NBER WORKING PAPER SERIES

CALCULATING DISEASE-BASED MEDICAL CARE EXPENDITURE INDEXES FOR MEDICARE BENEFICIARIES: A COMPARISON OF METHOD AND DATA CHOICES

Anne E. Hall
Tina Highfill

Working Paper 19720
http://www.nber.org/papers/w19720

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2013

We would like to thank Ana Aizcorbe, Abe Dunn, Richard Frank, Virginia Henriksen, participants in the NBER/CRIW conference on modeling health care costs, Tom Selden and other staff at the Agency for Healthcare Research and Quality, and an anonymous reviewer for helpful comments and advice. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Bureau of Economic Analysis. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Calculating Disease-Based Medical Care Expenditure Indexes for Medicare Beneficiaries: A Comparison of Method and Data Choices

Anne E. Hall and Tina Highfill

NBER Working Paper No. 19720
December 2013, Revised August 2014
JEL No. I1

ABSTRACT

Diseasebased medical care expenditure indexes are currently of interest to measurement economists and have been the subject of several recent papers. These papers, however, produced widely different results for medical care inflation and also varied in the datasets and methods used, making comparison difficult. In this paper, using two data sources and two different methods for calculating expenditure indexes for the Medicare population, we compare the indexes produced and establish some results that will help guide policymakers in choosing indexes for this population. We compare two methods: the primary diagnosis method and a regressionbased method. The former is preferable because of its transparency but makes stringent demands of the data. We find that when the methods are applied to the same datasets, the primary diagnosis method produces higher average annual aggregate growth rates. The difference implies that the regressionbased method should therefore be employed with caution and only when necessary. We also compare medical care expenditure indexes produced from the Medicare Current Beneficiary Survey and the Medical Expenditure Panel Survey. The MEPS is the only dataset with diagnoses attached to drug events, which significantly affects the resulting indexes. On balance, however, the MCBS is probably the preferable dataset for Medicare beneficiaries because of its greater sample size and its inclusion of nursing home residents. The optimal index may be a hybrid of the primary diagnosis method applied to Medicare claims and a regressionbased index for pharmaceutical spending. We discuss further avenues for research, such as comparing our results with indexes created with commercial groupers, and what data to use for Medicare private plan enrollees.

Anne E. Hall
Bureau of Economic Analysis
anne.hall@bea.gov

Tina Highfill
Bureau of Economic Analysis
tina.highfill@bea.gov

1 Introduction

Inflation in the health-care sector is usually measured by tracking the costs of patient goods and services and how these costs change over time. Indexes to measure the cost of services are the indexes used by the Bureau of Labor Statistics in the Consumer and Producer Price Indexes for services such as doctor visits and goods such as prescription drugs. The Committee on National Statistics of the National Research Council, however, has recommended that government statistical agencies investigate methods and data for measuring inflation in health care by measuring the costs of treating medical conditions on the grounds that measurement in this way will better capture changes in productivity in the health-care sector. Papers such as Aizcorbe and Nestoriak (2011), Dunn et al. (2012), and our previous work (Hall and Highfill 2013) followed these recommendations and found average annual growth rates for health-care inflation that ranged from 2.9 percent to 6.9 percent. However, the papers used different datasets on different populations and employed different methods for splitting up health-care expenditure by condition, making it difficult to analyze why the results varied so much.

In this paper, we are following up on our previous work by comparing medical expenditure indexes for the Medicare population calculated from different datasets and using different methods. The Medicare program plays a very large role in the US healthcare system, in the federal budget, and in the US economy. In calendar year 2012, the program covered 50.8 million people (42.2 million aged and 8.6 million disabled). Medicare spending comprised about one-fifth of national health care spending in the 2012 National Health Expenditure Accounts (Center for Medicare and Medicaid Services). According to the Congressional Budget Office (CBO), outlays on Medicare were 17% of total federal outlays in 2013 and the CBO expects the aging of the population and rising health-care costs, through their effect on Medicare spending, to be major factors in the projected increase in the deficit later this decade (CBO 2014, CBO 2013). Finally, outlays on Medicare were 3.5% of US gross domestic product in 2013 (CBO 2014). Accurately measuring inflation in the spending of Medicare beneficiaries is therefore important both for more precise measurement of the economy and for better understanding of the value of a large fraction of public spending.

Several Medicare program datasets are available, each with unique characteristics that present measurement challenges. In this paper, we compare two major data sources on Medicare beneficiaries: the Medicare Current Beneficiary Survey (MCBS) and the Household Component of the Medical Expenditure Panel Survey (MEPS). Both surveys are conducted by the Department of Health and Human Services. The MCBS surveys Medicare beneficiaries exclusively while the MEPS Household Component surveys US residents living in the community (that is, not in institutions such as nursing homes) about their health care and spending.

By necessity, our comparison focuses on the potential indexes for Medicare beneficiaries enrolled in fee-for-service (FFS) Medicare, for whom the most complete data exists. Medicare private-plan enrollees are, however, making up an increasing share of Medicare enrollment but the data for them is much less comprehensive. In the MCBS, the only source for their spending, diagnoses and medical events is the in-person survey but as we shall see, the survey asks about a limited set of diagnoses and some important ones would be omitted if we based their expenditure index on the MCBS survey. However, the Medicare private plan sample in the MEPS is quite small, only about 300 to 400 beneficiaries per year.

Using the MCBS and MEPS, we compare two methods for calculating medical expenditure indexes: the primary diagnosis method and a regression-based approach. The primary diagnosis method simply assigns spending to the illness associated with the diagnosis code (or first diagnosis code in the case of multiple diagnosis codes) of each claim or survey-collected medical event. The regression-based approach regresses individual annual health-care spending on dummy variables for a beneficiary's diagnosed medical conditions, and divides up each beneficiary's health-care spending on that basis.

The paper is organized as follows. Section 2 gives background on the concept of a disease-based medical care expenditure index. Section 3 introduces the various methods available for producing medical care expenditure indexes. Section 4 introduces the datasets available for Medicare beneficiaries. Section 5 introduces our full matrix of comparisons, compares the methods, and draws some conclusions about the methods. Section 6 compares the datasets and draws some conclusions about the datasets. Section 7 concludes.

2 Medical Care Expenditure Indexes

Medical care expenditure indexes (MCE indexes) measure inflation in health-care spending by measuring the changes in the costs of treating individual illnesses. This approach contrasts with that of the official health-care price indexes, such as the Consumer Price Index (CPI) and the Producer Price Index (PPI), which measure the change in price over time of specific health-care services, such as doctor's visits. Early papers that used MCE indexes such as Cutler et al. (1998), Shapiro et al. (2001), and Berndt et al. (2002) argued that measuring health-care inflation with MCE indexes was more meaningful since it took into account shifts in utilization from one service category to another as the technology of treating a particular disease improves and becomes more efficient. Examples include shifting from psychotherapy to prescription drugs in the treatment of psychiatric illnesses and shifting surgeries from inpatient to outpatient procedures. Health-care inflation, when measured in this way, often turns out to be lower than inflation as measured by service price indexes (SPI) such as the CPI or PPI. For example, Cutler et al. found that the measured price change of treating a heart attack was lower than an SPI by three percentage points annually. For more on these issues, see Berndt et al. (2000) and National Research Council (2002).

Cutler et al. and Berndt et al. constructed quality-adjusted expenditure indexes for individual conditions (heart attacks and major depression respectively). Quality adjustment of expenditure indexes in the health-care context entails connecting shifts in treatments with changes in health outcomes. This adjustment requires high-quality data on individual treatments and outcomes and papers such as Aizcorbe and Nestoriak (2010) and Dunn et al. (2012) that constructed expenditure indexes for a broad set of illnesses did not adjust for changes in health outcomes when comparing MCE indexes to SPIs like the CPI. Both papers found that health-care inflation was lower when measured by MCE indexes than by SPIs even though the MCE indexes were not quality-adjusted. In this paper, we will only present MCE indexes that are not quality-adjusted.

3 Methods

Individuals seeking medical care are often diagnosed with more than one disease; a central problem in the creation of MCEs is how to divide expenditures among multiple diagnoses. Several different methods have been proposed and used in prior papers. It should be noted that economic theory is no guide on methodology in this area and there is no true way of validating any of the methods. Brief descriptions of the methods available to us follow here and the results of the previous papers are summarized in Table 1.

3.1 Regression-based method: This method models total individual annual health-care spending as a function of each individual's diagnoses during that year and uses the parameter values to divide the individual's health-care spending among his or her diagnoses. Then individual spending on conditions is averaged to give mean expenditures which are the inputs to an MCE index. Health-care spending has certain characteristics that make it challenging to model econometrically: spending is non-negative, there are a large number of observations with zero spending, and the distribution of spending is typically skewed with a long right-hand tail. Modeling health-care spending is the subject of an extensive literature (Duan 1983, Manning 1998, Jones 2000, Manning and Mullahy 2001, Buntin and Zaslavsky 2004). The consensus is that health-care spending should be transformed (with a log or square root transformation) to accommodate its skewness. Either a one-part or two-part model, where the probability of having spending at all is modeled in a separate step, may be used. The parameters may be estimated with either ordinary least squares (OLS) or a generalized linear model (GLM). Buntin and Zaslavsky (2004) investigate the modelling of the spending of Medicare beneficiaries in particular. Their recommendation is to estimate with GLM with the mean-variance relationship established with a Park test (Park 1966). Following the results of a Park test, we estimate the following equation with a one-step GLM, with a log link and the standard deviation proportional to the mean (a gamma distribution):

$$\ln(y_i) = \beta_0 + \sum_{j=1}^{J} \beta_j D_{ij} + \varepsilon_i$$

In this equation, i indexes the individual beneficiary, j indexes the conditions, y_i is each beneficiary's annual health-care spending, and D_{ij} is an indicator variable for whether individual i has condition j.

Since we are fitting log spending, we cannot use the resulting coefficients to directly find the average spending on each condition. Instead, we use a method for using parameter estimates from a model of the log of health-care spending to assign spending to conditions originally proposed by Trogdon et al. (2008). In this method, each individual's spending is divided up into their diagnosed conditions in proportion to the regression coefficients. A share of spending is calculated for each beneficiary-illness combination as follows:

$$S_{ij} = \frac{\left[\exp\left(\hat{\beta}_j\right) - 1\right] * D_{ij}}{\sum_{j=1}^{J}\left\{\left[\exp\left(\hat{\beta}_j\right) - 1\right] * D_{ij}\right\}}.$$

The shares are then applied to each individual's spending to give that individual's spending on each condition with which he or she is diagnosed. Then the average expenditure for each condition is calculated across individuals. These average expenditures are analogous to prices in a price index and are the inputs to the MCE indexes. We combine them with the diagnosed prevalences for each condition (which are analogous to quantities in a price index) and calculate the MCE indexes as Fisher indexes in the usual manner.

The regression-based method has the advantage that it does not make huge demands of the data, compared to the other methods discussed here: it only requires individual annual medical spending and dummy variables for whether or not the individual was diagnosed with a certain condition that year. The primary disadvantage is that the method of assignment of spending to particular diagnoses is not based on any theory or model that relates health-care spending to conditions diagnosed.

3.2 Primary diagnosis: This method is used in Aizcorbe et al. (2011). With this method, the spending attached to a claim or medical event is assigned to the diagnosis or the first diagnosis, if there are multiple diagnoses, attached to that claim or event. In some sense, the primary diagnosis method is preferable to all others because the connection between spending and

diagnosis is transparent and not dependent on an econometric model of health-care spending. This method requires, however, that every claim or event have a diagnosis attached to it. Other than the MEPS, the datasets being considered for the Medicare population do not have a diagnosis attached to every claim or event; the survey events collected by the Medicare Current Beneficiary Survey that are not in the Medicare claims (such as drug events and all medical events for Medicare private plan enrollees) do not have diagnoses attached to them and the Medicare Part D claims do not have diagnoses attached to them either.

A further difference between the two methods to keep in mind is that they measure slightly different costs. The regression-based method measures the net cost of a condition; that is, the average difference in spending between a beneficiary with the condition and one without. The net cost of a condition can be negative and we found in Hall and Highfill (2013) that several conditions from the MCBS in-person survey, such as Alzheimer's disease, had negative coefficients in the regression and were therefore assigned negative costs. With the primary diagnosis method, all conditions have positive costs. If there are more beneficiaries with, for example, Alzheimer's disease, the per-patient expenditures of other conditions that are comorbid with Alzheimer's disease will be lower in the primary diagnosis while they would, in theory, be unaffected in the regression-based method. This is another reason for preferring the primary diagnosis method since price indexes in other sectors generally do not allow goods or services with negative prices.

3.3 Commercial grouper: Several private companies have developed commercial software for grouping medical spending by episode of illness based on clinical knowledge. These packages are used in Aizcorbe and Nestoriak (2011), Dunn et al. (2012), and Aizcorbe et al. (2011) to create expenditure indexes for medical care. Aizcorbe and Nestoriak (2011) use one on medical claims data from private employer-sponsored health insurance plans collected by Pharmetrics, Inc., and Dunn et al. (2012) use one on the MarketScan database, a similar dataset of medical claims collected by Truven Health Analytics. Aizcorbe et al. (2011) use a commercial grouper on the MEPS data. The main advantage of these packages over the primary diagnosis method is that they are able to assign spending associated with claims, such as drug claims, that do not

have a diagnosis attached directly to them. In addition, unlike the annual regression-based method, they are able to separate out multiple episodes of the same illness occurring in one year. Depending on the package, they are also able to assign a severity level to the illness. However, their methods are proprietary and therefore completely opaque to the economist using them. We will not be considering groupers in our comparisons in this paper but may study them in future work.

Table 1: Previous papers calculating medical care expenditure indexes					
Paper	**Dataset**	**Population**	**Years**	**Method**	**Results (AAGR of price index in percentage points)**
Aizcorbe and Nestoriak 2011	Pharmetrics, Inc. (medical claims from private employer-sponsored health insurance plans)	Beneficiaries of private employer-sponsored insurance	2003-2005	Commercial grouper (Symmetry)	3.7
Dunn et al. 2012	MarketScan (medical claims from private employer-sponsored health insurance plans)	Beneficiaries of private employer-sponsored insurance	2003-2007	Commercial grouper (Symmetry)	3.6
Aizcorbe et al. 2011	Medical Expenditure Panel Survey (MEPS)	US non-institutionalized civilian residents	2001-2005	Primary diagnosis, proportional diagnosis, commercial grouper (Truven Health Analytics)	6.9, 6.6, 6.8
Hall and Highfill 2013	Medicare Current Beneficiary Survey	Medicare beneficiaries	2001-2005	Regression-based	5.8

4 Datasets

Table 2 summarizes the features of datasets available for calculating medical care expenditure indexes for Medicare beneficiaries. In general, to create a medical care expenditure index, we need variables for total spending and for diagnoses at at least an annual level.

Table 2: Summary of datasets covering Medicare beneficiaries

Dataset	Coverage	Annual sample size of Medicare beneficiaries	Data available for FFS beneficiaries	Data available for private plan enrollees
Medicare Current Beneficiary Survey (MCBS)	All Medicare beneficiaries	12,000	Annual demographic and conditions survey, all medical events and spending, Medicare Part A (hospital) and Part B (physician) claims	Annual demographic and conditions survey, all medical events and spending
Medicare claims	FFS Medicare beneficiaries	~2 million	Part A and Part B claims; Part D (pharmaceutical) claims for about 50-60% of 5% sample	n/a
Medical Expenditure Panel Survey (MEPS)	Non-institutionalized Medicare beneficiaries	4,600	All medical events and spending, with diagnoses attached (collected by survey)	All medical events and spending, with diagnoses attached (collected by survey)

4.1.1 Medicare Current Beneficiary Survey (MCBS): The MCBS is a survey of the demographics, diagnosed conditions, health status, and total medical spending of a representative sample of Medicare beneficiaries. It is conducted by the Center for Medicare and Medicaid Services, the agency that operates Medicare. As it samples from the universe of Medicare beneficiaries, it includes both FFS Medicare beneficiaries and those enrolled in Medicare private plans, and

both beneficiaries residing in the community and in institutions such as nursing homes. The medical conditions portion of the survey takes place once a year, towards the end of the year, and in it, the respondent is asked whether they have been told by a doctor if they have each of about 30 conditions. Health-care spending and medical events are also collected directly from the respondent on a regular basis. For FFS beneficiaries, the Part A and Part B claims with dollar amounts and diagnosis codes are also attached to the survey so there are two sources of diagnoses and spending for these beneficiaries. The MCBS reconciles the orally reported events and the claims so that spending and events are not duplicated in the final version of the dataset. For private plan enrollees, the only source of information is the spending, events, and diagnoses reported in the in-person survey.

4.1.2 Medicare claims: Part A (hospital) and Part B (doctor) claims are available for a 5% random sample of Medicare beneficiaries from CMS for research purposes. In addition, starting in 2006, a sample of Part D claims are available for those 50-60% of FFS beneficiaries in the 5% who are on Part D. In this paper, rather than showing calculations from the full 5% sample, we will evaluate Medicare claims data by using the Medicare claims data tied to the MCBS.

4.1.3 Medical Expenditure Panel Survey (MEPS): The MEPS is a nationally representative survey of healthcare coverage, utilization, and expenditures for the civilian non-institutionalized U.S. population. It is conducted by the Department of Health and Human Services' Agency for Healthcare Research and Quality (AHRQ). The survey sample is drawn from the respondents of the prior year's National Health Interview Survey (NHIS) and includes both fee-for-service and private plan Medicare beneficiaries living in the community (that is, not in a nursing home or other institution). Using an overlapping panel design, each household is surveyed over the course of two years in five rounds of interviews. The family member most knowledgeable about the entire household's health and health care use is interviewed. Observations are collected and reported for every medical event and may contain up to four diagnoses each. The MEPS also collects data from a sample of respondents' providers to verify use of services, charges and sources of payments, and diagnoses.

Table 3 compares the spending of all Medicare FFS beneficiaries in the MCBS, Medicare FFS beneficiaries in the MCBS who are living in the community, and Medicare FFS beneficiaries in the MEPS (who all reside in the community). As it shows, mean spending by Medicare beneficiaries in the MEPS is quite a bit lower than that of similar Medicare beneficiaries in the MCBS. Zuvekas and Olin (2009) conduct a detailed comparison of full-year Medicare beneficiaries residing in the community in the MEPS, those same beneficiaries' Medicare claims, and similar beneficiaries in the MCBS in the years 2002-2003. They find that the ratio of mean spending of beneficiaries in the MEPS to that of similar beneficiaries in the MCBS to be 0.81. We find a similar result; as table 3 shows, the ratio of the spending of full-year Medicare beneficiaries living in the community in our analytic sample from the MEPS to that of similar beneficiaries in our sample from the MCBS ranges between 0.69 and 0.80 over the period studied. Our MCBS sample includes claims by separately billing labs, which that of Zuvekas and Olin does not, which may explain the slightly lower ratio. Zuvekas and Olin find that about half of the gap is due to underreporting of spending by respondents to the MEPS, and about half is due to the absence of higher-expenditure cases in the MEPS. We also find an absence of higher-expenditure cases in the MEPS as compared to the MCBS; as table 3 shows, both the skewness coefficient for the distribution of spending and the value of the maximum observation are generally much higher in the MCBS, implying a longer tail in that distribution. Unfortunately, we have no way of adjusting for underreporting and the absence of more expensive cases in the MEPS but simply note that they may be factors in any differences we observe between the MEPS and the MCBS.

4.2 Prevalence

In this section, we discuss issues relating to how treated prevalence of medical conditions is measured in the datasets we are using since it has important implications for the resulting expenditure index. To see this, consider that total health-care spending in a population can be expressed as a sum over conditions:

$$Total\ health-care\ spending = \sum_{j=1}^{J} \bar{P}_j N_j$$

where j indexes conditions, P_j is the average amount spent to treat condition j and N_j is the number of people treated for condition j. N_j divided by the total population is the treated prevalence for that condition.

P_j is the object of concern when we are calculating an expenditure index; however, the data we use have total spending and N_j as given and we must infer P_j with one of the various methods that will be discussed further down. In addition, as is standard in a Fisher index formula, treated prevalences are used to weight the indexes for the individual conditions when they are aggregated into an expenditure index for medical care as a whole. Differences in methods for collecting conditions from respondents will therefore lead to different results in the resulting expenditure indexes.

4.2.1 Treated prevalence in the Medicare Current Beneficiary Survey

The first issue we note is the differences in the MCBS between survey-based prevalences and claims-based prevalences. The survey-based prevalences in the MCBS are based on the annual survey administered to all respondents. In this survey, respondents who live in the community are asked if they have been told by a doctor in the past year if they have each of about 30 conditions. For respondents who are residing in nursing homes, the MCBS has a nursing home staff member fill out a questionnaire about conditions based on the respondent's nursing home and medical records. The claims-based prevalences are based on the respondent's Part A (hospital) and Part B (physician) Medicare claims. The claims are only available for respondents who are enrolled in fee-for-service (FFS) Medicare and not for respondents enrolled in Medicare private plans.

For the FFS beneficiaries in the MCBS therefore, we have both sets of diagnoses for the same beneficiaries and can examine the degree of agreement between them. Table 4 reports prevalences in 2001 from the MCBS survey and from the MCBS claims, as well as the percent of beneficiaries who are indicated as having an illness in both the survey and the claims.[2] First, note that in general claims-based prevalence is higher than survey-based prevalence. There is

[2] For how the survey diagnoses were translated into ICD-9 codes for purposes of comparison with the claims, see Appendix Table 1 in Hall and Highfill 2013.

also quite a bit of disagreement between the two sources of diagnoses and the agreement rate is often less than half of either the survey-based prevalence or the claims-based prevalence. When weighted by prevalence, the overlap rate divided by the claims-based prevalence averages about 40 percent and the overlap rate divided by the survey-based prevalence averages about 62 percent. One possible cause of disagreement in prevalence between the survey and the claims may come from the lack of drug-related diagnoses in the claims data. As mentioned above, the claims data attached to the MCBS only include doctor and hospital claims. There are no prescription drug claims, even for beneficiaries enrolled in Medicare Part D after 2006 and, in any case, prescription drug claims generally do not include diagnoses. Beneficiaries may therefore have chronic illnesses for which they are taking prescription drugs regularly and which they report on the survey but either they did not have a non-drug event related to the condition during the survey year or these conditions are not recorded by their health-care providers. However, some of the highest rates of agreement between survey and claims are in illnesses for which this issue would most be of concern. There are five chronic illnesses for whom the share of spending on pharmaceuticals is over 50 percent of total spending on the illness: diabetes, mental/psychiatric disorder, Alzheimers/dementia, osteoporosis, and hypertension. Agreement rates between survey-based and claims-based prevalence are in fact noticeably higher for this group: the overlap rate divided by the claims-based prevalence has a weighted average of 63 percent and the overlap rate divided by the survey-based prevalence has a weighted average of 72 percent. It seems possible therefore that the act of taking a daily drug for an illness may actually improve survey respondents' knowledge and memory of what diagnoses they have.

There are multiple other potential causes of the discrepancies between survey-based and claims-based prevalence. The reasons that a condition might be reported in the survey but not in the claims include beneficiary errors in the existence of illnesses, beneficiary errors in the timing of illnesses, providers omitting to code a condition, and that the condition was treated but the treatment was not paid for by Medicare. The reasons that a condition might be reported in the claims but not in the survey include again beneficiary or provider error and the provider upcoding a diagnosis. Given that claims-based prevalence is generally higher than

survey-based prevalence and that the disagreement is noticeably lower for conditions for which beneficiaries are taking a prescription drug on a regular basis, it seems likely that beneficiary error is the largest source of disagreement. Claims-based prevalence may therefore be a more reliable measure of treated prevalence in the MCBS. However, as noted above, there is a notable disadvantage to relying solely on claims data as the source data for Medicare beneficiaries when constructing an MCE index, namely that we do not have claims for the enrollees in Medicare private plans.

4.2.2 Treated prevalence in the Medical Expenditure Panel Survey

The second problem with measuring treated prevalence in the datasets we are examining relates to a change in the Household Component of the MEPS in 2007 in its method for relating health care events and spending to medical conditions. In the MEPS, medical conditions are collected up front from respondents in their initial survey. Before 2007, in the initial interview, conditions were only reported if the respondent volunteered them in response to a general question about medical conditions. In 2007, MEPS introduced the Priority Conditions Enumeration section of the survey, in which respondents were asked whether they had certain priority conditions. (This method is similar to how conditions are collected in the MCBS throughout the entire period.) When medical events are later collected, the respondent is prompted to relate these events to conditions that had been reported in the initial survey.

The treated prevalences from the MEPS that we use to calculate an MCE are calculated from the medical events files and not from the conditions survey because we are primarily interested in conditions as they relate directly to health care expenditure. However, the methodology change in 2007 resulted in a dramatic increase in the event-based treated prevalences of certain conditions. Table 5 shows the treated prevalences of the priority conditions in the Priority Conditions Enumeration file from the MEPS by year from 2001-2009. The total prevalence in 2007 is the top number of the three reported for each condition for that year. As it shows, there are increases in nearly all the conditions between 2006 and 2008. There are particularly dramatic rises in the prevalences of heart disease, arthritis, heart attacks, and stroke, which are clearly unrelated to any pre-existing trend in treated prevalence.

The Priority Conditions Enumeration section was phased in with the introduction of Panel 12 of the MEPS.[3] Table 5 also shows treated prevalence by condition in 2007 split out by panel. Treated prevalence in Panel 11 in 2007 for most conditions is comparable to treated prevalence in 2006, while treated prevalence of most conditions in Panel 12 in 2007 is more similar to treated prevalence as measured in 2008, reflecting the methodology change that affected Panel 12 but not Panel 11. The differences between Panels 11 and 12 in treated prevalence are especially noticeable for the conditions mentioned above that have the greatest increases in treated prevalence. For example, the treated prevalence of heart disease is more than four times higher in Panel 12 than in Panel 11, and that of arthritis is more than five times higher. Similarly, the treated prevalence of myocardial infarction is more than four times higher in Panel 12 than Panel 11. The methodology change also helps explain part of the increases in treated prevalence of hypertension and hypercholesterolemia, two conditions with treated prevalence that are both high and trending up over this period. Panel 12 has a treated prevalence of hypertension in 2007 that is over five percentage points higher than that of Panel 11, and for hypercholesterolemia the difference is just over ten percentage points.

In our MEPS-based MCE indexes, in order to have growth rates that are at least measured over consistent samples, we use the growth rate calculated from Panel 11 alone for the change from 2006 to 2007, and the growth rate calculated from Panel 12 alone for the change from 2007 to 2008. The expenditure indexes based on the MEPS are therefore not strictly comparable before and after 2007, but this solution is the best we can do with the data available. The effect of implementing this change on one of the MCE indexes we calculate is shown in Figure 1. This MCE index is calculated from the MEPS on all Medicare beneficiaries using the primary diagnosis method. (See below for a discussion of this method for calculating MCE indexes.) As the figure shows, using the growth rate from Panel 11 for the change from 2006 to 2007 and the growth rate from Panel 12 for the change from 2007 to 2008 removes an unusual drop in the MCE index in 2007 that is out of line with the underlying trend.

[3] We are grateful to Tom Selden of AHRQ for this information and for proposing the solution that follows.

5 Comparisons of Different Methods for Calculating Medical Care Expenditure Indexes

As was shown in the previous two sections, we have a choice of methods and datasets available to us to calculate medical expenditure indexes for the Medicare population, although not every method will work with every dataset. Table 6 lays out the possible combinations of methods and datasets and shows the average annual growth rates of Fisher medical care expenditure indexes calculated from those methods and datasets for the years 2001-2009. The columns represent different combinations of samples (FFS and private plan, FFS only), datasets (MCBS or MEPS) and sets of illnesses, and the rows of the table represent different methods (regression-based or primary diagnosis, with or without drug spending). The illnesses used are either the 27 illnesses from the MCBS survey (see Hall and Highfill 2013 for a list and detailed discussion) or the 260 categories from the Clinical Classifications System (CCS), a system devised by the AHRQ for classifying the 10,000 or so ICD-9 diagnosis codes used in claims datasets into medical conditions. In one column we use 27 CCS diagnoses that correspond to the MCBS survey diagnoses; see Appendix Table 1 in Hall and Highfill (2013) for a crosswalk that translates survey diagnoses into claims diagnoses. For ease of comparison across years, when creating the FFS-only sample, we restricted the sample to beneficiaries enrolled in Medicare for the full calendar year. The downside of this approach is that we lose beneficiaries who pass away during the year and who are responsible for a good part of total Medicare spending.

The cells in the table are labeled with letters for ease of reference. As described above in section 3, the regression-based indexes are based on a GLM model of the log of health-care spending as a function of diagnoses. For most of the regression-based indexes, a beneficiary is coded as having a condition if the associated diagnosis code appears in any diagnosis variable but for the indexes whose references end in 2 (H2, I2, and M2), the diagnosis codes were only counted if they appeared as a primary diagnosis in order to make those indexes more comparable to the primary diagnosis analyses. As described in section 4, the growth rates of the MEPS indexes from 2006 to 2007 are only calculated from Panel 11 while those of the MEPS

indexes from 2007 to 2008 are only calculated from Panel 12, in order to adjust for the methodology change in collecting conditions from respondents in the MEPS in 2007.

As Table 6 shows, the MCE indexes show a range of average annual growth rates, from 1.4 percent to 6.4 percent per year. In general, average annual growth rates are higher in the indexes based on the MEPS than in those based on the MCBS, and are higher when calculated with the primary diagnosis method than with a regression-based method. The growth rates in the indexes based on the MCBS are lower when drug spending is omitted, but those in indexes based on the MEPS are higher when drug spending is omitted; we will discuss this result further below.

Table 7 shows the growth rates of some selected indexes by year. There is a significant amount of volatility in the growth rates from year to year. In addition, there is surprisingly little correlation in yearly movements between the datasets with the exception that all indexes have strongly positive growth rates in 2003. There is some correlation in yearly movements between the two indexes constructed from the MCBS claims as the two indexes always move in the same direction but little correlation among the indexes constructed from the MEPS.

The first step in our comparisons is to compare the price indexes obtained using the regression-based method with the primary diagnosis method when we use them on the same samples and the same illnesses. As noted above, the primary diagnosis method is probably preferable but can only be used when the data meet its stringent requirements. The main purpose of this comparison therefore is to see if the regression-based method gives similar or very different results from the primary diagnosis method and if it is an adequate substitute for the primary diagnosis method when the data do not have a diagnosis attached to each event or claim.

There are three pairs of analyses to compare, as summarized in Table 8 which shows their average annual growth rates from 2001-2009 as reported in Table 6:

Table 8: Average annual growth rates for selected medical expenditure indexes 2001-2009		
Data	**Regression-based**	**Primary diagnosis**
MEPS FFS beneficiaries	3.6 (H2)	5.0 (P)
MEPS FFS and private plan beneficiaries	3.8 (I2)	5.0 (Q)
Medicare claims (omits drug spending)	2.0 (M2)	2.8 (S)

As the table shows, the primary diagnosis method produces higher growth rates for the MCE indexes. The difference is over a percentage point for the indexes based on the MEPS and almost a percentage point for the index based on the MCBS.

Table 9 examines the similarity of the two methods in measuring the per-patient expenditures of individual conditions. It shows the correlations in per-patient expenditure assigned to conditions by the two methods, both in their level (averaged over 2001 and 2009) and in their annualized growth rates from 2001 to 2009. As the table shows, the correlations between the levels of per-patient expenditures are strongly positive. They are especially high in the two MEPS comparisons, at over 80%, but slightly lower for the MCBS comparison where the correlation is about 68%. The correlations between the growth rates in per-patient expenditures are much lower. They are still positive in the two MEPS comparisons but are essentially zero for the MCBS comparison.

Table 10 shows the annualized net growth rates of per-patient expenditures from 2001 to 2009 for a selection of individual conditions. The conditions shown are the union of the sets of the top ten conditions by per capita spending in 2001 and 2009 for each dataset and method combination shown. As suggested by the results in Table 9, there is little similarity in growth rates across methods applied to the same data. The table shows that, in general, the regression-based method produces more extreme growth rates, both positive and negative, than the primary diagnosis method. The regression-based method can produce particularly extreme growth rates when combined with the MCBS claims data, as in the case of "deficiency and other anemia", "aortic and peripheral arterial embolism or thrombosis", and "other

Table 10

Annualized growth rates of indexes for selected individual conditions 2001-2009

Conditions*	MCBS claims		MEPS (FFS only)		MEPS (FFS and private plan)	
	Regression-based (M2)	Primary diagnosis (S)	Regression-based (H2)	Primary diagnosis (P)	Regression-based (I2)	Primary diagnosis (Q)
Medical examination/evaluation	9.0%	4.2%	-14.6%	-2.1%	-15.8%	-2.1%
Other lower respiratory disease	1.6%	4.1%	-11.8%	-0.9%	-9.7%	-1.1%
Other connective tissue disease	18.3%	5.7%	-0.4%	1.3%	4.1%	1.8%
Nonspecific chest pain	5.1%	6.2%	36.5%	9.9%	52.9%	10.2%
Other aftercare	7.1%	7.3%	12.8%	7.1%	5.9%	7.2%
Chronic renal failure	-9.7%	-4.6%	1.4%	-1.5%	-0.7%	-1.3%
Complication of device, implant or graft	5.2%	-1.4%	-14.0%	24.3%	-13.8%	25.8%
Deficiency and other anemia	102.9%	8.9%	7.9%	-4.0%	9.5%	-3.6%
Coronary atherosclerosis and other heart disease	-0.3%	0.1%	5.4%	4.1%	6.1%	3.5%
Spondylosis, intervertebral disc disorders, other back problems	9.4%	9.4%	2.4%	7.9%	1.2%	7.2%
Residual codes, unclassified	20.1%	9.1%	6.9%	11.8%	6.3%	9.6%
Diabetes mellitus without complication	4.5%	1.9%	2.8%	0.7%	2.3%	0.9%
Pneumonia (except that caused by tuberculosis or sexually transmitted disease)	3.9%	6.6%	6.7%	6.4%	5.2%	6.3%
Cardiac dysrhythmias	-4.2%	2.8%	7.0%	-6.2%	7.2%	-5.3%
Chronic obstructive pulmonary disease and bronchiectasis	-3.5%	2.6%	-5.1%	5.1%	-4.8%	5.3%
Acute myocardial infarction	3.1%	3.5%	-8.0%	-6.8%	-8.5%	-7.4%
Congestive heart failure, nonhypertensive	14.3%	9.0%	-2.4%	2.4%	-3.8%	3.0%
Osteoarthritis	3.0%	4.0%	-7.9%	-7.1%	-8.6%	-6.8%
Cataract	-0.5%	3.7%	-1.6%	5.1%	-2.0%	4.5%
Rehabilitation care, fitting of prostheses, and adjustment of devices	-2.4%	3.2%	-6.5%	-1.1%	106.1%	-1.3%
Other non-traumatic joint disorders	-7.2%	8.0%	17.4%	14.2%	14.7%	12.3%
Mood disorders	-7.6%	-1.0%	-9.6%	-8.5%	-9.4%	-8.3%
Acute cerebrovascular disease	-4.8%	-3.6%	-2.9%	0.6%	-2.6%	1.1%
Other and ill-defined heart disease	-7.6%	1.2%	9.3%	2.5%	9.9%	2.9%
Essential hypertension	-10.3%	4.0%	6.2%	2.9%	4.6%	3.2%
Anxiety and personality disorders	28.7%	7.5%	47.3%	-3.0%	31.1%	-3.8%
Aortic and peripheral arterial embolism or thrombosis	915.7%	37.8%	11.0%	-4.8%	12.0%	-3.9%
Acute and unspecified renal failure	12.7%	3.7%	1.6%	5.5%	-0.5%	4.3%
Other disorders of stomach and duodenum	-129.0%	2.0%	-9.0%	2.3%	-6.6%	3.6%
Disorders of lipid metabolism	3.4%	6.8%	-2.3%	1.2%	0.3%	1.4%

*Ranked by 2009 per capita spending in M2

disorders of stomach and duodenum." In addition, as we noted above, table 9 shows that the correlation in results between the two methods is noticeably smaller for the MCBS claims data than for the MEPS data. These results are probably due to the greater sensitivity of the regression-based method to outliers combined with the greater presence of outliers in the MCBS claims data as shown above in table 3. As we discussed in the introduction to the methods section, there is no reason to use the regression-based method with data where it is possible to use the primary diagnosis method, and the evidence in this table supports that claim. In situations where it is necessary to use the regression-based method, the individual and aggregate growth rates of the MCE indexes should be interpreted with caution.

It is unclear why the primary diagnosis method consistently produces higher growth rates than the regression-based method. There is no reason to expect it a priori to do so. The difference is widespread across conditions; as Table 10 shows, the primary diagnosis method produces higher growth rates for just over half of the thirty conditions, and the relationship holds for both chronic and acute conditions. Of the full set of conditions, about half have a higher growth rate in their MCEs when estimated with the primary diagnosis method.

In conclusion, it seems that the regression-based method should be employed with caution, given its sensitivity to outliers and propensity for producing volatile indexes. In addition, as we discussed above, the regression-based method has the undesirable feature that it can assign negative costs to some conditions. However, it may be necessary to use it with data where events are not assigned diagnoses individually. It should be noted that, while we presented one regression-based approach, there is a considerable amount of discretion in how this approach can be applied, both in the modeling of spending as a function of diagnoses and in how the coefficients are used to divide up individuals' health-care spending. One potential approach going forward therefore may be to adjust the regression-based approach until the results match the primary diagnosis method within a certain level of tolerance in the MEPS data, and then apply that adjusted approach to the larger sample and larger targeted population of the MCBS. Another approach may be to combine the two methods: use the primary diagnosis

method on the Medicare Part A and B claims in the MCBS but use a regression-based method on the drug spending from the survey portion of the MCBS.

6 Comparison of Medical Care Expenditure Indexes Calculated from Different Datasets

As Table 6 shows, we also calculated price indexes with the same methods and on comparable populations from the MCBS and the MEPS, for the purposes of comparing the results. In general, the MCBS or the Medicare claims are preferable datasets for analyzing the Medicare population to the MEPS since they have larger samples, also include the nursing home population, and do not seem to have the same underreporting issues as the MEPS. However, comparing the datasets offers a useful check on the MEPS which has been used on other populations in the medical-care expenditure index literature.

Table 11: Average annual growth rates for selected MCE indexes 2001-2009

Method	MCBS	MEPS
Regression-based including drug spending	2.1 (D)	2.8 (F)
Regression-based omitting drug spending	1.4 (M)	4.8 (N)
Primary diagnosis omitting drug spending	2.8 (S)	6.4 (T)

Table 11 shows the pairs of cells from Table 6 that directly compare datasets with the same method and comparable populations. The population in both cases is FFS Medicare beneficiaries who are enrolled in Medicare for the full calendar year. In the MCBS sample, we use claims for diagnoses so as to be able to compare on a condition-by-condition basis with the MEPS, in which conditions are coded in ICD-9 diagnosis codes. As the table shows, when drug spending is included and we use a regression-based method, the MCE index based on the MEPS is slightly higher, by about 0.8 percentage point, than that based on the MCBS. When drug

spending is omitted, however, the difference is over three percentage points, whether a regression-based method or the primary diagnosis method is used.

Table 12 shows the correlations in per-patient expenditures across conditions between the expenditure indexes calculated from the MCBS and the MEPS with the same methods. The first column shows the correlations in the level of per-patient expenditure averaged across 2001 and 2009; as they show, the correlations are positive but not overwhelmingly strong. The correlations are strongest for the two indexes calculated using a primary diagnosis method without drug spending, which is the pair with the largest difference between their annual average growth rates. It is important to keep in mind therefore that the same method, similar data, and similar results can still produce substantively different average annual growth rates for an MCE index. The next column shows the correlations in the growth rates of the expenditure indexes for individual conditions from 2001 and 2009 across the two datasets; these are much lower and often close to zero.

Returning to table 11, we see that omitting spending and events related to pharmaceuticals from the MCE indexes has the opposite effect on the indexes from the two datasets. The growth rates of MCE indexes without pharmaceuticals are lower than those with pharmaceuticals when they are calculated from the MCBS but omitting pharmaceuticals raises the growth rates of MCE indexes based on the MEPS considerably. The difference arises from the way pharmaceutical events are recorded in the two surveys. Pharmaceutical events are recorded together with a diagnosis in the MEPS while in the MCBS, pharmaceutical events are collected in the oral survey portion and are not associated with a diagnosis. In the MEPS, dropping the pharmaceutical events can therefore change the treated prevalence of conditions, since some conditions are only reported in pharmaceutical events. As discussed above in the section on prevalence, a change in the growth rate of treated prevalence can affect the growth rate of an MCE index, all other things equal, both because the per-patient expenditure will change and because the aggregation weights in the Fisher index will change.

When we compare an unweighted average of treated prevalences across conditions based on drug events and the same average based on non-drug events in the MEPS, we find that the

former grows considerably faster over the period studied. Treated prevalence based on drug events has an average annual growth rate between 2001 and 2009 of 4.8% while that based on non-drug events has an average annual growth rate of 2.8%. Omitting drug events would therefore tend to lower the growth rate of prevalence and raise the growth rate of the MCE index, all other things equal. The other main factor affecting the difference between MCE indexes with and without pharmaceutical spending and events is pharmaceutical spending itself. Omitting pharmaceutical spending but not the associated diagnoses would tend to lower the growth rate of the MCE index since pharmaceutical spending generally grows more quickly than other categories of medical spending. In the case of these indexes based on the MEPS, however, it appears that the changes in prevalence resulting from removing the pharmaceutical events more than offset the change to the index resulting from removing pharmaceutical spending.

From these comparisons we may conclude that the inclusion of pharmaceutical events in the treated prevalence in the data can have a substantive effect on an MCE index based upon that data; this has repercussions for the use of the MCBS as a data source for Medicare beneficiaries. As the MCBS lacks diagnoses related to pharmaceutical events, it will omit some conditions that beneficiaries are diagnosed with entirely and a regression-based index will misattribute their pharmaceutical spending to other conditions.

7 Conclusion

We have compared medical care expenditure indexes calculated from different datasets and using different methods. Our belief going into this research was that the primary diagnosis method was the best method for dividing up health-care expenditure by disease but that the Medicare Current Beneficiary Survey was the best dataset for analyzing Medicare beneficiaries, as it has the widest coverage and the most information on them. However, the primary diagnosis method cannot be used with the drug spending and events in the MCBS. We therefore compared the primary diagnosis method with a regression-based method for estimating expenditures by disease and found that, when they are used to calculate MCE

indexes on the same datasets, the primary diagnosis method produces higher average annual aggregate growth rates. The difference is relatively small for the analyses with the MCBS claims and much larger with the analyses using the MEPS. The annual indexes show some correlation between methods in yearly movements with the MCBS data but little correlation across the MEPS analyses, possibly because of the smaller sample size of the MEPS. There is strongly positive correlation in the levels of per-patient expenditures across conditions and between the two methods but almost no correlations in the growth rates of the individual condition indexes. The regression-based method produces much more volatile individual indexes when applied to the MCBS claims data, probably because its results are more sensitive to outliers and there are more outliers in the MCBS claims data. The regression-based method should be employed with caution and only when necessary therefore.

We also compared medical care expenditure indexes for Medicare beneficiaries produced from the MCBS and the MEPS. As noted above, the MCBS has greater coverage and a larger sample size and the MEPS appears to have some problems with underreporting and underwent a methodology change in collecting diagnoses during the time period we are studying. When drug spending is included, however, the two datasets produce MCE indexes with very similar average annual growth rates, the one from the MEPS being a bit higher. When drug spending is omitted, however, the indexes from the MEPS have much higher growth rates, by more than three percentage points. As we discuss, the difference probably comes from the fact that when drug events are dropped from the analyses, the measurement of prevalence is affected in the MEPS but not in the MCBS, since the MCBS lacks diagnoses for its drug events. That the omission of drug-related diagnoses affects the MCE indexes for Medicare beneficiaries to such a degree is an argument in favor of using the MEPS, which attaches diagnoses to individual drug events. On balance, however, the MCBS is probably the preferable dataset for Medicare beneficiaries because of its greater sample size and its inclusion of nursing home residents.

The best solution for FFS Medicare beneficiaries, in the end, may be a hybrid index: one that combines the primary diagnosis method applied to the Part A and Part B claims in the MCBS with a regression-based index for pharmaceutical spending. For private-plan beneficiaries, however, the solution is not as clear. For this population, our choices are between a regression-based method run on the MCBS with only the 27 diagnoses included in the MCBS survey, or the primary diagnosis or regression-based method run on the small group (about 300-400 a year) of Medicare private-plan enrollees in the MEPS. Exploring these options will probably be the subject of some of our future work.

While we believe this comparison has covered the most important methods and datasets, it is not comprehensive. In particular, we did not cover the use of commercial groupers and we did not explore the larger sample of Medicare claims that are available for research (including the Part D pharmaceutical claims). We hope to address these gaps in future work.

References

Aizcorbe, A., R. Bradley, R. Greenaway-McGrevy, B. Herauf, R. Kane, E. Liebman, S. Pack, and L. Rozental (2011), "Alternative Price Indexes for Medical Care: Evidence from the MEPS Survey", BEA working paper.

Aizcorbe, A. and N. Nestoriak (2011), "Changing mix of medical care services: Stylized facts and implications for price indexes", *Journal of Health Economics* 30: 568-574.

Berndt, E. R., A. Bir, S. H. Busch, R. G. Frank, and S. T. Normande (2002), "The Medical Treatment of Depression, 1991-1996: Productive Inefficiency, Expected Outcome Variations, and Price Indexes", *Journal of Health Economics*, 21: 373-396.

Berndt, E. R., D. Cutler, R. Frank, Z. Griliches, J. Newhouse, and J. Triplett (2000), "Medical Care Prices and Output", in *Handbook of Health Economics,* vol 1. (A. J. Cuyler and J. P. Newhouse, eds.), Elsevier.

Buntin, M. B. and A. M. Zaslavsky (2004), "Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures", *Journal of Health Economics* 23: 525-542.

Center for Medicare and Medicaid Services, National Health Expenditure Data, http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html.

Congressional Budget Office (2013), "The Budget and Economic Outlook: Fiscal Years 2013 to 2023", Feb. 13 (http://www.cbo.gov/publication/43861).

Congressional Budget Office (2014), "Updated Budget Projections: 2014 to 2024", April (http://www.cbo.gov/publication/45229).

Cutler, D. M., M. McClellan, J. P. Newhouse, D. Remler (1998), "Are Medical Prices Declining? Evidence from Heart Attack Treatments", *Quarterly Journal of Economics* 113: 991-1024.

Duan, N. "Smearing estimate: a nonparametric retransformation method", *Journal of the American Statistical Association* 78 (1983): 605-610.

Dunn, A., E. Liebman, S. Pack, and A. Shapiro (2012), "Medical Care Price Indexes for Patients with Employer-Provided Insurance: Nationally-Representative Estimates from MarketScan Data", Health Services Research, October.

Hall, A. and T. Highfill (2013), "A Regression-Based Medical Care Expenditure Index for Medicare Beneficiaries", BEA working paper (http://www.bea.gov/papers/pdf/a_regression_base_medical_care_expenditure_index_for_medicare.pdf).

Jones, A. M. (2000), "Chapter Six: Health Econometrics" in *Handbook of Health Economics,* vol 1. (A. J. Cuyler and J. P. Newhouse, eds.), Elsevier.

Manning, W. G. (1998), "The logged dependent variable, heteroscedasticity, and the retransformation problem", *Journal of Health Economics* 17: 283-295.

Manning, W. G. and J. Mullahy (2001), "Estimating log models: to transform or not to transform?", *Journal of Health Economics* 20: 461-494.

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

Park, R. (1966), "Estimation with heteroscedastic errors", *Econometrica* 34:888.

Shapiro, I., M. D. Shapiro, and D. W. Wilcox (2001), "Measuring the Value of Cataract Surgery", in *Medical Care Output and Productivity* (D. M. Cutler and E. R. Berndt, eds.), University of Chicago Press.

Trogdon, J., E. A. Finkelstein, and T. J. Hoerger (2008), "Use of Econometric Models to Estimate Expenditure Shares", *Health Services Research* 43:4 (August): 1442-145.

Zuvekas, S. H., and G. L. Olin (2009), "Accuracy of Medicare Expenditures in the Medical Expenditure Panel Survey", *Inquiry* 46 (Spring): 92-108.