Preliminary and Incomplete
Please do not circulate

Changes in Treatment Intensity, Treatment Substitution and Price Indexes for Health Care Services

Ana Aizcorbe[1]
Nicole Nestoriak[2]

September 30, 2007

[1] BEA Thank Forrest McClure, folks at Pharmetrics, Ernie Berndt, Ralph Bradley, David Cutler and Allison Rosen and colleagues at BEA, Hoxie Ackerman wrote the code for the standard errors.
[2] BLS

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 by 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 expenditures on health.

This sector poses formidable measurement problems, as many of the assumptions underlying traditional price measurement do not hold here.[3] There are three broad types of issues. The first issue is how to define the "good"—that is, how to group transactions (or medical events) into homogeneous classes whose price can be tracked over time.[4] The second issue—not directly addressed here—is how to measure changes in the quality of treatments (i.e., outcomes).[5] Finally, there is the issue of aggregation: how to form some sort of average of the price changes associated with the homogeneous diseases to obtain an overall index for the cost of health care.[6]

This paper deals with the first issue. While statistical agencies have traditionally measured prices for health care services along treatment lines (e.g., separate price indexes for outpatient services, prescription drugs), the consensus among public health experts is that one must measure health care prices by disease (e.g., price indexes for the treatment of diabetes, treatment of depression). As discussed in the National Academy report "At What Price?" the issue is not merely one of convention or interpretation. Measurement by treatment types ignores potential substitutions across treatments that may reduce the cost of care to the patient. As explained in Berndt, Busch, Frank and Normand (2005):

> …Even though both the per session costs of psychotherapy and the costs of prescriptions for antidepressant pharmaceuticals may be increasing over time, to

[3] See "The Special Case of Medical Services" in National Research Council, *At What Price: Conceptualizing and Measuring Cost-of-Living and Price Indexes* (2002).
[4] This issue first arose in the context of prescription drugs and generics. See Fisher and Griliches (1995) and Griliches and Cockburn (2005) for a treatment of the problem in the context of cost of living index theory.
[5] This issue is a particularly thorny one, upon which there is no consensus. However, we see nothing in the quality adjustment issue that would overturn our work here. Instead, the quality issue is an added layer of complexity, separate from the treatment intensity issue.
[6] How one should aggregate is an open question. Much of the theory underlying standard aggregation techniques assumes consumers that maximize utility subject to budget constraints. Health is very different in that one doesn't choose when to get sick, once sick, one doesn't choose the treatment (doctors do) and, finally, one doesn't pay for these services directly.

> the extent higher cost psychotherapy sessions are substituted by increased use of lower cost antidepressant pharmaceuticals, it is possible that this substitution offset" compositional change results in net lower total episode treatment costs. In contrast, the BLS' fixed input bundle approach fails to incorporate the potential for such substitutability among treatment inputs." (P. 10)

Moreover, the intensity of treatments can also change over time. For example, the arrival of SSRIs could have prompted many to add drug therapy to their existing talk therapy visits to achieve a better outcome in the treatment of depression, rather than substituting from one treatment type to another. In this case, the cost of treating depression has risen but so has the well-being of patients. Our focus is on getting the change in costs right under the assumption that the outcomes are the same. To the extent that the arrival of new treatments increases the price and outcomes of treatment, our price indexes should be viewed as an upper bound to the cost of treating disease.[7]

Both of these issues can generate numerical differences between disease- and treatment-based price indexes. The importance of these issues has been studied for heart attacks (Cutler et al), cataract surgery (Shapiro and Wilcox) and various mental conditions (Berndt et al). The only attempt to assess whether this issue is important across a broader array of diseases is by Song et al, where they randomly sampled 40 or so diseases in three cities and compared indexes for that subset of diseases to treatment-based indexes. They failed to reject the hypothesis that the indexes were different in a statistical sense. However, that study has been viewed as inconclusive, as several have expressed reservations with the power of that test, given the relatively small sample size (Berndt).

We build on this body of work in two ways. First, we provide an algebraic treatment of the problem that formally exposits points that have been made in the literature. We are able to derive expressions that link the disease-based indexes advocated in the health economics literature to standard price indexes commonly used in the measurement literature. We use these expressions to derive necessary conditions for the validity of standard price indexes and apply those tests to a large database of health insurance claims. Moreover, we develop a decomposition of differences in the two types of indexes into contributions from the different treatment classes. This allows one to

[7] We are implicitly assuming that reductions in treatment that reduce outcomes do not happen.

identify the presence of treatment substitution even when the indexes are the same as well as to explain any divergences in the indexes. Our empirical contribution is the construction of disease-based price indexes over a comprehensive set of diseases. Extending the previous empirical work to cover more diseases allows us to assess the numerical importance of these issues over all diseases. We also construct bootstrapped standard errors under a more general set of assumptions than those previously used in the literature.

Our test for the validity of standard price indexes is rejected, demonstrating that shifts in treatment intensity have occurred in our relatively short time period: 2001:1 – 2003:4.. Second, those shifts do drive gaps between the disease-and treatment-based indexes; the average gap in our data is 1 percentage point on the compound annual growth rates. Third, we show that these gaps are generated mostly by treatment substitution (sometimes towards less expensive treatments sometimes not) and not by absolute changes in treatment intensity, where the number of treatments per patient all increase or decrease at the same time.

Our work on constructing bootstrapped standard errors to test whether the differences we find in these indexes are statistically significant and the construction of industry deflators that are consistent with the disease-based indexes is still to come.

Differences in Treatment- and Disease-Based Price Indexes for Diseases

Health economists have long argued in favor of price indexes that track the cost of treating disease. Exactly how this differs from a standard index number treatment for the price of health is not well understood. We begin with a discussion of the two indexes to highlight the differences in the two approaches and to obtain algebraic expressions to facilitate our empirical work.

Indexes Based on Price Per Treatment

Statistical agencies typically measure price change for health using treatment-based price indexes. In particular, they form Laspeyres indexes for each treatment type, $I_t(T)$, (like inpatient, outpatient, office visits and drugs) and then form an overall

Laspeyres index of those four price indexes. The notation I(T) denotes that the overall price index is a treatment-based index:

$$(1) \qquad I(T) = \sum_t \frac{\sum_d p_{d,t}^0 \, q_{d,t}^0}{\sum_t \sum_d p_{d,t}^0 \, q_{d,t}^0} \, I_t(T)$$

where the $p_{d,t}^0$ and $q_{d,t}^0$ denote the price and quantity of treatment, t, of disease d over some time period, 0, and $I_t(T)$ is the treatment-based index for treatment t:

$$(2) \qquad I_t(T) = \sum_d \frac{p_{d,t}^0 \, q_{d,t}^0}{\sum_d p_{d,t}^0 \, q_{d,t}^0} \left(p_{d,t}^1 / p_{d,t}^0 \right)$$

In what follows, it will be useful to note that while (1) states this overall price index as an average of price indexes for *treatments*, one can also rewrite the overall price index as an average of price indexes for *diseases*:

$$(1') \qquad I(T) = \sum_d \frac{\sum_t p_{d,t}^0 \, q_{d,t}^0}{\sum_t \sum_d p_{d,t}^0 \, q_{d,t}^0} \, I_d(T)$$

where $I_d(T)$ is the treatment-based index for disease d:

$$(2') \qquad I_d(T) = \sum_t \frac{p_{d,t}^0 \, q_{d,t}^0}{\sum_t p_{d,t}^0 \, q_{d,t}^0} \left(p_{d,t}^1 / p_{d,t}^0 \right) = \sum_t w_{d,t}^0 \left(p_{d,t}^1 / p_{d,t}^0 \right)$$

Notice that $I_d(T)$ is still a treatment-based index; it measures price change for individual treatments and does not capture any potential shifts across treatment classes. So, although it is possible to re-weight price relatives and form price indexes by disease as in (2')—they are still treatment-based price indexes and do not address the treatment substitution issue. The importance of this distinction is demonstrated below, where we build a mapping from the indexes in (2') to disease-based price indexes.

Indexes Based on Price Per Patient

Consider a price measure that tracks the cost of treating some disease, d, where different types of treatments, t, may be required. We assume that the diseases can be defined so that they are "clinically homogeneous." That is, the medical conditions within the disease bucket are the same across patients and over time.

Health economists have advocated the use of episode-based indexes to measure changes in the cost of treating disease. 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 includes costs accrued in three previous quarters. This presents problems for national income accounting because our goal is to form price indexes to deflate *current* spending. Therefore, we will consider the cost of treating disease over some period of time, a quarter. And, instead of tracking the cost of episodes, we will track the cost of treating patients over a quarter. This is essentially what Cutler and Rosen are doing in their National Health Accounts study.

If P_d^0 is the number of patients under treatment for condition d at time 0, then we define the average cost of treating that condition in period 0 as the number of dollars of treatment spent on all treatments divided by the number of patients receiving the treatment:

$$p_d^0 = \Sigma_t (p_{d,t}^0 \, q_{d,t}^0) / P_d^0$$

The change in this price over time, called the *price relative*, is simply the ratio of the two prices:

$$(3) \qquad I_d(D) = \frac{\Sigma_t (p_{d,t}^1 \, q_{d,t}^1) / P_d^1}{\Sigma_t (p_{d,t}^0 \, q_{d,t}^0) / P_d^0}$$

The idea is that medical conditions are normally treated with bundles of treatments. Tracking the cost of treating disease, therefore, is best done by tracking the price of the bundle of treatments rather than tracking prices of treatments separately. In order to view this price relative as a *constant-quality* price relative, one must assume that the treatment bundles provide the same medical outcomes. To the extent that the outcomes associated with new treatments are at least as good as those of the old treatments, our measures may be viewed as upper bounds to true price change.

The price relative in (3) will vary with changes in the prices of treatments as well as changes in the intensity of treatment. With regard to changes in intensity, the index will increase as the number of treatments per patient increases; since we've assumed that outcomes are the same, more treatment to achieve the same outcome means the price of treatment has gone up. The index also captures changes in the mix of treatments. Though we normally think of patients as gravitating towards lower-cost treatments that provide the same medical outcomes, one can imagine cases where patients move towards new, more-costly treatments that provide better outcomes. Again, in this case, our price index would be viewed as an upper bound to constant-quality price change, since it ignores any potential improvements in outcomes.

<u>Comparison of the two types of indexes</u>

It turns out that the price relative in (3) may be restated as a Laspeyres price index that is adjusted for treatment intensity. We, therefore, rename the price relative in (3) as a *disease-based price index* for condition d, and restate it as:

$$(3') \qquad I_d(D) = \Sigma_t \frac{p_{d,t}^0 \, q_{d,t}^0}{\Sigma_t \, p_{d,t}^0 \, q_{d,t}^0} \; \frac{q_{d,t}^1 / P_d^1}{q_{d,t}^0 / P_d^0} \; (p_{d,t}^1 / p_{d,t}^0)$$

$$= \Sigma_t \, w_{d,t}^0 \, \gamma_{d,t} \, (p_{d,t}^1 / p_{d,t}^0)$$

Each term in the summation corresponds to a particular treatment. For each treatment, the expenditure shares ($w_{d,t}^0$) and price relatives ($p_{d,t}^1 / p_{d,t}^0$) are those found in the usual Laspeyres index for the price of treatments in (2'). Thus, we can view the disease-based index as a special Laspeyres, where the $\gamma_{d,t}$'s adjust for changes in treatment intensity, the $\gamma_{d,t}$ terms; they are literally the change in the number of treatments of type t for patients under treatment for condition d.[8]

One way to think about this index is as a Laspeyres that uses a different price relative: i.e., it uses changes in the price per patient $(p_{d,t}^{1} q_{d,t}^{1} / P^{1})/(p_{d,t}^{0} q_{d,t}^{0} / P^{0})$ rather than changes in the prices of individual treatments $(p_{d,t}^{1} / p_{d,t}^{0})$. Again, if the outcome of treatment is the same, but it now requires more dollars per patient than in the base period, then the cost of achieving that outcome has risen.

Comparing (3) and (2'), it is clear that a sufficient condition for the equivalence of the disease- and treatment-based price indexes for disease d is if the treatment intensity terms equal 1 for all treatments:

Proposition 1. The treatment-based price index in (2') reduces to the disease-based price index in (3) if $\gamma_{d,t} = 1$ *for all t.*[9]

The implications of this for the validity of treatment-based price indexes used by statistical agencies is clear: if (2') reduces to (3), then the overall treatment-based price index reduces to the overall disease-based price index.

This essentially tests that the translation from treatments (inputs) to patients (outputs) is Leontief, a necessary and sufficient condition for the absence of treatment substitution but only a sufficient condition for there to be a gap between the two indexes. This is because there could be increasing intensity in some treatments that offset the

---

[8] A more intuitive way to define the price relative—more in line with the way that prices are defined in standard price indexes—would be with prices defined per treatment, not per patient. It turns out, however, that using the number of treatments as a basis for the price introduces a units problem, not unlike adding apples and oranges. To see this, suppose that P measured the total number of treatments (e.g., $P_{d}^{1}=\Sigma_{t} q_{d,t}^{1}$). If we changed the way we measure drug treatments from number of prescriptions to the number of days of treatment, we are changing the units of only some of the q's in the denominator and the two ways of measuring drugs would yield different measures of price change. Perhaps an easier way to see it is to consider what happens if we measure the number of hospital stays in thousands without changing the way we measure other treatments. That change in units would not cancel out in the above price relative.

[9] There are other conditions under which this price index reduces to a Laspeyres index of the type used at the BLS. One is that if, instead of the shares being constant, if the prices move in constant proportions, $(c_{d}^{1} / c_{d}^{0}) = (c_{t}^{1} / c_{t}^{0})$ then the two indexes are the same. An important difference in the two expressions, however, has to do with the treatment of new goods. As is well known, the matched-model index in (2) ignores any price relatives where either price is missing – either a new good enters the market, in which case the expenditure weights are zero or a good exits, in which case the price in period 1 is missing. In contrast, the price index in (1) does not require that one exclude treatments where there has been turnover.

effects of decreasing intensity in other treatments on the price index. In that case, the price indexes could be the same, despite the presence of substitution.

A Decomposition

This mapping between disease- and treatment-based indexes is also useful in identifying sources of differences in the two price indexes. In particular, we show that differences in the indexes do not require treatment substitution and that the presence of treatment substitution does not guarantee that the indexes will diverge.

To see this, consider the following decomposition of differences in the two indexes into treatments:

$$(4) \qquad I_d(D) - I_d(T) = \Sigma_t \frac{p_{d,t}^0 \, q_{d,t}^0}{\Sigma_t \, p_{d,t}^0 \, q_{d,t}^0} \left[\frac{q_{d,t}^1 / P_d^1}{q_{d,t}^0 / P_d^0} - 1 \right] \left(p_{d,t}^1 / p_{d,t}^0 \right)$$

$$= \Sigma_t \, w_{d,t}^0 \, (\gamma_{d,t} - 1) \left(p_{d,t}^1 / p_{d,t}^0 \right)$$

With four treatment types, the differences in the disease- and treatment-based price indexes for each disease can be parsed out into the four summations in (4) in order to interpret the sources underlying the differences in the indexes. To see that differences in the indexes do not require substitution, consider a case where $\gamma_{d,t} > 1$ for all treatments. In that case, $I_d(D) > I_d(T)$ because the price per patient has increased. Second, 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: as $\gamma_{d,t} < 1$ for office visits and $\gamma_{d,t} > 1$ for drugs. This, all else held equal, would cause $I_d(D) < I_d(T)$. However, with more than two treatment types, it is possible that there could also have been a shift from outpatient to the more expensive inpatient care that could dominate the substitution towards drugs and ultimately cause $I_d(D) > I_d(T)$. These examples suggest the following definitions.

Definition: *Treatment substitution occurs when at least one of the treatment intensity terms, ($\gamma_{d,t}$ – 1), has a different sign than the rest.*

Definition: *<u>Absolute changes in treatment intensity</u> occurs when all of the treatment terms ,(* $\gamma_{d,t} - 1$*) have the same sign.*

We use the decomposition in (4) in our empirical work to identify shifts in treatment intensity and their impact on price indexes. We also note that this decomposition is just a formal restatement of ideas in Cutler's work. (cite)

Differences in Treatment- and Disease-Based Price Indexes for Treatments

While all can agree that disease-based indexes are the way to go, there is a practical problem in that one may still want to have price indexes by industry (what we have, until now, been calling "treatment") for all sorts of purposes (to measure productivity by industry, for example). The algebraic mapping we used above can be applied to address this issue. In particular, we form price indexes for each treatment class that are based on price per patient (not price per treatment).

Specifically, we exploit the following identity to rewrite the overall disease-based index in terms of treatment subindexes:

$$(5) \qquad I(D) = \sum_d \frac{\sum_t p_{d,t}^0 \, q_{d,t}^0}{\sum_t \sum_d p_{d,t}^0 \, q_{d,t}^0} \left[I_d(D) \right]$$

$$= \sum_t \frac{\sum_d p_{d,t}^0 \, q_{d,t}^0}{\sum_t \sum_d p_{d,t}^0 \, q_{d,t}^0} \left[I_t(D) \right]$$

where

$$I_t(D) = \sum_d \frac{p_{d,t}^0 \, q_{d,t}^0}{\sum_d p_{d,t}^0 \, q_{d,t}^0} \, \frac{q_{d,t}^1 / P_d^1}{q_{d,t}^0 / P_d^0} \left(p_{d,t}^1 / p_{d,t}^0 \right)$$

While the usual practice is to deflate industry (what we've been referring to as treatments) using I(T)'s, using the I(D)'s instead will yield an overall price index that will line up exactly with that obtained from the disease side.

In our empirical work, we will construct these indexes and compare them with the standard treatment indexes (i.e., the $I_t(T)$'s).

Aggregation over diseases

The differences in the two types of indexes for each disease may be aggregated up to a measure of the differences in the overall indexes using the Laspeyres formula:

$$I(D) - I(T) = \sum_d \frac{\sum_t p_{d,t}^0 \, q_{d,t}^0}{\sum_t \sum_d p_{d,t}^0 \, q_{d,t}^0} \left[I_d(D) - I_d(T) \right] \tag{6}$$

This measures the difference between a Laspeyres of the disease-based indexes in (3) over all diseases and a Laspeyres of the treatment-based price indexes in (1). The former has the interpretation of changes in the price per patient of treating all existing diseases while the latter reduces to precisely the type of index that statistical agencies use (as shown in (1').

Empirical Implementation

We apply the formulas above to a large dataset of claims for the commercially-insured. It is a 10 percent sample of a larger database maintained by Pharmetrics, Inc. Unfortunately, the data are not representative and, so, the estimates shown here are only suggestive. We plan to design a representative sample that can then be pulled from their larger database.

Our current sample contains over 140 million claims records for over 40 health insurance plans covering the period 2001 to 2003. (there will be a data appendix with a fuller description of the data). We exclude patients over 65 years of age, since many of their claims are filed through Medicare and are not in our dataset. The data have been processed with the Ingenix episode grouper and we use those definitions to identify the medical condition associated with each claim. 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.[10]

We use four types of treatments—inpatient, care, outpatient care, office visits and prescription drugs—the same classes traditionally used by statistical agencies.[11] We define the elementary prices as, loosely speaking, the price per encounter—an inpatient stay (not the price per day), an outpatient stay, an office visit and a prescription (not the number of pills or the number of days).

We do not attempt to mimic the BLS procedures for our empirical work. Instead, we assume that the prices for the four types of treatments are constructed in a manner consistent with the disease-based indexes. In particular, we use unit values (expenditures divided by number of treatments) in each ETG class as the elementary price. For inpatient, outpatient, and office visits, this is not too different from what is usually done, except that we assume that the elementary price is formed by disease (not just by treatment).

For drugs, this is a very different concept from what is typically done. 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. The way we do it counts shifts from cheaper to expensive drugs as in increase in price whereas the usual way does not.

Testing the Stability of Treatment Intensity

We use proposition 1 and a simple procedure to test for shifts in treatment intensity at a disaggregate level. For each of the 600 or so ETGs, we form a time series of the treatment intensity for each of the treatment types. Then, we test, for each treatment, whether the mean of the time series equals one. Our quarterly data covers the period 2001:1-2003:4 so that each time series contains 12 observations. One would think

[10] 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."

that using so few observations would yield a weak test but, as can be seen in table 1, the null hypothesis that the translation from treatments to patients is Leontief is soundly rejected for most ETGs. The null hypothesis of no change in treatment intensity is rejected in about 90 percent of ETGs, with 90 percent confidence; it is rejected for about 80 percent of the ETGs at the 99 percent confidence level.

For these ETGs, we conclude that treatment-based indexes will likely not provide an accurate read for changes in the cost of treating illness.

Comparison of Price Indexes by Disease

Do these shifts in treatment intensity cause the disease- and treatment-based indexes to diverge? Chart 1 shows that they do and that the differences can be large. The plots show overall disease- and treatment-based indexes, constructed as Laspeyres aggregates of the treatment-based indexes in (as in (1)) and a Laspeyres of the disease-based indexes in (3). The indexes are constructed relative to 2001:1 and not as quarter-to-quarter indexes that are then chained over time.

As may be seen, the overall disease-based index typically shows slower price growth than the overall treatment-based index though, in these data, the gap is narrowed by 2003:4. Because the indexes are relative to 2001:1, the plot suggests that something occurred over 2001 that caused the indexes to diverge, and the divergence lasted through 2002 and began to narrow in 2003. Over most of this time period, then, the cost of treating patients rose less fast than the average change in the cost of individual treatments.

To assess the statistical significance of these gaps, we constructed standard errors using bootstrap techniques. For the overall indexes, we constructed 1) an overall treatment-based price index—over all treatments and diseases—as well as 2) an overall disease-based price index and 3) the difference between the two. Using the 250 replicates yields 250 such calculations and we report the mean and standard deviation for these distributions as a test for whether the two types of indexes differ in a statistical sense. [still to come]

[11] The BLS PPI program actually lumps inpatient and outpatient care in one category. However, it is not clear how the price relative is formed.

Decomposition

To explore the sources of differences in these two indexes, we take the 2003:3 value of the disease-specific deflators and calculate the decomposition of the differences in the two indexes (in (4)) at the ETG level. Table 2 shows these results for ETGs corresponding to three diseases that have been studied extensively in the literature—heart attacks (AMI), cataract and mental illness conditions. The 2003:3 values for the disease- and treatment-based indexes for each ETG are shown in the first two columns and their difference is given in the third column. For depression, anxiety disorders and many of the AMI conditions, the disease-based price indexes typically rose slower than the treatment-based price indexes from 2001:1 to 2003:4, that is, the differences in the third column are negative and can be sizable. In contrast, the effect works in the other direction for cataract.

The remaining columns of the table give the contributions to these differences from the four types of treatments. For minor depression and anxiety disorders, for example, a shift in treatments from inpatient, outpatient, and office visits towards drugs causes the disease-based index to fall more rapidly than the treatment-based indexes. That is, the number of treatments per patient (the treatment intensity terms) fell for all treatment types except drugs from 2001:1 to 2003:3 while that for drugs increased. This is consistent with the findings in Ernie's study that the arrival of drugs generated substitution from talk-therapy towards drug therapy.

There are examples where the differences in the indexes do not involve treatment substitution, but absolute changes in treatment intensity. For the first ETG in the cardiovascular disease category, the differences in the indexes are generated by a decline in the use of inpatient visits to treat patients with little change in the intensity of other treatments. Similarly, the ETG for cataract, with surgery, shows an absolute increase in treatment intensity (all in the outpatient treatment class).

For these diseases, then, we see three types of patterns, absolute increases, decreases, and substitution of treatments. Looking over the entire 600 or so ETGs, table 3 provides summary statistics for treatment shifts. The first two columns show that relatively few ETGs involve absolute changes in treatment intensity; only 61 ETGs show increases in some treatment with no attendant decline in others and 62 other ETGs show declines in treatment without a treatment type showing an increase. The remaining ETGs show some degree of substitution—some treatments became more intensive, others less so.

The table also provides information on three types of treatment substitution that many would argue are taking place over this period. First, the arrival of microsurgeries and anesthesias that allow surgery to occur without an inpatient stay are thought to have generated shifts from inpatient to outpatient stays. We see some evidence of that in our data, as 209 of the ETG classes show a decline in the number of inpatient stays at the same time that there is an increase in the number of outpatient stays per patient. Second, ambulatory surgical centers have become a substitute for for elective and relatively minor surgical procedures that can also be performed as an outpatient. The ASC association documents a large increase in the number of procedures performed at these centers that has allowed lower prices, since these centers do not have the high overhead costs typically seen at traditional hospitals. Consistent with this, 133 of the ETG classes in our data show reductions in the intensity of outpatient treatments accompanied by increases in the number of office visits per patient. Finally, innovation in the pharmaceutical industry has generated prescription drugs that allow for treatment regimens that involve fewer office visits. 106 of the ETG's in our data show patterns consistent with that trend.

Finally, table 4 provides summary information by broad disease classes. The decomposition for orthopedic conditions is consistent with a shift from inpatient to outpatient care. The data for psychiatric conditions are consistent with a shift towards drug therapy. Working in the other direction, there are large differences in the indexes for the nephrology and obstetric classes, owing to increases in intensity of outpatient treatment (nephrology) and all hospital stays (obstetrics).

Comparison of Price Indexes by Treatment

[still to come]

Conclusion

We have developed an algebraic framework that yields a statistical test for the validity of standard, treatment-based price indexes. Our application of that test to our health claims data rejects the validity of those indexes. A comparison of disease- and treatment-based indexes constructed over a comprehensive list of medical conditions shows nontrivial gaps in the two indexes. In our data, the disease-based index rises slower than the standard one. Third, our algebraic framework provides a decomposition that one may use to identify the sources of any differences in the two indexes and to test for the presence of treatment substitution.

In future work, we hope to construct these measures on nationally-representative data and extend our analysis to volume measures.

Table 1. Percent of ETG Classes that reject H0: $\gamma_i = 1$, by level of significance

	Inpatient	Outpatient	Office Visits	Pharmacy
> 90%	90.9%	89.4%	94.4%	96.1%
> 95%	87.5%	86.4%	91.6%	94.1%
> 99%	83.0%	78.8%	85.5%	88.0%

Chart 1. Price Indexes for Health Care Services Paid by Commercially Insured Patients, 2001:1 - 2003:4

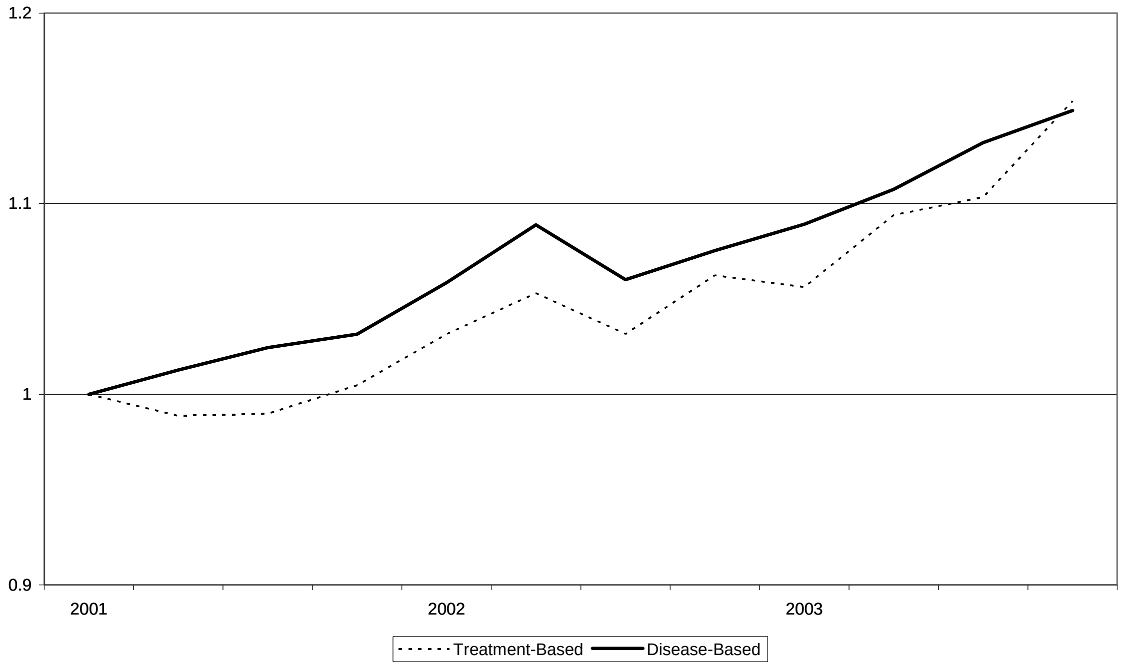

Table 2. Decomposition of differences in price indexes, selected ETGs, 2001:1-2003:3.

	ETG Class	Price Indexes Disease	Treatment	Difference	Contributions to differences from: Inpatient	Outpatient	Office Visits	Drugs	
Cardiovascular	Coronary heart disease, with AMI, with coronary artery bypass graft	1.01	1.34	-0.33	-0.37	0.04	0.00	0.00	$723,974
	Coronary heart disease, with AMI or acquired defect,with valvular procedure	1.48	2.95	-1.48	0.00	0.00	0.00	-1.48	$82
	Coronary heart disease, with AMI, with angioplasty	1.02	1.08	-0.06	-0.09	0.03	0.00	0.00	$1,749,233
	Coronary heart disease, with AMI, with cardiac catheterization	0.56	1.12	-0.56	-0.55	-0.01	0.00	0.00	$294,519
	Coronary heart disease, with AMI, anterior wall with complication	4.29	1.44	2.85	0.00	3.25	-0.05	-0.35	$1,456
	Coronary heart disease, with AMI, anterior wall w/o complication	0.49	0.56	-0.07	-0.05	0.04	-0.03	-0.02	$27,712
	Coronary heart disease, with AMI, inferior wall with complication	2.63	1.27	1.36	1.57	-0.16	-0.04	-0.01	$42,741
Eye and Adnexa	Cataract, with surgery	1.19	1.08	0.11	0.00	0.11	-0.01	0.00	$633,031
	Cataract, w/o surgery	1.15	1.06	0.09	0.00	0.08	0.03	-0.02	$92,086
Psychiatric	Major depression	0.94	0.98	-0.04	-0.06	0.00	0.00	0.01	$2,841,809
	Minor depression	0.95	1.03	-0.08	-0.02	-0.02	-0.06	0.03	$3,049,439
	Anxiety disorder or phobias, major	1.03	1.08	-0.05	0.00	0.00	-0.04	-0.01	$222,484
	Anxiety disorder or phobias, minor	1.01	1.10	-0.09	0.00	-0.05	-0.10	0.06	$454,659

Table 3. Changes in Treatment Intensity.

	Number	Percent
Absolute changes		
Increase ($\gamma_i<1$, for all i)	66	11.3%
Decrease ($\gamma_i<1$, for all i)	59	10.1%
Relative changes		
IP--> OP	186	31.7%
OP--> OV	82	14.0%
OV--> PH	177	30.2%

Table 4. Alternate Price Indexes by Major Disease Group

Disease Category	Price Index			Contribution from:				Expenditures
	Disease	Treatment	Difference	Inpatient	Outpatient	OfficeVisits	Pharmacy	2001:1
	(1)	(2)	(1)-(2)					
Cardiovascular	0.99	1.10	-0.11	-0.11	0.01	-0.01	0.00	$2,451,657
Drug Dependence	0.91	0.91	-0.01	-0.08	0.06	0.01	0.00	$651,770
Endocrine	1.14	1.19	-0.05	-0.06	-0.01	-0.02	0.03	$2,031,106
Eye and Adnexa	1.12	1.08	0.04	0.00	0.05	0.00	-0.01	$607,514
Gastroenterology	1.10	1.12	-0.02	-0.02	0.00	-0.01	0.01	$1,634,174
Genitourinary	1.17	1.18	-0.01	-0.01	0.00	-0.01	0.00	$1,076,817
Gynecology	1.08	1.14	-0.06	-0.06	0.00	-0.01	0.00	$1,951,853
Hemotology	1.18	1.18	0.00	-0.03	0.01	0.00	0.03	$1,473,076
Hepatology	1.37	1.33	0.04	0.00	0.02	0.00	0.02	$1,300,476
Infectious Diseases	1.16	1.15	0.01	-0.01	0.03	-0.01	-0.01	$595,270
Neonatal	1.36	1.37	0.00	-0.01	0.00	0.00	0.00	$5,731,436
Nephrology	1.39	0.96	0.43	-0.05	0.42	0.06	0.00	$2,456,443
Neurological	1.08	1.16	-0.08	-0.09	0.00	-0.01	0.01	$922,574
Obstetric	1.21	1.12	0.09	0.06	0.03	-0.01	0.00	$2,886,619
Orthopedic	1.05	1.08	-0.03	-0.04	0.02	-0.01	0.00	$4,479,414
Otolaryngology	1.11	1.17	-0.06	0.00	0.01	-0.04	-0.02	$1,444,546
Preventative and Administrative	1.07	1.09	-0.02	0.00	-0.03	0.00	0.01	$8,054,222
Psychiatric	0.90	1.01	-0.11	-0.08	-0.02	-0.02	0.01	$2,741,290
Pulmonary	1.15	1.18	-0.03	0.01	0.01	-0.04	-0.01	$852,292
Skin Disorders	1.12	1.14	-0.01	0.00	0.00	-0.01	-0.01	$2,573,299