

The Importance of Pricing the Bundle of Treatments*

Ana Aizcorbe
Nicole Nestoriak

July 2008

*Ana Aizcorbe (Ana.Aizcorbe@bea.gov) is affiliated with the Bureau of Economic Analysis and Nicole Nestoriak (Nestoriak.Nicole@bls.gov) with the Bureau of Labor Statistics. Both would like to thank Ernie Berndt, David Cutler, Anne Hall, Richard Kane, Allison Rosen, Louise Sheiner and participants of the NBER Conference on Research in Income and Wealth Summer Institute, Productivity and Health Economics workshops for useful comments. LeRoynda Brooks and Sarah Pack provided research assistance.

I. Introduction and Summary

The health care services sector is a large and growing segment of the economy. The share of GDP devoted to health care services has doubled over the past 25 years, to 16 percent in 2005. Looking ahead, this growth is expected to continue, as aging baby boomers continue to increase the share of elderly in the population, a segment of the population that accounts for a disproportionately high level of expenditures on health.

The rising health care costs have raised questions about whether these medical expenditures are, in some sense, worth it. A natural starting point is to ask how much of the rising costs represent increases in real services vs. price inflation, a question that can be addressed using price deflators or indexes. However, studies in the health economics literature have raised questions about the use of standard price indexes for this purpose and their empirical findings suggest that some of what is currently recorded as price increases actually represents increases in services.

Currently-available price indexes define the “good” or the “output” of health care as the treatment (i.e., an office visit or prescription drug) and track the prices of those treatments over time. Health economists have long advocated an alternative definition of output as the *bundle* of treatments received by a patient for the treatment of some condition (Scitovsky (1964)). The existing empirical evidence suggests that how one defines the good matters. Detailed case studies of important diseases show that, for these conditions, the price indexes that health economists advocate show slower price growth than standard indexes.¹ This result is consistent with the view that indexes that track the prices of individual treatments tend to miss any shifts towards lower-cost treatments that may occur over time. For example, in the treatment of depression, there has been a shift away from talk-therapy and towards (the lower cost) drug therapy that has reduced the cost of treating depression. Because standard indexes track prices for these two types of treatments separately, they miss this substitution and overstate the cost of treating depression.

These studies also account for changes in the quality of health care by measuring the health outcomes associated with treatment. For example, Cutler, in his study of heart

¹ See, for example, studies on heart attacks (Cutler, McClellan, Newhouse, and Remler(2001)), depression (Berndt, Busch and Frank (2001)) and schizophrenia (Frank, Berndt, Busch and Lehman (2004)).

attacks, found that the substitution pattern there was towards more costly treatments. But, because the costlier treatments provided better outcomes—in terms of health improvements—the quality-adjusted price still showed slower growth than standard indexes.

In this paper, we explore the importance of shifts in treatments for explaining increases in the cost of health care services. We begin by implementing a price index of the type advocated by health economists—the “preferred” index—over a broader range of conditions than previously done. Although the list of conditions is comprehensive, our estimates are based on a sample of claims data for only a subset of patients—those covered by certain types of commercial insurance—and, thus, our estimates are not representative of the entire population of patients. Given that caveat, our sample of over 700 million health claim records covering the period 2003-2005 shows substantial cost savings in the treatment of disease generated by shifts in treatment regimens. While the price of treating diseases grew an average of 12% over this period, costs would have risen even faster, 17%, if the mix of treatments in 2005 had been the same as that in 2003. In terms of compound annual growth rates (CGARs), the differences are 4.4 vs. 6.1 percentage points a year.

To explore the sources of those cost savings, we develop an algebraic expression for the contributions of shifts in different treatments to the cost savings by linking the preferred index to one that tracks fixed baskets of treatments, the type of index typically produced by statistical agencies. In our empirical work, the decomposition confirms the presence of treatment substitution for several important disease classes: shifts from office visits towards drugs in many psychiatric conditions, shifts from care at hospitals towards care at ambulatory surgical centers for orthopedic and gastroenterological conditions, and similar shifts in endocrinology (a disease class that contains diabetes and obesity). However, the decomposition also reveals other patterns associated with these cost savings. In cardiology, for example, the data literally show a large decline in the use of inpatient care with little change in the intensity of other treatments. We take this to mean that although patients appear to do as many office visits and purchase as many prescriptions as they did in 2003, perhaps the treatments they receive in 2005 are better, obviating the need for inpatient care and, thus, giving rise to cost savings. Finally, there

are disease classes where the cost of treating conditions rises faster than the prices of the underlying treatments. Two notable cases are obstetrics and neonatology; where the increased intensity of treatment is associated with complications while normal pregnancies and uncomplicated neonatal management show little to no cost savings, respectively.

To make inferences about any potential cost savings and how rising costs are split between changes in prices vs. real services, we must assume that the health outcomes associated with the bundle of treatments are constant; if, instead, the trend is for outcomes to improve with time, we must interpret the resulting price change as an upper bound. This assumption is required because, unlike in previous case studies, our focus on a comprehensive list of conditions does not allow us to account for potential changes in the outcome of treatment (quality). This assumption represents another major caveat of our work.

The paper is organized as follows. After defining the "price" that we will track in our price index, section III provides the decomposition that we use to decompose cost savings into contributions from individual treatments. Section IV provides empirical results and section V concludes.

II. Defining the price of the treatment bundle

A key issue in constructing a price deflator is how one defines the good and the attendant price. We follow the health economics literature and define the "good" as the treatment of a medical condition--fixing a bad knee, for example. The idea is that medical conditions are normally treated with bundles of treatments so that tracking the cost of treating the disease is best done by tracking the price of the bundle rather than tracking prices of treatments separately. Moreover, the arrival of new treatments often generates shifts in treatments that can change the price of the bundle without changing the prices of the individual treatments and, as shown in the literature, this effect will not be picked up with standard treatment-based indexes.

But, how should one define the price? If one takes the patient's perspective, one would define the price as whatever the patient pays for the service. This is the perspective taken in the Consumer Price Index, an index that aims to track payments for

health insurance and out of pocket payments for treatments. We take a provider perspective and define the "price" as the amount of revenues received by providers from all payers. This price is not intended as a measure of the "resources" used in the production of medical care, or its "value" in terms of welfare, but rather as a measure of the amount of financial resources that were devoted to health care services—i.e., the 16% of GDP.²

The preferred index is a unit value that divides the total dollars received by "the health care system"—all providers taken together—for the treatment of some condition. "Unit values" are one way to define a price for homogeneous goods and services, in our case, homogeneous conditions or diseases. In previous work, health economists have tended to use the notion of "an episode of care" as the homogeneous good. Specifically, they track changes in the cost of *completed* episodes, where the episode can last a long time. For example, the cost of a pregnancy that ends in the first quarter can include costs accrued in previous quarters. We consider the cost of treating disease over some period of time, a quarter, in order to develop a price measure suitable for deflating a quarterly nominal series.

Formally, consider a price measure that tracks the cost of treating some disease, d , where different types of treatments, i , may be required. If P_d^1 is the number of patients under treatment for condition d in some time period $t=1$, then we define the average cost of treating that condition as the dollars spent on all treatments for that condition divided by the number of patients treated for the condition:

$$(1) \quad c_d^1 = \sum_i (c_{d,i}^1 x_{d,i}^1) / P_d^1$$

where $c_{d,i}^1$ measures the cost of treatment i for condition d and $x_{d,i}^1$ is the number of such treatments. By "treatment" we mean broad aggregates like "an inpatient confinement" so that $c_{d,i}^1$ measures all the costs incurred during the confinement and $x_{d,i}^1$ measures the number of confinements.

A counterintuitive feature of the index is that one only wants to control for the homogeneity of the condition, not the insurance type or other types of heterogeneity that

² Although the Producer Price Index also takes this perspective, it defines outputs as individual treatments and is, therefore, not useful for our purposes.

can affect the prices received by providers. To see why this is important, consider a simple example where Medicare Part D was introduced; Chart 1 assumes that uninsured patients pay more than those with Part D coverage, the prices of both are constant, and that at the time of introduction, uninsured patients began to switch to Part D and paying less for drugs. This shift lowers nominal expenditures on drugs even if the number of prescriptions is the same. If one uses price indexes that “control” for insurance coverage, one would track the price of drugs for the uninsured separately from the price of drugs to those with Part D coverage; an aggregate price index over the two types of transactions is just a weighted average of changes in the individual prices and, so, would show no price change. The problem, for our purposes, is that using that price index to deflate nominals would result in a decline in real services, even if the same number of prescriptions were purchased in both periods. Instead, we want to record these changes in costs that arise from changes in insurance coverage as a change in the price, not a change in real services. To do so, the average price in (1) should be calculated for all patients (regardless of insurance coverage). Other sources of price variation—like differences across regions—should be accounted for in the same way.

A ratio of the unit values in (1) is a price relative that tracks changes in the price of treating a homogeneous condition from $t=0$ to $t=1$.³ We define this unit value index, $UV_d^{1,0}$, as:

$$(2) \quad UV_d^{1,0} = c_d^1 / c_d^0 = \frac{\sum_i (c_{d,i}^1 x_{d,i}^1) / P_d^1}{\sum_i (c_{d,i}^0 x_{d,i}^0) / P_d^0}$$

Note that this index accounts for new treatments the period they are introduced: Unlike standard price indexes, this index does not require that treatments exist in both periods in order to measure changes in the price of treating disease d . For example, the arrival of

³ In other work, we've shown a set of assumptions under which this index has an equivalent variation interpretation as the amount of money a central planner would have to give health providers to keep them indifferent between treating today's patients at today's vs. yesterday's prices. The assumptions are that the slow diffusion of new treatments arises from a fixed cost borne by providers and that this fixed cost declines over time. As long as the counterfactual involves yesterday's fixed cost, providers will use the same treatments in the counterfactual as they used yesterday and comparing the average prices in the two periods appropriately accounts for any providers that switched treatments over the period.

microscopic surgeries at t=1 as an alternative to the more traditional surgeries could create a new treatment and when it is used to treat disease d, the associated cost will be reflected in the numerator of (2), even if it is not included in the denominator. In that sense, the index fixes the “new goods problem.” Moreover, costs associated with relatively rare procedures that appear only occasionally, will be included in (2) but not in standard indexes. Note, though, that the arrival of a new disease in period t=1 still presents problems, since, in that case, c_d^0 is undefined and the disease cannot be included in the index. So, while the index solves the “new goods problem” for treatments, it introduces a “new disease problem.”

We can use this index to obtain a measure of "real" spending, or real services, for individual conditions and doing so yields a volume index that tracks the number of patients treated for condition d. Equation (2) gives a price index that can be used to translate nominal spending to treat d at t=1, say, into a measure of real spending; i.e., into the cost of treating the P_d^1 patients at some base period (t=0) prices:

$$\sum_t (c_{d,i}^1 x_{d,i}^1) / UV_d^{1,0} = P_d^1 [\sum_t (c_{d,i}^0 x_{d,i}^0) / P_d^0]$$

So, the growth in real spending from t=0 to t=1 reduces to the growth in the number of patients (the basis of volume measures):

$$[\sum_t (c_{d,i}^1 x_{d,i}^1) / UV_d^{1,0}] / [\sum_t (c_{d,i}^0 x_{d,i}^0) / UV_d^{1,0}] = P_d^1 / P_d^0$$

For individual conditions, then, the price measure advocated by health economists (to capture the change in costs that arise from changes in price) implies a volume measure of the number of patients treated to capture the changes in costs that come from changes in real services.

This unit value index applies to individual, homogeneous, conditions. Measuring changes in the price of treatment bundles across a broader range of conditions requires an

aggregate. Following Diewert (1999), we construct overall measures that average over a broad range of conditions using the Fisher Ideal formula.⁴

II. Calculating cost savings and its sources

Empirical work in this area has traditionally used a case study method, where information about potential cost savings were known a priori and the empirical work was used to estimate the cost savings that were already known to exist. The strategy in these studies was to compare the preferred index with treatment-based price indexes to quantify cost savings associated with known substitution trends.

Because our work involves a comprehensive list of diseases, we do not know *a priori* which diseases have become more or less costly or the reasons for those changes in costs. We want to calculate changes in cost of treating diseases, how much of those changes in cost can be attributed to changes in underlying prices and if there are cost savings generated by switching treatments, we want to know the extent to which those gains were generated by substitution towards lower cost treatments vs. something else.

In this section, we develop an expression for calculating any potential cost savings and the contribution of shifts in treatment intensity to those cost savings. Following the earlier literature, the cost savings are calculated as the differences between the preferred index and one that holds the mix of treatments fixed. The latter is a Laspeyres index that tracks prices of individual treatments.

The two indexes diverge when the types of treatments used to treat diseases change over time in one of two ways. First, as mentioned above, it is possible that treatments used in one period are not used in the other, in which case the treatment index will exclude the treatment while the preferred index will not. We call this difference a “selection” problem in the standard price index, related to the well-known problem that standard indexes to account for new and disappearing goods. The other difference in the two indexes is that any changes in costs arising from shifts in existing treatments will be

⁴ One potential theoretical basis for superlative indexes like the Fisher—the cost of living index literature—does not provide useful guidance in the case of health care spending. That theory is based on a representative consumer optimally choosing goods in response to changes in relative prices. In the context of health care, with goods defined as “the treatment of disease,” the conditions are exogenous and not chosen in response to changes in relative prices. Thus, the theory is not relevant in our case.

reflected in the unit value index but not in standard indexes (the latter assume the same distribution of treatments in both periods).

Formally, redefine the subscript i to refer to continuing treatments—those that were used in both periods—and use subscripts “ n ” and “ x ” to denote treatments that were only used in $t=1$ and $t=0$, respectively. Using that notation, we can rewrite the unit value index in (2) as the product of two terms:

$$(2') \quad UV_d^{1,0} = c_d^1 / c_d^0 = \frac{(\omega_d^n)}{(\omega_d^x)} \frac{\sum_i (c_{d,i}^1 x_{d,i}^1) / P_d^1}{\sum_i (c_{d,i}^0 x_{d,i}^0) / P_d^0}$$

$$= (\sigma_d^{1,0}) (UVC_d^{1,0})$$

The first term, $\sigma_d^{1,0}$, measures the effect on costs from non-continuing treatments: distortions in the standard index from excluding treatments that exist only at $t=1$ is given by the numerator: $\omega_d^n = (\sum_i c_{d,i}^1 x_{d,i}^1 + \sum_n c_{d,n}^1 x_{d,n}^1) / \sum_i (c_{d,i}^1 x_{d,i}^1)$; treatments that existed only in period $t=0$ are handled the same way. This term is greater than one when the expenditure share associated with new treatments exceeds that of exiting treatments and vice versa. In our empirical application, we define “treatments” coarsely—inpatient care, outpatient care, for example—and anticipate that any deviation of (ω_n / ω_x) from 1 will represent extremely rare events or rare conditions. For example, when the number of observations in a disease class is small—rare conditions—the chances that treatments used at $t=0$ and $t=1$ differ can be large. Similarly, although simple infections typically involve only an office visit and a prescription for antibiotics, for some patients they result in a costly inpatient stay—a rare, but costly, event. If the unusual inpatient stay appears in one period but not the other, $\sigma_d^{1,0} \neq 1$ for that condition. Our framework allows for these possibilities and allows us to calculate their numerical importance in our empirical work.

The second way that the indexes can diverge occurs when there are shifts in treatment intensity among treatments that exist in both periods. We measure this effect by further breaking down the unit value index that includes only continuing treatments,

$UVC_d^{1,0}$. We provide the link between that index and the standard treatment-based index with the following expression:

$$(3) \quad UVC_d^{1,0} = \sum_i \frac{c_{d,i}^0 x_{d,i}^0}{\sum_i c_{d,i}^0 x_{d,i}^0} \frac{x_{d,i}^1 / P_d^1}{x_{d,i}^0 / P_d^0} (c_{d,i}^1 / c_{d,i}^0)$$

$$= \sum_i w_{d,i}^0 \gamma_{d,i} (c_{d,i}^1 / c_{d,i}^0)$$

Working from right to left, $(c_{d,i}^1 / c_{d,i}^0)$ measures the change in the cost of treatment i in treating condition d and is the type of price relative used in the standard index. Note that this term is only defined when treatment i for condition d is observed in both periods (i.e., when all treatments are continuing). The middle term, $\gamma_{d,i}$, captures changes in treatment intensity and literally measures changes in the number of treatments per patient treated for condition d . This can represent either changes in the number of patients that receive the treatment or changes in the number of treatments for each patient that received them. All else held equal, an increase (decrease) in utilization of all treatments translates into higher (lower) costs: For example, $\gamma_{d,i} > 1$ with $\gamma_{d,j} \geq 0$ for all other treatments and no change in prices will raise c_d^1 / c_d^0 . Often, though, intensity increases for some treatments and decreases for others. Cutler and McClellan (1998) call the first case an increase in treatment *intensity* and the second *treatment substitution*, which will increase or decrease costs depending on whether treatments are shifting towards higher-cost vs lower-cost alternatives.

Finally, the first term is a base-period expenditure share ($w_{d,i}^0$) that measures the relative importance of treatment i in the treatment of condition d .

To assess the numerical importance of shifts in treatments, we compare the index in (3) to one where we fix treatment intensities at $t=0$ levels. That is, we set all the $\gamma_{d,i}=1$ to obtain:

$$(4) \quad I_d^{1,0} = \sum_i w_{d,i}^0 (c_{d,i}^1 / c_{d,i}^0)$$

This is precisely the type of price index typically provided by statistical agencies. While equation (3) measures changes in the cost of treating disease d, equation (4) measures what's happening to prices of the individual treatments used to treat that disease.

The difference in the two indexes quantifies any cost savings from shifting treatment intensities and allows us to parse out the contribution of each treatment to the overall cost savings as follows:

$$(5) \quad UVC_d^{1,0} - I_d^{1,0} = \sum_i w_{d,i}^0 (\gamma_{d,i} - 1) (c_{d,i}^1 / c_{d,i}^0)$$

Three possibilities will highlight how this works:

- Suppose $\gamma_{d,i} > 1$ for all treatments. Here, treatment intensity increased for all treatment types, increasing the price of treating disease d faster than the increases in the prices of the individual treatments: $UVC_d^{1,0} > I_d^{1,0}$. This illustrates that, in theory, the disease based price index can rise faster than the usual treatment-based index.
- The usual case that's been documented in the literature is one where treatment substitution causes a gap in the indexes. In this case, some $\gamma_{d,i} < 1$ and some $\gamma_{d,i} > 1$. When the shift is toward a lower-cost treatment, then $UVC_d^{1,0} < I_d^{1,0}$.
- But, to see that treatment substitution can exist without generating a gap in the indexes, consider the well-known example of treatment substitution from talk to drug therapy: $\gamma_{d,i} < 1$ for office visits and $\gamma_{d,i} > 1$ for drugs. This, all else held equal, would cause $UVC_d^{1,0} < I_d^{1,0}(T)$. But, with more than two treatment types, it is possible that there could also have been a shift in the other treatments—from outpatient to the more expensive inpatient care, for example—that could offset the substitution towards drugs and ultimately cause $\geq I_d^{1,0}$.

As the health care system finds ways to provide the same medical care at a lower cost by changing the mix of treatments towards lower-cost treatments, the two indexes will diverge and the difference provides a measure of the cost savings (productivity gains). Numerically, we can use 2) to calculate changes in the cost of treating condition d,

$UV_d^{1,0}$, use (3) to calculate the unit value index for continuing treatments only, $UVC_d^{1,0}$ and use (5) to quantify the effect of changes in treatment intensity on costs.

Aggregation

This decomposition relates to individual diseases. How do we obtain summary measures of these contributions over a broad range of diverse conditions? Our interest is in assessing the numerical importance of shifts in treatments for changes in costs, as measured in a Fisher aggregate of the unit values for individual all conditions. Ideally, we would find an algebraic expression that links that Fisher index to the standard Laspeyres that tracks prices of individual treatments.⁵ Instead, we use a Laspeyres average to do the aggregate decomposition, noting that there may be differences in the Fisher and Laspeyres. As will be seen shortly, the Fisher and Laspeyres aggregates are virtually identical in our data so this will not be an issue for us.

That expression for the average cost savings over several individual diseases is:

$$(5) \quad UVC^{1,0} - I^{1,0} = \sum_d \frac{\sum_i c_{d,i}^0 x_{d,i}^0}{\sum_i \sum_d c_{d,i}^0 x_{d,i}^0} [UV_d^{1,0} - I_d^{1,0}]$$

$$= \sum_d w_d^0 [UV_d^{1,0} - I_d^{1,0}]$$

This is just a weighted average of the cost savings for individual diseases, where cost shares from the base period gives the relative importance given to each condition.

This aggregation has the advantage that a Laspeyres of the treatment indexes has the same structure as the price indexes used by statistical agencies; a Laspeyres of all the disease Laspeyres price indexes equals a Laspeyres of the T treatment types:⁶

$$I^{1,0} = \sum_d \frac{\sum_i c_{d,i}^0 x_{d,i}^0}{\sum_i \sum_d c_{d,i}^0 x_{d,i}^0} I_d^{1,0} = \sum_i \frac{\sum_d c_{d,i}^0 x_{d,i}^0}{\sum_i \sum_d c_{d,i}^0 x_{d,i}^0} I_i^{1,0}$$

⁵ See Silver(2008) for one possible decomposition of the Fisher index that might be useful in our setting.

⁶ Note that simply reweighting a treatment index into disease classes will not provide an index for the prices of the bundle. For that, one needs to redefine the "good" as the bundle of treatments.

Moreover, the use of Laspeyres weights provides a clean way to calculate the contributions of shifts in different treatments to the overall cost savings; combining (5) and (4) and rearranging, we can express differences in aggregate indexes in terms of the underlying treatments:

$$\begin{aligned}
 (6) \quad UV^{1,0} - I^{1,0} &= \sum_d w_d^0 \left[\sum_i w_{d,i}^0 (\gamma_{d,i} - 1) (c_{d,i}^1 / c_{d,i}^0) \right] \\
 &= \sum_i \left[\sum_d w_d^0 w_{d,i}^0 (\gamma_{d,i} - 1) (c_{d,i}^1 / c_{d,i}^0) \right]
 \end{aligned}$$

Thus, the difference in the indexes can be decomposed into contributions from each of the treatments. This formula forms the basis for the decompositions we report in our empirical work.

III. Empirical Implementation

Our sample, from the Pharmedics, Inc., contains over 700 million claims records from HMO, PPO, and POS plans covering the period 2003 to 2005. The data have been processed with the Symmetry/Ingenix episode grouper and we use those definitions of diseases to identify the medical condition associated with each claim. There are potential problems with defining diseases in this way, namely it is not clear exactly how the groupers deal with comorbidities—the fact that claims typically have more than one diagnosis associated with them—or how they allocate spending on claims that have no diagnosis listed—pharmacy claims. These are potentially important issues under investigation in other work.⁷ For now, we simply take a literal read of the data as a laboratory to explore the potential numerical importance of how one defines the good.

In our data, the Symmetry grouper allocates the bulk of spending on health care (87%) into 542 disease groups—called episode treatment groups (ETGs)—and the rest to

⁷ Ana Aizcorbe is working with David Cutler and Allison Rosen in a study of different methods for reporting health care spending by disease, including allocations made by several existing episode groupers.

catch-all, non-disease categories such as routine office visits, immunizations, and miscoded claims.⁸ Because the construction of price indexes requires homogeneous conditions, we only include spending that was allocated into disease classes. Despite the large number of claims in our data, a surprising amount of spending is allocated to “thin” cells, defined as cells that included less than 1000 patients over a quarter. Our sample contains ETG classes that contain as little as \$36 of spending in some quarters and as few as two patients. In all, there are 195 ETGs that contain data for fewer than 1000 patients but they make up a relatively small amount of spending—eight percent. Excluding them raises the minimum level of spending in the remaining ETGs to just over \$1000. In our empirical work, we assess the sensitivity of our results to these limitations in the data.

There are three types of “prices” reported in the dataset: the charge (a mostly-fictitious list price that is only paid by the uninsured), the allowed amount (the negotiated price that the provider receives from both the insurance company and the patient) and the paid amount (the amount paid by the insurance company). For our purposes, the relevant price is the value of the service and is typically measured as the total receipts taken in by the provider, without regard to payee. We, therefore, use the allowed price to construct expenditures from each claim, a variable that has a high response rate in our data (over 94 percent).⁹

As detailed above, we use unit values (expenditures divided by number of patients) to define the price associated with each ETG. Similarly, we use unit values (expenditures divided by number of treatments) for each condition as the elementary price. For health care that does not involve an overnight stay, this is not too different from what is usually done, except that we assume that the elementary price is stratified by

⁸ Although these so-called “groupers” also make some attempt to identify the beginning and end of individual episodes of illness, we use only the allocation to disease “buckets” and use annual costs of treatment as our basis for the price indexes. As noted earlier, many episodes span longer than one year. Our goal is to obtain deflators with which to translate changes in nominal expenditures in some period into changes in real quantities. Thus, nominal expenditures will be quarterly, say, while the time associated with grouper-defined episodes could reach back further than the current quarter. Because it does not make sense, for our purposes, to include services provided outside of the current quarter, we do not use concepts like “completed episodes.”

⁹ One reason that the BLS is fairly negative on the use of claims data is that their preferred price measure is the portion paid by the consumer. To determine this amount, one must subtract the amount paid by the insurance company from the allowed amount, and the response rates on the paid amount can be quite low. This does not present a problem for us, since we only need the allowed amount. Some of the prices in the

disease (not just by treatment). The way we handle drugs is very different from what is typically done in standard price indexes. The BLS prices drugs by NDC code; academics have priced drugs by molecule. We are pricing drugs by disease; our unit value is price per prescription for all drugs given to patients in a given ETG class. This way, shifts from cheaper to expensive drugs are properly recorded as an increase in price.

Finally, we construct a price for inpatient stays as expenditures over the entire hospital stay, broken out by ETG. The grouper keeps track of coexisting conditions and allocates spending among ongoing episodes. For example, a depressed patient in the hospital for a heart attack might be given antidepressants that are unrelated to the heart attack. The episode grouper attempts to place the treatments for depression in a separate bucket despite the fact that they occurred during a heart-related confinement. We track those dollars separately.¹⁰ With regard to timing issues, we allocate all of the spending from the confinement in the period where the confinement ended; this assigns the expenditures in a consistent manner across indexes.¹¹

To break out spending by “treatment,” we use coarse categories like inpatient, outpatient, etc., as defined by a “place of service” variable in the claims data. Table 1 lists the 9 classes that contained over one percent of total spending in 2003:1; in constructing the indexes, we group the remaining claims in the “rest” category. Column 2 gives the cost per visit and, as may be seen, it varies widely across treatments. Inpatient care, defined as the cost of an average inpatient confinement, averages about \$4,000 per hospital stay. The categories “inpatient hospital,” outpatient, emergency room and office visits, as well as visits to ambulatory surgical centers, do not involve an overnight stay and are substantially less costly. The category “pharmacy” contains the average price of a prescription, among the least costly encounters.

Estimates of Cost Savings and their Sources

dataset are imputed, however, and we need to do a full investigation of the extent to which the allowed prices were imputed, how often invalid prices appear in the data, etc.

¹⁰ We are not entirely comfortable doing this since the usual practice is to attribute all the dollars of spending during a hospital confinement to the primary diagnosis. In future work, we’ll consider other ways to break out the spending.

¹¹ Because our data begin in 2003:1, we will understate the cost of confinements that were in progress over the turn of that year. However, to the extent that our goal is to compare results from the two types of

Chart 2 shows the growth in the price of treating diseases—represented by a Fisher Index of the preferred, unit value indexes in (2)—and that of the prices of individual treatments—the usual Laspeyres index of the treatment indexes in (4). The indexes are constructed relative to 2003:1 and are not chained indexes of quarter-to-quarter indexes.

As may be seen, the cost of treating diseases—the dashed line—grew slower than the prices of the underlying treatments—the solid line. The difference in the two indexes is substantial; over the three year period, the treatment indexes growth nearly 18 percent while the preferred index only grows about 13 percent. The differences amount to about 1-1/2 percentage points on the compound annual growth rates—6.1 vs 4.4 percentage points. Because our data are not representative, it is perilous to generalize to the economy-wide level. But, just to gain some sense for the potential importance of these differences, health spending makes up 16% of GDP, so if a difference of this magnitude held across all types of patients (i.e, the uninsured, Medicare and Medicaid patients), changing from the current deflator—a treatment-based index—to the preferred index in the national accounts would increase measured GDP growth by a substantial amount—about a quarter percentage point a year.

To explore the sources of these differences, the left panel of table 2 compares the growth in the two types of price indexes across 19 major disease groups. The growth rates shown are for the entire 2003-2005 period and represent averages of the growth rates for the individual conditions underlying each group—the appendix provides similar tables for the individual conditions.

The price of treating disease—shown in the first column—increased for all major groups over this period, with conditions under the cardiology category showing the slowest growth—about 2 percent—and infectious diseases showing the fastest—nearly 40 percent. As shown in the third column, for most major groups, the preferred index shows slower growth than the treatment index, suggesting the presence of cost savings in the underlying conditions. The exceptions are four disease categories that, combined, make up about 10% of total spending and, so, do not have much influence on the top line:

indexes, and that the understatement is the same in both of the indexes, our inferences about treatment shifts and their effect on cost savings should be valid.

infectious diseases, chemical dependency, obstetrics and neonatology.¹² Nonetheless, these are examples where the cost of treating diseases rose faster than the cost of the individual treatments. For the remaining groups, the reverse is true, with cardiology showing the largest difference—the preferred index grew about 15 percentage points slower than the treatment index.

The contributions of shifts in treatments to the cost savings are shown in the middle panel of the table. There, a positive sign reflects an increase in treatment intensity and a negative sign the opposite. There appear to be significant treatment substitution for several of the disease group. In the orthopedic and rheumatology group, there appears to be a shift from conventional treatment at hospitals and doctors offices towards home care and treatment at ambulatory surgical centers that lowered the growth in the cost of treating these conditions from 18 to 12 percent over the period. Similarly, for conditions in the gastroenterology and ophthalmology classes, shifts towards care at ambulatory surgical centers appear to have held down costs. Finally, two categories that show similar shifts towards drugs that held down cost increases are psychiatry (depression, anxiety disorders, e.g.) and endocrinology (which contains diabetes and obesity).

For the other classes, the story is more nuanced. For example, the conditions in the cardiology group show, on average, large declines in inpatient care that, taken with declines in the intensity of other hospital treatment and office visits, are not offset by increases elsewhere. For many of the conditions making up this group, the numbers literally show a decrease in intensity of several of the important treatments—like inpatient care—with little change in the intensity of other treatments. Similar issues pertain to conditions in the hematology, neurology and gynecology disease groups.

One possible conclusion is that patients are receiving less care, perhaps because many of the treatments received in 2003 were, in some sense, unnecessary. This would be consistent with the findings of the Dartmouth group, which suggest that patients are receiving more care than necessary so that declines in treatments could be achieved without reducing outcomes. Another possibility, however, is that patients still do the same number of office visits and purchase the same number of prescriptions but the care

¹² For obstetrics and neonatology conditions, a look at data for the underlying conditions shows cost savings for uncomplicated conditions—normal pregnancies and uncomplicated neonatal management—and higher costs for conditions that involve complications.

is better (the drugs are better or the procedures performed at the doctors' offices are better) and, ultimately, obviate the need for inpatient care and it is the resulting reduction in (costly) inpatient care holds down costs. It is not possible to distinguish between these two possibilities without accounting for outcomes.

We close with some technical issues. First, the table below shows that the choice of aggregation formula is not numerically important—indexes that use a Laspeyres formula show virtually the same growth as those that use the Fisher Ideal formula. This supports our use of a Laspeyres formula in constructing the aggregate decompositions. Second, we note that while the dominant effect is for the treatment index to grow faster than the preferred index, this is not always the case. For example, from the first to second quarters of 2005, the preferred index grows a bit faster than the treatment index and there are other periods where the two indexes show parallel growth. Finally, we did these calculations separately for the three different types of health insurance and obtained the same qualitative results—the preferred indexes grow slower than the treatment indexes—but the individual growth rates can differ substantially across treatment plans. These differences underscore the importance of using comprehensive data, rather than data on one type of insurance to cover all patients with commercial insurance.

Table 3 applies these price indexes to obtain measures of real services (ie., real spending) and compares them to a measure of real services obtained using a volume index of patients. Total spending on health care grew 31 percent in our sample over 2003-2005, or at a CAGR of about 10 percent. The first two lines of the table show how the preferred and treatment price indexes parse out this growth into price and service components: they directly give a measure of price change that, once divided into the total growth of spending, implies a growth rate for real services. As can be seen, the preferred index attributes 5 percentage points more to services growth than the treatment index, or about 1-1/2 percentage points on the CAGR. The last line shows that a calculation done using a direct volume index is very similar to the one using the preferred index: This similarity is not surprising, given that the results hold exactly at the individual disease level.

VI. Conclusion

Our empirical work suggests that there have been shifts in treatment intensity that have an important effect on costs and that, on average, those treatment shifts served to lower the cost of treating disease for patients in our sample. These cost savings appear to be numerically important and pervasive. As noted by health economists, standard price indexes provided by statistical agencies do not capture these cost savings and, thus, overstate how much of rising health care costs can be attributed to rising prices of treatments.

There are three major caveats to our work. First, our data set, though large, represents a particular type of patient—patients with commercial health insurance often provided by their (large) employers—and are, thus, not general. In future work, we will construct similar measures using claims data for patients covered by Medicare and Medicaid to increase our coverage; we are still exploring data sources for patients that are uninsured and/or institutionalized.

The second caveat relates to our inability to account for any changes in health outcomes. Under the assumption that outcomes have not, on average, declined over time, the preferred index gives an upper bound for the rise in health care costs that can be attributed to increases in the price of treating diseases. However, little is known on how new, better treatments are diffused across patients and how that translates into changes in average health outcomes.

Finally, the price indexes we constructed require that one parse out dollars of spending into different disease classes and there is no consensus on how best to do that allocation. We have used a commercial algorithm that "groups" spending into fairly granular disease categories. Although these groupers are used in a variety of settings, little is known about exactly how they deal with two important stumbling blocks in this area: assigning dollars of spending when patients have more than one disease and allocating dollars for claims that do not contain diagnoses.

Nonetheless, our work points to the potential importance of this issue and underscores the importance of further work to form more precise estimates with sufficiently broad coverage of patients to make inferences about the cost of treating diseases for the nation as a whole.

Bibliography

Aizcorbe, Ana, and Nicole Nestoriak. 2007. "Changes in Treatment Intensity, Treatment Substitution, and Price Indexes for Health Care Services." Paper presented at the National Bureau of Economic Research Productivity Workshop, Cambridge, MA, December 5.

Aizcorbe, Ana and Nicole Nestoriak (2008) "An Equivalent Variation Interpretation of Price Deflators for Health Care Spending," mimeo.

Aizcorbe, Ana, Bonnie Retus and Shelly Smith (2008) Health Satellite Accounts, *Survey of Current Business*, May.

Berndt, ER, AB Busch, RG Frank and SL Normand (2006) "Real Output in Mental Health Care During the 1990s" NBER Working Paper No. 11557.

Berndt, Ernst R., Anupa Bir, Susan H. Busch, Richard G. Frank and Sharon-Lise T. Normand [2002], "The Medical Treatment of Depression, 1991-1996: Productive Inefficiency, Expected Outcome Variations and Price Indexes," *Journal of Health Economics*, 21(3):pp. 373-396.

Berndt, Ernst R., Susan H. Busch and Richard G. Frank [2001], "Treatment Price Indexes for Acute Phase Major Depression," ch. 12 in David M. Cutler and Ernst R. Berndt, eds., *Medical Care Output and Productivity*, Studies in Income and Wealth, Vol. 61, Chicago: University of Chicago Press for the National Bureau of Economic Research, pp. 463-505.

Berndt, Ernst R., David M. Cutler, Richard G. Frank, Joseph E. Newhouse and Jack E. Triplett [2000], "Medical Care Prices and Output," in Anthony Cutler and Joseph E. Newhouse, eds., *Handbook of Health Economics*, Amsterdam: Elsevier, pp. 119-180.

Crawford, Greg and Matthew Shum (2005). Uncertainty and Learning in Pharmaceutical Demand (with Greg Crawford). *Econometrica*, 73: 1137-1174 (July 2005).

Cutler, David M., Mark McClellan, Joseph E. Newhouse and Dahlia Remler [1998], "Are Medical Prices Declining? Evidence from Heart Attack Treatments," *Quarterly Journal of Economics*, 113(4):pp. 991-1024.

Diewert, W. Erwin [1993], "The Economic Theory of Index Numbers: A Survey," ch. 7 in W. Erwin Diewert and Alice O. Nakamura, eds., *Essays in Index Number Theory*, Vol. 1, London: North-Holland, pp. 177-221.

Frank, Richard G., Ernst R. Berndt and Susan M. Busch [1999], "Price Indexes for the Treatment of Depression," in Jack E. Triplett, ed., *Measuring the Prices of Medical Treatments*, Washington DC: The Brookings Institution, pp. 72-102.

Frank RG, Berndt ER, Busch AB, Lehman AF. Quality-constant prices for the ongoing treatment of schizophrenia. An exploratory study. *Quarterly Review of Economics and Finance*. July 2004;44(3):390-409.

National Research Council, *At What Price: Conceptualizing and Measuring Cost-of-Living and Price Indexes*, Panel on Conceptual, Measurement and Other Statistical Issues in Developing Cost-of-Living Indexes, Charles L. Schultze and Christopher Mackie, eds., Committee on National Statistics, Division of Behavioral and Social Sciences and Education, Washington DC: National Academy Press, 2002:178-190.

Scitovsky, Anne A. 1964. "An Index of the Cost of Medical Care—A Proposed New Approach." In *The Economics of Health and Medical Care*, ed. by Solomon J. Axelrod, 128–142. Ann Arbor: University of Michigan, Bureau of Public Health Economics.

Shapiro, Irving L., Matthew Shapiro and David Wilcox [2001], "A Price Index for Cataract Surgery," in David M. Cutler and Ernst R. Berndt, eds., *Medical Output and Productivity*, Chicago: University of Chicago Press for the National Bureau of Economic Research, pp. 411-438.

Silver, Mick [2008] An Index Number Formula Problem: The Aggregation of Broadly Comparable Items. Paper presented at Second International Index Measures Congress, May 13-17, 2008, Washington, D.C.

Song X, Marder WD, Baser O, Houchens R, Conklin JE, and Bradley R. Can health care claims data improve the estimation of the Medical CPI? Paper presented at the National Bureau of Economic Research/Conference on Research in Income and Wealth Conference, Vancouver BC, June 28, 2004; also presented at the National Bureau of Research Summer Institute, Health Care Program, August 6, 2004.

Trajtenberg, M. (1990) *Economic Analysis of Product Innovation: The Case of CT Scanner*. Cambridge, Mass.: Harvard University Press.

Chart 1. Example of change in revenues with the Introduction of Medicare Part D

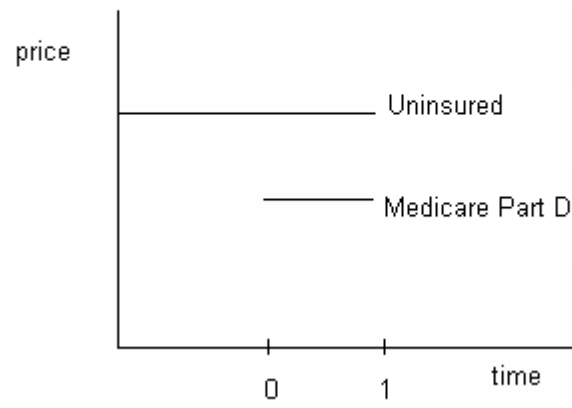
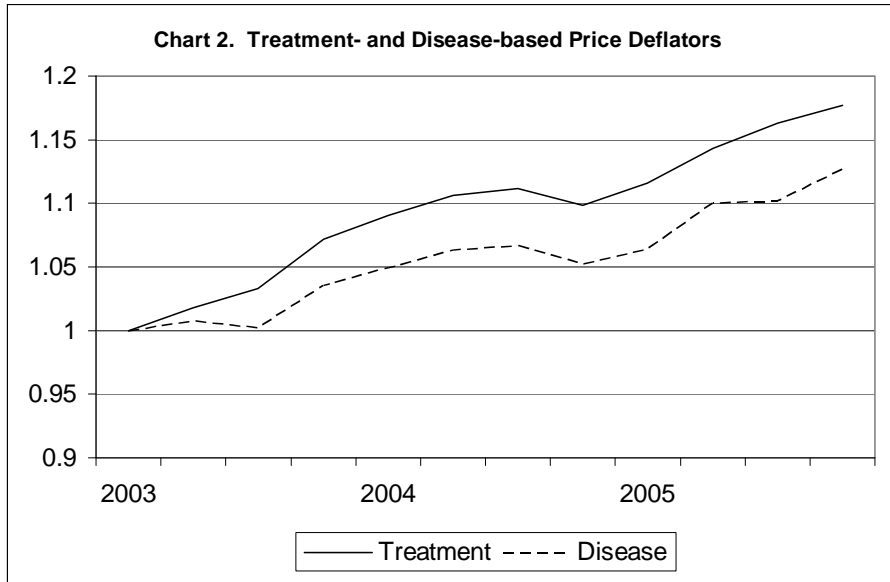


Table 1. Health Care Revenues by Place of Service, 2003:1

Place of Service	Dollars per Visit	Total Spending	Percent of Total
Ambulatory Surgical Center	\$1,079	\$53.3	1.4
Emergency Room-Hospital	\$370	\$134.1	3.6
Independent Lab	\$54	\$37.6	1.0
Inpatient Hospital	\$432	\$41.7	1.1
Inpatient Hospital Confinement	\$4,332	\$906.0	24.1
Office	\$105	\$935.3	24.9
Outpatient Hospital	\$337	\$482.5	12.8
Patient's Home	\$281	\$44.0	1.2
Pharmacy	\$87	\$960.9	25.6
Unknown	\$314	\$137.3	3.7
Rest	\$211	\$23.3	0.6
Total		\$3,755.9	100.0



Year	Quarter	Preferred		Treatment	
		Laspeyres	Fisher	Laspeyres	Fisher
2003	1	1.00	1.00	1.00	1.00
2003	2	1.01	1.01	1.02	1.02
2003	3	1.00	1.00	1.03	1.03
2003	4	1.04	1.03	1.07	1.07
2004	1	1.05	1.05	1.09	1.09
2004	2	1.06	1.06	1.11	1.10
2004	3	1.07	1.06	1.11	1.11
2004	4	1.05	1.05	1.10	1.10
2005	1	1.06	1.06	1.12	1.11
2005	2	1.10	1.10	1.14	1.14
2005	3	1.10	1.10	1.16	1.16
2005	4	1.13	1.12	1.18	1.17
Memo: CAGR		4.4%	4.2%	6.1%	6.0%

Table 2. Cost Savings from Treatment Shifts and Contributions, 2003:1--2005:4.

Major Disease Category	Preferred	Treatment	difference	Hospital		Office Visits	Drugs	ER	Lab	Home			Share of total cost
				Inpatient	Outpatient					Care	ASC	Other	
1 INFECTIOUS DISEASES	37.3%	32.9%	4.3%	3.4%	-0.7%	-0.8%	1.2%	0.3%	0.2%	0.6%	0.0%	0.0%	1.0%
2 ENDOCRINOLOGY	11.8%	14.9%	-3.1%	-2.8%	-1.0%	-2.2%	3.0%	-0.1%	-0.1%	0.5%	-0.1%	-0.4%	6.2%
3 HEMATOLOGY	18.8%	22.2%	-3.5%	-0.7%	-2.3%	-1.7%	-0.4%	0.0%	0.1%	0.0%	0.0%	1.4%	2.7%
4 PSYCHIATRY	3.1%	8.0%	-4.9%	-1.0%	-0.3%	-5.3%	2.3%	0.0%	0.0%	0.0%	0.0%	-0.7%	5.4%
5 CHEMICAL DEPENDENCY	18.3%	12.3%	6.0%	2.4%	-2.4%	-2.0%	3.4%	2.9%	0.1%	0.0%	0.1%	1.5%	0.7%
6 NEUROLOGY	15.4%	21.3%	-5.9%	-0.5%	-1.9%	-2.9%	0.5%	-0.3%	0.0%	0.0%	0.0%	-0.8%	5.9%
7 OPHTHALMOLOGY	8.4%	10.8%	-2.4%	-0.1%	-2.1%	-0.6%	-0.5%	-0.1%	0.0%	0.3%	0.8%	-0.2%	1.9%
8 CARDIOLOGY	1.7%	17.5%	-15.7%	-11.6%	-1.6%	-1.5%	-0.1%	0.1%	0.1%	0.1%	-0.1%	-1.0%	10.6%
9 OTOLARYNGOLOGY	9.2%	14.8%	-5.6%	0.1%	-2.6%	-2.0%	-0.8%	-0.2%	0.1%	0.2%	0.0%	-0.3%	8.3%
10 PULMONOLOGY	16.3%	18.9%	-2.6%	0.7%	-1.7%	-1.8%	0.0%	-0.5%	0.0%	0.3%	0.0%	0.4%	5.3%
11 GASTROENTEROLOGY	16.3%	21.6%	-5.2%	-1.3%	-2.7%	-2.0%	-0.5%	-0.1%	0.2%	0.0%	0.7%	0.4%	8.5%
12 HEPATOLOGY	9.4%	11.6%	-2.3%	0.3%	-1.6%	-0.5%	-1.7%	0.2%	0.0%	0.0%	0.2%	0.9%	3.3%
13 NEPHROLOGY	3.6%	10.2%	-6.6%	-0.2%	-5.9%	-0.5%	0.4%	0.0%	0.1%	-0.2%	-0.1%	-0.3%	1.2%
14 UROLOGY	7.0%	15.8%	-8.8%	-3.0%	-3.4%	-1.9%	0.2%	-0.2%	0.1%	0.1%	-0.1%	-0.6%	3.1%
15 OBSTETRICS	19.1%	16.1%	3.0%	3.1%	-0.5%	0.0%	-0.4%	0.2%	0.2%	0.1%	-0.1%	0.3%	5.1%
16 GYNECOLOGY	11.2%	21.0%	-9.8%	-3.0%	-2.8%	-3.0%	-0.5%	0.1%	-0.1%	0.0%	-0.4%	0.0%	7.4%
17 DERMATOLOGY	16.4%	19.3%	-3.0%	0.7%	-1.3%	-1.7%	-1.0%	-0.3%	0.3%	0.3%	-0.6%	0.6%	4.5%
18 ORTHOPEDICS & RHEUMATOLOGY	11.8%	18.0%	-6.2%	-1.1%	-2.8%	-1.4%	-0.1%	-0.2%	0.0%	0.6%	0.2%	-1.5%	16.0%
19 NEONATOLOGY	30.8%	28.7%	2.2%	2.1%	-0.1%	0.6%	-0.1%	0.0%	0.0%	-0.3%	0.0%	0.0%	2.9%

Table 3. Decomposition of Growth in Health Care Cost

	Price Change	Growth in Services	Growth in Costs
	(1)	(2)	(1)(2)
<u>Index Levels (2003:1=1.00)</u>			
Preferred Index	1.13	1.16	1.31
Treatment Index	1.18	1.11	1.31
Volume Measure	1.12	1.17	1.31
<u>Compound Annual Growth Rates</u>			
Preferred Index	4.4%	5.6%	10.3%
Treatment Index	6.1%	3.9%	10.3%
Volume Measure	4.1%	6.0%	10.3%

APPENDIX

This attached excel spreadsheet is an appendix that contains two types of tables for individual conditions, as defined by the Symmetry/Ingenix episode grouper.

Table A1 shows differences in the unit value indexes that include new and disappearing treatments (UV) and those that only include continuing treatments (UVC). Although these differences were negligible at a high level of aggregation (i.e., the 19 disease groups), this table shows that the differences can be large at the individual disease level. The fourth column of the table gives a ratio of the spending used in the two indexes and shows that diseases where the UV and UVC indexes differ tend to use different levels of spending in the two indexes. The last two columns show that the differences in spending levels are related to "thin" cells with little spending for few patients.

Table A2 gives contributions to cost savings (UVC/I) at the individual disease level. The first column contains the UVC indexes, the second the Laspeyres treatment indexes, and the third gives the difference in the two. The remaining columns give the contribution of each treatment class to the cost savings in the third column.