Volume Title: Price Index Concepts and Measurement

Volume Author/Editor: W. Erwin Diewert, John S. Greenlees and Charles R. Hulten, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-14855-6

Volume URL: http://www.nber.org/books/diew08-1

Conference Date: June 28-29, 2004

Publication Date: December 2009

Chapter Title: Can A Disease-Based Price Index Improve the Estimation of the Medical Consumer Price Index?

Chapter Author: Xue Song, William D. Marder, Robert Houchens, Jonathan E. Conklin, Ralph Bradley

Chapter URL: http://www.nber.org/chapters/c5083

Chapter pages in book: (329 - 368)

8

Can A Disease-Based Price Index Improve the Estimation of the Medical Consumer Price Index?

Xue Song, William D. Marder, Robert Houchens, Jonathan E. Conklin, and Ralph Bradley

8.1. Introduction

This chapter examines the effects of two separate factors that make it particularly challenging to construct health care price indexes in the United States. The first challenge is to obtain real prices for representative medical treatments. The widespread use of third-party reimbursement for services covered by health insurance plans puts the consumer of health care, the patient, in the unusual position of having another institution pay for the bulk of services consumed. Third-party reimbursement is characterized by complicated price negotiations that are not visible to the consumer at the time of purchase. It is also not visible to the Bureau of Labor Statistics (BLS) data collection efforts that depend on point-of-purchase surveys and follow-up monthly price checks. Thus, questions can be raised about the accuracy of the Medical Care component of the Consumer Price Index (MCPI), which relies on the general approach of the BLS to price the same bundle of goods and services that would be purchased by consumers. Which health plan reimbursement negotiations would be relevant or accessible to the data collector? In the absence of special efforts, the data collector is most likely to capture a (possibly discounted) list price, rather than an appropriately sampled real transaction price.

The second challenge is to keep pace with treatment innovations. Like

Xue Song is research leader at Thomson Reuters. William D. Marder is senior vice president and general manager of Thomson Reuters. Robert Houchens is senior director of Thomson Reuters. Jonathan E. Conklin is vice president of Thomson Reuters. Ralph Bradley is a research economist at the U.S. Department of Labor, Bureau of Labor Statistics.

Support for this work was provided to Thomson Reuters under contract to the Bureau of Labor Statistics. Opinions expressed are those of the authors and do not represent the official position of the Bureau of Labor Statistics.

many parts of the economy, the health care sector in recent years has experienced rapid technological change. New drugs have been introduced that can radically alter the style of treatment available for many common and rare conditions. New surgical and medical techniques have been developed and put into widespread use. Consequently, the treatment of many conditions has moved away from inpatient settings to outpatient settings or prescription drugs.

The nature of demand for health care services provides opportunities for measurement that are not applicable in most other sectors. Following Grossman (1972), physician visits, prescription drugs, and overnight stays in hospitals are not viewed as direct arguments in a consumer's utility function. The demand for health care services is a derived demand generated from an underlying demand for health, not health care. Health can be produced through preventive services in advance of illness or with curative services in the event of illness. By examining episodes of care for carefully selected illnesses, a number of authors (e.g., Berndt et al. 2002; Cutler, McClellan, and Newhouse 1998, 1999) have successfully examined the changing price of treatment for specific illnesses, such as depression and acute myocardial infarction. These studies look at the types of treatments patients receive to help them recover from illness. The ultimate demand is for recovery. As the technology available to health care providers improves, the inputs used in an episode of care will change. By measuring the total cost of the restructured episode, these authors were able to track the price of care.

Based largely on this evidence, a Committee on National Statistics (CNSTAT) panel recommended that the BLS develop an experimental version of the MCPI that derives prices for the total treatment costs of randomly sampled diagnoses.[1] Additionally, CNSTAT suggested that instead of collecting price quotes directly from providers, the MCPI could use the reimbursement information on retrospective claims databases. Pricing based on diseases and treatment episodes allows for medical care substitution across medical inputs in the treatment of patients. Because it does not rely on subjective response, claims-based pricing also eliminates respondent burden and may have the advantages of larger sample size and greater data validity.

This study uses medical insurance claims data to investigate both issues: (a) obtaining real prices for representative medical treatments to examine the impact of third-party reimbursement on measured trends in health care inputs of prescription drugs, physician services, and hospital services; and (b) capturing the substitution effects of health care inputs on the trend in medical care prices captured by episodes of care for some randomly selected conditions.

In section 8.2, we describe the data that are employed. Section 8.3 focuses on the replication analysis of the current BLS methodology. Section 8.4 provides the analysis of episodes of care, and the results are summarized

1. See Schultze and Mackie (2002).

in section 8.5. Section 8.6 discusses potential improvements that could be applied to studies in this area and the limitations of relying solely on claims data to produce medical CPI.

8.2 Data

Data for this study come from the MarketScan® Research Databases from Thomson Reuters. These databases are a convenience sample reflecting the combined health care service use of individuals covered by Thomson Reuters employer clients nationwide. Personally identifiable health information is sent to Thomson Reuters to help its clients manage the cost and quality of health care they purchase on behalf of their employees. MarketScan is the pooled and deidentified data from these client databases. Two MarketScan databases are used in this MCPI study: the Commercial Claims and Encounters (Commercial) Database and the Medicare Supplemental and Coordination of Benefits (COB; Medicare) Database.

The Commercial Claims and Encounters Database contains the health care experience of approximately four million employees and their dependents in 2002. These individuals' health care is provided under a variety of fee-for-service, fully capitated, and partially capitated health plans, including preferred provider organizations, point-of-service plans, indemnity plans, and health maintenance organizations. The database consists of inpatient admissions, inpatient services, outpatient services (including physician, laboratory, and all other covered services delivered to patients outside of hospitals and other settings where the patient would spend the night), and outpatient pharmaceutical claims.

The 2002 Medicare Supplemental and COB Database contains the health care experience of almost nine hundred thousand individuals with Medicare supplemental insurance paid for by employers. Both the Medicare-covered portion of payment (represented as the COB amount) and the employer-paid portion are included in this database. The database also consists of inpatient admissions, inpatient services, outpatient services, and outpatient pharmaceutical claims.

Our analysis is limited to three metropolitan areas that serve as primary sampling units (PSUs) for the BLS MCPI and that have significant numbers of covered lives captured in MarketScan databases. These metropolitan areas are New York City (CPI area A109), Philadelphia (A102), and Boston (A103). While the number of covered lives in each of the cities varies by year, MarketScan has many more respondents in Boston (146,000 in 1998) than in Philadelphia (104,901) or New York (43,520).

8.3 Replication of the Medical CPI

The BLS CPI is constructed using a two-stage process. In the first stage, price indexes are generated for 211 different item categories for each of thirty-eight urban areas. The indexes in the first stage are then used to generate an all-items-all-cities index. The overall medical CPI is an expenditure weighted average of such item indexes. Although the medical CPI includes eight of the 211 item categories—including, for example, dental services, nonprescription drugs, and medical supplies—this study only constructed price indexes for prescription drugs, physician services, and hospital services.

The initial BLS sample at the item-area level is implemented with two surveys. The first is a telephone point-of-purchase survey (TPOPS), where randomly selected households are asked where they purchase their medical goods and services and how much they spend at each outlet. In the second survey, the results of TPOPS are used to select outlets where the probability of selection for a particular outlet is proportional to its expenditure share in TPOPS.

Once an outlet is drawn, the BLS field representative goes to the outlet to select either a good or a service that falls within a certain item category. There is a detailed checklist of important characteristics of the item. The field representative determines the expenditure share for each characteristic, and the probability that an item is drawn is proportional to the expenditure share of its characteristics within the outlet. For pharmaceuticals, a key characteristic is the National Drug Code (NDC); for physicians, it is the Current Procedure Terminology (CPT) code; and for hospitals, it is based on the Diagnosis Related Group (DRG).

Once the outlets and items are selected, they stay in the BLS sample for four years.[2] The implicit assumption of this fixed sample is that the inputs used to treat each specific disease are constant. As Cutler, McClellan, and Newhouse (1998, 1999) and Shapiro and Wilcox (1996) argued, if less expensive inputs are substituted for more expensive ones, this will not be reflected as a decrease in the BLS price index.

On a monthly or bimonthly basis, the BLS reprices the items in its sample.[3] For all medical items except pharmaceuticals, the BLS generates an arithmetic mean (Laspeyres-type) price index in each area. For pharmaceuticals, a geometric mean index is computed. The Laspeyres formula is then used to aggregate the area indexes to the national level.

No claims database contains the information needed to precisely mimic these procedures. Appendix A provides the eleven detailed steps we took to create analytic files that would provide as much of the information just described as possible. All outlets and items were selected using probability in proportion to size with replacement, the same method that the BLS uses to collect its samples.

2. Beginning in 2001, the BLS began reselecting prescription drugs within its outlet sample at two-year intervals—that is, midway between outlet resamplings.

3. Most areas have on-cycle and off-cycle months. For some areas, the on-cycle months are the even ones, and for others, they are the odd ones. Repricing is only done in the on-cycle months, and the price index represents the price change over a two-month period.

We developed two sets of input-based indexes. One is based on the same sample sizes as those of the BLS MCPI,[4] and the other is based on much larger sample sizes (ten times as large as the BLS sample sizes wherever possible). We used the small-sample index to investigate if the price distribution was statistically different between the claims database and the BLS sample. The large-sample index was intended to examine whether the sample sizes had a significant impact on indexes.[5]

8.4 Episode-Based Price Indexes

A number of studies previously cited have studied the changing cost of treating specific illnesses by examining episodes of care for those illnesses and how the cost of a treatment episode changed over time. Based on that literature, the CNSTAT recommended a study of a generalization of this approach (Schultze and Mackie 2002, 6–9):

> BLS should select between 15–40 diagnoses from the ICD (International Classification of Diseases), chosen randomly in proportion to their direct medical treatment expenditures and use information from retrospective claims databases to identify and quantify the inputs used in their treatment and to estimate their cost. On a monthly basis, the BLS could re-price the current set of specific items (e.g., anesthesia, surgery, and medications), keeping quantity weights temporarily fixed. Then, at appropriate intervals, perhaps every year or two, the BLS should reconstruct the medical price index by pricing the treatment episodes of the 15 to 40 diagnoses—including the effects of changed inputs on the overall cost of those treatments. The frequency with which these diagnosis adjustments should be made will depend in part on the cost to BLS of doing so. The resulting MCPI price indexes should initially be published on an experimental basis. The panel also recommends that the BLS appoint a study group to consider, among other things, the possibility that the index will "jump" at the linkage points and whether a prospective smoothing technique should be used.

8.4.1 Description of Medstat Episode Grouper

In order to implement the committee's recommendation with the data available for this study, we used the Medstat Episode Grouper (MEG) to transform a stream of claims data into episodes of care for the full range

4. The BLS sample sizes are:

City	Drug	Physician	Hospital
Philadelphia	34	32	31
Boston	42	27	46
New York City	41	35	59

5. McClelland and Reinsdorf (1999) found that the small sample bias of the geometric means index was larger than that of the seasoned index.

of conditions covered by the ICD system. The MEG is predicated on the Disease Staging patient classification system, developed initially for the Healthcare Cost and Utilization Project. The MEG uses sophisticated logic to create clinically relevant, severity-rated, and disease-specific groupings of claims. There are 593 episode groups. Episodes can be of several types:

- *Acute Condition* type includes episodes of care of acute conditions, which are generally reversible, such as an episode of sinusitis or otitis media.
- *Chronic Maintenance* episodes refer to episodes of routine care and management for a chronic, typically nonreversible condition or lifelong illness, such as diabetes mellitus episodes. All cancers are considered chronic.
- *Acute Flare-Up* type includes episodes of acute, generally reversible, and ideally preventable exacerbations of chronic conditions—such as an episode of diabetes with gangrene.
- *Well Care* type includes administrative and preventative care provided to a patient for ongoing health maintenance and wellness.

For the acute conditions and flare-ups identified in the claims, we define clean periods that mark the beginning or end of an episode of care. For chronic maintenance episodes, the first occurrence of the diagnosis can open an episode, and the calendar year is used to define endpoints.

Figure 8.1 illustrates how a stream of claims can be transformed into three episodes of care for a fifty-five-year-old male patient. In this example, episodes of care occur for two conditions: acute prostatitis and a herniated disc.

An episode for the care of the herniated disc (Episode 1) begins with an office visit on January 10. It includes all services related to an identified health problem of low back pain, including diagnostic imaging and a hospitalization. The episode ends with a follow-up physician office visit on May 8.

The treatment of acute prostatitis is divided into two episodes (Episodes 2 and 3). First, the patient is seen in his physician's office for acute prostatitis on February 4. The length of time between the February 4 visit and the May 18 visit is sufficiently long enough to begin a new episode, rather than continue the first episode. Consequently, a second episode (Episode 3) is initiated with the office visit for acute prostatitis on May 18. A complication of prostatitis, pyelonephritis, occurs within a short time, so the June 1 visit is a continuation of the second prostatitis episode.

This example also illustrates the difference between complications and comorbidities. A disease complication arises from the progression of an underlying disease. For example, pyelonephritis is a complication of acute prostatitis and is therefore a part of the episode for acute prostatitis. Disease comorbidities are diseases that are concurrent but not related to one another.

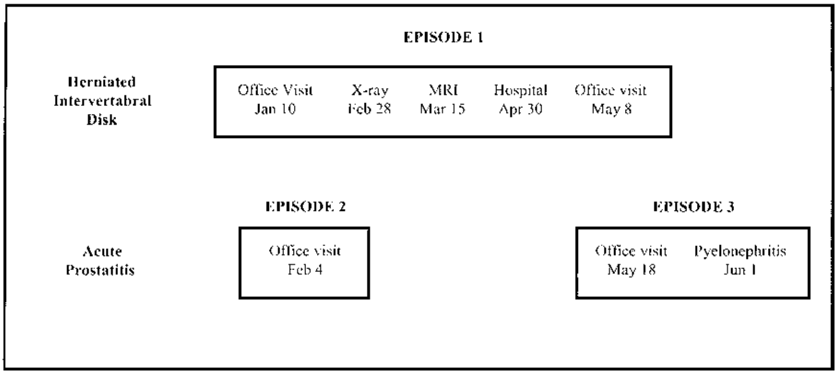


Fig. 8.1 Three episodes of care

For instance, the acute prostatitis and the herniated disc are comorbidities unrelated to one another. Therefore, separate disease episodes are created for the two comorbidities.

An episode of care is initiated with a contact with the health delivery system. In a claims-based methodology, the beginning of an episode is the first claim received for an episode grouping. The MEG methodology allows physician office visits and hospitalizations to open or extend patient episodes. As the coding of claims for laboratory tests and x-rays are not always reliable, these services can join existing episodes but cannot open an episode. Frequently in the practice of medicine, a physician will order a test prior to seeing a patient. To recognize this, a look-back mechanism has been incorporated in MEG. When a lab or x-ray service is encountered that occurred prior to the date of the claim that established an episode, MEG checks to see if an episode with the same episode group number has been opened within fifteen days following the test. If so, the lab or x-ray will be added to the episode.

An episode ends when the course of treatment is completed. Because the end of an episode is not designated on a claim, the clean period decision rule has been employed to establish the end date. Clean periods represent the period of time for a patient to recover from a disease or condition. If a subsequent visit for a disease occurs within the clean period, then it is assumed to be a part of the episode containing previous visits for that disease. If a visit for a disease occurs later than the clean period, then it defines the beginning of a new episode. The duration of clean periods was empirically and clinically reviewed and varies by disease.

Nonspecific initial diagnoses are relatively common in the billing of treatments of patients. For instance, an initial visit may be coded as abdominal pain but later be classified as appendicitis. The MEG incorporates logic to link nonspecific diagnoses and costs to specific episodes. The linkage occurs when a nonspecific claim has a date close in time to the specific episode and the linkage makes clinical sense.

The MEG incorporates drug claims into episode groups, even though drug claims themselves do not contain diagnostic information. The process of integrating pharmacy information into MEG begins with obtaining NDC information from Micromedex, a Thomson Reuters affiliate. Micromedex staff, made up of recognized pharmacological experts, map NDC codes from product package inserts to ICD-9-CM codes. This information is then reviewed by Thomson Reuters clinical and coding experts and mapped to MEG episode groups.

8.4.2 Construction of Episode-Based Disease Indexes

To construct episode-based disease indexes, we identified all claims for patients residing in the three metropolitan areas. We processed this group of claims with the episode software and created a file containing all of the episodes of care. Less than ten episode groups computed by MEG were excluded because they represent a collection of disparate conditions. This group contains only a small dollar amount.

Because diseases with low incidence (for example, cancer and kidney failure) usually command a much higher expenditure share than population share, it is possible that expenditure-based indexes and population-based indexes are very different. The cost-of-living theory is based on the cost functions of the individual consumer, but disease incidence and medical care spending are very skewed, both across individuals and over time for any given individual. Disease selection based on expenditure share increases the chances that less common but more severe diseases are selected; thus, the sample of selected diseases will not be representative of a typical consumer's experience in a given year. For example, in 2002, 5.7 percent of the national medical expenditure went to the treatment of acute myocardial infarction, while only 0.2 percent of the national population had this disease. Therefore, it is interesting to contrast indexes based on expenditure weighting with those based on population weighting.

To investigate the differences between the expenditure-based price index and the population-based price index, we randomly selected forty episodes with probability in proportion to their direct medical expenditures and another forty episodes with probability in proportion to the frequency of their occurrence in the population. Both sets of episodes were selected with replacement. All sample selection was carried out independently in each metropolitan area using MarketScan 1998 data. Because there could be more than one episode of a specific type chosen in this random selection, for the conditions represented in the selected episodes, all episodes of the same type in the city were selected, and the inputs used in these episode types were identified. For each selected episode, the volumes of inputs were updated at yearly intervals, and prices were estimated monthly from January 1999 to December 2002.

Appendixes B and C present the characteristics of the specific episode types that comprise the expenditure-based samples and the population-based samples in each city. For the expenditure-based samples, acute myocardial infarction, angina pectoris chronic maintenance, type 2 diabetes, and osteoarthritis were selected in all three cities. Neoplasm (with different types) also showed up in all cities. Only three diseases were commonly selected into the population-based samples in all three cities: aneurysm, thoracic; asthma, chronic maintenance; and tibial, iliac, femoral, or popliteal artery disease. Again, different types of neoplasm were sampled in all cities.

Standard grouping methods were utilized to compute the inputs into each episode type. For inpatient stays, we examined DRGs. For physician services and hospital outpatient services, we used the Berenson-Eggers Type of Service codes (BETOS, a transformation of the CPT-4 codes) developed by the Center for Medicare and Medicaid Services. For prescription drugs, we used Red Book therapeutic classes. The motivating factor in the decision to use grouped data was the desire to examine the full range of services that might appear in the episode and the concern with the magnitude of the detail that would need to be captured. The more detailed data we use, the bigger the concern with adequate cell size for monthly reporting. That is, grouping helps avoid months with no observations on price for detailed inputs that are rarely used. As we use grouped data, however, we introduce the potential for month-to-month changes within the group service mix.

For each year t, we identified all the inpatient discharges (DRGs), physician services (BETOS), and prescription drugs (therapeutic classes) used to treat episodes of care of each type in each city. This captures local variation in practice patterns that have been the subject of much discussion. Given the mix of inputs in year $t - 1$, we captured monthly prices for each input in each city in year t and computed a Laspeyres index. We allowed the mix of inputs to vary from year to year to capture the substitution effect. Because the total number of episodes of a specific type could also differ from one year to another, we used the average volume of inputs for each episode type, which was the total volume of each DRG, BETOS, or therapeutic class, divided by the total number of episodes in that group.

The hospital prices driving the hospital index in each city were city-specific average prices in MarketScan. We were concerned that there would be a large number of months with no observation of a discharge in specific DRGs that occasionally appeared in the treatment episode. Our general strategy for months with no relevant observation on price was to assume that the price was the same as the last month with a valid observation.

We first constructed component indexes for prescriptions, outpatient, and inpatient, and then we calculated their relative expenditure share within each episode as weight. The overall disease index was constructed as a weighted sum of these component indexes.

The expenditure shares that we calculated for experimental price indexes were different from those of BLS MCPI; in particular, the weight for prescription drugs was much smaller for the experimental price indexes than for the BLS index. The difference in the expenditure shares was much larger in New York City than in Philadelphia and Boston. For example, in 1999, the expenditure shares for inpatient and physician office visits were 71 percent and 28 percent in New York City for the experimental price index, while the corresponding shares were 40 percent and 44 percent for the BLS MCPI; the expenditure share of prescription drugs was around 16 percent for the BLS index but less than 1 percent for the experimental price index. The low share for prescription drugs could be explained by the following: (a) In the MarketScan database, drugs administered in hospitalizations are not recorded separately from other inpatient costs, which would lower the expenditure share for drugs and raise the expenditure share for hospitalizations. (b) The MEG grouper did not assign all prescription claims with an episode number. (c) The BLS drug weight comes from the Consumer Expenditure Survey (CEX), which includes individuals who are not insured and who are publicly insured; because the uninsured have a low inpatient utilization rate, their inpatient expenditure share might be extremely low and their drug share relatively high. (d) The CEX includes all prescription purchases, regardless of whether they are reimbursed, but the claims database only includes prescription purchases made by privately insured individuals that are reimbursed by health plans.

8.4.3 Bootstrapping Method

To decompose the differences between the episode-based price index and the BLS MCPI and to test their statistical significance, we need to estimate the mean and standard errors from the original sample first, and then use a parametric model (random walk with normal errors) to generate bootstrap samples. The standard error from the original sample was estimated using bootstrapping. In each month, we bootstrapped the ratio of the prices in that month and prices in the month prior to obtain the standard errors for prescription, outpatient, and inpatient separately. The monthly price change and standard error in Month 1 were set to one and zero, respectively. This section describes how bootstrapping was carried out for the cumulative indexes (across forty-eight months in 1999 to 2002) in this study.

Let $I_{i,a,t}$ be the month-to-month percentage change for index i, city a, and month t. The forty-eight-month cumulative index is

$$I_{i,a} = \prod_{t=1}^{48} I_{i,a,t}.$$

The individual $I_{i,a,t}$ is a mean with the variance of $\sigma^2_{i,a,t}$, which is the square of the standard error of the original sample. We assume a random walk with a drift, and the random variable is $I_{i,a,t} + \varepsilon_{i,a,t}$, where $\varepsilon_{i,a,t}$ is drawn from $N(0, \sigma^2_{i,a,t})$.[6] To bootstrap, we took 999 samples. For each sample $b = 1$ to 999, we drew $\varepsilon_{i,a,t}$ from $N(0, \sigma^2_{i,a,t})$ and added it to $I_{i,a,t}$ to get $\hat{I}_{i,a,t}$ for $t = 1$ to 48, which represents the forty-eight months from 1999 to 2002. This was done for prescription, outpatient, and inpatient services separately. The overall index was a weighted mean of these component indexes using their relative expenditure as weights $R_{i,a}$. So, the overall index for each of the 999 samples became

$$I_{b,t} = \Pi_{i,a} \hat{I}_{i,a,t,b} R_{i,a}.$$

We replicated each index one thousand times to obtain an estimate of variances for all MarketScan indexes. The BLS variances are the square of the BLS standard errors provided by the BLS. With the variances, we could calculate the difference between the claims-based index and the BLS index and its 95 percent confidence intervals. If zero falls between the confidence intervals, then the difference is not statistically significant.

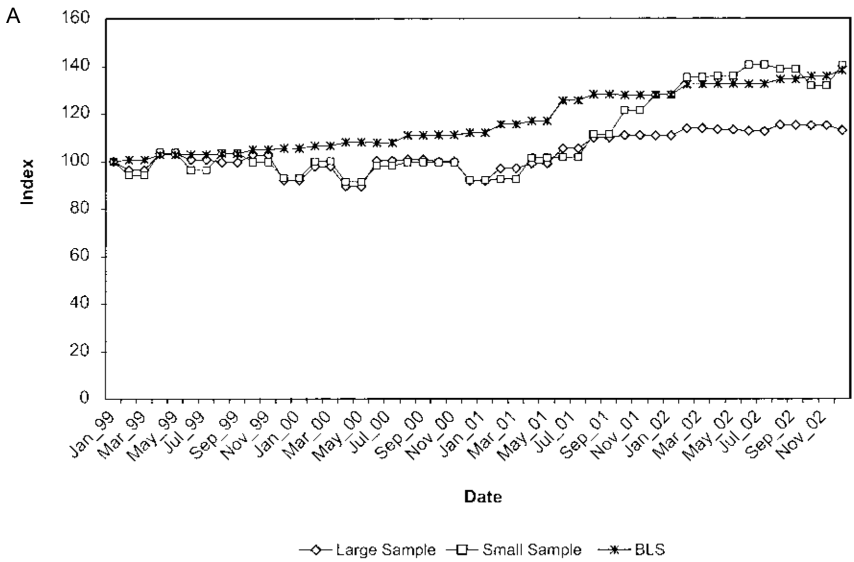
8.5 Results

8.5.1 Price Trends

The small-sample and large-sample indexes reported in figure 8.2 suggest a slower price increase than that suggested by the BLS city-specific medical care indexes over the period 1999 to 2002 (January 1999 = 100). The BLS indexes presented in figure 8.2 include only drugs, physicians, and hospital services to be comparable with the experimental price indexes that we have calculated. Except in New York City, the BLS indexes are bimonthly, with the Boston index repriced in odd months and the Philadelphia index repriced in even months.

In Philadelphia, the trends of large-sample and small-sample prices are very similar, and both are below the BLS trend most of the time; in Boston, the small-sample index presents a much larger price variation in 2001 and early 2002 than the large-sample index, and in most of the months, the small-sample index is above the BLS index, while the large-sample index is below the BLS index. In New York City, the trends of large-sample and small-sample prices are very similar, and both are below or above the BLS trend in about the same months: from the end of 1999 to mid-2001, both large-sample and small-sample indexes show a price decrease and are well

6. This assumption is similar to those often used in modeling of financial asset prices.



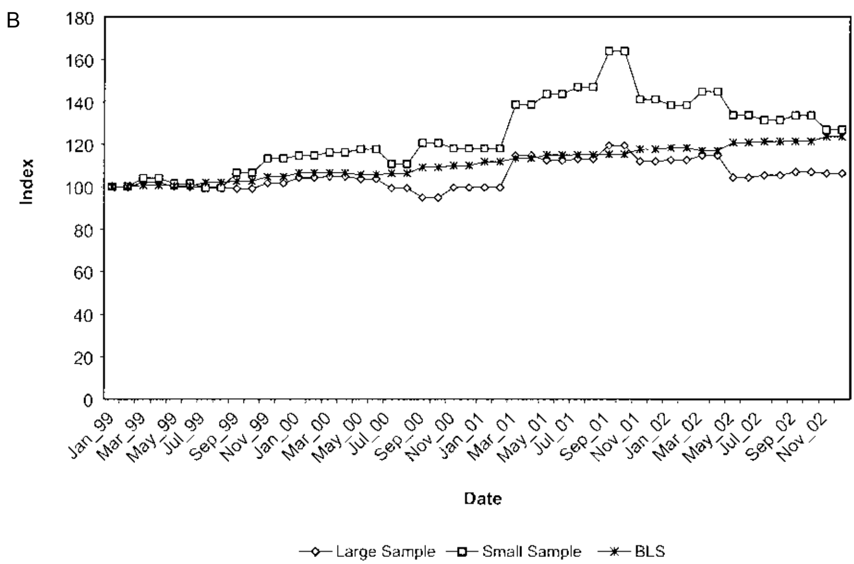


Fig. 8.2 Large-sample index vs. small-sample index vs. BLS index: *A*, Philadelphia; *B*, Boston; *C*, New York City

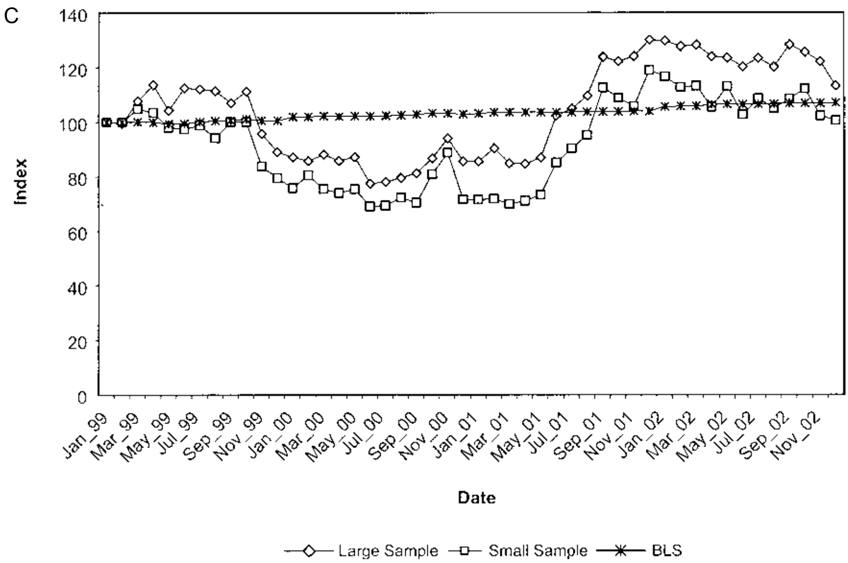


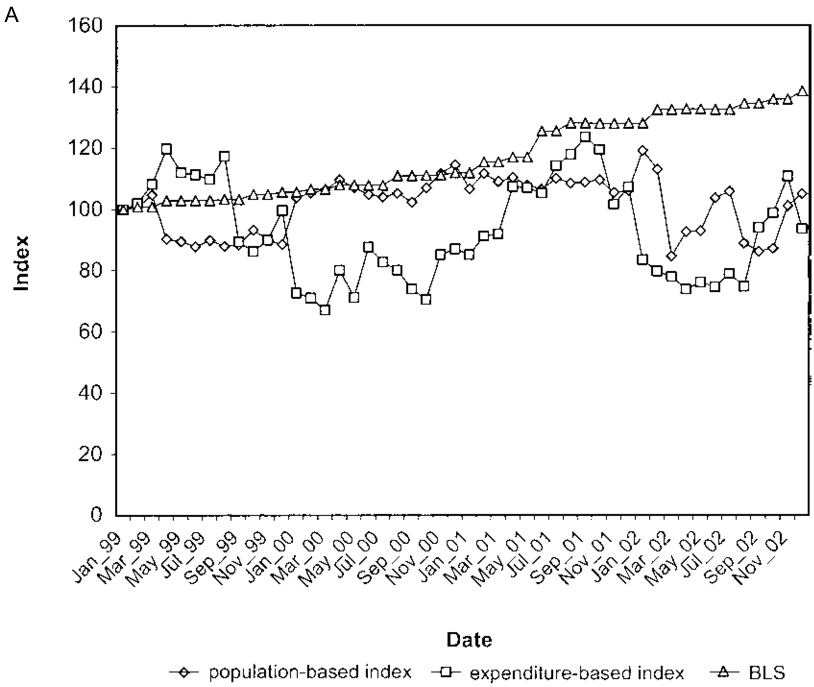
Fig. 8.2 (cont.)

below the BLS trend. As the sample sizes for the small sample are quite limited, the price variation we have found might just be random.

The episode-based indexes reported in figure 8.3 demonstrate that between January 1999 and December 2002, the cost of treatment has declined in all three cities for all indexes, except the population-based index in Philadelphia, which has risen, but much less than the corresponding BLS index. In fact, the correlation between the BLS index levels and the expenditure-based episode index levels is –0.68 in Boston, –0.57 in New York City, and –0.03 in Philadelphia, and the correlation between the BLS index levels and the population-based episode index levels is –0.06, –0.19, and 0.11, respectively.

As expected, the expenditure-based and population-based indexes present a different, although not statistically significant, price trend. In general, the expenditure-based index is lower than the population-based index in Philadelphia and Boston but is higher in New York City. In spite of these differences, these two indexes do give the three cities the same rank when considering the relative magnitude of the cumulative price changes from 1999 to 2002: New York City experiences the largest price decline, and Philadelphia sees the smallest price decline (expenditure-based index) or even a price increase (population-based index).

Overall, episode-based indexes fluctuate much more than the BLS MCPI, and one of the reasons is that we allowed the mix of inputs of treatment to



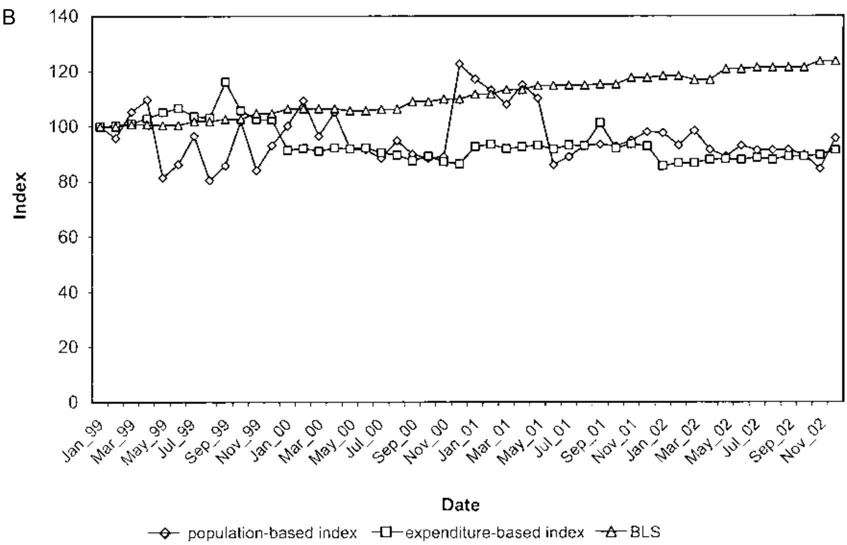


Fig. 8.3 Population-based disease index vs. expenditure-based disease index vs. BLS index: *A*, Philadelphia; *B*, Boston; *C*, New York City

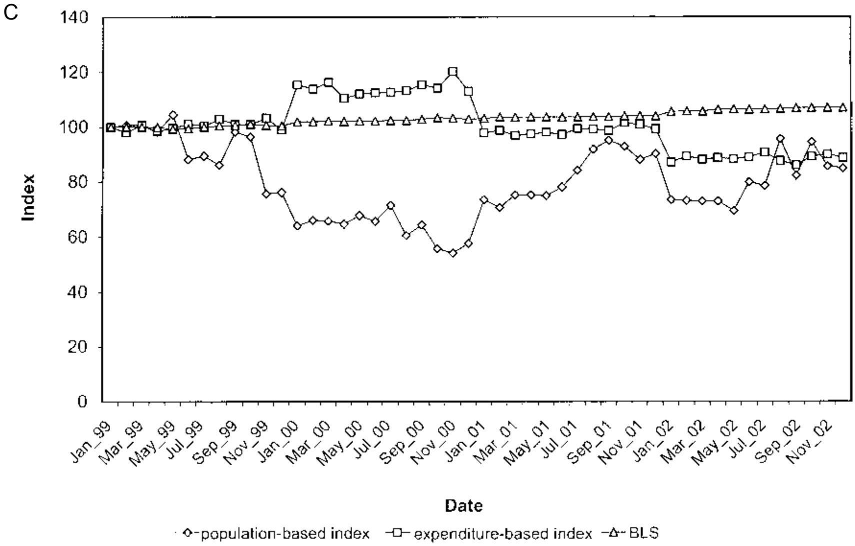


Fig. 8.3 (cont.)

change from year to year. Because the volume of inputs was updated in January of each year, the largest price jump usually occurs between December and January. Song et al. (2004) provide a close look at the mix of inputs of treating two specific episodes of angina pectoris with chronic maintenance and malignant neoplasm of female breast. We find that it is the change in volumes, not in prices, that produces such a dramatic jump.

To examine the statistical significance of the differences between the BLS index and the experimental price indexes, we bootstrapped their forty-eight-month cumulative changes and standard errors, as discussed in section 8.4.3. Tables 8.1 and 8.2 present the month-to-month percentage changes and estimated standard errors of the large-sample index, small-sample index, BLS MCPI, expenditure-based disease index, and population-based disease index in each city. Based on these statistics, we derived the forty-eight-month cumulative change for each index, their differences, and the lower and upper bound of the 95 percent confidence intervals of these differences.

The comparison of disease indexes for expenditure-based episodes and population-based episodes is reported in table 8.3.[7] From January 1999 to December 2002, we found a consistent decrease in the overall episode-based index in Boston and New York City: –8 percent in Boston and –10 percent in New York City for expenditure-based episodes, and –9 percent and

7. The cumulative index levels in tables 8.3 and 8.4 differ slightly from those shown in figures 8.2 and 8.3 as a result of the bootstrapping process.

Table 8.1 **Comparison of MarketScan large-sample index and small-sample index**

Months	Large-sample index: Month-to-month percentage change	Large-sample index: Standard errors	Small-sample index: Month-to-month percentage change	Small-sample index: Standard errors
		Philadelphia		
Jan_99	0.0000	0.0000	0.0000	0.0000
Feb_99	–0.0344	0.0113	–0.0573	0.0332
Mar_99	0.0000	0.0000	0.0000	0.0000
Apr_99	0.0702	0.0218	0.1001	0.0505
May_99	0.0000	0.0000	0.0000	0.0000
Jun_99	–0.0260	0.0148	–0.0706	0.0389
Jul_99	0.0000	0.0000	0.0000	0.0000
Aug_99	–0.0094	0.0192	0.0751	0.0756
Sep_99	0.0000	0.0000	0.0000	0.0000
Oct_99	0.0289	0.0168	–0.0375	0.0541
Nov_99	0.0000	0.0000	0.0000	0.0000
Dec_99	–0.1021	0.0342	–0.0662	0.0694
Jan_00	0.0000	0.0000	0.0000	0.0000
Feb_00	0.0623	0.0170	0.0744	0.0584
Mar_00	0.0000	0.0000	0.0000	0.0000
Apr_00	–0.0845	0.0305	–0.0857	0.0341
May_00	0.0000	0.0000	0.0000	0.0000
Jun_00	0.1194	0.0327	0.0727	0.0411
Jul_00	0.0000	0.0000	0.0000	0.0000
Aug_00	0.0049	0.0210	0.0160	0.0509
Sep_00	0.0000	0.0000	0.0000	0.0000
Oct_00	–0.0082	0.0164	–0.0012	0.0629
Nov_00	0.0000	0.0000	0.0000	0.0000
Dec_00	–0.0815	0.0284	–0.0753	0.0466
Jan_01	0.0000	0.0000	0.0000	0.0000
Feb_01	0.0585	0.0133	0.0053	0.0484
Mar_01	0.0000	0.0000	0.0000	0.0000
Apr_01	0.0198	0.0255	0.0972	0.0619
May_01	0.0000	0.0000	0.0000	0.0000
Jun_01	0.0640	0.0212	0.0016	0.0395
Jul_01	0.0000	0.0000	0.0000	0.0000
Aug_01	0.0432	0.0270	0.0940	0.0702
Sep_01	0.0000	0.0000	0.0000	0.0000
Oct_01	0.0082	0.0182	0.0902	0.0629
Nov_01	0.0000	0.0000	0.0000	0.0000
Dec_01	–0.0006	0.0196	0.0546	0.0782
Jan_02	0.0000	0.0000	0.0000	0.0000
Feb_02	0.0273	0.0141	0.0573	0.0301
Mar_02	0.0000	0.0000	0.0000	0.0000
Apr_02	–0.0054	0.0152	0.0049	0.0381
May_02	0.0000	0.0000	0.0000	0.0000
Jun_02	–0.0050	0.0312	0.0336	0.0633
Jul_02	0.0000	0.0000	0.0000	0.0000
Aug_02	0.0225	0.0234	–0.0120	0.0447
Sep_02	0.0000	0.0000	0.0000	0.0000
Oct_02	–0.0014	0.0289	–0.0513	0.0367
Nov_02	0.0000	0.0000	0.0000	0.0000
Dec_02	–0.0172	0.0192	0.0664	0.0282

Table 8.1 (continued)

Months	Large-sample index: Month-to-month percentage change	Large-sample index: Standard errors	Small-sample index: Month-to-month percentage change	Small-sample index: Standard errors
		Boston		
Jan_99	0.0317	0.0125	0.0016	0.0405
Feb_99	0.0000	0.0000	0.0000	0.0000
Mar_99	0.0200	0.0119	0.0397	0.0353
Apr_99	0.0000	0.0000	0.0000	0.0000
May_99	–0.0186	0.0113	–0.0243	0.0303
Jun_99	0.0000	0.0000	0.0000	0.0000
Jul_99	–0.0031	0.0134	–0.0190	0.0380
Aug_99	0.0000	0.0000	0.0000	0.0000
Sep_99	–0.0075	0.0225	0.0709	0.0473
Oct_99	0.0000	0.0000	0.0000	0.0000
Nov_99	0.0264	0.0205	0.0630	0.0513
Dec_99	0.0000	0.0000	0.0000	0.0000
Jan_00	0.0252	0.0116	0.0112	0.0254
Feb_00	0.0000	0.0000	0.0000	0.0000
Mar_00	0.0060	0.0100	0.0139	0.0318
Apr_00	0.0000	0.0000	0.0000	0.0000
May_00	–0.0116	0.0133	0.0132	0.0347
Jun_00	0.0000	0.0000	0.0000	0.0000
Jul_00	–0.0409	0.0187	–0.0601	0.0438
Aug_00	0.0000	0.0000	0.0000	0.0000
Sep_00	–0.0447	0.0418	0.0897	0.0651
Oct_00	0.0000	0.0000	0.0000	0.0000
Nov_00	0.0495	0.0246	–0.0220	0.0497
Dec_00	0.0000	0.0000	0.0000	0.0000
Jan_01	0.0000	0.0000	0.0000	0.0000
Feb_01	0.0000	0.0000	0.0000	0.0000
Mar_01	0.1509	0.0456	0.1767	0.0938
Apr_01	0.0000	0.0000	0.0000	0.0000
May_01	–0.0211	0.0204	0.0357	0.0296
Jun_01	0.0000	0.0000	0.0000	0.0000
Jul_01	0.0063	0.0175	0.0233	0.0290
Aug_01	0.0000	0.0000	0.0000	0.0000
Sep_01	0.0570	0.0540	0.1145	0.1122
Oct_01	0.0000	0.0000	0.0000	0.0000
Nov_01	–0.0629	0.0414	–0.1383	0.0575
Dec_01	0.0000	0.0000	0.0000	0.0000
Jan_02	0.0065	0.0104	–0.0197	0.0280
Feb_02	0.0000	0.0000	0.0000	0.0000
Mar_02	0.0172	0.0119	0.0470	0.0371
Apr_02	0.0000	0.0000	0.0000	0.0000
May_02	–0.0890	0.0199	–0.0769	0.0364
Jun_02	0.0000	0.0000	0.0000	0.0000
Jul_02	0.0105	0.0178	–0.0180	0.0396
Aug_02	0.0000	0.0000	0.0000	0.0000
Sep_02	0.0141	0.0169	0.0167	0.0502
Oct_02	0.0000	0.0000	0.0000	0.0000
Nov_02	–0.0072	0.0167	–0.0498	0.0323
Dec_02	0.0000	0.0000	0.0000	0.0000

(continued)

Table 8.1 (continued)

Months	Large-sample index: Month-to-month percentage change	Large-sample index: Standard errors	Small-sample index: Month-to-month percentage change	Small-sample index: Standard errors
		New York City		
Jan_99	0.0483	0.0167	0.0039	0.0340
Feb_99	–0.0061	0.0149	–0.0001	0.0403
Mar_99	0.0838	0.0253	0.0488	0.0446
Apr_99	0.0556	0.0199	–0.0127	0.0236
May_99	–0.0820	0.0133	–0.0532	0.0274
Jun_99	0.0779	0.0159	–0.0069	0.0138
Jul_99	–0.0036	0.0101	0.0143	0.0239
Aug_99	–0.0067	0.0141	–0.0468	0.0309
Sep_99	–0.0389	0.0111	0.0636	0.0436
Oct_99	0.0386	0.0220	–0.0010	0.0411
Nov_99	–0.1393	0.0282	–0.1629	0.0472
Dec_99	–0.0702	0.0286	–0.0518	0.0358
Jan_00	–0.0209	0.0087	–0.0451	0.0291
Feb_00	–0.0152	0.0100	0.0613	0.0357
Mar_00	0.0259	0.0124	–0.0614	0.0256
Apr_00	–0.0252	0.0105	–0.0209	0.0400
May_00	0.0154	0.0104	0.0206	0.0417
Jun_00	–0.1114	0.0322	–0.0847	0.0509
Jul_00	0.0087	0.0128	0.0048	0.0393
Aug_00	0.0189	0.0185	0.0413	0.0274
Sep_00	0.0208	0.0229	–0.0246	0.0433
Oct_00	0.0676	0.0520	0.1473	0.0794
Nov_00	0.0845	0.0491	0.0968	0.0597
Dec_00	–0.0907	0.0340	–0.1926	0.0393
Jan_01	0.0000	0.0000	0.0000	0.0000
Feb_01	0.0570	0.0145	0.0036	0.0401
Mar_01	–0.0622	0.0099	–0.0272	0.0284
Apr_01	–0.0019	0.0119	0.0160	0.0591
May_01	0.0278	0.0236	0.0319	0.0513
Jun_01	0.1761	0.0668	0.1594	0.0778
Jul_01	0.0259	0.0169	0.0615	0.0399
Aug_01	0.0451	0.0185	0.0553	0.0348
Sep_01	0.1301	0.0454	0.1828	0.0538
Oct_01	–0.0130	0.0128	–0.0329	0.0437
Nov_01	0.0155	0.0192	–0.0301	0.0365
Dec_01	0.0475	0.0181	0.1273	0.0494
Jan_02	–0.0030	0.0112	–0.0197	0.0307
Feb_02	–0.0153	0.0135	–0.0342	0.0264
Mar_02	0.0053	0.0117	0.0049	0.0372
Apr_02	–0.0338	0.0102	–0.0689	0.0259
May_02	–0.0046	0.0126	0.0714	0.0415
Jun_02	–0.0258	0.0199	–0.0908	0.0304
Jul_02	0.0254	0.0146	0.0568	0.0569
Aug_02	–0.0255	0.0171	–0.0327	0.0423
Sep_02	0.0673	0.0306	0.0328	0.0484
Oct_02	–0.0212	0.0189	0.0341	0.0396
Nov_02	–0.0265	0.0166	–0.0885	0.0423
Dec_02	–0.0741	0.0231	–0.0173	0.0425

Table 8.2 Comparison of BLS MCPI and episode-based index

Months	BLS		Expenditure-based disease index		Population-based disease index	
	Month-to-month percentage change	Standard errors	Month-to-month percentage change	Standard errors	Month-to-month percentage change	Standard errors
			Philadelphia			
Jan_99	0.0000	0.0000	–0.1005	0.0241	0.1145	0.1018
Feb_99	0.0085	0.0033	0.0189	0.0210	0.0070	0.0218
Mar_99	0.0000	0.0000	0.0613	0.0638	0.0398	0.1560
Apr_99	0.0203	0.0125	0.1067	0.0591	–0.1379	0.1073
May_99	0.0000	0.0000	–0.0638	0.0552	–0.0089	0.0081
Jun_99	0.0001	0.0013	–0.0065	0.0232	–0.0188	0.0133
Jul_99	0.0000	0.0000	–0.0141	0.0484	0.0235	0.0298
Aug_99	0.0028	0.0039	0.0689	0.0412	–0.0219	0.0192
Sep_99	0.0000	0.0000	–0.2385	0.0495	0.0062	0.0525
Oct_99	0.0168	0.0114	–0.0348	0.0342	0.0540	0.0561
Nov_99	0.0000	0.0000	0.0434	0.0483	–0.0337	0.0522
Dec_99	0.0065	0.0062	0.1069	0.0891	–0.0177	0.0480
Jan_00	0.0000	0.0000	–0.2710	0.0863	0.1733	0.3120
Feb_00	0.0094	0.0065	–0.0230	0.0742	0.0147	0.0310
Mar_00	0.0000	0.0000	–0.0552	0.0469	0.0107	0.0828
Apr_00	0.0127	0.0035	0.1950	0.1284	0.0302	0.0288
May_00	0.0000	0.0000	–0.1112	0.1018	–0.0246	0.0185
Jun_00	–0.0007	0.0049	0.2315	0.1104	–0.0199	0.0164
Jul_00	0.0000	0.0000	–0.0568	0.0844	–0.0082	0.0412
Aug_00	0.0286	0.0079	–0.0306	0.0425	0.0119	0.0131
Sep_00	0.0000	0.0000	–0.0787	0.0882	–0.0269	0.0229
Oct_00	0.0012	0.0060	–0.0457	0.0411	0.0445	0.0449
Nov_00	0.0000	0.0000	0.2080	0.0980	0.0456	0.0664
Dec_00	0.0073	0.0077	0.0228	0.0253	0.0237	0.0300
Jan_01	0.0000	0.0000	–0.0219	0.1094	–0.0679	0.1421
Feb_01	0.0321	0.0234	0.0713	0.0469	0.0460	0.0553
Mar_01	0.0000	0.0000	0.0066	0.0466	–0.0234	0.0859
Apr_01	0.0128	0.0068	0.1695	0.1074	0.0125	0.0757
May_01	0.0000	0.0000	–0.0038	0.0236	–0.0220	0.0325
Jun_01	0.0725	0.0603	–0.0162	0.0640	–0.0136	0.0427
Jul_01	0.0000	0.0000	0.0851	0.0555	0.0341	0.0406
Aug_01	0.0211	0.0029	0.0321	0.0623	–0.0138	0.0236
Sep_01	0.0000	0.0000	0.0487	0.0220	0.0026	0.0085
Oct_01	–0.0019	0.0024	–0.0339	0.0223	0.0064	0.0138
Nov_01	0.0000	0.0000	–0.1494	0.0533	–0.0379	0.0318
Dec_01	0.0010	0.0045	0.0563	0.0327	0.0069	0.0135
Jan_02	0.0000	0.0000	–0.2223	0.1054	0.1224	0.5559
Feb_02	0.0346	0.0268	–0.0442	0.0363	–0.0512	0.0504
Mar_02	0.0000	0.0000	–0.0232	0.0666	–0.2509	0.1659
Apr_02	0.0015	0.0023	–0.0516	0.0743	0.0943	0.1308
May_02	0.0000	0.0000	0.0297	0.0456	0.0025	0.0124
Jun_02	–0.0011	0.0023	–0.0208	0.0659	0.1161	0.0789
Jul_02	0.0000	0.0000	0.0603	0.0358	0.0208	0.0181
Aug_02	0.0151	0.0005	–0.0539	0.0290	–0.1600	0.0743

(*continued*)

Table 8.2 (continued)

Months	BLS: Month-to-month percentage change	BLS: Standard errors	Expenditure-based disease index: Month-to-month percentage change	Expenditure-based disease index: Standard errors	Population-based disease index: Month-to-month percentage change	Population-based disease index: Standard errors
Sep_02	0.0000	0.0000	0.2575	0.2232	–0.0300	0.0175
Oct_02	0.0097	0.0064	0.0508	0.0575	0.0114	0.0419
Nov_02	0.0000	0.0000	0.1213	0.2059	0.1594	0.1535
Dec_02	0.0194	0.0154	–0.1552	0.0930	0.0386	0.0795
			Boston			
Jan_99	0.0190	0.0073	0.0042	0.0033	0.1136	0.0511
Feb_99	0.0000	0.0000	0.0033	0.0013	–0.0424	0.0493
Mar_99	0.0100	0.0081	0.0079	0.0049	0.0998	0.1241
Apr_99	0.0000	0.0000	0.0188	0.0110	0.0413	0.0777
May_99	–0.0023	0.0015	0.0208	0.0093	–0.2572	0.0987
Jun_99	0.0000	0.0000	0.0142	0.0088	0.0577	0.0396
Jul_99	0.0122	0.0140	–0.0279	0.0137	0.1206	0.0763
Aug_99	0.0000	0.0000	–0.0051	0.0051	–0.1662	0.0652
Sep_99	0.0069	0.0043	0.1267	0.0547	0.0650	0.1775
Oct_99	0.0000	0.0000	–0.0902	0.0262	0.1898	0.0774
Nov_99	0.0198	0.0080	–0.0299	0.0178	–0.1762	0.1515
Dec_99	0.0000	0.0000	–0.0014	0.0065	0.1076	0.0828
Jan_00	0.0155	0.0076	–0.1081	0.0344	0.0771	0.1130
Feb_00	0.0000	0.0000	0.0065	0.0026	0.0912	0.1344
Mar_00	0.0013	0.0089	–0.0105	0.0027	–0.1173	0.1451
Apr_00	0.0000	0.0000	0.0128	0.0066	0.0888	0.0946
May_00	–0.0070	0.0101	–0.0036	0.0043	–0.1256	0.0553
Jun_00	0.0000	0.0000	0.0053	0.0057	–0.0054	0.0341
Jul_00	0.0053	0.0031	–0.0205	0.0057	–0.0328	0.0249
Aug_00	0.0000	0.0000	–0.0091	0.0024	0.0716	0.0488
Sep_00	0.0233	0.0135	–0.0227	0.0094	–0.0508	0.0562
Oct_00	0.0000	0.0000	0.0181	0.0071	–0.0178	0.0362
Nov_00	0.0080	0.0112	–0.0219	0.0136	0.0075	0.0463
Dec_00	0.0000	0.0000	–0.0095	0.0027	0.3784	0.3084
Jan_01	0.0157	0.0043	0.0724	0.0337	–0.0446	0.1016
Feb_01	0.0000	0.0000	0.0094	0.0094	–0.0347	0.0408
Mar_01	0.0147	0.0073	–0.0181	0.0092	–0.0454	0.0649
Apr_01	0.0000	0.0000	0.0079	0.0015	0.0660	0.0632
May_01	0.0117	0.0043	0.0058	0.0045	–0.0421	0.1024
Jun_01	0.0000	0.0000	–0.0143	0.0026	–0.2203	0.1693
Jul_01	0.0005	0.0030	0.0155	0.0041	0.0346	0.0454
Aug_01	0.0000	0.0000	–0.0017	0.0036	0.0452	0.0446
Sep_01	0.0027	0.0034	0.0893	0.0526	0.0056	0.0389
Oct_01	0.0000	0.0000	–0.0921	0.0377	–0.0078	0.0514
Nov_01	0.0198	0.0056	0.0180	0.0022	0.0223	0.0610
Dec_01	0.0000	0.0000	–0.0092	0.0041	0.0340	0.0844
Jan_02	0.0059	0.0079	–0.0777	0.0314	–0.0050	0.1089
Feb_02	0.0000	0.0000	0.0123	0.0040	–0.0458	0.0751

Table 8.2 (continued)

Months	BLS		Expenditure-based disease index		Population-based disease index	
	Month-to-month percentage change	Standard errors	Month-to-month percentage change	Standard errors	Month-to-month percentage change	Standard errors
Mar_02	–0.0124	0.0100	–0.0001	0.0024	0.0570	0.0595
Apr_02	0.0000	0.0000	0.0153	0.0018	–0.0697	0.0394
May_02	0.0317	0.0372	0.0028	0.0038	–0.0279	0.0319
Jun_02	0.0000	0.0000	–0.0034	0.0035	0.0447	0.0614
Jul_02	0.0042	0.0027	0.0081	0.0076	–0.0185	0.0436
Aug_02	0.0000	0.0000	–0.0077	0.0018	0.0011	0.0218
Sep_02	0.0006	0.0010	0.0131	0.0021	0.0011	0.0346
Oct_02	0.0000	0.0000	0.0012	0.0023	–0.0203	0.0451
Nov_02	0.0175	0.0031	0.0046	0.0022	–0.0571	0.0393
Dec_02	0.0000	0.0000	0.0200	0.0021	0.1322	0.0593
			New York City			
Jan_99	0.0092	0.0038	0.0485	0.0074	–0.0066	0.0376
Feb_99	0.0000	0.0000	–0.0206	0.0076	0.0078	0.0253
Mar_99	0.0007	0.0007	0.0287	0.0076	–0.0014	0.0065
Apr_99	–0.0010	0.0007	–0.0220	0.0039	–0.0247	0.0230
May_99	–0.0056	0.0031	0.0108	0.0047	0.0643	0.0797
Jun_99	0.0005	0.0003	0.0155	0.0065	–0.1563	0.0941
Jul_99	0.0062	0.0137	–0.0071	0.0067	0.0141	0.0209
Aug_99	0.0039	0.0053	0.0250	0.0058	–0.0370	0.0408
Sep_99	0.0013	0.0015	–0.0186	0.0042	0.1401	0.1567
Oct_99	0.0047	0.0062	–0.0009	0.0048	–0.0175	0.0131
Nov_99	–0.0042	0.0020	0.0227	0.0040	–0.2146	0.1699
Dec_99	–0.0009	0.0010	–0.0415	0.0047	0.0058	0.0084
Jan_00	0.0127	0.0053	0.1671	0.1223	–0.1595	0.1465
Feb_00	0.0000	0.0000	–0.0142	0.0080	0.0305	0.0289
Mar_00	0.0033	0.0063	0.0211	0.0059	–0.0030	0.0237
Apr_00	–0.0020	0.0045	–0.0494	0.0092	–0.0174	0.0090
May_00	0.0036	0.0015	0.0138	0.0058	0.0493	0.0275
Jun_00	–0.0022	0.0032	0.0043	0.0092	–0.0329	0.0226
Jul_00	0.0032	0.0023	0.0011	0.0083	0.0891	0.1732
Aug_00	0.0005	0.0005	0.0055	0.0039	–0.1540	0.0647
Sep_00	0.0037	0.0019	0.0189	0.0070	0.0644	0.0481
Oct_00	0.0045	0.0028	–0.0111	0.0066	–0.1355	0.1113
Nov_00	–0.0009	0.0009	0.0530	0.0073	–0.0281	0.0205
Dec_00	–0.0031	0.0017	–0.0599	0.0038	0.0653	0.1122
Jan_01	0.0023	0.0021	–0.1336	0.1111	0.2766	0.2492
Feb_01	0.0053	0.0025	0.0091	0.0059	–0.0394	0.0202
Mar_01	–0.0008	0.0115	–0.0190	0.0043	0.0646	0.0592
Apr_01	0.0004	0.0013	0.0048	0.0041	0.0001	0.0022
May_01	0.0000	0.0000	0.0064	0.0045	–0.0032	0.0045
Jun_01	–0.0012	0.0007	–0.0079	0.0038	0.0420	0.0363
Jul_01	0.0013	0.0031	0.0223	0.0065	0.0772	0.0733

(*continued*)

Table 8.2 (continued)

	BLS		Expenditure-based disease index		Population-based disease index	
Months	Month-to-month percentage change	Standard errors	Month-to-month percentage change	Standard errors	Month-to-month percentage change	Standard errors
Aug_01	0.0009	0.0034	–0.0023	0.0069	0.0922	0.1489
Sep_01	0.0000	0.0000	–0.0051	0.0049	0.0350	0.0423
Oct_01	0.0011	0.0041	0.0293	0.0058	–0.0236	0.0343
Nov_01	0.0012	0.0028	–0.0075	0.0050	–0.0514	0.0516
Dec_01	–0.0010	0.0017	–0.0148	0.0037	0.0232	0.0172
Jan_02	0.0173	0.0062	–0.1228	0.0731	–0.1851	0.1781
Feb_02	0.0020	0.0057	0.0248	0.0087	–0.0042	0.0216
Mar_02	–0.0005	0.0103	–0.0128	0.0070	–0.0021	0.0235
Apr_02	0.0071	0.0044	0.0064	0.0064	0.0006	0.0006
May_02	0.0007	0.0016	–0.0041	0.0051	–0.0478	0.0222
Jun_02	–0.0006	0.0011	0.0078	0.0037	0.1506	0.1338
Jul_02	0.0009	0.0005	0.0182	0.0043	–0.0187	0.0558
Aug_02	0.0007	0.0007	–0.0337	0.0074	0.2182	0.1595
Sep_02	0.0030	0.0026	–0.0173	0.0031	–0.1405	0.0676
Oct_02	0.0004	0.0010	0.0363	0.0055	0.1505	0.1514
Nov_02	0.0005	0.0004	0.0091	0.0099	–0.0944	0.0581
Dec_02	0.0002	0.0001	–0.0134	0.0027	–0.0091	0.0502

–16 percent for population-based episodes. In both cases, New York City experiences a much larger price decline than Boston. The expenditure-based index and population-based index display an opposite trend in Philadelphia: the former has dropped by 4 percent, while the latter has gone up by 8 percent.

The component indexes are not consistent across cities, either. In fact, the expenditure-based and population-based indexes often show an opposite trend. Both the expenditure-based index and the population-based index have moved in the same direction for prescription drug prices: they have gone up in Philadelphia and Boston but have gone down in New York City. In fact, the prescription price index has gone up by 97 percent in Philadelphia and 10 percent in Boston, but it has dropped by 39 percent in New York City for the expenditure-based episodes. The outpatient prices have increased in Boston and decreased in New York City for both expenditure-based and population-based episodes; in Philadelphia, it has gone up for the expenditure-based episodes but dropped for the population-based episodes. The inpatient index has dropped in all cases, except for population-based episodes in Philadelphia. It is difficult to know what factors might explain the differences between the cities. Part of the story could relate to the size of the claims database in each city; for example, Boston constitutes the largest

Table 8.3 **Decomposition of the difference between expenditure-based disease index and population-based disease index using raw payments: 48-month cumulative effect**

	Expenditure-based disease index		Population-based disease index			95% CI for total difference	
	Percentage change	SE	Percentage change	SE	Difference	Lower bound	Upper bound
Philadelphia							
RX	0.9742	0.4864	1.3281	0.4190	–0.3539	–1.6122	0.9043
OP	0.8074	0.7541	–0.0639	0.0455	0.8713	–0.6093	2.3519
IP	–0.2571	0.6124	0.0147	0.9816	–0.2718	–2.5394	1.9958
All-item	–0.0422	0.5435	0.0815	0.8260	–0.1237	–2.0616	1.8142
Boston							
RX	0.1021	0.0995	0.2341	0.0822	–0.1320	–0.3850	0.1210
OP	0.3791	0.4697	0.0107	0.0493	0.3684	–0.5573	1.2940
IP	–0.2308	0.0654	–0.1526	0.8023	–0.0782	–1.6558	1.4995
All-item	–0.0803	0.1055	–0.0861	0.6123	0.0058	–1.2120	1.2236
New York City							
RX	–0.3941	0.2070	–0.2827	0.2232	–0.1114	–0.7080	0.4851
OP	–0.1001	0.2087	–0.2354	0.0841	0.1353	–0.3057	0.5763
IP	–0.1313	0.2062	–0.1263	0.7582	–0.0051	–1.5451	1.5350
All-item	–0.1005	0.1710	–0.1592	0.5382	0.0586	–1.0482	1.1654

Note: “SE” = standard errors; “RX” = prescriptions; “OP” = outpatient treatment; “IP” = inpatient treatment; “CI” = confidence interval.

of the three city samples in MarketScan but is the smallest of the three BLS PSUs. Discrepancy in the health delivery systems in the three cities could be another potential explanation.

Despite the different trends that expenditure-based and population-based indexes demonstrated in some cases, we found that there is no significant difference between these two disease indexes. For all three cities, zero falls inside the 95 percent confidence intervals of the differences for the overall index, as well as all component indexes.

8.5.2 Decomposition Analysis

An initial look at the monthly index difference showed no statistical significance at the monthly level, so we examined the cumulative forty-eight-month indexes from 1999 to 2002. Three potential sources could contribute to the difference between the forty-eight-month cumulative BLS index and disease index: different index construction methods, different sample sizes, and different price distributions. To identify the importance of these sources, we decomposed the difference according to the following formula:

$$\begin{aligned} \mathrm{DPIMDT}_{m,y} - \mathrm{MPIBLS}_{m,y} = {} & (\mathrm{DPIMDT}_{m,y} - \mathrm{MPIMDTL}_{m,y}) \\ & + (\mathrm{MPIMDTL}_{m,y} - \mathrm{MPIMDTS}_{m,y}) \\ & + (\mathrm{MPIMDTS}_{m,y} - \mathrm{MPIBLS}_{m,y}). \end{aligned}$$

That is, TotalDifference = Method + SampleSize + DifferentPriceDistributions, where m,y = index month and year, DPIMDT = the disease index generated with claims data, MPIBLS = the BLS Medical CPI index with BLS data, MPIMDTL = the large-sample BLS CPI index with claims data, and MPIMDTS = the BLS CPI index with claims data using BLS sample sizes.

Table 8.4 reports the differences in the forty-eight-month cumulative changes between the expenditure-based disease index, the BLS index, the large-sample index, and the small-sample index. From January 1999 to December 2002, the BLS index shows a 38 percent increase in Philadelphia, a 23 percent increase in Boston, and a 7 percent increase in New York City, while our expenditure-based disease index presents a consistent decline in Philadelphia (–4 percent), Boston (–8 percent), and New York City (–10 percent).

The differences between the overall expenditure-based disease index and the BLS MCPI are –42 percent, –31 percent, and –17 percent in Philadelphia, Boston, and New York City, respectively, but only the –31 percent is statistically different from zero. Most episode-based component indexes are not significantly different from the BLS medical component indexes, either. In fact, the only significant difference is the difference in the inpatient index in Boston and the difference in the prescription index in New York City.

In sum, the decomposition results suggest that differences between the

Table 8.4 **Decomposition of differences between BLS index, expenditure-based disease index, large-sample index, and small-sample index using raw payments: 48-month cumulative effect**

	Expenditure-based disease index		BLS MCPI		95% CI for total difference		
	Percentage change	SE	Percentage change	SE	Total difference	Lower bound	Upper bound
Philadelphia							
RX	0.9742	0.4864	0.1650	0.0525	0.8092	–0.1496	1.7681
OP	0.8074	0.7541	0.2959	0.2003	0.5116	–1.0176	2.0408
IP	–0.2571	0.6124	0.6486	0.1600	–0.9057	–2.1462	0.3349
All-item	–0.0422	0.5435	0.3803	0.1002	–0.4224	–1.5055	0.6607
Boston							
RX	0.1021	0.0995	0.1893	0.0700	–0.0872	–0.3257	0.1513
OP	0.3791	0.4697	0.0400	0.0470	0.3391	–0.5861	1.2643
IP	–0.2308	0.0654	0.5055	0.1832	–0.7362	**–1.1175**	**–0.3549**
All-item	–0.0803	0.1055	0.2291	0.0627	–0.3093	**–0.5499**	**–0.0687**
New York City							
RX	–0.3941	0.2070	0.1294	0.0457	–0.5235	**–0.9389**	**–0.1081**
OP	–0.1001	0.2087	0.0178	0.0407	–0.1179	–0.5346	0.2988
IP	–0.1313	0.2062	0.1012	0.0666	–0.2326	–0.6573	0.1922
All-item	–0.1005	0.1710	0.0701	0.0320	–0.1706	–0.5116	0.1703

	Large sample size: Replication			95% CI for method difference	
	Percentage change	SE	Method difference	Lower bound	Upper bound
Philadelphia					
RX	0.3129	0.0071	0.6613	–0.2921	1.6147
OP	–0.0697	0.1589	0.8771	–0.6333	2.3875
IP	0.1502	0.3161	–0.4073	–1.7581	0.9434
All-item	0.1328	0.1268	–0.1749	–1.2687	0.9188
Boston					
RX	0.1191	0.1731	–0.0170	–0.4084	0.3745
OP	0.1051	0.1286	0.2740	–0.6805	1.2285
IP	–0.1833	0.1962	–0.0475	–0.4529	0.3579
All-item	0.0571	0.1279	–0.1373	–0.4623	0.1876
New York City					
RX	0.1830	0.0142	–0.5771	**–0.9837**	**–0.1705**
OP	–0.1640	0.1562	0.0639	–0.4470	0.5748
IP	–0.0189	0.3537	–0.1124	–0.9148	0.6899
All-item	0.1220	0.1786	–0.2225	–0.7072	0.2622

(*continued*)

Table 8.4 (continued)

	Small sample size: Replication			95% CI for sample size difference	
	Percentage change	SE	Sample size difference	Lower bound	Upper bound
Philadelphia					
RX	0.3283	0.0203	–0.0153	–0.0575	0.0268
OP	0.2860	0.7094	–0.3557	–1.7806	1.0692
IP	0.2654	0.5761	–0.1152	–1.4031	1.1728
All-item	0.4037	0.3528	–0.2710	–1.0058	0.4638
Boston					
RX	0.1144	0.0234	0.0046	–0.3378	0.3470
OP	–0.0115	0.3576	0.1166	–0.6283	0.8614
IP	0.3405	0.5411	–0.5238	–1.6520	0.6044
All-item	0.2535	0.2959	–0.1964	–0.8283	0.4354
New York City					
RX	0.1719	0.1076	0.0110	–0.2017	0.2238
OP	–0.3538	0.3809	0.1898	–0.6170	0.9966
IP	–0.0512	0.3734	0.0323	–0.9757	1.0403
All-item	0.0022	0.2901	0.1198	–0.5480	0.7875
	BLS MCPI			**95% CI for sample size difference**	
	Percentage change	**SE**	**Price difference**	**Lower bound**	**Upper bound**
Philadelphia					
RX	0.1650	0.0525	0.1633	**0.0531**	**0.2736**
OP	0.2959	0.2003	–0.0098	–1.4547	1.4350
IP	0.6486	0.1600	–0.3832	–1.5550	0.7887
All-item	0.3803	0.1002	0.0235	–0.6954	0.7423
Boston					
RX	0.1893	0.0700	–0.0749	–0.2195	0.0698
OP	0.0400	0.0470	–0.0515	–0.7584	0.6554
IP	0.5055	0.1832	–0.1650	–1.2847	0.9548
All-item	0.2291	0.0627	0.0244	–0.5685	0.6173
New York City					
RX	0.1294	0.0457	0.0426	–0.1865	0.2717
OP	0.0178	0.0407	–0.3716	–1.1223	0.3791
IP	0.1012	0.0666	–0.1524	–0.8959	0.5910
All-item	0.0701	0.0320	–0.0679	–0.6400	0.5041

Note: Numbers in bold are significantly different from 0. "SE" = standard errors; "RX" = prescriptions; "OP" = outpatient treatment; "IP" = inpatient treatment; "CI" = confidence interval.

large- and small-sample indexes are never significant, which is not a surprise, as the large sample and the small sample were both drawn from the same MarketScan population data file. The majority of differences due to methods and the majority of differences due to price distributions are not significant, either. However, it is important to keep in mind that because the city-specific indexes are measured with only limited precision, the differences between the methods may reflect random differences.

In addition to differences in sample sizes, methods, and price distributions, another reason for the difference in the all-item indexes is the different relative weighting of prescription drugs, as we have discussed.

8.6 Summary and Discussion

The findings reported here suggest that using medical claims data to measure price changes in health care based on episodes of care is feasible, although claims data alone are not sufficient to replace the current medical CPI.

To summarize the finding from this study, the analysis of trends in treatment costs for a randomly selected set of diseases yields a different picture than the BLS overall medical care price index. Where the current methods indicate consistent price increases over time, the disease-based indexes suggest that treatment prices (i.e., cost for an episode of care) have dropped in Philadelphia, Boston, and New York City during 1999 to 2002. These results on the trends in treatment costs are similar to a generalized version of the findings in cataract surgery, depression, and acute myocardial infarction as reported by Berndt, Cockburn, and Griliches (1996), Berndt et al. (2002), Busch, Berndt, and Frank (2001), Cutler, McClellan, and Newhouse (1998, 1999), and Shapiro, Shapiro, and Wilcox (2001). In addition, in this case, the finding of a substantially different trend in price change is for forty diagnoses randomly selected from a sampling frame that contains virtually all potential diagnoses. However, despite the different trends, the forty-eight-month cumulative changes of the expenditure-based disease index and the BLS index are not significantly different from each other in Philadelphia or New York City.

The results we have obtained suggest that the disease-based index may measure the real price changes better than the current MCPI, because the disease index allows for the substitution effect among treatment inputs. The percentages of the total expenditures on prescriptions, outpatient, and inpatient treatment of the forty randomly selected expenditure-based episodes have changed considerably in all three cities during 1999 to 2002. In Philadelphia, the share of prescription expenditure went up from 2.1 percent in January 1999 to 4.6 percent in December 2002, the share of outpatient expenditure increased from 16.9 percent to 34.1 percent, and the inpatient expenditure share dropped from 81.0 percent to 61.4 percent. In Boston, these expenditure shares were 3.6 percent, 22.7 percent, and 73.7 percent in January 1999 and became 4.4 percent, 34.6 percent, and 61.0 percent in December 2002. In New York City, the outpatient expenditure share rose from 28.4 percent to 34.4 percent, the hospitalization share dropped from 71.0 percent to 65.2 percent, and the prescription share decreased slightly from 0.6 percent to 0.4 percent. A similar pattern was also observed for population-based episodes over the same time period. Overall, the treatment pattern of disease episodes seems to move away from inpatient hospitalizations to outpatient settings.

It is important to note that all results presented in this chapter are based on raw payments in the claims database, which could help explain the large variance we observed in claims data indexes. To avoid the small-sample issues with hospital stays and procedures, one could use the nationwide database and a two-level random-effect model to produce a Bayes estimate of the monthly payment for each DRG and BETOS at the city level. Song et al. (2004) report disease indexes that are constructed using Bayes-estimated prices for BETOS and DRGs for the same two sets of forty episodes. The overall trend of payments is determined from the overall MarketScan trend, and an adjustment is made to the intercept of each city. Indexes based on Bayes-estimated prices present a more consistent trend and reveal less fluctuation than indexes based on raw payments. However, depending on how big a value should be placed on consistency, it is not clear whether the addition of analytic complexity is worth the computational burden for the BLS.

The sampling method taken in this chapter selected drugs, physician office visits, hospitals, and disease episodes using probability in proportion to size with replacement, as the BLS does for the MCPI. However, sampling without replacement is more efficient than sampling with replacement (Foreman 1991). A further advantage to sampling without replacement is that the episode groupers can be randomly ordered within a body system, and then the body systems can be randomly ordered in the MEG list. This would cause the sample of episodes to be implicitly stratified by body systems, ensuring that the sample of episodes tended to be representative of the various body systems, so there is no chance of selecting only metabolic diseases, for example. We could select diseases from each body system in proportion to the expenditures or frequency of occurrence for treatment of that body system.

Disease-based price indexes rely heavily on MEG in this study. In addition to MEG, there exist several other proprietary episode grouper software products. Rosen and Mayer-Oakes (1999) compared four such episode groupers based on characteristics such as purpose, case-mix adjustment, comprehensiveness, and clinical flexibility. Although it would be interesting to see whether different episode groupers would generate different trends in treatment costs, we believe that correcting the information technology failure in the medical market is more important in the calculation of the cost of episode care using claims databases than trying to choose the best episode grouper. The current medical record-keeping system does not adequately keep track of all the inputs that are used to treat a patient disease or patient episode. The lack of sufficient record-keeping and the existence of incomplete claims are two examples of the information technology failure. For instance, in the Medical Expenditure Panel Survey 2003 data, about 8 percent of medical expenditures were due to orphan records (i.e., records with a dollar amount for the use of a service but no diagnoses). In each year of 1998 to 2003, orphan records had the highest expenditure share. No episode grouper can correctly bundle orphan records into a particular disease. We do not think we can generate any type of accurate price index from claims data until this information technology failure is corrected. To achieve this, all physicians, public insurers, and private insurers must be responsible for maintaining an audited record-keeping system that is consistently updated for the inputs used to treat diseases, for corrections or changes in diagnosis, and for an established beginning and end date established by the physician for every acute disease.

In addition to the information technology failure, there are four limitations in using claims data to generate a medical CPI. One limitation of the price index developed in this study is that it does not include health insurance premiums. A true CPI needs to account for the role of health insurance, because it represents a major medical purchase for most consumers. Unfortunately, information on health insurance premiums and characteristics is not available in a medical claims database. Secondly, it is important to point out that all indexes constructed in this study are indexes only for those covered by health plans in the United States. We did not estimate price indexes for the uninsured population, who may face different incidence of diseases, and who, for a particular disease, may consume different inputs. A third limitation of using the claims data set is that treating a disease may require more types of inputs than those reimbursed by an insurer. For example, over-the-counter medicines may play an important role, and products such as sunscreen, gym memberships, and dental floss are often used to prevent disease and should be considered as part of the mix of goods used to stay healthy. Finally, whether the insured people in a claims database are representative of the whole privately insured population in the United States remains to be seen. Thus, a medical CPI cannot be generated solely on claims databases.

Appendix A

Analytic File Construction for BLS MCPI Replication Analysis

The analytic file was built from the MarketScan databases, following the steps summarized below.

1. Using the first three digits of providers' ZIP codes, we selected all inpatient admissions, inpatient services, outpatient services, and pharmacy claims for the following metropolitan areas from the Commercial and Medicare Databases between January 1, 1998, and December 31, 2002: New York City (A109), Philadelphia (A102), and Boston (A103).
2. We combined the resulting data sets from the Commercial and Medicare Databases.
3. To sample pharmacies, we randomly selected a given number of pharmacy IDs in proportion to their expenditure share within a city. Because MarketScan databases do not record the annual expenditure of any pharmacy, we summed up all payment to a given pharmacy in a year recorded in MarketScan to calculate the probability of selecting that pharmacy. The computed total payment to a pharmacy could differ from its actual annual revenue, as some large pharmacies may have a small number of patients in MarketScan databases.
4. For each selected pharmacy ID, we randomly selected one NDC in proportion to its expenditure share within that pharmacy at yearly intervals. All drugs and medical supplies dispensed by prescription, including prescription-dispensed over-the-counter drugs, were included in this random selection. Inpatient hospital prescriptions and prescriptions paid by Medicaid or worker's compensation were ineligible for the medical price index. For each NDC selected, both the insurance reimbursement and the patient co-pay, if any, were included to arrive at the total reimbursement for that prescription.
5. Hospitals that are owned and operated by health maintenance organizations (HMOs) should be excluded, because they are not eligible for CPI pricing; but because hospital ownership is not included in the MarketScan databases, these hospitals cannot be identified directly. Instead, we excluded all services that are paid by the capitation method, and by default, these hospitals were excluded from our sample.
6. We relied on the provider type variable (STDPROV) to exclude ophthalmologists, dentists, podiatrists, and other medical practitioners who are not medical doctors or osteopaths from our sample, because they are not eligible for medical price indexes. We also excluded services reimbursed by capitation.
7. To calculate physician indexes, we first randomly selected a given number of physicians in proportion to their expenditure share within a city,

and then we randomly selected one CPT in proportion to its expenditure share for that physician. As MarketScan databases do not record the annual revenue of any physician, we summed up all payment to a given physician in a year recorded in MarketScan to calculate the probability of selecting that physician. It is important to note that the computed total payment to a physician could differ from his or her actual annual revenue.

8. MarketScan outpatient services database does not contain the same hospital ID that is contained in the inpatient admissions and inpatient services databases; therefore, we could not link inpatient stays and outpatient visits that occur within the same hospital. We used hospital IDs (UNIHOSP) in the inpatient data sets to identify hospitals, and we used provider IDs (PROVID) in the outpatient data set to identify hospitals.

9. To sample a given number of hospitals for the hospital indexes, we randomly selected the same number of hospitals in proportion to their expenditure share within a city. As MarketScan databases do not record the annual revenue of any hospital, we summed up all payment to a given hospital in a year recorded in MarketScan to calculate the probability of choosing that hospital. It is important to note that the computed total payment to a hospital could differ from its actual annual revenue, as some large hospitals may have a small number of patients in MarketScan database.

10. For each selected hospital ID, we randomly chose one hospital stay in proportion to its expenditure share within all inpatient hospital stays; for each selected provider ID, we randomly selected one outpatient visit in proportion to its expenditure share within all outpatient services in that hospital. Thus, for each hospital ID, we selected one inpatient stay; for each provider ID, we selected one outpatient visit. All random selection occurred at yearly intervals. Hospital outpatient services were identified using the place of service variable (STDPLAC).

11. We calculated the final reimbursements for each selected NDC, CPT, and hospital stay/visit in each month. The PAY variable in MarketScan measures total payment reimbursed from all sources.

Appendix B

Table 8B.1 **Expenditure-based sampling characteristics: Conditions sampled with probability proportional to expenditure with replacement**

Episode group number	Episode label	Total payments ($)	Number of times drawn	Expected number of times drawn
	Philadelphia			
10	Angina pectoris, chronic maintenance	16,594,049	6	2.965
187	Renal failure	9,538,311	2	1.704
11	Acute myocardial infarction	8,740,579	2	1.562
374	Osteoarthritis	8,349,565	1	1.492
397	Cerebrovascular disease with stroke	6,531,667	2	1.167
426	Complications of surgical and medical care	5,092,074	1	0.910
212	Neoplasm, malignant: Breast, female	4,802,802	1	0.858
336	Neoplasm, malignant: Prostate	3,671,757	2	0.656
274	Cholecystitis and cholelithiasis	2,868,658	1	0.513
51	Diabetes mellitus with complications	2,393,106	1	0.428
189	Urinary tract infections	2,231,736	2	0.399
405	Injury: Spine and spinal cord	1,955,954	1	0.349
50	Diabetes mellitus type 2 and hyperglycemic states, maintenance	1,779,928	1	0.318
535	Infections of skin and subcutaneous tissue	1,359,153	1	0.243
1	Aneurysm, abdominal	1,103,927	1	0.197
411	Neoplasm: Central nervous system	967,351	1	0.173
357	Fracture or sprain: Ankle	899,572	1	0.161
285	Pancreatitis	812,379	1	0.145
149	Functional digestive disorders	805,771	1	0.144
556	Injury: Other	690,768	1	0.123
138	Appendicitis	591,180	1	0.106

204	Dysfunctional uterine bleeding	387,773	1	0.069
366	Infectious arthritis	372,816	1	0.067
386	Anomaly: Musculoskeletal system	336,946	2	0.060
206	Endometriosis	288,676	1	0.052
220	Pelvic inflammatory disease	125,016	1	0.022
547	Adverse drug reactions	119,754	1	0.021
304	Herpes simplex infections	77,815	1	0.014
58	Neoplasm, benign: Adenoma, parathyroid, or hyperparathyroidism	63,178	1	0.011
	Boston			
10	Angina pectoris, chronic maintenance	27,424,386	4	2.690
374	Osteoarthritis	16,971,880	3	1.665
11	Acute myocardial infarction	16,192,922	1	1.588
13	Essential hypertension, chronic maintenance	13,013,202	2	1.277
397	Cerebrovascular disease with stroke	11,187,732	2	1.097
187	Renal failure	10,737,384	2	1.053
92	Cataract	8,905,881	1	0.874
500	Chronic obstructive pulmonary disease	8,752,026	1	0.859
6	Arrhythmias	8,653,448	1	0.849
212	Neoplasm, malignant: Breast, female	8,542,333	1	0.838
348	Fracture: Femur, head, or neck	7,267,359	2	0.713
426	Complications of surgical and medical care	6,681,475	1	0.655
50	Diabetes mellitus type 2 and hyperglycemic states, maintenance	5,312,060	1	0.521
24	Tibial, iliac, femoral, or popliteal artery disease	5,176,366	1	0.508
203	Delivery, vaginal	4,850,220	1	0.476
398	Dementia: Primary degenerative (Alzheimer's or Pick's disease)	2,540,752	1	0.249
536	Neoplasm, malignant: Carcinoma, basal cell	2,236,696	1	0.219
209	Neoplasm, benign: Breast	1,994,974	1	0.196

(*continued*)

Table 8B.1 (continued)

Episode group number	Episode label	Total payments ($)	Number of times drawn	Expected number of times drawn
361	Fracture, dislocation, or sprain: Humerus (head) or shoulder	1,968,915	2	0.193
164	Peptic ulcer disease	1,916,376	1	0.188
88	Sinusitis	1,915,297	1	0.188
23	Thrombophlebitis	1,848,311	2	0.181
357	Fracture or sprain: Ankle	1,589,749	1	0.156
149	Functional digestive disorders	1,564,679	1	0.153
491	Schizophrenia	1,021,690	1	0.100
2	Aneurysm, thoracic	888,311	1	0.087
355	Fracture: Tibia	624,201	1	0.061
516	Pulmonary embolism	515,951	1	0.051
387	Injury: Other and ill-defined musculoskeletal sites	427,259	1	0.042
	New York City			
203	Delivery, vaginal	3,580,789	5	1.816
10	Angina pectoris, chronic maintenance	2,737,849	3	1.388
374	Osteoarthritis	2,379,435	1	1.207
212	Neoplasm, malignant: Breast, female	1,868,716	2	0.948
11	Acute myocardial infarction	1,407,383	1	0.714
411	Neoplasm: Central nervous system	1,173,222	2	0.595
508	Neoplasm, malignant: Lungs, bronchi, or mediastinum	1,150,206	1	0.583
6	Arrhythmias	989,755	1	0.502
209	Neoplasm, benign: Breast	926,921	1	0.470
341	Bursitis	880,225	1	0.446
510	Pneumonia: Bacterial	632,472	1	0.321

211	Neoplasm, benign: Uterus (leiomyomas)	574,949	1	0.292
152	Hernia, external	552,161	1	0.280
427	Encounter for chemotherapy	533,949	1	0.271
158	Neoplasm, benign: Adenomatous polyps, colon	530,215	1	0.269
274	Cholecystitis and cholelithiasis	528,935	1	0.268
85	Otitis media	433,414	1	0.220
50	Diabetes mellitus type 2 and hyperglycemic states, maintenance	425,695	1	0.216
317	Rheumatic fever	357,445	1	0.181
173	Gastroenteritis	330,471	1	0.168
213	Neoplasm, malignant: Cervix uteri	307,638	1	0.156
370	Injury, open wound, or blunt trauma: Lower extremity	209,390	1	0.106
163	Neoplasm, malignant: Stomach	177,219	1	0.090
289	Neoplasm, malignant: Other hepatobiliary tract	166,445	1	0.084
398	Dementia: Primary degenerative (Alzheimer's or Pick's disease)	152,349	1	0.077
114	Macular degeneration	136,377	1	0.069
434	Neoplasm, benign: Other sites	122,331	1	0.062
487	Eating disorders: Anorexia nervosa	85,993	1	0.044
443	Anomaly: Defects of kidney	67,282	1	0.034
190	Neoplasm, benign: Urinary tract	40,339	1	0.020
307	Infectious mononucleosis	22,333	1	0.011
343	Dislocation: Knee	14,330	1	0.007

Appendix C

Table 8C.1 **Population-based sampling characteristics: Episodes sampled with probability proportional to population with replacement**

Episode group number	Episode label	Number of patients	Number of times drawn	Expected number of times drawn
	Philadelphia			
425	Abnormal lab, x-ray, and clinical findings	2,438	2	1.0079
331	Benign prostatic hypertrophy	2,275	2	0.9405
371	Injury, open wound, or blunt trauma: Upper extremity	2,206	1	0.9120
402	Headache	2,125	1	0.8785
496	Asthma, chronic maintenance	2,116	1	0.8748
535	Infections of skin and subcutaneous tissue	2,069	1	0.8554
158	Neoplasm, benign: Adenomatous polyps, colon	1,427	1	0.5899
24	Tibial, iliac, femoral, or popliteal artery disease	1,385	3	0.5726
536	Neoplasm, malignant: Carcinoma, basal cell	1,368	1	0.5656
173	Gastroenteritis	1,327	2	0.5486
49	Diabetes mellitus type 1 maintenance	1,036	1	0.4283
489	Generalized anxiety disorder	1,016	4	0.4200
150	Gastritis	972	1	0.4018
51	Diabetes mellitus with complications	945	1	0.3907
267	Anemia: Other	906	1	0.3746
361	Fracture, dislocation, or sprain: Humerus (head) or shoulder	888	1	0.3671
199	Ante- and postpartum complications	770	1	0.3183
368	Injury, knee, ligamentous	741	1	0.3063
156	Irritable bowel syndrome	725	1	0.2997
98	Detachment of the retina	683	1	0.2824
152	Hernia, external	622	1	0.2571
541	Psoriasis vulgaris	398	1	0.1645
363	Gout	388	1	0.1604

354	Fracture: Radius, lower end	380	1	0.1571
339	Neoplasm, benign: Male reproductive system	285	1	0.1178
79	Neoplasm, benign: Sinuses	240	1	0.0992
497	Asthma with complications	228	1	0.0943
213	Neoplasm, malignant: Cervix uteri	211	1	0.0872
304	Herpes simplex infections	159	1	0.0657
172	Complications of gastrointestinal treatment	94	1	0.0389
2	Aneurysm, thoracic	69	1	0.0285
552	Laceration: Esophagus	37	1	0.0153
	Boston			
496	Asthma, chronic maintenance	5,174	1	1.0959
189	Urinary tract infections	4,746	1	1.0053
539	Neoplasm: Atypical nevus	4,511	1	0.9555
425	Abnormal lab, x-ray, and clinical findings	4,364	1	0.9244
24	Tibial, iliac, femoral, or popliteal artery disease	3,731	3	0.7903
370	Injury, open wound, or blunt trauma: Lower extremity	3,585	1	0.7594
536	Neoplasm, malignant: Carcinoma, basal cell	3,402	1	0.7206
153	Hernia, hiatal, or reflux esophagitis	3,385	3	0.7170
51	Diabetes mellitus with complications	3,036	1	0.6431
158	Neoplasm, benign: Adenomatous polyps, colon	2,718	1	0.5757
149	Functional digestive disorders	2,550	1	0.5401
543	Neoplasm, benign: Skin or subcutaneous tissue	2,162	1	0.4579
93	Conjunctivitis: Bacterial	2,033	1	0.4306
378	Osteoporosis	1,985	1	0.4205
204	Dysfunctional uterine bleeding	1,842	1	0.3902
49	Diabetes mellitus type 1 maintenance	1,781	1	0.3772
226	Vulvovaginitis	1,726	1	0.3656
362	Fracture, dislocation, or sprain: Wrist, hand, or fingers	1,551	1	0.3285
365	Herniated intervertebral disc	1,302	1	0.2758
151	Hemorrhoids	1,288	1	0.2728

(*continued*)

Table 8C.1 (continued)

Episode group number	Episode label	Number of patients	Number of times drawn	Expected number of times drawn
305	Herpes zoster	723	1	0.1531
421	Obesity	661	1	0.1400
477	Bipolar disorder: Manic episode	654	3	0.1385
404	Injury: Craniocerebral	628	1	0.1330
253	Neoplasm, malignant: Lymphoma, diffuse large cell	384	1	0.0813
75	Labyrinthitis	371	1	0.0786
476	Bipolar disorder: Major depressive episode	359	1	0.0760
283	Hepatitis, chemical	292	1	0.0619
275	Cirrhosis of the liver	246	1	0.0521
490	Obsessive-compulsive neurosis	216	1	0.0458
356	Fracture or dislocation: Patella	180	1	0.0381
522	Tuberculosis	137	1	0.0290
138	Appendicitis	132	1	0.0280
2	Aneurysm, thoracic	101	1	0.0214
	New York City			
209	Neoplasm, benign: Breast	1,029	1	1.135
539	Neoplasm: Atypical nevus	989	1	1.09
204	Dysfunctional uterine bleeding	923	1	1.018
496	Asthma, chronic maintenance	821	1	0.905
173	Gastroenteritis	768	1	0.847
50	Diabetes mellitus type 2 and hyperglycemic states, maintenance	765	2	0.843
405	Injury: Spine and spinal cord	576	2	0.635
543	Neoplasm, benign: Skin or subcutaneous tissue	576	1	0.635
150	Gastritis	514	2	0.567
93	Conjunctivitis: Bacterial	491	1	0.541

331	Benign prostatic hypertrophy	437	1	0.482
153	Hernia, hiatal, or reflux esophagitis	409	1	0.451
151	Hemorrhoids	407	2	0.449
211	Neoplasm, benign: Uterus (leiomyomas)	378	1	0.417
519	Rhino, adeno, and corona virus infections	349	1	0.385
368	Injury, knee, ligamentous	311	1	0.343
125	Strabismus	246	1	0.271
436	Neoplasm, malignant: Unspecified primary site	240	3	0.265
361	Fracture, dislocation, or sprain: Humerus (head) or shoulder	228	1	0.251
506	Influenza	226	1	0.249
24	Tibial, iliac, femoral, or popliteal artery disease	217	2	0.239
23	Thrombophlebitis	147	1	0.162
5	Aortic stenosis	134	1	0.148
2	Aneurysm, thoracic	101	1	0.0214
210	Neoplasm, benign: Ovary	100	1	0.11
477	Bipolar disorder: Manic episode	84	1	0.093
137	Anorectal suppuration	72	1	0.079
339	Neoplasm, benign: Male reproductive system	54	1	0.06
395	Cerebrovascular disease, chronic maintenance	47	1	0.052
71	Foreign body: Nasopharynx, throat, or bronchus	42	1	0.046
435	Neoplasm, malignant: Nonspecific sites	37	1	0.041
214	Neoplasm, malignant: Endometrium	31	1	0.034
284	Neoplasm, malignant: Pancreas	29	1	0.032

References

Berndt, E. R., A. Bir, S. Busch, R. Frank, and S. Normand. 2002. The treatment of medical depression, 1991–1996: Productive inefficiency, expected outcome variations, and price indexes. *Journal of Health Economics* 21 (3): 373–96.

Berndt, E. R., I. Cockburn, and Z. Griliches. 1996. Pharmaceutical innovations and market dynamics: Tracking effects on price indexes on anti-depressant drugs. *Brookings Papers on Economic Activity, MicroEconomics:* 133–88.

Busch, S. H., E. R. Berndt, and R. G. Frank. 2001. Creating price indexes for measuring productivity in mental health care. In *Frontiers in health policy research,* vol. 4, ed. A. Garber, 115–47. Cambridge, MA: MIT Press.

Cutler, D., M. McClellan, and J. Newhouse. 1998. Are medical prices declining? Evidence from heart attack treatments. *Quarterly Journal of Economics* 113 (4): 991–1024.

———. 1999. The costs and benefits of intensive treatment for cardiovascular disease. In *Measuring the prices of medical treatments,* ed. J. Triplett, 34–72. Washington, DC: Brookings Institution.

Foreman, E. K. 1991. *Survey sampling principles.* New York: Marcel Dekker.

Grossman, M. 1972. On the concept of health capital and the demand for health. *Journal of Political Economy* 80 (2): 223–55.

McClelland, R., and M. Reinsdorf. 1999. Small sample bias in geometric mean and seasoned CPI component indexes. BLS Working Paper no. 324. U.S. Department of Labor, Bureau of Labor Statistics.

Rosen, A. K., and A. Mayer-Oakes. 1999. Episodes of care: Theoretical framework versus current operational realities. *Joint Commission Journal on Quality and Patient Safety* 25 (3): 111–28.

Schultze, C., and C. Mackie, eds. 2002. *At what price? Conceptualizing and measuring cost-of-living and price indexes.* Washington, DC: National Academies Press.

Shapiro, I., M. Shapiro, and D. Wilcox. 2001. Measuring the value of cataract surgery. In *Studies in income and wealth,* vol. 62, *Medical care output and productivity,* ed. D. M. Cutler and E. R. Berndt, 411–39. Chicago: University of Chicago Press.

Shapiro, M., and D. Wilcox. 1996. Mismeasurement in the Consumer Price Index: An evaluation. In *NBER macroeconomics annual 1996,* ed. B. S. Bernanke and J. J. Rotemberg, 93–142. Cambridge, MA: MIT Press.

Song, X., D. W. Marder, O. Baser, R. Houchens, E. J. Conklin, and R. Bradley. 2004. Can health care claims data improve the estimation of the medical CPI? Available at: http://www.ipeer.ca/papers/Song,Marder,Baser,%20Houchens,Conklin, Brakley,%20Oct.2004CRIW%20Paper_Oct1404.doc.