CPI for hospital services: Concepts and procedures. *Monthly Labor Review* 119 (7): 34–42.
- . 1996b. Revision of the CPI hospital services component. *Monthly Labor Review* 119 (12): 40–48.
- Catron, Brian, and Bonnie Murphy. 1996. Hospital price inflation: What does the new PPI tell us? *Monthly Labor Review* 119 (7): 24–31.
- Cohen, Joel W., Alan C. Monheit, Karen M. Beauregard, Steven B. Cohen, Doris C. Lefkowitz, D. E. B. Potter, John P. Sommers, Amy K. Taylor, and Ross H. Arnett III. 1996. The medical expenditure panel survey: A national health information resource. *Inquiry* 33 (winter): 373–89.
- Council on Wage and Price Stability. 1977. *The Wholesale Price Index: Review and evaluation*. Washington, D.C.: U.S. Government Printing Office.
- Cutler, David M., Mark B. McClellan, Joseph P. Newhouse, and Dahlia Remler. 1998. Are medical prices declining? *Quarterly Journal of Economics* 113:991–1024.
- Daugherty, James C. 1964. Health insurance in the revised CPI. *Monthly Labor Review* 87 (11): 1299–1300.
- Diewert, W. Erwin. 1983. The theory of output price index and the measurement of real output change. In *Price level measurement: Proceedings from a conference sponsored by Statistics Canada*, ed. W. Erwin Diewert and Claude Montmarquette, 1049–1113. Ottawa: Statistics Canada.
- Dranove, David, Mark Shanley, and William D. White. 1991. Does the Consumer Price Index overstate hospital price inflation? *Medical Care* 29 (August): 690–96.
- Eldridge, Lucy P. 1998. The role of prices in measuring productivity for the business sector of the U.S. economy. Washington, D.C.: U.S. Bureau of Labor Statistics, Office of Productivity and Technology. Draft manuscript, 29 July.
- Ellis, Randall P., Gregory C. Pope, Lisa I. Iezzoni, John Z. Ayanian, David W. Bates, Helen Burstin, and Arlene S. Ash. 1996. Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financing Review* 17 (3): 101–28.
- Ellison, Sara Fisher, Iain Cockburn, Zvi Griliches, and Jerry Hausman. 1997. Price competition among pharmaceutical products: An examination of four cephalosporins. *RAND Journal of Economics* 28 (3): 426–46.
- Feldstein, Martin S. 1969. Improving medical care price statistics. *1969 Proceedings of the Business and Economics Statistics Section*, 361–65. Washington, D.C.: American Statistical Association.
- . 1970. The rising price of physicians' services. *Review of Economics and Statistics* 52 (2): 121–33.
- Feldstein, Paul J. 1979. *Health care economics*. New York: Wiley.
- . 1983. *Health care economics*. 2d ed. New York: Wiley.
- . 1988. *Health care economics*. 3d ed. New York: Wiley.

- Fisher, Franklin M., and Zvi Griliches. 1995. Aggregate price indexes, new goods, and generics. *Quarterly Journal of Economics* 110 (1): 229–44.
- Fisher, Franklin M., and Karl Shell. 1972. The pure theory of the national output deflator. In *The economic theory of price indexes*, ed. Franklin M. Fisher and Karl Shell, 49–113. New York: Academic.
- Fixler, Dennis. 1996. The treatment of the price of health insurance in the CPI. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics. Unpublished manuscript.
- Ford, Ina Kay. 1995. Memorandum to Mary Lynn Schmidt, 22 June. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics.
- Ford, Ina Kay, and Philip Sturm. 1988. CPI revision provides more accuracy in the medical care services component. *Monthly Labor Review* 111 (4): 17–26.
- Frank, Richard G., and Julie R. Lave. 1985. The psychiatric DRGs: Are they different? *Medical Care* 28 (11): 1145–55.
- Fuchs, Victor. 1974. *Who shall live? Health, economics and social choice*. New York: Basic.
- . 1983. *How we live*. Cambridge, Mass.: Harvard University Press.
- Getzen, Thomas E. 1992. Medical care price indexes: Theory, construction and empirical analysis of the U.S. Series 1927–1990. In *Advances in health economics and health services research*, 83–128. Greenwich, Conn.: JAI.
- Gilbert, Milton. 1961. The problem of quality changes and index numbers. *Monthly Labor Review* 84 (9): 992–97.
- . 1962. Quality change and index numbers: The reply. *Monthly Labor Review* 85 (5): 544–45.
- Ginsburg, Daniel H. 1978. Medical care services in the Consumer Price Index. *Monthly Labor Review* 101 (8): 35–39.
- Griliches, Zvi. 1962. Quality change and index numbers: A critique. *Monthly Labor Review* 85 (5): 542–44.
- . 1988. *Technology, education, and productivity*. New York: Basil Blackwell.
- . 1992. Introduction to *Output measurement in the service sectors*, ed. Zvi Griliches. NBER Studies in Income and Wealth, vol. 56. Chicago: University of Chicago Press.
- . 1997. The commission report on the Consumer Price Index, *Federal Reserve Bank of St. Louis Review* 79 (3): 169–73.
- Griliches, Zvi, and Iain Cockburn. 1994. Generics and new goods in pharmaceutical price indexes. *American Economic Review* 84 (5): 1213–32.
- Grossman, Michael. 1972a. *The demand for health: A theoretical and empirical investigation*. New York: Columbia University Press.
- . 1972b. On the concept of health capital and the demand for health. *Journal of Political Economy* 80 (2): 223–55.
- Gruber, Jonathan. 1994. The incidence of mandated maternity benefits. *American Economic Review* 84 (3): 622–41.
- . 2000. Health insurance and the labor market. In *Handbook of health economics*, ed. Joseph P. Newhouse and Anthony J. Culyer. Amsterdam: Elsevier Science.
- Hoechst Marion Roussel. 1997. *Managed care digest series 1997*. Kansas City, Mo.: Hoechst Marion Roussel.
- Hoover, Ethel D. 1961. The CPI and problems of quality change. *Monthly Labor Review* 84 (11): 1175–85.
- Jensen, Gail A., and Michael A. Morrissey. 1990. Group health insurance: A hedonic price approach. *Review of Economics and Statistics* 72 (1): 38–44.
- Kanoza, Douglas. 1996. Supplemental sampling in the PPI Pharmaceuticals Index. *Producer Price Indexes, Detailed Price Report* (January): 8–10.

- Kelly, Gregory G. 1997. Improving the PPI samples for prescription pharmaceuticals. *Monthly Labor Review* 120 (10): 10–17.
- Kendrick, John W. 1991. Appraising the U.S. output and productivity estimates for government: Where do we go from here? *Review of Income and Wealth* 37 (2): 149–58.
- Lane, Walter. 1996. Changing the item structure of the Consumer Price Index. *Monthly Labor Review* 119 (12): 18–25.
- Langford, Elizabeth A. 1957. Medical care in the CPI: 1935–1956. *Monthly Labor Review* 80 (9): 1053–58.
- Lawson, Ann M. 1997. Benchmark input-output accounts for the U.S. economy, 1992: Make, use, and supplementary tables. *Survey of Current Business* 78 (11): 36–82.
- Leaver, Sylvia G., William H. Johnson, Robert Baskin, Samuel Scarlett, and Robert Morse. 1997. Commodities and services sample redesign for the 1998 Consumer Price Index revision. Washington, D.C.: U.S. Bureau of Labor Statistics. Unpublished memo.
- Levit, Katharine R., Helen C. Lazenby, Bradley R. Braden, and the National Health Accounts Team. 1998. National health spending trends in 1996. *Health Affairs* 17 (1): 35–51.
- Meltzer, David. 1997. Accounting for future costs in medical care cost-effectiveness analysis. *Journal of Health Economics* 17 (4): 33–64.
- Moulton, Brent R. 1996. Bias in the Consumer Price Index: What is the evidence? *Journal of Economic Perspectives* 10 (4): 159–77.
- Moulton, Brent R., and Karin E. Moses. 1997. Addressing the quality change issue in the Consumer Price Index. *Brookings Papers on Economic Activity*, no. 1, 305–49.
- Moulton, Brent R., and Kenneth J. Stewart. 1997. An overview of experimental U.S. Consumer Price Indexes. Washington, D.C.: U.S. Department of Labor, Bureau of Labor Statistics. Unpublished paper, March.
- Murray, Richard. 1992. Measuring public-sector output: The Swedish report. In *Output measurement in the service sectors*, ed. Zvi Griliches, 517–42. NBER Studies in Income and Wealth, vol. 56. Chicago: University of Chicago Press.
- Newhouse, Joseph P. 1989. Measuring medical prices and understanding their effects. *Journal of Health Administration Education* 7 (1): 19–26.
- Newhouse, Joseph P., and the Insurance Experiment Group. 1993. *Free for all? Lessons from the RAND Health Insurance Experiment*. Cambridge, Mass.: Harvard University Press.
- Nordhaus, William D. 1998. Quality change in price indexes. *Journal of Economic Perspectives* 12 (1): 59–68.
- Pauly, Mark V. 1997. *Health benefits at work: An economic and political analysis of employment-based health insurance*. Ann Arbor: University of Michigan Press.
- . 1999. Costs, effects, outcomes, and utility: Concepts and usefulness of medical care price indexes. In *Measuring the prices of medical treatments*, ed. Jack E. Triplett. Washington, D.C.: Brookings Institution.
- Persky, Joseph. 1998. Retrospectives: Price indexes and general exchange values. *Journal of Economic Perspectives* 12 (1): 197–206.
- Pollak, Robert A. 1980. Group cost-of-living indexes. *American Economic Review* 70 (2): 273–78.
- . 1998. The Consumer Price Index: A research agenda and three proposals. *Journal of Economic Perspectives* 12 (1): 69–78.
- Prescott, Edward C. 1997. On defining real consumption. *Federal Reserve Bank of St. Louis Review* 79 (3): 47–53.

- Prospective Payment Assessment Commission. 1995. *Report and recommendations to the Congress*. Washington, D.C.: U.S. Government Printing Office.
- Reder, Melvin W. 1969. Some problems in the measurement of productivity in the medical care industry. In *Production and productivity in the service industries*, ed. Victor Fuchs. New York: Columbia University Press.
- Rice, Dorothy P., and Loucele A. Horowitz. 1967. Trends in medical care prices. *Social Security Bulletin* 30 (7): 13–28.
- Ristow, William. 1996. IMS presentation to the BLS. Plymouth Meeting, PA: IMS America. Mimeo.
- Scitovsky, Anne A. 1964. An index of the cost of medical care—A proposed new approach. In *The economics of health and medical care: Proceedings of the Conference on the Economics of Health and Medical Care, May 10–12, 1962*. Ann Arbor: University of Michigan.
- . 1967. Changes in the costs of treatment of selected illnesses, 1951–65. *American Economic Review* 57 (5): 1182–95.
- Shapiro, Matthew D., and David W. Wilcox. 1996. Mismeasurement in the Consumer Price Index: An evaluation. In *NBER Macroeconomics Annual*, vol. 11, ed. Ben S. Bernanke and Julio J. Rotemberg. 93–142. Cambridge, Mass.: MIT Press.
- . 1997. Alternative strategies for aggregating prices in the CPI. *Federal Reserve Bank of St. Louis Review* 79 (3): 113–25.
- Summers, Lawrence H. 1989. Some simple economics of mandated benefits. *American Economic Review* 79 (2): 177–83.
- Swedish Ministry of Finance. Budget Department. 1997. *Public sector productivity in Sweden*. Papers on Public Sector Budgeting and Management in Sweden, vol. 3. Stockholm: Ministry of Finance.
- Trajtenberg, Manuel. 1990. *Economic analysis of product innovation: The case of CT scanners*. Cambridge, Mass.: Harvard University Press.
- Triplett, Jack E. 1983. Escalation measures: What is the answer? What is the question? In *Price level measurement: Proceedings from a conference sponsored by Statistics Canada*, ed. W. E. Diewert and C. M. Montmarquette, 457–87. Ottawa: Statistics Canada.
- . 1990. The theory of industrial and occupational classifications and related phenomena. In *Proceedings of the Bureau of the Census 1990 Annual Research Conference*. Washington, D.C.: U.S. Department of Commerce.
- . 1999. Accounting for health care: Integrating price index and cost-effectiveness research. In *Measuring the prices of medical treatments*, ed. Jack E. Triplett. Washington, D.C.: Brookings Institution.
- U.S. Congress. Joint Economic Committee. 1961. *Government price statistics, hearings before the Subcommittee on Economic Statistics of the Joint Economic Committee, Congress of the United States. 87th Cong., 1st sess., pursuant to Sec. 5(a) of P.L. 304 (79th Cong.)*, pt. 1, 24 January.
- U.S. Department of Health, Education, and Welfare. 1967. *A report to the president on medical care prices*. Washington, D.C.: U.S. Government Printing Office, February.
- U.S. Department of Health and Human Services. 1980. *International classification of diseases (ICD-9-CM)*, 2d ed. (PHS)-80-1260. Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Labor. Bureau of Labor Statistics. 1987. *The Consumer Price Index: 1987 revision*. Report 736. Washington, D.C.: U.S. Government Printing Office, January.
- . 1992. *Handbook of methods*. Bulletin 2414. Washington, D.C.: U.S. Government Printing Office, April.

- . 1995. *Consumer expenditure survey 1995*. Washington, D.C.: U.S. Government Printing Office.
- . 1996a. *Producer price index coverage expansion plan*. Washington, D.C.: U.S. Government Printing Office, December.
- . 1996b. *Relative importance of components in the Consumer Price Index 1995*. Bulletin 2476. Washington, D.C.: U.S. Government Printing Office, February.
- . 1997a. The experimental CPI using geometric means (CPI-U-XG). Washington, D.C.: U.S. Government Printing Office, 10 April.
- . 1997b. Measurement issues in the Consumer Price Index. BLS response to letter from Jim Saxton, chairman of the Joint Economic Committee, to Katharine Abraham, commissioner of the Bureau of Labor Statistics, June.
- . 1997c. *Relative importance of components in the Consumer Price Indexes: U.S. city average*. Washington, D.C.: U.S. Government Printing Office, December.
- . 1998. Planned changes in the Consumer Price Index formula. Washington, D.C.: U.S. Government Printing Office, 16 April.
- . N.d.-a. A description of the PPI physician services initiative. Washington, D.C.: U.S. Government Printing Office.
- . N.d.-b. A description of the PPI hospital services initiative. Washington, D.C.: U.S. Government Printing Office.
- . N.d.-c. Improvements to CPI procedures: Prescription drugs. Washington, D.C.: U.S. Government Printing Office.
- U.S. General Accounting Office. 1996. *Consumer Price Index: Cost-of-living concepts and the housing and medical care components*. Report to the Ranking Minority Member, Committee on Banking and Financial Services, House of Representatives. GAO/GGD-96-166, August. Washington, D.C.: U.S. Government Printing Office.
- U.S. Senate Finance Committee. 1996. *Final report from the advisory commission to study the Consumer Price Index*. Washington, D.C.: U.S. Government Printing Office, 4 December.
- Weiner, Jonathan P., Allen Dobson, Stephanie L. Maxwell, Kevin Coleman, Barbara Starfield, and Gerald Anderson. 1996. Risk-adjusted Medicare capitation rates using ambulatory and inpatient diagnoses. *Health Care Financing Review* 17 (3): 77–100.
- Weiss, Samuel. 1955. The development of index numbers in the BLS. *Monthly Labor Review* 78 (1): 20–25.

Comment Brent R. Moulton

Berndt, Cutler, Frank, Griliches, Newhouse, and Triplett have presented a comprehensive examination of the current state of knowledge on mea-

Brent R. Moulton is associate director for national income, expenditure, and wealth accounts at the Bureau of Economic Analysis, U.S. Department of Commerce.

The opinions expressed in this paper are those of the author and do not represent an official policy of the Bureau of Economic Analysis or the views of other BEA staff.

Copyright is not claimed for this contribution.

suring inflation in the medical care sector. They summarize the key market characteristics that make price and welfare measurement for this set of activities difficult. They then present a detailed and thorough discussion of the issues and problems associated with the consumer and producer price indexes for this sector. They briefly summarize the empirical studies that have attempted to design improved measures of price change for medical care. I note that these studies are very recent indeed; most have been written within the last three years. The authors conclude by making a series of recommendations for improving price measurement, most of which involve further cooperative efforts between the federal statistical agencies and the wider research community.

With respect to most of this paper, I found the presentation to be excellent and I have little to add. I would like to focus my remaining comments on the small section of the paper in which the measurement target, the conceptual cost-of-living index, is formulated.

Berndt et al. take a rather unusual approach to deriving a cost-of-living index. Usually, the cost-of-living index has been derived from the consumer's cost (or expenditure) function (Konüs 1939). Berndt et al., however, form the cost-of-living index from the direct utility function. The reasons for this choice are clear. The cost function is based on an assumption of consumer optimization, whereas Berndt et al. intend to allow for nonoptimal choices due to moral hazard and principal-agent problems, which are important features of medical care markets.

The resulting cost-of-living index, shown as their equations (3) and (4), has some unusual features. Although it measures the changes in the cost of living, it is not a price index; that is, it is a function of expenditures (specifically, "nonmedical expenditures divided by numeraire price"), rather than prices. Presumably, the numeraire price is the price index for nonmedical care expenditures.

The other thing I would point out is that the information requirements for this index are extremely high. These include, at a minimum, knowing (1) the marginal rate of substitution between health status and nonmedical care expenditures (which must be estimated without assuming that consumers are optimizing); (2) an index of health status that must be comprehensive enough to represent all of the medical conditions for which medical services are obtained; and (3) an index of the time allocated to medical treatments. In addition, a more comprehensive measure is suggested that would entail knowing the effects of changes in medical care knowledge and the environment on health status.

Contemplating the difficulties that would be associated with directly implementing this type of cost-of-living measure leads me to appreciate the relative simplicity of the traditional, Konüs-type approach. Berndt et al. should be complimented for comprehensively laying out the problems and

suggesting some important improvements. They have also shown, however, that for those of us who would like to develop medical care price indexes that would be useful as a proxy for a well-established cost-of-living index, the task before us is daunting indeed.

Reference

Konüs, A. A. 1939. The problem of the true index of the cost of living. *Econometrica* 7:10–29.