

ALTERNATIVE PRICE INDEXES FOR MEDICAL CARE: EVIDENCE FROM THE MEPS SURVEY¹

Ana Aizcorbe^{*}
Ralph Bradley^{}**
Ryan Greenaway-McGrevy^{*}**
Brad Herauf^{*}**
Richard Kane^{*}**
Eli Liebman^{*}**
Sarah Pack^{*}**
Lyubov Rozental^{}**

A B S T R A C T

Spending on medical care is a large and growing component of GDP. There are well-known measurement problems that are estimated to overstate inflation and understate real growth for this sector by as much as 1-1/2 percentage points per year. Because of its size, this would translate into an overstatement of inflation for the overall economy of about 1/4 percentage point with an equal understatement in real GDP growth. In this paper, we use data from the Medical Expenditure Panel Survey to obtain new, more comprehensive estimates for this bias and to explore a possible adjustment to existing official price indexes. The MEPS data show an upward bias to price growth in this sector of 1 percentage point, which translates into an overstatement of overall inflation of .2 percentage point and an understatement of GDP growth of the same amount. We also find that an adjustment recently used in Bradley et al provides a useful approximation to the indexes advocated by health economists.

¹ We are grateful to AHRQ for access to unpublished diagnosis codes that allowed us to do this study. We also thank Ray Kunze for helpful comments and guidance, Stephen Heffler for comments and Matt Lebowitz for excellent research assistance. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of their respective agencies.

^{*} Corresponding author: ana.aizcorbe@bea.gov, Bureau of Economic Analysis, 1441 L Street, NW, Washington DC

^{**} Division of Price and Index Number Research, Bureau of Labor Statistics, Washington, DC

^{***} Bureau of Economic Analysis, 1441 L Street, NW, Washington DC

I. INTRODUCTION

The historical rise in health care expenditures is a major national concern. In 1960, the ratio of national health care expenditures to GDP was 5.2%. Almost fifty years later, in 2008, this ratio had increased more than threefold to 16.2% and is expected to continue to increase to more than 19% by 2019 (Hartman et al 2010).

In light of this growing presence in the economy, it is essential that statistical agencies be able to accurately measure what part of this growth is inflationary, and what part is an increase in output. The generation of biased medical price indexes would generate a biased decomposition of health care expenditure growth into inflation and output growth, and this could misinform health care policy discussions.

This study focuses on this important issue of constructing of health care price indexes that are used to measure health care inflation. There is wide empirical evidence that the traditional methods used by the Bureau of Labor Statistics (BLS) to compute medical price indexes lead to upward bias. This issue was first addressed over forty years ago by Scitovsky (1967) when she computed alternative medical indexes for a single clinic in Palo Alto. She found that a price index for the treatment of conditions generated different results than a price index for individual medical services. After her study, this issue remained dormant for over twenty years, and no study by another author tried to apply her methods used in a single clinic to a national setting. However, during the past twenty years, there have been studies such as Shapiro and Wilcox (1996), Berndt et al. (1996, 1998, 2000), and Cutler et al. 1998 that focus on deriving a price index for a single disease rather than for a service. All these studies conclude that the traditional BLS method overstated the true price indexes for the specific diseases covered in their studies.

These studies prompted the Committee on National Statistics to call on statistical agencies to report national expenditures, price indexes, and output by disease rather than by medical service.²

Building on the more recent studies that focused on one disease, newer studies have attempted to find results to an index that covered all diseases. Song et. al 2009 find that indeed traditional methods did lead to upward bias. This study is based on forty

² See National Research Council (2010), Mackie (2009), and Schultze C. and Mackie C., (2002).

randomly selected diseases for patients in three cities and used a non representative claims data base for only privately administered plans. Aizcorbe and Nestoriak 2010 and Dunn et al 2010 use a claims data that are more comprehensive for geographically areas and medical conditions. They also find bias in existing BLS methods. Bradley et. al 2010 combine the Medical Expenditure Panel Survey (MEPS) with the BLS databases to explore possible adjustments to the BLS price indexes that account for this bias and found large numerical differences in the adjusted vs unadjusted price indexes.³

This study continues to compare the traditional BLS method to the disease based approach. Unlike previous studies, this study uses only MEPS data to determine i) the effect on published health inflation data if the price index were to change from a service price index (SPI) to an alternative index that accounts for changes in service mix when treating a disease (hereafter the “medical care indexes (MCE)), ii) the difference between using the adjustment method used in Bradley et al and the MCE indexes, iii) and how using different disease classifications and comorbidity methods impact the MCE indexes.

This study finds that traditional price index methods generate an annual 1 percentage point upward bias on the medical price indexes that are computed with traditional methods. This translates into an overall .2 percentage point overstatement for overall inflation with an attendant understatement of the same amount to GDP growth. In addition to these distortions to macroeconomic statistics, the traditional indexes and the MCE indexes give us different reasons for the threefold rise of the ratio of medical spending to GDP from 1960 to 2008. As such, existing measures do not properly characterize growth in health care markets. This issue and these results are also critically important for our understanding of health care markets.

This paper is organized as follows: Section II describes the price indexes and the three methods for allocating spending by disease; Section III details the data and

³ The reason that price imputation is evaluated is that BLS publishes indexes on a monthly basis and MEPS data is only available annually. Therefore, the monthly prices used to generate a monthly index can come from two sources. The first source is a medical insurance claims database. But, these databases are expensive and BLS would take on delivery risk if the vendor could not supply this data before the index publication date. The second source, price imputation, using existing price indexes would not require any additional budgetary outlays. BLS is working towards publishing experimental indexes that use this adjustment strategy (Horrihan 2010).

definitions; Section IV presents the study's findings and price indexes produced by each method; Section V concludes.

II. METHODS

We explore the potential bias in official price indexes by using the MEPS data to construct a price index of the same structure as the official price indexes provided by BLS and compare it to the type of index preferred by health economists. Both of these indexes have the problem that they do not account for changes in outcomes associated with care. However, it is widely held that the latter provide a better measure of price change than the traditional service price (SPI) indexes and that is our focus in this paper.⁴

The conventional price indexes are treatment-based, detailing spending according to specific treatments and procedures, such as a doctor's office visit or a particular drug. They reflect what is happening to the provider prices of a fixed basket of goods and services. By design, these indexes will only capture the effect of increased provider prices, not shifts to lower-cost services. Following Cutler et al. (1998), we call these fixed-basket indexes service price indexes, or SPIs; these SPIs are the official Laspeyres indexes produced by the Bureau of Labor Statistics (BLS). They answer the question "What would expenditures be if patients received the same services today as they did in the past?"

Health economists have long advocated pricing "episodes of care" or the "treatment of a condition" rather than the specific medical services provided (Scitovsky 1964).⁵ These indexes track the actual expenditures associated with an episode of care, without holding fixed the service mix. For example, if chronic episodes of depression are now treated with drug therapy, rather than the more costly talk therapy, the alternative index takes into account any cost reductions associated with this switch when quantifying what has happened to the cost of treating depression. We call these "medical care expenditure indexes" (MCEs) to emphasize that they track the overall cost of care for a condition (all expenditures), not the costs of the individual services.

⁴ See, for example, Schultze and Mackie (2002) and National Research Council (2010).

⁵ See Berndt et al. (2000) and Schultze and Mackie (2002) for a full discussion of the issues.

Below we describe these indexes in detail. We also discuss a major stumbling block to constructing MCE indexes: they require one to measure healthcare spending by disease. Because patients often suffer from more than one illness, or experience comorbidities, allocating spending to specific diseases becomes difficult.⁶ Because there is no consensus on which of the existing allocation methods is best, we do the allocations in three ways to explore the robustness of our estimates.

A. CONSTRUCTING PRICE INDEXES

The conventional indexes currently available, SPIs, *separately* track expenditures for individual services, such as inpatient stays, outpatient visits, or prescriptions. Official price indexes obtain representative bills (e.g. visits to a physician for a particular condition) and track prices of similar bills over time. We mimic this procedure in our data by tracking prices for a fixed basket of conditions treated using particular services (a bill).

The standard formula for an SPI index for service s , SPI_s that holds the basket of conditions and types of encounters in period 2 to that which was provided in period 1 is:

$$SPI_s = \frac{\sum_d c_{d,s}^2 x_{d,s}^1}{\sum_d c_{d,s}^1 x_{d,s}^1}$$

where we denote expenditures for a service used to treat condition d in period 2 using service s as $c_{d,s}^2$ and the associated quantities with $x_{d,s}^2$.

Holding the market basket fixed requires holding the x 's at period 1 levels. The numerator indicates how much the services provided to patients treated in period 1 would have cost at period 2 prices.

⁶ A similar issue arises elsewhere in the national accounts when revenues for establishments are allocated to industry classes. In this case, the revenues for individual establishments are assigned to an industry according to their primary economic activity. Thus, if a business produces goods that fall under two or more industries, the business is classified according to its major output. Thus, if a business produces goods that fall under two or more industries, the business is classified according to its major output.

In contrast, the preferred MCE index begins by considering the cost of treating individual diseases. Operationally, that cost is calculated by totaling dollars spent on all services to treat the condition (e.g. office visit and antibiotics for ear infections) and dividing those dollars by the number of cases treated, N_d : $\sum_s c_{d,s}^2 x_{d,s}^2 / N_d^2$. The ratio of this price in period 2 to that of period 1 gives disease d's component for an overall MCE and tracks changes in what is spent to treat disease d:

$$MCE_d = \frac{\sum_s (c_{d,s}^2 x_{d,s}^2) / N_d^2}{\sum_s (c_{d,s}^1 x_{d,s}^1) / N_d^1}$$

One can define aggregate versions of the SPI and MCE indexes and compare them. An aggregate SPI for all services that is similar to the BLS indexes takes a weighted average of the price indexes for each service using Laspeyres expenditure weights: $SPI = \sum_s w_s^1 SPI_s$, where $w_s^1 = \sum_d (c_{d,s}^1 x_{d,s}^1) / \sum_s \sum_d (c_{d,s}^1 x_{d,s}^1)$. Similarly, an aggregate MCE aggregates over diseases and conditions using Laspeyres expenditure weights. For example, an overall MCE takes averages of the MCEs for individual conditions using $w_d^1 = \sum_s (c_{d,s}^1 x_{d,s}^1) / \sum_s \sum_d (c_{d,s}^1 x_{d,s}^1)$, so that the overall MCE is: $MCE = \sum_d w_d^1 MCE_d$.

As discussed in the literature, differences in the aggregate MCE and SPI indexes arise from the fact that the SPI holds the market basket fixed whereas the MCE index does not. Aizcorbe and Nestoriak (2010) derive an expression showing that differences in the aggregate MCE and SPI as defined above result from changes in utilization rates over time, $dU_{d,s} = (x_{d,s}^2 / N_d^2) / (x_{d,s}^1 / N_d^1)$:

$$MCE = SPI + \sum_d w_d^1 \sum_s \{ SPI_{d,s} (dU_{d,s} - 1) \}$$

where $SPI_{d,s} = c_{d,s}^1 x_{d,s}^1 / \sum_d c_{d,s}^2 x_{d,s}^1 (c_{d,s}^2 / c_{d,s}^1)$, the contribution of disease d to SPI_s .

Intuitively, this formula says that changes in utilization rates create a wedge between the overall SPI and MCE indexes. For example, with no changes in utilization, $dU_{d,s} = 0$ for all diseases and services, the second term equals zero and the two indexes

coincide (MCE=SPI). If, instead, there were increases in the utilization of all services, all the dU terms would be greater than one and the MCE would show faster price growth—the overall cost of treating diseases would be increasing faster than the cost of the underlying services.

The data requirements for constructing MCE indexes are substantially greater than those for constructing SPIs.⁷ Because SPIs are specific to a type of service (or industry), it is possible to collect these data from the usual sources for official statistics: hospitals, pharmacies, and other establishments. Moreover, all one needs is expenditures and number of visits or encounters. In contrast, the MCE indexes require data that allow one to add up all the spending associated with the care of a patient, or data that can link services to patients. If such data were available at monthly frequencies, the BLS could construct a monthly MCE. However, the patient-level data typically only allow one to construct annual indexes and, so, cannot be used to generate official price indexes in real time.

Given these data constraints, Bradley et al. (2010) use the MEPS survey to study potential adjustments to the official Consumer Price Indexes (CPI). Below, we use the MEPS data to assess the usefulness of this strategy by comparing an adjusted SPI to the unadjusted SPI and MCE indexes.

The proposed index, SPI_s^{ADJ} may be constructed as follows: for each service, construct a weighted average of changes in utilization rates and apply the result to the SPI index for that service. Formally, SPI_s^{ADJ} for each service, s, may be written:

$$SPI_s^{ADJ} = SPI_s \sum_d \{ (c_{d,s}^1 x_{d,s}^1 / \sum_d c_{d,s}^1 x_{d,s}^1) \} \{ dU_{d,s} \}$$

and the aggregate is $SPI^{ADJ} = \sum_s w_s^1 SPI_s^{ADJ}$, where $w_s^1 = \sum_d c_{d,s}^1 x_{d,s}^1 / (\sum_s \sum_d c_{d,s}^1 x_{d,s}^1)$.

⁷ Currently, BLS conducts several surveys to collect the necessary data to construct its medical price indexes. None of these household or outlet surveys collect any information on disease conditions. Therefore, it is not possible with current BLS surveys to construct MCE indexes. Changing each of these surveys would be prohibitively expensive. Absent data on each of the components of the MCE indexes, approximating an MCE would require i) the use of MEPS to obtain the utilizations, $x_{d,d}$, and ii) the imputation of $c_{d,s}^2 / c_{d,s}^1$, as discussed below.

To see how this compares with the service indexes implicit in the MCE indexes, consider the components of the MCE indexes for services (rather than conditions). For a given service, s , this term may be written,

$$MCE_s = \sum_d \left\{ (c_{d,s}^1 x_{d,s}^1 / \sum_d c_{d,s}^1 x_{d,s}^1) \right\} \{ dU_{d,s} \} \{ c_{d,s}^2 / c_{d,s}^1 \}$$

This expression says that an MCE index for service s is an expenditure-weighted average (1st term) of the price relative (last term) multiplied by changes in utilization (middle term). This is the ideal index that the BLS should approximate.

When comparing the SPI_s^{ADJ} and the MCE_s , the SPI_s^{ADJ} is the MCE_s with $\{c_{d,s}^2 / c_{d,s}^1\}$ imputed with SPI_s . Identifying the assumptions necessary to justify this imputation is useful in understanding any differences we find in the SPI_s^{ADJ} and MCE indexes in our empirical work. There are two assumptions that could be applied to this MCE_s to obtain the SPI_s^{ADJ} . First, one could assume that the price relatives in MCE_s move according to the economy-wide SPIs that average over all conditions. In that case, we would replace $\{c_{d,s}^2 / c_{d,s}^1\}$ with SPI_s and obtain the SPI_s^{ADJ} directly. This assumption is unrealistic, however. As one example, drug prices show different growth rates depending on the entry of new drugs and generics so prices in segments with rapid product introductions (like psychotropic medications in recent decades) will likely show very different pricing patterns than those in segments with relatively slow innovation.

Alternatively, a sufficient condition for $MCE_s = SPI_s^{ADJ}$ is that $\{dU_{d,s}\}$ and $\{c_{d,s}^2 / c_{d,s}^1\}$ be independent in a particular sense. To see this, rewrite SPI_s^{ADJ} with SPI_s written in terms of the underlying cost and utilization data:

$$SPI_s^{ADJ} = \left\{ \sum_d (c_{d,s}^1 x_{d,s}^1 / \sum_d c_{d,s}^1 x_{d,s}^1) (c_{d,s}^2 / c_{d,s}^1) \right\} \left\{ \sum_d (c_{d,s}^1 x_{d,s}^1 / \sum_d c_{d,s}^1 x_{d,s}^1) (dU_{d,s}) \right\}$$

This is a product of weighted averages whereas the aggregate MCE is a weighted average of products.

To find a condition that equates $SPI_s^{ADJ} = MCE_s$ in expectation, let $\hat{E}(\cdot)$ denote the expectations operator under the frequency weight $(c_{d,s}^1 x_{d,s}^1 / \sum_d c_{d,s}^1 x_{d,s}^1)$. Then, SPI_s^{ADJ} will equal MCE_s in expectation if:

$$\hat{E}(c_{d,s}^2 / c_{d,s}^1) \hat{E}(dU_{d,s}) = \hat{E}(c_{d,s}^2 / c_{d,s}^1 \times dU_{d,s}) \text{ for all } s.$$

A sufficient condition is that the $(c_{d,s}^2 / c_{d,s}^1)$ and $dU_{d,s}$ are independent with respect to the frequency distribution for all s . This is satisfied if changes in utilization, $dU_{d,s}$, are not based on the ratio of $c_{d,s}^2/c_{d,s}^1$. For example if s is inpatient hospital services and the disease is Acute Myocardial Infraction (AMI) then the changes to hospital utilizations can not be based on the hospital price ratio for AMI and vice versa.

In our empirical work, we obtain growth rates using this adjusted SPI and compare them to the growth rates of the MCE index to assess how well the adjusted SPI approximates the MCE indexes. We evaluate the differences between the price indexes by calculating bootstrapped confidence intervals using the percentile method.

B. ALLOCATING SPENDING BY DISEASE

As seen above, the building blocks for these indexes are expenditures and utilization for the various services, further broken down by disease. There are three existing methods to allocate expenditures by disease; in our empirical work, we compare price indexes constructed with each of the three methods to assess the sensitivity of the indexes to how one allocates spending.

Many studies that attempt to measure expenditures on healthcare by disease have traditionally used the concept of “primary diagnosis” to assign spending to disease categories. More recently, these allocations have been tried using proportional methods and episode grouping algorithms (known as “groupers”). While there is no consensus about which of these alternative methods is best, work by Cutler and Rosen (2007) is underway to assess the relative merits of these methods for different purposes.

1. PRIMARY DIAGNOSIS METHOD

The simplest method allocates all spending from a medical encounter to the first-listed diagnosis on an insurance claim, which is assumed to be the “primary diagnosis.”

The primary diagnosis is considered the condition that prompted the encounter with the healthcare system, resulting in the spending being allocated entirely to that disease. Of course, this method does not account for the contribution of comorbidities to expenditures. The nature and magnitude of this omission, as well as the specific diagnoses affected, remain unknown (Cohen 2002).

In addition to the precedent set in national accounting practice for other sectors, this is the method used in the traditional cost of illness studies to measure spending for particular diseases.⁸ One advantage to this method moving forward is that the U.S. Census Bureau plans to collect data on spending by disease using this method beginning with the 2012 Economic Census.

2. PROPORTIONAL ALLOCATION METHOD

An alternative to the primary diagnosis method is allocating some of the spending from a single medical encounter to all of the diagnoses reported for the encounter. The particular amount allocated to each condition is based on encounters where only one diagnosis is reported. As described in Thorpe et al. (2004):

“We tabulated spending per event for those reporting a single medical condition (for example, heart disease and no other condition). We then tabulated spending per event for those reporting two or more medical conditions associated with the event (for example, heart disease and hypertension). We calculated the ratio of these two spending totals and used it to determine how much of the spending associated with heart disease plus other conditions should be attributed to heart disease.”

This method was also used recently by Roehrig et al. (2009) to form time series of spending on 18 disease groups and by Bradley et al. (2010) to construct price indexes. The advantage of this method over the primary diagnosis method is that it allocates some spending to all of the conditions reported with the encounter, thereby beginning to address the issue of comorbidities. The drawback is that it is not clear if spending from encounters with only one diagnosis listed is a good way to allocate spending for other encounters.

⁸ Rice (1967) is the seminal study; see Hodgson and Cohen (1999) for more recent work.

3. EPISODE GROUPING ALGORITHMS⁹

Unlike the primary diagnosis and proportional approaches, episode grouping algorithms use information from patients' histories—not just from individual encounters—to allocate spending to disease groups. This method has been used to construct price indexes by Berndt et al. (2001), Song et al. (2009), Aizcorbe and Nestoriak (2010), and Dunn et al. (2010). The major drawback to this method is that it is relatively new and is viewed by many as a black-box because the allocation method is not readily transparent.

Similar to other grouping algorithms, Thomson Reuters' Medical Episode Grouper (MEG) uses all encounters a patient has with the medical care system over a period of time to create what is referred to as an episode of care; the information gained after the grouping process allows one to assign costs to specific episodes. The grouping process considers each claim to be a single record, using diagnosis information to determine assignments of costs to episodes (MaCurdy et al. 2008).

Each episode is classified as chronic, acute, or preventative. The latter two classifications typically have clearly defined start and end dates, whereas chronic episodes have a user-defined episode length, usually of one year. The MEG grouper assigns each episode to one of 560 disease categories, and also accounts for the severity of a disease (MaCurdy et al. 2008).

The primary goal of grouping algorithms is constructing episodes of care and assigning a unique disease classification to each episode; however we only use the disease allocation assigned to each claim in order to count the number of cases treated for particular conditions over a given time period (essentially, number of patients).

III. DATA

The methods described in the previous section are all applied to the Medical Expenditure Panel Survey (MEPS) data for the years 2001-2005 in order to create the

⁹ See MaCurdy et al. (2008), Rosen and Cutler (2009) and National Research Council (2010) for

various types of price indexes. We focus on this period in part because the earlier years (1998-1999) of the MEPS survey contain substantially fewer observations; furthermore, the recommended application of the MEG episode grouper calls for dropping the first and last years of the sample (2000 and 2006).

MEPS, which is conducted by the Department of Health and Human Services' Agency for Healthcare Research and Quality (AHRQ) is a survey of the healthcare utilization and expenditures that may be used to obtain nationally representative estimates for the civilian non-institutionalized U.S. population. As a data source, MEPS is a well-known, nationally representative sample, and is generally regarded as a high quality source of data on high-prevalence health conditions. Another important strength of the MEPS data is its ability to directly link expenditures from all services (across all types of providers) to patient care events (Mackie 2009; Sing et al. 2006). Finally, MEPS is the only data set available to capture the expenditures of the uninsured (Cohen 2009).

The main drawback to these data stems from its small sample size. Although 15,000 families (35,000 individuals) are surveyed per year (Cohen, Cohen, and Banthin 2009), it has been shown that the MEPS survey undercounts spending for many conditions and misses high-cost cases (Aizcorbe et al. 2010).¹⁰ Depending on how many disease groups one wants to consider, this could lead to very few observations for many of the cells that we use as building blocks for our indexes.¹¹

With regard to variables, the MEPS survey provides both household- and patient-level data on personal healthcare expenditures. The survey contains data on health services used as well as the frequency with which households use them, their cost, and how they are paid for.¹² An observation in the data corresponds to a medical encounter.

discussions and perspectives on the usefulness of these episode grouping algorithms.

¹⁰ The survey sample is drawn from the prior year's National Health Interview Survey (NHIS) sampling frame. The survey uses an overlapping panel design in which the data are collected through a series of five rounds of interviews; the data from the overlapping panels are then used to produce annual estimates. For each household surveyed, MEPS interviews a single respondent – the family member most knowledgeable about the entire household's health and healthcare use (Zuvekas and Olin 2009a).

¹¹ Despite these disadvantages, we find MEPS to be a suitable dataset for adjusting the SPI. While claims data has superior information on prices, these data are expensive to obtain and there is a risk of the vendor failing to supply the data at any given time.

¹² MEPS actually consists of a family of three interrelated surveys: the Household Component (HC), the Medical Provider Component (MPC), and the Insurance Component (IC). The Household Component of the survey interviews individuals and families; the Medical Provider Component supplements this information by verifying prices, but not quantities, from medical providers and pharmacies. The final

For each encounter, the public use file provides up to four diagnoses, the type of service (inpatient confinement, outpatient care, either at outpatient hospitals or physician offices, or pharmacies), and the expenditure; expenditures are measured as the amount received by all providers of the services (including both out-of-pocket payments and amounts paid by insurance firms).

We use events from the following files: inpatient confinements, outpatient hospital care, office visits, emergency room care, and prescription drugs. We sum the number of encounters assigned to a each condition to obtain the number of encounters, designated with an “x” as discussed above. We measure the number of cases treated for disease *d* as the number of patients that received treatment for a disease *d* in a given period.

In the MEPS data, the overwhelming majority of events only contain one diagnosis (84 percent). The fact that very few comorbidities are reported and that the three methods rely so heavily on primary diagnoses means these data may not yield significant differences between the price indexes calculated from the three methods.

Processing the data with the MEG grouper requires 5-digit ICD-9 codes on medical records, valid NDC codes for prescription drugs, and valid dates for each event. For medical events, the MEPS public-use file only reports 3-digit ICD-9 codes , which the MEG grouper cannot process. We obtained the 5-digit codes from AHRQ and merged them with the event files. For drug events, about half of the NDC codes associated with drug events were not valid and could not be assigned a disease category by the MEG grouper; however, because drug events in MEPS have self-reported ICD-9 codes, the other methods are able to allocate the spending. With regard to dates, when there was a response for the month of service but no specific date, we arbitrarily assigned the event to the first day of the reported month. For drug events, the month of service was often missing. In those cases, we assigned those drug records to December 31 of the year in which they appeared. Although this does not affect how these records are assigned to disease classes (the MEG grouper does not use timing of services to assign

component is the Insurance Component, which collects data from employers regarding the employers' characteristics and the insurance they offer their employees (Sing et al.2006; Zuvekas and Olin 2009b; and Cohen, Cohen, and Banthin 2009).

drug events to disease classes), this means we can only produce indexes with an annual frequency.¹³

The resulting data provide the building blocks for constructing the aforementioned indexes. An important practical issue is ensuring that the data are sufficiently rich to support these calculations. The MCE indexes require measures for each disease in each period and one needs to ensure that these cells are sufficiently populated to support precise estimates of price growth. For example, the MEG grouper uses 560 disease categories. If patients were distributed uniformly across disease classes, this would place about 60 of the 30,000 or so patients each year into each disease cell.

However, patients are not distributed uniformly and this leads to thin cells with too few observations to support valid price indexes. As seen in Table 1, using the 560 MEG disease categories, only 413 of them are populated at all (we do not observe any patients for the rest) and among those, the median number of observations is 15, with 25 percent of the cells containing three or fewer observations.

It is possible to aggregate over these disease classes to categories that are coarser and, hence, will contain more observations. The Summary Diagnostic Category (SDC) schema folds the 560 MEG disease groups into 195 disease groups; the Medical Diagnostic Category (MDC) further groups these into 23 classes. Similarly, ICD-9-based groupings—like those from the primary and proportional approaches—can be scrolled up to the Clinical Classification System (CCS) level or the CCS chapter level.

The obvious tradeoff is that coarser disease groups have patients with less homogeneous conditions. In our empirical work, we construct MEG indexes at the SDC level and the primary and proportional methods at the 200+ CCS level. For ease of reporting, we take weighted averages of these disaggregate indexes to report them at the chapter level or overall.

IV. STUDY FINDINGS

For each of the three allocation methods, we construct measures of spending by disease and the corresponding implied price indexes. We then consider the adjusted SPI

¹³ If BLS applied these methods to their CPI and PPI data, they would be able to produce monthly indexes.

method and assess how well it approximates the MCE indexes. All estimates are nationally representative in that all indexes are based on data weighted by the sampling weights provided in the MEPS data.

A. SPENDING BY DISEASE

We compared how the three methods allocated spending to disease categories in a number of ways. First, the primary and proportional methods are easy to compare because they are on the same disease schema (ICD-9, CCS, and chapters). At the CCS level, the allocations by these two methods are very similar— a correlation coefficient over the 200 or so disease classes is .9. We attribute this similarity to the fact that the only difference in the allocations for these two methods would come from the presence of more than one diagnosis on each record and that happens only rarely in the MEPS data (18 percent of records).

At the CCS level, we assess the similarity between the MEG allocations to the primary method by assessing how well each of the MEG classes maps into a CCS class. We do this according to the CCS where most of that MEG's spending is allocated. If the allocations are identical, each MEG would map into only one CCS class. In our data, we find that the allocations are very similar in that about half of the MEG disease groups map over 90 percent of their spending into one CCS class. We attribute this similarity to the fact that both methods rely heavily on the primary diagnosis to allocate spending. However, there are differences: the other half of the disease groups contain spending that the episode grouper algorithm assigns to many different CCS classes. This is evidence that the MEG and algorithm methods do in fact allocate spending differently.

We also compare the allocations at the broader CCS chapter level by mapping each MEG directly to the chapters, rather than the CCS classes. The mapping at the chapter level is much more precise, with 90 percent of the MEG disease classes allocating at least 82 percent of spending to one CCS category. To compare spending at the chapter level, we assign each MEG to the CCS class where most of the spending goes.

The left panel of Table 2 shows how the three methods allocate spending into CCS chapters. As seen at the bottom, MEG does not allocate about 10 percent of spending to disease categories; these are virtually all drug claims. About one-half of the unallocated spending happens when the NDC codes that MEG uses are invalid; the rest is spending that MEG cannot allocate because it is for patients that have no episodes of care in the data to which MEG can associate the drug spending (called orphan records).¹⁴ In contrast, the primary and proportional approaches use ICD-9 codes to allocate spending and use the self-reported diagnosis codes on the drug claims to allocate drug spending to diseases.

Excluding drug events, the spending allocations by the three methods to CCS chapters are much more similar (see right panel of Table 2). For example, in the case of mental illness—a group with substantial drug spending—the three methods have very similar allocations once drug spending is excluded but the spending allocated by MEG is 25 percent less than the other two methods when drug spending is included.

These differences in the spending levels attributed to diseases may or may not translate into differences in price indexes.

B. PRICE INDEXES

We construct an SPI index based on the MEG allocation of spending and compare that index to three MCEs constructed using the three allocation methods: principal diagnosis, proportional, and MEG. To explore how sensitive price indexes are to the underlying allocation of spending, we calculate price indexes based on each of these methods and compare them.

The resulting indexes—averaged over all diseases—are shown in Figure 1. The overall SPI index is the solid line while the three MCE indexes are the dashed lines. The difference between each of the MCEs and the SPI represents an estimate of the bias present in the official prices.

¹⁴ One could fool MEG by processing the drug events as if they were ancillary services, in which case MEG would use the ICD-9 code to allocate the spending. This would take care of those claims that had invalid NDC codes but would not take care of the orphan records.

Over the period 2001-2005, all indexes show rapid price growth but the SPI index is consistently higher than the other three indexes, growing faster than the MCE indexes in three of the four years. The average compound annual growth rates are 7.8 percentage points for the SPI index and 6.9, 6.6, and 6.8 percentage points for the principal diagnosis, proportional, and MEG indexes, respectively. The estimated bias is, therefore, around one percentage point per year, depending on the spending allocation method. This estimate of the bias is lower than that reported in Aizcorbe and Nestoriak (2010) but higher than that in Dunn et al. (2010) and Bradley et al. (2010).

That the three MCE indexes show very similar rates of growth reflects how similar the spending allocations are in the MEPS data, where the vast majority of MEPS event records have only one reported ICD-9 code. This result, therefore, may not hold up in datasets that report more than one diagnosis on event records. Moreover, the exclusion of drug spending from the MEG index does not appear to generate substantial differences in the MCE indexes.

We next construct the adjusted SPI index and compare it to the unadjusted SPI and MEG indexes. As seen in Figure 2, the adjusted SPI is very similar to the MEG index: both grow around 6.8 percent per year, about one percentage point less than the unadjusted SPI. At first blush, SPIADJ does a reasonable job of approximating a disease-based index.¹⁵

V. CONCLUSION

This study uses MEPS data to determine i) the effect on published health inflation data if the price index were to change from a service price index (SPI) to an alternative index that accounts for changes in service mix when treating a disease (MCE), ii) the difference between using the adjustment method used in Bradley et al and the MCE indexes, iii) and how using different disease classifications and comorbidity methods impact the MCE indexes.

¹⁵ One caveat to our results lies in the precision of the estimates. We constructed bootstrapped confidence intervals for the SPIADJ and MCE indexes and could not reject the null that the growth in these indexes was the same as that in the SPI index. This underscores the importance of finding ways to combine the

Comparing SPI and MCE indexes, this study finds that traditional price index methods generate an annual 1 percentage point upward bias on the medical price indexes that are computed with traditional methods. This translates into an overall .2 percentage point overstatement for overall inflation with an attendant understatement of the same amount to GDP growth, *per year*. This is a substantial bias that is consistent with results found in earlier studies. The results also suggest that one can approximate the MCE indexes using the method studied in Bradley et al; the resulting SPIADJ grows at rates comparable to the MCE indexes. Finally, these results are robust to how one allocates spending by disease; the three alternative methods considered yield very similar growth rates.

More work is needed to explore the robustness of this result to choice of dataset and methods. With regard to the data, it would be useful to confirm that this result holds up in other datasets where more diagnoses are reported. With regard to the MCE indexes that we use as our benchmark, it would be useful to see how they compare to indexes constructed using other methods for allocating cost by disease, particularly the methods currently under development by Cutler and Rosen.

richness of claims data with representative surveys like the MEPS in order to improve the precision of these indexes.

References

- Aizcorbe A., Nestoriak N. (2010) [Changing Mix of Medical Care Services: Stylized Facts and Implications for Price Indexes](#) Bureau of Economic Analysis: Working Paper WP2010-08.
- Aizcorbe A., Liebman E., Pack S., Cutler D.M., Chernew M.E., Rosen A.B. (2010) Measuring Health Care Costs of Individuals with Employer-Sponsored Health Insurance in the U.S.: A Comparison of Survey and Claims Data. Bureau of Economic Analysis: Working Paper WP2010-06.
- Berndt E.R., Busch S.H., Frank R.G. (2001) Treatment Price Indexes for Acute Phase Major Depression, in: D. M. Cutler and E. R. Berndt (Eds.), *Medical Care Output and Productivity, Studies in Income and Wealth*. University of Chicago Press Chicago. pp. 463-505.
- Berndt E.R., Cutler D.M., Frank R.G., Newhouse J.P., Triplett J.E. (2000) Medical Care Prices and Output, in: A. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, Elsevier, Amsterdam. pp. 119-180.
- Bradley R., Cardenas E., Ginsburg D.H., Rozental L., Velez F. (2010) Producing disease-based price indexes. *Monthly Labor Review* 133:20-28.
- Cohen J.W. (2002) The Medical Care Expenditure Panel: An Overview. *Effective Clinical Practice*: 5(3 Suppl):152
- Cohen J.W., Cohen S.B., Banthin J.S. (2009) The Medical Expenditure Panel Survey: A National Information Resource to Support Healthcare Cost Research and Inform Policy and Practice. *Medical Care* 47:S44-S50.
- Cutler, D.M., McClellan M., Newhouse J.P., Remler D. (1998) Are Medical Prices Declining? Evidence from Heart Attack Treatments. *Quarterly Journal of Economics* 113:991-1024.
- Dunn A. Liebman, E., Pack, S., Shapiro A. (2010) Alternative Price Indexes for Medical Care: A Comparison of MarketScan and Pharmetrics Data. Bureau of Economic Analysis, mimeo.
- Hodgson, T.A., Cohen A.J. (1999). Medical Expenditures for Major Diseases, 1995. *Health Care Financing Review* 21 (Winter): 119–64.
- Mackie C. (2009) (Ed.) *Strategies for a BEA Satellite Health Care Account: Summary of a Workshop*. The National Academies Press, Washington, DC.
- MaCurdy, T., Theobald, N., Kerwin, J., Ueda, K. (November 2008b) [Prototype Medicare Resource Utilization Report Based on Episode Groupers](#). Acumen, LLC, Burlingame, CA. (Accessed 2009.)
- National Research Council (2010) *Accounting for Health and Health Care: Approaches to Measuring the Sources and Costs of their Improvement*. Panel to Advance a Research Program on the Design of National Health Accounts, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Rice, D.P. (1967) Estimating the Cost of Illness. *American Journal of Public Health* 57 (March): 424–440.
- Roehrig C., Miller G., Lake C., Bryant J. (February 2009) National Health Care Spending by Medical Condition, 1996-2005. *Health Affairs*. 28: w358-w367.
- Rosen AB, Cutler DM. Challenges in building disease-based national health accounts. *Med Care*. 2009;47(suppl 7):S7-S13.

- Rosen A.B., Cutler D.M. (2007) Measuring Medical Care Productivity: A proposal for U.S. National Health Accounts. *Survey of Current Business* 87:54-58.
- Schultze C., Mackie C. (2002) (Eds.) *At What Price: Conceptualizing and Measuring Cost-of-Living and Price Indexes*, National Academies, Washington, DC.
- Scitovsky A.A. (1964) An Index of the Cost of Medical Care - A Proposed New Approach, in: S. J. Axelrod (Ed.), *The Economics of Health and Medical Care*, University of Michigan, Bureau of Public Health Economics, Ann Arbor. pp. 128-142.
- Scitovsky, Anne A. 1967, "Changes in the Costs of Treatment of Selected Illnesses, 1951-65," *American Economic Review*, 57:5, December, 1182-1195.
- Shapiro, Matthew and David Wilcox [1996], " Mismeasurement in the Consumer Price Index:An Evaluation," *NBER Macroeconomics Annual* , Vol. 11, pp. 93-142.
- Shapiro I.L., Shapiro M., Wilcox D. (2001) Measuring the Value of Cataract Surgery, in: D. M. Cutler and E. R. Berndt (Eds.), *Medical Output and Productivity*, University of Chicago Press Chicago. pp. 411-438.
- Sing M., Banthin J.S., Seldin T.M., Cowan C.A., Keehan S.P. (2006) Reconciling Medical Expenditure Estimates from the MEPS and NHEA, 2002. *Health Care Financing Review* 28:25-40.
- Song X., Marder W.D., Houchens R., Conklin J.E., Bradley R. (2009) Can A Disease-Based Price Index Improve the Estimation of the Medical Consumer Price Index?, in: W. E. Diewert, et al. (Eds.), *Price Index Concepts and Measurements*. The University of Chicago Press for the National Bureau of Economic Research, Chicago.
- Thomas F., Caplan C., Levy, J., Cohen, M., Leonard J., Caldis T., Mueller C. (2009) Clinician Feedback on Using Episode Groupers with Medicare Claims Data. *Health Care Financing review/web exclusive/October 2009*.
- Thorpe K.E., Florence C.S., Joski P. (2004) Which Medical Conditions Account for the Rise in Health Care Spending? *Health Affairs Web Exclusive*.
- Zuvekas S.H., Olin G.L. (October 2009a) Validating Household Reports of Health Care Use in the Medical Expenditure Panel Survey. *Health Services Research* 44:1679-1700.
- Zuvekas S.H., Olin G.L. (Spring 2009b) Accuracy of Medicare Expenditures in the Medical Expenditure Panel Survey. *Inquiry* 46:92-108.

Table 1. Cell Counts for MCE and SPI Indexes

| | MEG Episode Grouper | | | Primary Diagnosis | | Proportional Approach | |
|------------------------------------|---------------------|--------|----------|-------------------|-----------|-----------------------|-----------|
| | MEG | SDC | MDC | CCS | Chapter | CCS | Chapter |
| Number of Possible Disease Classes | 560 | 195 | 23 | 249 | 18 | 249 | 18 |
| Number Populated with Data | 413 | 191 | 23 | 234 | 18 | 235 | 18 |
| Within Populated Cells | | | | | | | |
| Median Number of Patients | 15 | 60 | 1814 | 66 | 3165 | 71 | 3250 |
| IQR | 3-80 | 16-190 | 319-2156 | 16-264 | 1354-4951 | 18-295 | 1346-5155 |
| Median Number of Services | 14 | 30 | 367 | 28 | 569 | 29 | 577 |
| IQR | 3-64 | 8-129 | 98-1619 | 6-124 | 137-3018 | 6-137 | 154-3133 |

Table 2. Alternative Allocations of Spending By Disease, 2005

| | All Spending | | | Excluding Drugs | | |
|--|--------------|------------|------------|-----------------|------------|------------|
| | EA | PA | MEG | EA | PA | MEG |
| 1 Infectious and parasitic diseases | \$ 12,125 | \$ 12,658 | \$ 10,443 | \$ 7,202 | \$ 7,801 | \$ 6,886 |
| 2 Neoplasms | \$ 70,811 | \$ 71,336 | \$ 72,506 | \$ 67,414 | \$ 67,924 | \$ 70,965 |
| 3 Endocrine; nutritional; and metabol | \$ 59,270 | \$ 58,982 | \$ 41,848 | \$ 19,631 | \$ 19,422 | \$ 19,574 |
| 4 Diseases of the blood and blood-for | \$ 9,675 | \$ 10,060 | \$ 4,048 | \$ 3,924 | \$ 4,288 | \$ 3,941 |
| 5 Mental Illness | \$ 46,608 | \$ 46,480 | \$ 35,621 | \$ 23,051 | \$ 22,853 | \$ 23,714 |
| 6 Diseases of the nervous system and | \$ 52,570 | \$ 51,273 | \$ 45,724 | \$ 41,343 | \$ 40,125 | \$ 37,281 |
| 7 Diseases of the circulatory system | \$ 120,894 | \$ 121,774 | \$ 120,238 | \$ 88,462 | \$ 89,270 | \$ 96,081 |
| 8 Diseases of the respiratory system | \$ 61,666 | \$ 60,526 | \$ 53,660 | \$ 41,132 | \$ 40,062 | \$ 41,983 |
| 9 Diseases of the digestive system | \$ 63,723 | \$ 64,161 | \$ 51,972 | \$ 48,570 | \$ 48,976 | \$ 46,544 |
| 10 Diseases of the genitourinary syste | \$ 51,472 | \$ 51,863 | \$ 52,131 | \$ 44,525 | \$ 44,990 | \$ 43,939 |
| 11 Complications of pregnancy; childbi | \$ 34,110 | \$ 34,302 | \$ 35,569 | \$ 32,254 | \$ 32,421 | \$ 35,057 |
| 12 Diseases of the skin and subcutaneo | \$ 16,587 | \$ 14,693 | \$ 14,306 | \$ 13,565 | \$ 11,659 | \$ 12,790 |
| 13 Diseases of the musculoskeletal sys | \$ 73,121 | \$ 73,584 | \$ 68,471 | \$ 57,741 | \$ 58,194 | \$ 57,197 |
| 14 Congenital anomalies | \$ 5,411 | \$ 5,624 | \$ 5,265 | \$ 5,121 | \$ 5,346 | \$ 5,086 |
| 15 Certain conditions originating in t | \$ 1,823 | \$ 1,907 | \$ 1,830 | \$ 1,823 | \$ 1,907 | \$ 1,823 |
| 16 Injury and poisoning | \$ 68,577 | \$ 69,796 | \$ 69,701 | \$ 66,982 | \$ 68,203 | \$ 67,515 |
| 17 Symptoms; signs; and ill-defined co | \$ 21,741 | \$ 21,191 | \$ 8,807 | \$ 10,940 | \$ 10,399 | \$ 5,959 |
| 18 Residual codes; unclassified; all E | \$ 8,944 | \$ 8,919 | \$ 2,119 | \$ 4,666 | \$ 4,508 | \$ 1,731 |
| Unallocated | \$ - | \$ - | \$ 84,876 | \$ - | \$ - | \$ 280 |
| | \$ 779,130 | \$ 779,130 | \$ 779,135 | \$ 578,346 | \$ 578,346 | \$ 578,346 |

Figure 1. Alternative Price Indexes For Medical Care Expenditures, 2001-2005

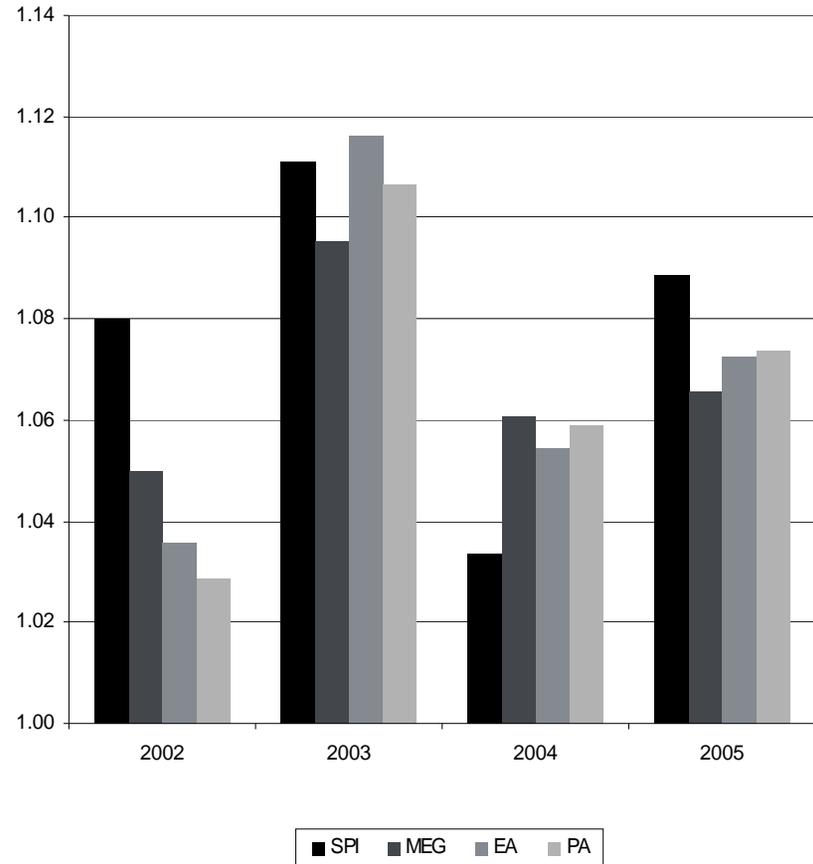
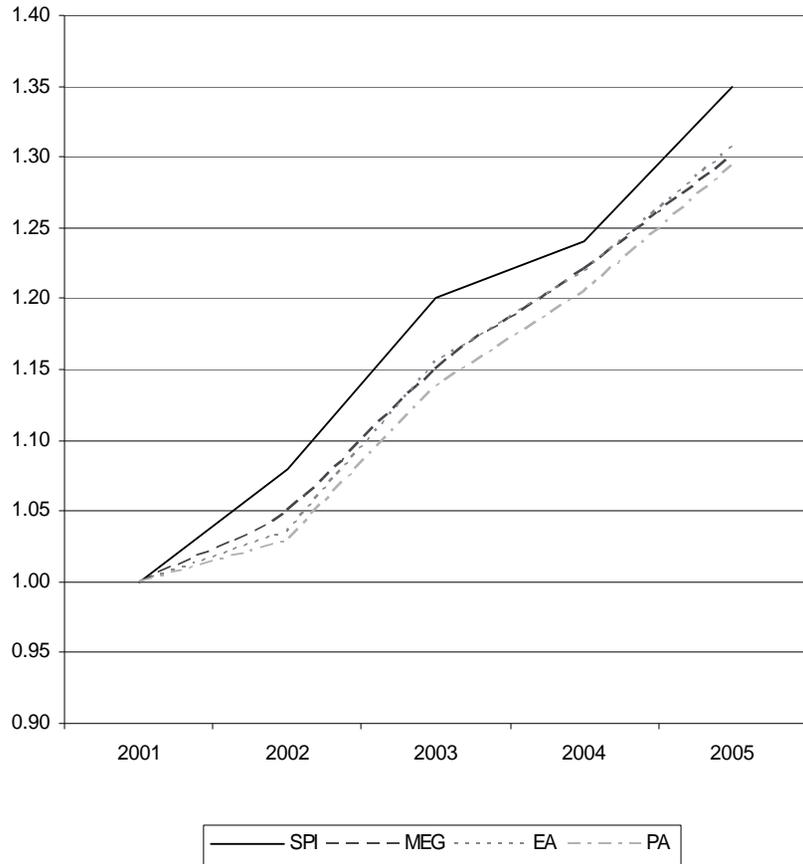


Figure 2. Effect of Adjusting SPI Index For Utilization Shifts

