

Developing a Framework for Decomposing Medical-Care Expenditure Growth: Exploring Issues of Representativeness*

Abe Dunn, Eli Liebman, and Adam Hale Shapiro

November 20, 2012

Abstract

Medical care expenditures have been rising rapidly over time and in 2009 health care accounted for 17.9 percent of GDP, but there are many areas where we have an incomplete understanding of spending growth in this sector. This is especially true of the commercial sector, where our primary data sources are often non-random convenience samples (i.e. available claims data from contributing insurers and employers). The goal of this paper is to better understand issues related to using convenience samples to obtain nationally representative estimates of the various components of expenditure growth. Using a multitude of weighting strategies, including weighted and unweighted estimates, we find similar qualitative results with higher prevalence and increases in medical care service prices being the key drivers of spending growth. Utilization per episode - which was relatively flat for both weighted and unweighted estimates - made no significant contribution to growth in spending. However, applying population weights provide quantitative results that align better with aggregate price and expenditure measures, including

*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 Federal Reserve Bank of San Francisco, or the Board of Governors of the Federal Reserve System.

BEA price measures and commercial expenditure per capita estimates from the NHEA, and may be more useful for certain policy and projection purposes.

1 Introduction

The large and growing share of GDP allocated to medical care has prompted greater focus on producing health statistics that provide more detailed information on the sources of expenditure growth and the value of those expenditures. One shortcoming of current national statistics is that they contain no information on medical expenditures by disease, even though the primary aim of purchasing medical-care services is disease treatment. This gap in our understanding of health expenditures has been noted by numerous academics and policy makers that have called for additional research in this area and the development of a National Satellite Account for Health Care that would help fill this void ((See Berndt et al. (2000) and Accounting for Health and Health Care (2011))). There have been a number of case studies on the value of health spending that explore the costs and benefits of different treatments and technologies for particular diseases (e.g., heart attacks, depression, cataracts and high cholesterol). However, only relatively recent research in this area has started to formulate statistical indexes that track the components of health expenditures by disease for a broad set of health conditions. For example, studying the full range of diseases, Aizcorbe and Nestoriak (2011) decompose expenditure per episode and service prices, and Roehrig and Rousseau (2011) focus on decomposing prevalence and expenditure per episode.¹ More recently, Dunn, Liebman, and Shapiro (2012a), a companion piece to this paper, decomposes each of these dimensions of health expenditure growth in a single framework.

This paper focuses on health care expenditures in the commercial insurance market, which is an economically important segment—accounting for more enrollees than

¹There are a number of other studies that look at expenditures per episode and service prices. Dunn et al. (2012) follow the methodology of Aizcorbe and Nestoriak (2011), but use an alternative data source. Aizcorbe et al. (2011) look at the decomposition between service price and expenditure per episode using national survey data. Other studies have also looked at the decomposition of prevalence and expenditures per case including Thorpe, Florence, and Joski (2004), Roehrig and Rousseau (2011), Roehrig et al. (2009).

Medicare and Medicaid combined. The importance of the commercial sector is set to grow with the implementation of the Affordable Care Act that is predicted to result in millions of uninsured individuals entering the commercially-insured segment of the market. Many of the richest data sources for studying the commercial sector are convenience samples comprised of insurers and employers that contribute their employee or enrollee claims information.² Measurement issues may arise if these convenience samples are not representative and are evolving in a non-random fashion. For instance, estimates may not be representative of the U.S. population if the sample is disproportionately from enrollees living in a particular geographic area. The work by Dunn, Shapiro, and Liebman (2012) shows there is large variation in medical care expenditure levels across the U.S., so it is conceivable that growth rates could also vary geographically. The age and sex distribution of the convenience sample may also be different from the actual U.S. commercial population, leading to inaccurate estimates of expenditure growth. Finally, differences in the data contributors may also impact our estimates, as different groups with distinct populations (e.g., different employers or insurers) enter and leave the sample. In this setting, the application of population weights and an appropriately selected sample may be critical for obtaining meaningful estimates that are nationally representative. This paper studies issues related to representativeness of the sample by examining how various weighting strategies and samples affect the components of spending growth .

To study this topic, we apply the full decomposition methodology outlined in Dunn, Liebman, and Shapiro (2012a) that looks at the various components of expenditure growth at the disease level. This decomposition starts with per capita spending, which is further broken into prevalence and expenditure per disease episode. Expenditure per disease is further broken out into a service utilization and service price component. Here we apply this methodology to study how different population weights and samples could impact the components of medical care expenditure growth. Our study employs the MarketScan commercial claims data that is a convenience sample from health insurers

²Although several studies use the Medical Expenditure Panel Survey data, which is a representative sample of the civilian noninstitutionalized U.S. population. The key limitation is that the sample size is relatively small (32,000 per year) which makes estimates from this data source to be greatly impacted by outliers and, given that this is a household survey, many diseases are potentially under-reported.

and large employers for the years 2003 to 2007. There is no guarantee that the data that is provided by these contributors will reflect the overall commercial population. Moreover, additional data is added to the sample over time, with the number of enrollees increasing from around 7 million in 2003 to more than 13 million in 2007, which could influence our estimates.

Our primary strategy for dealing with the potentially non-representative aspects of the data is to apply population weights, so that the weighted sample reflects the commercially insured population. We find that the unweighted MarketScan data produces qualitatively similar results to our weighted estimates along many dimensions. As in the work of Dunn, Liebman, and Shapiro (2012a) we find that the main trends in spending growth are characterized by both increases in the underlying services price (e.g., price for a 15 minute office visit) and the growth in the prevalence of disease episodes. The utilization per episode of treatment is flat or is falling slightly over the period of study, which implies that conditional on having a disease episode, individuals are receiving approximately the same intensity of treatment over time.

We also investigate how changes in the data contributors (i.e., the insurers and employers providing the data) may impact our estimates. To address this issue, the sample is limited to those contributors that provide data for the whole sample period. Overall trends in this subsample are similar, although certain aspects of the fixed contributor estimates appear more plausible than those of the full sample. We discuss this issue in greater detail in the text. Although many of the qualitative findings do not change with the use of weights or alternative samples, there is evidence that the application of population weights may be of practical importance. In particular, the application of population weights produces spending and price growth figures that are more aligned to national benchmark estimates.

The qualitative finding that expenditure growth is driven by prevalence and service prices is interesting, but it is also worth highlighting that after adjusting for overall inflation, real medical-care expenditure growth is almost entirely caused by an increase in treated prevalence. Specifically, more than three quarters of the “real” growth in spending may be attributed to treated prevalence, with the remainder accounted for by expenditures per episode. This finding is robust across numerous estimates. This result

contrasts sharply with estimates of Roehrig and Rousseau (2011) who find that growth is primarily driven by growth in expenditure per episode. Although we discuss some possible reasons for our different finding, more work is necessary to isolate the precise cause of this discrepancy.

The importance of population weights will partly depend on whether different demographic groups have distinct growth rates. For example, if younger individuals are over-represented and have faster than average expenditure growth, then estimates of expenditure growth will be over-stated. To see if the broadly observed trends apply to all segments of the population, we estimate the various components of expenditure growth for different subpopulations. Focusing first on subpopulations by age, we find the same general patterns of growth across all age groups, with expenditures primarily driven by service price and prevalence, but there also are some noteworthy differences among age groups. For instance, expenditure growth per person appears to be faster for children under 18 relative to other age segments. Also, service prices and utilization tend to grow more rapidly for younger populations, relative to older populations, although prevalence growth tends to be slower for younger individuals. Focusing next on the regional differences in growth, we find that spending patterns in three of the four regions follow a similar trend, but that spending in the South grows markedly slower. These findings hold for both the full sample and the sample with the fixed data contributors, although the differences among the growth rates in the different regions is less pronounced when using the information from the fixed data contributors.

There are several other methodological issues that arise when studying the components of expenditure growth that are not covered in this paper. Some of these topics are covered in companion pieces to this work: (1) Dunn et al. (2012b) examine different approaches for assigning medical services to disease categories and the effect on the components of expenditure growth; (2) Dunn, Liebman, and Shapiro (2012b) examine alternative strategies for separating utilization and price, and look at how this affects the decomposition; (3) Dunn, Shapiro, and Liebman (2012) study the geographic differences in expenditure levels across MSAs. Also, it should be noted that the primary focus of this paper is to discuss samples and the application of population weights, so several interesting economic trends observed in our indexes are not discussed in detail

here. See Dunn, Liebman, and Shapiro (2012a) for a more in-depth economic analysis of medical-care expenditure trends.

This paper is divided into five sections. First, we discuss our methodology for medical care expenditure construction. Next, we discuss the data used in our analysis and present some descriptive statistics. We then present our results and discuss the sensitivity of these results. In the last section, we conclude.

2 Methodology of Index Construction

The decomposition methodology of this paper borrows heavily from Dunn, Liebman, and Shapiro (2012a) looking at expenditure growth. To begin, we start with a measure of expenditure per capita for disease d for time period t , which is $C_{d,t} = \frac{\text{Total Expenditures}_{d,t}}{\text{Commercial Population}_t}$. A measure of medical care expenditure growth per capita from period 0 (the base period) to t is then the *expenditure per capita index* (ECI):

$$ECI_{d,t} = \frac{C_{d,t}}{C_{d,0}} \quad (1)$$

Since the denominator of the $C_{d,t}$ term is the full commercially insured population, this measure of expenditure growth does not take into account the health of the population. For instance, if expenditures are higher in the second period because more individuals develop heart disease, the $C_{d,t}$ will grow, even if the expenditure per heart disease episode do not change. Alternatively, $C_{d,t}$ may grow if the expenditure per heart disease episode increases, even if the population of individuals with heart disease remains unchanged.

2.1 Expenditure per Capita Decomposition: Expenditure Per Episode and Treated Prevalence

Given the expenditure per capita index, we next decompose ECI into the prevalence of the condition and the expenditure per episode. We start by dividing, $C_{d,t}$, into two components. One component of the expenditure per capita is the prevalence of treated disease, $prev_{d,t}$. Prevalence of treated disease d is the number of episodes treated

in the population for disease d , $N_{d,t}$, divided by the commercially insured population: $prev_{d,t} = \frac{N_{d,t}}{\text{Commercial Population}_t}$. Note that prevalence includes only those instances where there is awareness of a disease and treatment was provided. It therefore excludes those instances where the individual is unaware of their condition.³ The second component of expenditure per population is the expenditure per episode or average expenditure for treating disease d , $c_{d,t}$. The value $c_{d,t}$ may be calculated by dividing total expenditures of disease d by the number of episodes of disease d in period t , $c_{d,t} = \frac{\text{Total Expenditures}_{d,t}}{N_{d,t}}$. Distinct indexes may be constructed from each of these two components. One component is the growth in treated prevalence, relative to the base period:

$$PREV_{d,t} = \frac{prev_{d,t}}{prev_{d,0}}$$

The second component is the change in expenditures per case relative to the base period:

$$MCE_{d,t} = \frac{c_{d,t}}{c_{d,0}} \tag{2}$$

These two components of expenditure capture distinct elements of its growth. Changes in the prevalence of a treated condition capture the changing health of the population, such as the growth in diabetes due to obesity. It may also reflect a growing awareness of a condition, such as the increase in awareness and diagnosis of high cholesterol. The second component of care may be viewed as the price for treating the disease, which includes the prices of those services and also the mix of those services provided. Assuming that the quality of the underlying treatment mix remains constant, this treatment price reflects the productivity in the health sector for the treatment of disease d .

Using these equations, one can see that the expenditure per capita is then the cost per episode times the prevalence, $C_{d,t} = c_{d,t} \cdot prev_{d,t}$. From this we can see that the $ECI_{d,t}$ may be decomposed into the expenditure per episode index, $MCE_{d,t}$ and the

³Those individuals that have a condition but are unaware that they have a condition or do not seek medical attention for their condition would be considered in measuring the population's prevalence, but are not included in the treated prevalence figure.

treated prevalence index, $PREV_{d,t}$:

$$ECI_{d,t} = MCE_{d,t} + PREV_{d,t} + \frac{(prev_{d,t} - prev_{d,0})(c_{d,t} - c_{d,0})}{prev_{d,0}c_{d,0}} - 1.^4 \quad (3)$$

This equation makes it clear that a population-based measure of expenditure for a particular disease will rise if there is either an increase in the prevalence of the disease or an increase in the expenditures per episode.

The indexes presented here are directly related to a simple and often reported figure, total medical care expenditures per capita. To see this, we can create aggregate disease-specific indexes for the population-based measure, $ECI_{d,t}$. When $ECI_{d,t}$ is weighted by the national expenditure share for each disease in the base period, this becomes a measure of medical-care expenditures per person relative to the base periods's medical-care expenditures per person:

$$\begin{aligned} ECI_t &= \sum_D ECI_{d,t} \cdot (\text{Expenditure Share}_0) \\ &= \sum_D \frac{C_{d,t}}{C_{d,0}} \cdot \left(\frac{C_{d,0}}{\sum_D C_{d,0}} \right) = \frac{\sum_D C_{d,t}}{\sum_D C_{d,0}} \\ &= \frac{\text{Medical-Care Expenditures Per Capita}_t}{\text{Medical-Care Expenditures Per Capita}_0}. \end{aligned}$$

This measure *includes* any change attributable to the prevalence of certain diseases. Thus this measure will grow along with disease prevalence. The measure may also reflect the changing demographics of the population, such as, the growth from an aging population. As we will discuss later, we typically apply population weights that may change the meaning of these indexes. For instance, we may apply weights that hold the age, sex and location characteristics of the population constant, so that changes in prevalence do not affect these demographic changes. Population weights will be discussed in greater detail below. Next, we further decompose the disease expenditures per episode into a service price component and a service utilization component.

⁴A decomposition using logs is: $\log(ECI_{d,t}) = \log(MCE_{d,t}) + \log(PREV_{d,t})$.

2.2 Expenditure Per Episode Decomposition: Service Price and Service Utilization

The MCE index is a measure of the medical-care expenditures for the treatment of an episode of care for a certain disease, and is defined as the dollar amount of medical care used until treatment is completed.⁵ Since this index controls for the health of the population, it may be viewed as measuring the cost of treatment. Thus, if the $MCE_{d,t}$ is larger than one, it signifies that the expenditure for treating disease d is larger than the base period and if the index is less than one it signifies that the expenditure is less than the base period.

Our decomposition rests on the fact that the average expenditure, $c_{d,t}$, can be divided between a service price and service utilization component. This can be seen more easily by showing that the average expenditure is calculated by totaling dollars spent on all services to treat the condition and dividing those dollars by the number of episodes: $c_{d,t} = \sum_s p_{d,t,s} Q_{d,t,s} / N_{d,t}$, where $Q_{d,t,s}$ is the quantity of services for service type, s ; $p_{d,t,s}$ is the service price for service type s ; and $N_{d,t}$ is the number of episodes treated.

Measuring service utilization is not a straightforward task since the definition of a “service” is a bit ambiguous and there are a variety of ways that one could define it across various service types.⁶ The approach taken here to define service utilization closely follows the methodology of Dunn, Shapiro, and Liebman (2012). Ideally, we would like the definition of a specific service to depend on how the price of that service is typically set and paid. For example, for physician services, individuals pay a unique price for each procedure done to them (that is, the insurer and the patient together pay this amount). Therefore, we would like service utilization to reflect the amount of procedures done. Since not all procedures are equivalent, we weight each procedure by the average dollar amount paid for that procedure. This is a similar concept to a

⁵For example, for an individual with a broken foot, the episode of treatment will be defined by the dollar of medical services used to treat that condition from the first visit to a provider until the foot is healed. For medical conditions that are chronic, we interpret an episode as expenditure for services used to treat the chronic condition over a one year period.

⁶The key service types are inpatient hospital, outpatient hospital, general physician, physician specialist, and prescription drugs.

“Relative Value Unit” or “RVU”, which measures the approximate cost of each procedure and is used by Medicare to reimburse physicians for each procedure that is performed.⁷ For prescription drugs, we define the unit of service as a prescription filled, albeit this is a bit of a misnomer since a prescription is really a “good,” not a service. Since prescriptions vary depending on the active ingredient, the manufacturer, and strength, we weight each unique drug purchase by the average dollar amount we observe for that particular prescription across geographic areas. For hospital facility charges for inpatient stays, the prices paid to facilities are often set based on the disease and the number of visits to a facility. Therefore, for inpatient stays we define the unit of service as visit. For outpatient facility services we also define the service as the visit itself. The exact construction of these measures is explained in more detail later in this paper.

Given the definition of service and expenditure, the price for a particular service type and disease can be calculated by dividing its expenditure by the quantity of services provided: $p_{d,t,s} = \frac{c_{d,t,s}}{Q_{d,t,s}}$ where $c_{d,t,s}$ is the average expenditure on disease d for service type s at time t . For example, the price of an inpatient stay for treating heart disease is the total expenditure of inpatient treatment for heart disease in an area, divided by the quantity of inpatient services for heart disease in that area.

This decomposition allows us to create a service price and service utilization index. To simplify, let $q_{d,t}$ be a vector of services utilized for the typical treatment of diseases in an area, $q_{d,t} = Q_{d,t}/N_{d,t}$, where the elements of the utilization vector for service type s is, $Q_{d,t,s}/N_{d,t}$. Also, let $p_{d,t}$ be a vector of service prices, where the elements of the vector for service type s is, $p_{d,t,s}$. The service price index (SPI) is then calculated as:

$$SPI_{d,t} = \frac{p_{d,t} \cdot q_{d,0}}{c_{d,0}}$$

which holds the utilization of services fixed at a base period level, but allows prices to vary. Similarly, the service utilization index (SUI) may be defined as:

⁷This framework has also been adopted by the commercial market. In a survey of 20 health plans conducted by Dyckman & Associates, all 20 health plan fee schedules were influenced by a resource-based relative value scale (RBRVS). There are deviations from the basic RBRVS methodology, so taking the average of observed prices in the market for each procedure is one measure used for capturing the typical "resources" used for a procedure.

$$SUI_{d,t} = \frac{p_{d,0} \cdot q_{d,t}}{c_{d,0}}$$

which holds the price of services fixed while allowing the utilization of services to vary. Note that there is a precise relationship between these three indexes that is described by the following decomposition:

$$MCE_{d,t} = SPI_{d,t} + SUI_{d,t} + \frac{(q_{d,t} - q_{d,0})(p_{d,t} - p_{d,0})}{c_{d,0}} - \frac{p_{d,0} \cdot q_{d,0}}{c_{d,0}}$$

Here the MCE index is equal to the service price index, $SPI_{d,t}$, plus the service utilization index, $SUI_{d,t}$, plus a cross term, $\frac{(q_{d,t} - q_{d,0})(p_{d,t} - p_{d,0})}{c_{d,0}}$, and subtracting $\frac{p_{d,0} \cdot q_{d,0}}{c_{d,0}}$ (which is close to 1). The cross term accounts for joint changes in both price vectors and utilization vectors and, in practice, the term is near zero. In the case where there are very few changes in utilization over time, $SUI_{d,t}$ is fixed near 1, then the $MCE_{d,t}$ will entirely be determined by service prices. Similarly, if there are very few changes in service prices over time, $SPI_{d,t}$, is near 1, and the $MCE_{d,t}$ will entirely be determined by utilization.

3 Data

We use retrospective claims data for a sample of commercially-insured patients from the MarketScan[®] Research Database from Thomson Reuters. The specific claims data used is the “Commercial Claims and Encounters Database” which contains data from the employer and health plan sources containing medical and drug data for several million commercially-insured individuals, including employees, their spouses, and dependents. Each observation in the data corresponds to a line item in an “explanation of benefits” form; therefore each claim can consist of many records and each encounter can consist of many claims.

We use a sample of enrollees that are not in capitated plans from the MarketScan database for the years 2003 to 2007. We also limit our sample to enrollees with drug benefits because drug purchases will not be observed for individuals without drug coverage. The MarketScan database tracks claims from all providers using a nationwide convenience sample of enrollees. Each enrollee has a unique identifier and includes age,

sex and region information which may be used when calculating patient weights. All claims have been paid and adjudicated.⁸

The claims data has been processed using the Symmetry grouper from Ingenix. The grouper assigns each claim to a particular Episode Treatment Group (ETG) disease category.⁹ The grouper uses a proprietary algorithm, based on clinical knowledge, that is applied to the claims data to assign each record to a clinically homogenous episode. The episode grouper allocates all spending from individual claim records to a distinct condition; the grouper also uses other information on the claim (for example, procedures) and information from the patient’s history to allocate the spending. An advantage of using the grouper is that it can use patients’ medical history to assign diseases to drug claims, which typically do not provide a diagnosis. However, these algorithms are also considered a “black box” in the sense that they rely entirely on the expertise of those that developed the grouper software.

3.1 Service Price, Utilization, and Episodes

The number of episodes is a simple count of the total number of episodes of a medical disease that end in the sample period.¹⁰ Total episode expenditures are measured as the total dollar amount received by all providers for the services used to treat an episode of a specific disease (including both out-of-pocket payments and amounts paid by insurance firms).

Service utilization measures were created for each type of service based on the definition of a service within that service type. The service type categories are inpatient

⁸Additional details about the data and the grouper used in this paper are in Dunn et al. (2010).

⁹The ETG grouper allocates each record into one of over 500 disease groups. To ensure that we observe full episodes, we limit the sample to those enrollees that have a full year of continuous enrollment. In addition, we require that enrollees have one year of enrollment in the prior year and one year of enrollment in the following year to make sure that episodes occurring at the beginning or the end of a year are not truncated. This may be an overly conservative constraint on the sample of enrollees, and we are currently working on examining the sensitivity of our analysis to alternative assumptions on enrollment.

¹⁰For an episode to fall into the sample, the episode must end in the 2006 or 2007 year of the data. Episodes records that begin in 2005 are included in this study, while episodes that begin in 2007 and end in 2008 are not included.

hospital, outpatient hospital, physician, prescription drug, and other. Using the definitions of the unit of service for each service type, the price of the service is calculated as the total expenditures for a particular disease and service category, divided by the quantity of services performed for that disease and service category. Furthermore, service utilization for a particular category is defined as the quantity of services divided by the total number of episodes for a particular disease. Below is a listing of the service types and how the quantity of services is measured.

Physician Visits - The physician visits are based on procedures performed in a physician's office. We assign a measure comparable to a RVU for each procedure performed by the physician for that office visit. Specifically, for each CPT and modifier, we calculate a relative value unit by computing the simple average fee for that procedure performed in an office setting. The total amount of services performed in an office is calculated by summing over these calculated "RVU" units. Note that there is a simple interpretation of these amounts. For example, if the fees are the same as the average computed in our sample, then the total cost of office visit divided by the "amount" of the visit will be equal to 1.¹¹

Hospital Inpatient - Inpatient hospital stays consist of both facility fees paid to the hospital, but also fees paid to the physician. For the portion of fees paid to the hospital, the amount of services is measured as the average dollar amount for an inpatient stay for the observed disease. For the portion of fees paid to the physician, we assign a RVU in the same way that we calculate an RVU in an office setting. However, we average over procedure prices in an inpatient setting. The total amount of services performed in an inpatient setting is calculated by adding the physician and facility amounts.¹²

Hospital Outpatient - Outpatient hospital visits are calculated in an identical fashion to the inpatient hospital visits. That is, the facility amount is calculated based on the average outpatient visit for that disease, and the doctors portion of the total amount is

¹¹Although procedure codes are observed for 98 percent of physician office claim lines, in those cases that we don't observe a procedure code we calculate the average price for a missing procedure code for patients with a particular disease. The results of the paper do not change substantially if those claim lines missing procedure codes are dropped from the analysis.

¹²As an alternative, we have also examined changing this definition to consider the facility price per inpatient day. The results do not change significantly.

calculated based on the average payment for the procedure codes.

Prescription drugs - The amount of the prescription drug varies based on the molecule, the number of pills in the bottle, the strength of the drug, and the manufacturer. To capture these differences, we calculate the average price for each NDC code, since each prescription is given a unique NDC code. The average price for each NDC code represents the amount of the service used. If the expenditure on a prescription is greater than this amount, it suggests that prices have increased.¹³ There are two distinct categories for drugs, branded and generic.

All other - The other category primarily includes ambulatory care, independent labs, and emergency room visits. For these services, the amount of each category is measured as the average cost for a visit to that particular place of service, for example, the average cost of an ambulatory care visit to treat ischemic heart disease. For cases where procedure codes are available, we use the average cost of that procedure code for that place of service.

There are a few additional points to note. A small fraction of the procedures (less than 5% of the claims observations for non-facility claim lines) are missing procedure codes. For these procedures we take the average price of the missing procedure codes for that service and disease type.¹⁴

3.2 Population Weights and Samples

In an attempt to make the MarketScan convenience sample more representative, we apply post-stratification population weights. Some of our estimates apply *fixed demographic weights* that hold the age and location distribution constant. Fixed population weights are important if one is interested in isolating the performance of the medical care sector, rather than looking at the effects of an aging population on expenditures. As one example, we construct weights based on the four regions, age and sex of the individuals, so

¹³An 11 digit National Drug Code (NDC) uniquely identifies the manufacturer, the strength, dosage, formulation, package size, and type of package.

¹⁴The results presented here do not change when alternative methods for calculating utilization are used. For instance, we obtain similar results when we drop procedures that are missing procedure codes.

that when the weights are applied, the population distribution corresponds to the U.S. 2007 population in each year. The population estimates are specific to the privately insured population below 65, where the estimates are from the Current Population Survey (CPS).

In addition to the broadly defined area that apply regional weights, we also look at a more finely defined geographic area that fixes the population at the county level. Checking the estimates by applying county weights may be important, given that prior research has demonstrated significant variation in medical care service prices and utilization across markets, even within a region (See Dunn, Shapiro, and Liebman (2012)). When applying county weights, we use only those counties where we observe at least 2000 individuals in the sample in each year. The weights are applied so that every county included in the study has a age distribution identical to the 2007 U.S. population. Each county contributes to the U.S. and total estimates in proportion to the county's population.

While many researchers may be interested in the estimates from the fixed demographic weights, we also apply alternative population weights that match the changing population characteristics in each year. These weights are based on the location, age, and sex of individuals, so that the change in the weighted characteristics of the sample match the actual change in the population characteristics. We will refer to these as *changing demographic weights*. These weights will be important when trying to benchmark our spending estimates to other national estimates of health expenditure growth, such as commercial premium growth rates. Contrasting these estimates with the fixed population weights also helps us to better understanding how the changes in the demographics of the population may impact spending growth figures. As discussed in greater detail in Dunn, Liebman and Shapiro (2012a), the difference between the changing demographic weights and the fixed demographic weights may be used to isolate the contribution of the changing demographics on the expenditure growth estimates.

In addition to the application of different population weights, we also explore alternative subsamples in the MarketScan data. One concern with the MarketScan data is that the data contributors are changing over time and, more specifically, the overall sample is growing considerably. For this reason, we study an alternative sample that focuses on a fixed set of data contributors. That is, those insurers and employers that contribute

to the MarketScan data are selected if they contribute to the database in each year of the sample. When exploring this alternative sample, we also explore the use of different population weights on this subsample.

In all of our analysis, we exclude individuals that are in capitated plans and those that do not have drug benefits. These restrictions are important, since we have incomplete spending information on these individuals.

3.3 Descriptive Statistics

To better understand how the unweighted sample compares to the weighted sample, it is useful to compare the demographic characteristics of the actual commercial population with the unweighted MarketScan data. Table 1 reports a number of descriptive statistics for the commercial population and unweighted MarketScan data. The first thing to note about the MarketScan data is that the sample size is large and grows very rapidly over the period of study, with 7.0 million enrollees in the data in 2003 and 13.1 million in 2007.

The sample size is a major advantage of using the MarketScan data; compared to a nationally representative survey, such as the Medical Expenditure Panel Survey, which is a survey containing the response of just 30,000 individuals in each year. Since many important and costly medical conditions are relatively rare and heterogeneous (such as cancers or heart attacks), a survey that has just 30,000 individuals may not be sufficient to be representative of the disease costs of many conditions, not to mention the typical challenges and limitations that arise when conducting consumer surveys. Despite the sample size, it is important to keep in mind that the MarketScan data is still a limited sample and represents just a fraction of the overall commercial population, between 3.8 and 5 percent of the total commercial population, over the period of study. Therefore, ensuring that the population characteristics reflect the national population of the commercially insured individuals may be vital for obtaining reasonable estimates.

The distribution of the demographics in the MarketScan data is roughly similar to the commercial population estimates, based on the age and sex distributions. However, it is important to highlight a few key differences that potentially have an impact on our expenditure decomposition estimates. First, the location of individuals in the com-

mercial population and the unweighted population are quite distinct. The unweighted MarketScan data disproportionately draws enrollees from the South, with over 45 percent of the sample coming from that region, compared to 34 percent for the actual population. Second, the average age in the unweighted MarketScan data is higher than in the commercial population by two years, which could potentially lead to an upward bias on the unweighted data when looking at estimates of expenditure per capita. Third, the trends in the average age are distinct. The commercial population reflects the aging population in the U.S., with the average age growing by 0.6 years, while the average age in the unweighted MarketScan data actually declines by 0.4 years over the period, leading to a total difference in the age growth of 1 year. This may lead to a downward bias in spending growth when using the unweighted MarketScan data.

Table 1. Population Levels and Distributions for the Commercial Population and Unweighted MarketScan Data

	Commercial Population		Unweighted MarketScan	
	2003	2007	2003	2007
of Enrollees (millions)	180.6	182.5	7.0	13.1
<u>Gender</u>				
Male	49.5%	49.6%	47.9%	48.5%
Female	50.5%	50.4%	52.1%	51.5%
<u>Age</u>				
0 to 17	27.3%	26.3%	25.3%	26.4%
18 to 24	9.6%	9.6%	7.8%	7.6%
25 to 34	14.5%	14.7%	12.1%	11.9%
35 to 54	36.3%	35.6%	38.3%	37.6%
55 and over	12.2%	13.8%	16.6%	16.5%
Mean Age	32.3	32.9	34.5	34.1
<u>Region</u>				
NE	19.3%	18.9%	11.5%	12.0%
MW	24.4%	23.7%	29.4%	24.3%
S	34.0%	34.3%	45.5%	48.0%
W	22.3%	23.1%	13.7%	15.7%

Notes: Commercial population estimates are taken from the CPS estimates of the commercially insured population, while unweighted MarketScan estimates are enrollee counts from the MarketScan data for individuals in non-capitated plans with drug benefits that are enrolled for the entire year.

Table 1 shows that the demographics of the MarketScan sample differ in some ways

from that of the commercial population estimates. Next, we look at some basic per capita expenditure and expenditure growth estimates in Table 2, where we contrast estimates when population weights are applied (the first three columns) to estimates when population weights are not applied (the last three columns). The weights applied here are changing population weights, which allow the population distribution to reflect the characteristics of the actual population in each year. These expenditures are broken into Major Practice Categories (MPC), where they are listed in order of 2003 per capita expenditures, and the bottom row shows per capita spending. We see average expenditures tend to be greater for the unweighted population, with average per capita expenditures in the unweighted sample about 12 percent greater in the unweighted sample than in the weighted sample in 2003. We also see that expenditure growth is over 6 percent greater in the weighted sample compared to the unweighted sample. The higher growth rates appear for some important disease categories that tend to increase with age. The category that stands out the most is the growth rate for the cardiology conditions that are 10 percentage points greater in the weighted sample relative to the unweighted sample. However, more generally, the growth rates for most condition categories are greater in the weighted sample, relative to the unweighted sample.¹⁵

¹⁵In both the weighted and unweighted samples, about 13 percent of expenditures are not assigned to any ETG disease category. This includes screening for diseases and other records that cannot be assigned a disease category. Those claims that are not assigned disease categories are removed from our analysis. In most of the analysis we apply severity adjustment, which increases the share of ungrouped expenditures to 20 percent, since some episodes may be assigned a disease but not a severity level. As we will show later, similar results are found whether severity adjustment is applied or not, so removing those ungrouped claims that cannot be severity adjusted has little effect on our results. See Dunn, Liebman, and Shapiro (2012b) for additional discussion regarding disease classification.

Table 2. Total Annual Per Capita Expenditures, Shares, and Growth by Major Practice Category - Weighted and Unweighted

Major Practice Category	Changing Commercial Population Weights			Unweighted		
	2003 Spending Per Capita	2003 Share of Spending	Growth 2003-07	2003 Spending Per Capita	2003 Share of Spending	Growth 2003-07
Orthopedics & rheumatology	\$415.34	16.51%	1.32	\$460.72	16.41%	1.26
Cardiology	\$304.21	12.09%	1.12	\$382.64	13.63%	1.02
Gastroenterology	\$229.27	9.11%	1.30	\$260.39	9.27%	1.23
Gynecology	\$179.26	7.13%	1.21	\$201.71	7.18%	1.14
Endocrinology	\$166.12	6.60%	1.42	\$193.35	6.89%	1.35
Otolaryngology	\$159.57	6.34%	1.16	\$166.47	5.93%	1.16
Neurology	\$147.50	5.86%	1.30	\$161.03	5.74%	1.24
Pulmonology	\$121.08	4.81%	1.16	\$138.85	4.95%	1.10
Psychiatry	\$118.58	4.71%	1.23	\$115.23	4.10%	1.23
Dermatology	\$113.49	4.51%	1.30	\$116.08	4.13%	1.29
Obstetrics	\$111.29	4.42%	1.22	\$96.51	3.44%	1.18
Urology	\$90.53	3.60%	1.24	\$103.81	3.70%	1.17
Hematology	\$61.79	2.46%	1.34	\$67.65	2.41%	1.29
Hepatology	\$60.85	2.42%	1.08	\$69.35	2.47%	1.00
Preventive & administrative	\$57.25	2.28%	1.73	\$60.96	2.17%	1.70
Ophthalmology	\$40.00	1.59%	1.25	\$46.24	1.65%	1.18
Infectious diseases	\$35.45	1.41%	1.37	\$38.61	1.38%	1.35
Nephrology	\$34.32	1.36%	1.26	\$40.96	1.46%	1.20
Neonatology	\$25.26	1.00%	1.39	\$40.92	1.46%	1.30
Isolated signs & symptoms	\$18.67	0.74%	1.13	\$19.20	0.68%	1.11
Late effects, environmental trauma & poisonings	\$13.72	0.55%	1.24	\$14.83	0.53%	1.18
Chemical dependency	\$11.80	0.47%	1.58	\$12.09	0.43%	1.50
Total	\$2515.36	100.00%	1.265	\$2807.60	100.00%	1.204

Notes: Commercial population per capita spending estimates by disease are calculated by multiplying disease expenditures by changing population weights, summing over spending, and then dividing by the full population. Unweighted MarketScan estimates are per capita expenditure estimates by disease category.

To check whether the spending growth rates are in a reasonable range, we compare the growth rates above with benchmark growth rates from other sources. When making this comparison, one should keep in mind that the estimate that we offer from the MarketScan commercial claims data set are unique and independent of the other sources, so we should not expect the benchmark estimates to precisely match the estimates that we are computing. However, the underlying factors affecting growth here should correspond to expenditures by private insurers from the National Health Expenditure Accounts (NHEA) and estimates of premium growth rates. We find that the growth rate in our weighted estimates match very closely with NHEA expenditures for private insurers and premium growth rates from the MEPS data, which are arguably the two most relevant data sources in this study. Although the Kaiser Health Benefits Survey offers premium

estimates, the estimate is based on a much smaller sample size of around 2,000 firms, compared with a sample of about 40,000 firms in the MEPS-IC data. The other spending growth benchmarks (i.e., the NHEA - all categories and BEA - all categories) include other sources of payment, such as Medicare and Medicaid. Overall, the estimates from the weighted sample fall in a reasonable range to these benchmark estimates, while the unweighted sample falls a few percentage points below these benchmarks, which lends greater confidence to our weighted estimates. Matching expenditure growth rates to relevant expenditure benchmarks helps to bolster the case for applying population weights. In the next section, we focus on expenditure growth decompositions that apply age, sex and location weights, although we continue to contrast our results with unweighted estimates.

Table 3. Per Capita Spending and Premium Benchmarks

	2003-2007
<u>Spending Growth Benchmarks</u>	
NHEA - Private Insurance	1.261
NHEA - All Categories	1.246
BEA - All Categories	1.222
<u>Premium Growth</u>	
MEPS - Insurance Component	1.265
Kaiser Employer Health Benefit Survey	1.332

Notes: The 2007 premium figures for the MEPS - Insurance Component estimates are imputed from 2006 and 2008 estimates. Both the Kaiser premium estimates and the MEPS-IC estimates assume that 47 percent of employees are enrolled in a single plan and the remainder in family plans. The percentage was derived from the MEPS-Insurance Component 2003 estimates.

4 Results

Table 4 compares the expenditure growth decomposition using the unweighted data to weighted data that allows the population distribution to reflect national population estimates and changes in the distribution of the population. We see that the weighted ECI grows at 26.5 percent, about 6 percentage points slower using the unweighted estimates. This expenditure growth difference reflects per capita spending changes that are also reported at the bottom of Table 2. However, in Table 4 we observe the sources of the

expenditure growth differences in greater detail. The faster prevalence growth in the weighted data accounts for 2 percent of the difference in the ECI index and faster MCE growth (along with a cross-term difference) accounting for the remainder. Although we observe some differences in the weighted and unweighted estimates, there are also some interesting common patterns. In both sets of estimates we see utilization per episode remaining relatively flat or falling slightly, while expenditure growth is primarily driven by an increase in PREV and SPI.

Table 4. Decomposition of Growth Rates 2003-07 - Weighted and Unweighted

<u>Weighted - Changing Comm. Population</u>					
	ECI	PREV	MCE	SPI	SUI
2003	1.000	1.000	1.000	1.000	1.000
2004	1.070	1.039	1.033	1.028	1.007
2005	1.149	1.082	1.065	1.063	1.005
2006	1.211	1.106	1.100	1.104	1.003
2007	1.265	1.144	1.114	1.134	0.995

<u>Unweighted</u>					
	ECI	PREV	MCE	SPI	SUI
2003	1.000	1.000	1.000	1.000	1.000
2004	1.066	1.042	1.026	1.024	1.004
2005	1.130	1.080	1.051	1.059	0.996
2006	1.159	1.085	1.074	1.095	0.988
2007	1.204	1.123	1.080	1.120	0.977

Notes: Estimates are computed using ETG severity adjustments. To save space, the cross-terms from the different components of the decomposition are not reported.

Table 5 reports decomposition growth figures for some additional weighting strategies and samples, where each row of the table shows a distinct estimate of the growth decomposition for the 2003-07 period. The first two rows repeat the unweighted and weighted estimates reported in Table 4, but only shows the values of the index for 2007. The third row of the table holds the age, sex and regional distribution constant to 2007 levels. As mentioned previously, many researchers may be interested in growth estimates that hold demographic factors constant to better isolate the trends in treatment patterns for similar populations. In Dunn, Liebman, and Shapiro (2012a), this expenditure estimate is called the Demographically Adjusted Expenditure per Capita Index (DECI).

The difference in spending between the changing population weights and the fixed population weights is around 3.5 percentage points, with most of the differences in growth being driven by prevalence, as expected, since disease prevalence increases with age for many health conditions.¹⁶ One concern with the application of regional weights, is that focusing on regional weights may not capture the trends that are observed at a finer geographic level. Row 4 of Table 5 reports estimates that fix the population distribution for each county to 2007 levels, rather than fixing the regional population.¹⁷ The estimates applying county weights are nearly identical to the regional weight in every dimension, suggesting that the application of either county or regional weights may be appropriate for the study of the MarketScan claims data.

As mentioned previously, one might be concerned that changes in data contributors in the MarketScan data may have a measurable impact on our study in ways that are difficult to correct for. As an alternative estimate, we next focus on a subsample of the data that holds the data contributors fixed (including only those data contributors that are in the sample from 2003 to 2007). These estimates are shown in the bottom half of Table 5. The qualitative estimates of this subsample are quite similar to the full sample that applies population weights. The key difference is that the growth rate using the fixed contributors from the MCE is three percentage points higher by 2007, and the prevalence index is two percentage points lower over this period. Although these differences are notable, we calculate that based on CAGR estimates, the difference in the various components of the decomposition are less than 0.007 percentage points for each component of the decomposition across the two samples. The same aggregate patterns hold for these alternative sets of estimates. All of the estimates imply that prevalence and service price growth are the key contributors to expenditure growth.

¹⁶In the paper Dunn, Liebman, and Shapiro (2012a) the ECI reported in Table 5 that applies the fixed population weight is referred to as the demographically adjusted expenditure per capita index or DECI.

¹⁷Recall that the sample is slightly different than the regional estimates, since we only keep those estimates that have at least 2000 enrollees in each year of the sample.

Table 5. Decomposition of Growth Rates from 2003-07 for Different Weights and Samples

	ECI	PREV	MCE	SPI	SUI
<u>Full MarketScan Sample - Changing Contributors</u>					
Unweighted	1.204	1.123	1.080	1.114	0.971
Changing Regional Population Weights	1.265	1.144	1.114	1.134	0.995
Regional Weights, Fixed Demog.	1.231	1.118	1.110	1.132	0.993
County Weights, Fixed Demog.	1.235	1.118	1.110	1.130	0.996
<u>Fixed Contributors</u>					
Unweighted	1.247	1.114	1.124	1.153	0.989
Changing Regional Population Weights	1.283	1.128	1.143	1.160	0.998
Regional Weights, Fixed Demog.	1.250	1.103	1.139	1.159	0.996
County Weights, Fixed Demog.	1.231	1.093	1.132	1.145	1.004

Notes: Estimates are computed using ETG severity adjustments. To save space, the cross-terms from the different components of the decomposition are not reported.

Focusing on the estimates that apply changing population weights, note that the fixed contributor sample appears in a reasonable range to our national benchmark spending growth estimates reported in Table 3, but about 1.8 percentage points higher than growth from the NHEA expenditure estimates. Another important set of national statistics that we may benchmark against are service price measures. The key benchmark price estimate is the BEA GDP Health Services Price deflator, which shows a growth rate of 13.7 for the period of study. The SPI estimates that apply population weights centers around this figure with growth rates ranging from 13.0 percent to 16.0 percent.¹⁸

Table 5 presents a range of estimates, but researchers should be aware of the trade-off to using the fixed contributor sample is that the sample size shrinks significantly in each year, with about 1 million fewer enrollees in 2003 (losing about 20 percent of the sample) compared to the full sample, and 5 million fewer enrollees in 2007. There are trade-offs when choosing between the full sample or the sample with fixed data contributors. The full sample contains more enrollees and more data contributors in each year, but may be influenced by changes in the type of data contributors across years. On the other hand,

¹⁸It should be highlighted that an important difference between our estimates and those of the BEA Health Services price deflator, is that our estimates only contains private health insurance claims, while the GDP Health Services price deflator includes information on payments from all types of payers (e.g. private, Medicare and Medicaid.). This could account for some of the difference.

the sample with the fixed data contributors is a smaller sample size and may be more influenced by particular data contributors. Given this trade-off, in our work we look at both estimates and search for consistent patterns across each.

The results presented here may also be used to look at whether spending is growing due to treated prevalence or if spending is primarily rising due to growth in the MCE. Looking across the various estimates of Table 5 that hold population fixed, the MCE growth accounts for between 46 percent to 55 percent of the expenditure growth, with prevalence (along with a cross term) accounting for the remainder. Therefore, growing prevalence and expenditure per case, similarly contribute to overall spending growth. To compare these results to the analysis of Roehrig and Rousseau (2009), who perform a similar calculation, we first need to adjust spending and prices for overall inflation growth (which was 11.5 percent over this period). After making this adjustment the “real” share of growth attributed to expenditure per case would actually range from 17.5 percent to -4.0 percent.¹⁹ This result stands in stark contrast to the findings of Roehrig and Rousseau (2011) that use the Medical Expenditure Panel Survey data and find after adjusting for inflation that about 75 percent of spending growth may be attributed to expenditures per episode. Although Roehrig and Rousseau (2011) study a distinct period, 1996 to 2006, the work by Aizcorbe et al. (2011) that look at a more recent period (2001 to 2005) finds similarly rapid growth in expenditure per episode using the MEPS data. Additional work is necessary to better understand this discrepancy across the two data sources.

4.1 Heterogenous Trends in the Components of Expenditure Growth

The previous section focused on the aggregate trends in disease expenditure growth. However, there are differences in the growth rate for many disease conditions and their components that is reported in Table A1 in the appendix. Differences in growth rates across diseases is discussed and analyzed in greater detail in Dunn, Liebman and Shapiro

¹⁹The negative real growth arises because the expenditure per episode is rising slower than inflation for some estimates

(2012a).²⁰ For instance, in that paper we find unique trends for different disease categories, showing that utilization for cardiology conditions are falling on average, while the prevalence of endocrinology conditions (like diabetes or high cholesterol) is growing rapidly. Another dimension in which growth rates could potentially differ is by the age group of individuals, which is particularly relevant for the application of population weights. For instance, if older individuals are over-represented in the data, then trends will be more influenced by those diseases that tend to afflict older individuals, such as cardiology conditions.²¹ However, if the general trends in the components of spending growth are common across age categories, then applying population weights will have less of an impact on the overall trend.

More generally, looking at the components of spending growth by age is informative, since it offers a check on whether the broad trends we observe are in Table 5 are true for all segments of the population, or if there is a particular segment of the population that is driving spending growth that warrants further analysis. Table 6 reports trends in the different components of expenditure growth for the period 2003 to 2007 across different age groups, with each row representing a different age group. The top half of Table 6 reports the full sample using a fixed population, regional weights. Below the results with the full sample, Table 6 shows the components of spending growth using fixed contributors and a fixed population, regional weights. The left two columns of the table reports the population share in each age group, along with the expenditure share of that population.

The spending growth patterns for the different age groups in Table 6 are similar to the patterns observed in Table 5, for both the full sample and the fixed contributor sample. Utilization growth is relatively flat, and prevalence growth and price growth are the primary drivers of per capita expenditure growth for each age category. The common pattern in the components of expenditure growth is especially striking for age categories of 35 and above, which account for over 70 percent of the spending. Although there are

²⁰Table A1 of the appendix is taken from Dunn, Liebman and Shapiro (2012a) and reports the components of spending growth for different disease categories. The heterogeneity in disease trends, reported in Table A1 helps demonstrate the wide differences in the magnitudes of disease expenditures.

²¹Clearly the type of diseases treated change with the age of the individual as can be seen in Table A2 in the appendix, which charts expenditure shares as the population ages.

many similarities in growth patterns, there are some noteworthy differences. First, those age categories below 25 tend to experience faster overall spending growth relative to the those over 25, which appears to be caused by both higher service price growth and higher utilization growth. Second, the SUI is growing for those below 35, but declining for those in age categories of 35 and above. Therefore, it appears that younger populations are receiving more treatments for the same disease, relative to older individuals. This may partly reflect that younger individuals spend significantly less on cardiology conditions, a condition category that has seen a decline in utilization per episode.²²

Table 6. Components of Spending Growth for Different Age Categories Applying Regional Fixed Demographic Weights: 2003-07

Full MarketScan Sample - Changing Contributors							
Age Group	Population Share	Spending Share	ECI	PREV	MCE	SPI	SUI
0 to17	26%	12%	1.30	1.09	1.20	1.17	1.04
18 to 24	10%	5%	1.33	1.16	1.15	1.16	1.12
25 to 34	15%	12%	1.22	1.10	1.11	1.11	1.02
35 to 44	17%	17%	1.24	1.12	1.11	1.13	1.00
45 to 54	18%	26%	1.21	1.11	1.09	1.13	0.98
55 to 64	14%	29%	1.20	1.13	1.08	1.13	0.97

Fixed Contributors							
Age Group	Population Share	Spending Share	ECI	PREV	MCE	SPI	SUI
0 to17	26%	12%	1.28	1.06	1.21	1.19	1.04
18 to 24	10%	5%	1.33	1.13	1.19	1.41	1.04
25 to 34	15%	12%	1.21	1.07	1.14	1.16	1.02
35 to 44	17%	16%	1.24	1.10	1.14	1.20	1.00
45 to 54	18%	26%	1.24	1.11	1.13	1.16	0.99
55 to 64	14%	29%	1.25	1.13	1.11	1.15	0.98

Notes: The spending share is reported based on all 5 years of data. For both estimates, the population distribution is held fixed to 2007 levels, so regional shifts in population distribution have no effect on these estimates. These trends are computed based on the disease conditions for each age group. Many diseases are not observed across different age groupers.

As with trends in age, similar issues may arise when considering differences in regional growth rates. Some regions may drive growth in different ways relative to others, leading to a bias in national estimates if a particular region is over or under-weighted. Analogous to the estimates presented in Table 6, the expenditure growth rates and its components for each of the four regions are reported in Table 7 using both the full sample and the sample with fixed contributors. There are a number of interesting patterns. First, growth in overall spending, as reflected in the ECI, is quite different across regions, ranging from

²²See Table A2 that shows the expenditure share for each disease category by age group.

around 38 percent growth in the Northeast to 14 percent growth in the South. The lower growth rate in the South appears to be due to both falling utilization levels and lower price growth, although prevalence growth is similar to or larger than the other regions.

Table 7 also shows that the components of growth in the South depend greatly on whether the full sample is used or only the fixed contributors. Prevalence growth differs by 7 percentage points across the full sample and fixed contributor sample, while the MCE growth differs by 6 percentage points. This suggests at least a couple of possibilities. Either the fixed sample is not representative of the population in the South or there is a data contributor entering in the South which greatly affects prevalence and utilization. Using the sample with fixed contributors, the growth rate in the South appears more in line with the other regions and the service price growth rate in the South is closer to the benchmark price growth levels. These trends in the South using the full sample, appear far out of line with the other regional estimates and benchmark estimates, suggesting that the sample of fixed data contributors may produce more plausible figures. As a result, this is the focus of Dunn, Shapiro, and Liebman (2012). Although this is our current understanding of the data, it is difficult to be sure whether there is an actual bias without additional information.

Table 7. Components of Spending Growth for Different Regions: 2003-07

Full MarketScan Sample - Changing Contributors							
	Pop. Share	Spending Share	ECI	PREV	MCE	SPI	SUI
NE	19%	18%	1.38	1.10	1.26	1.24	1.04
MW	24%	24%	1.29	1.12	1.16	1.18	1.00
S	34%	35%	1.14	1.16	0.99	1.05	0.96
W	23%	23%	1.22	1.06	1.16	1.16	1.02

Fixed Contributors							
	Pop. Share	Spending Share	ECI	PREV	MCE	SPI	SUI
NE	19%	17%	1.34	1.11	1.21	1.21	1.04
MW	24%	24%	1.29	1.11	1.17	1.20	0.99
S	34%	36%	1.18	1.11	1.07	1.10	0.99
W	23%	23%	1.25	1.08	1.16	1.19	1.00

5 Some Alternative Approaches

Aside from the application of population weights, there are a number of other issues that researchers should keep in mind when studying expenditure growth. Here we focus briefly on two of these issues: (1) the classification of medical claims into disease episodes; and (2) using the panel structure of claims data.²³

5.1 Disease Classification

Throughout this paper, we have focused on a single approach for classifying medical claims into disease episodes (i.e., applying the ETG grouper with severity adjustment), but one may be concerned that a different classification strategy may have a large substantive impact on our analysis. Indeed, many research papers have proposed and applied a variety of strategies to classify medical claims into disease categories or disease episodes. In a companion piece to this paper, Dunn, Liebman, Shapiro, and Rittmueller (2012b), look at this issue in greater detail and explore the impact of numerous alternative classification strategies on the components of expenditure growth. They show that many of the key findings are similar across strategies and provide a range of estimates for disease expenditure growth. However, they primarily focus on a single weighting strategy. Similarly, in this study, it is difficult to tell if an alternative classification strategy may have a distinct effect on the estimates, depending on the weighting strategy that is applied.

To provide some range of estimates, we present results using a slightly different disease classification approach, ETG grouping without severity adjustment. Note that severity adjustment accounts for related complications, comorbidities and demographic factors that may influence the expected utilization of services needed to treat a condition of a particular severity. Therefore, removing severity adjustment produces more aggregate disease categories. Table 8 shows these results. There is a very clear and systematic effect from applying non-severity adjusted ETGs across all estimates: (1) the utilization

²³Although we highlight these two points, there are some additional robustness checks that we also looked at prior to reporting the estimates in this paper: (1) Removing outlier disease episodes and (2.) Focusing on the more frequently observed disease episodes (e.g. a minimum of 10,000 observed episodes in the data).

growth increases slightly (by about two percentage points); (2) the MCE grows by about two percentage points; and prevalence growth falls by about two percentage points. The likely reason for this difference is that there is a growth in the severity of illness within each broad ETG category, leading to more service utilization when collapsing across severity categories. Although we observe some differences, the main qualitative findings, remain unchanged.

Table 8. Decomposition of Growth Rates from 2003-07 for Different Weights and Samples: Not Severity Adjusted

	ECI	PREV	MCE	SPI	SUI
Full MarketScan Sample - Changing Contributors					
Unweighted	1.204	1.110	1.089	1.120	0.985
Changing Regional Population Weights	1.265	1.130	1.125	1.133	1.004
Regional Weights, 2007 Fixed Pop.	1.231	1.105	1.119	1.131	1.001
County Weights, 2007 Fixed Pop.	1.235	1.111	1.118	1.129	1.002
Fixed Contributors					
Unweighted	1.247	1.105	1.134	1.153	0.997
Changing Regional Population Weights	1.283	1.117	1.155	1.160	1.008
Regional Weights, 2007 Fixed Pop.	1.250	1.093	1.149	1.159	1.004
County Weights, 2007 Fixed Pop.	1.231	1.085	1.141	1.144	1.011

Notes: Estimates are computed using ETG without applying the severity adjustment. To save space, the cross-terms from the different components of the decomposition are not reported.

Since we are primarily interested in understanding the treatment for identical conditions, applying the severity adjustment is our preferred methodology.²⁴ It should also be noted that when applying a completely different methodology for grouping diseases, the MEG grouper, we find estimates that are consistent to those reported in Table 8. See Dunn, Liebman, Shapiro, and Rittmueller (2012b) for additional discussion.

5.2 Panel Analysis: Death and Selection Issues

The MarketScan data is a panel data set that tracks individuals over multiple years. This feature of the data is shared by other commercial claims data sets, and there are

²⁴As an additional check, we also look at alternative weights using the MEG grouper, and we obtain similar results.

potentially great advantages from exploiting the panel aspects of the data to study health expenditure growth, where the health condition of each individual may have unique idiosyncrasies that are specific to that individual.

Although the panel aspect of the data appears potentially useful, in this subsection we show how using the panel dimension of the data may actually lead to significant bias. The key problem is that in the first year that an individual enters the panel, we know that those individuals are selected to live at least one more year. In contrast, for the last year that an individual is in the panel, it is not clear whether the individual will be in the data the following year or not, and they could potentially exit the sample through death. In other words, the first year of the panel contains only those individuals that live one additional year, while the last year of the panel includes some individuals that may die in the following year. This fact, combined with the knowledge that the health care for individuals is typically much more expensive in the last year of life, leads to a potentially large and positive selection bias in expenditure growth.

To demonstrate this point, we estimate spending levels for two populations of individuals in 2006: (1) the continuing sample, which includes those individuals that are in the data for the additional full year of 2006; and (2) the exiting sample, which includes those individuals that do not have full enrollment in the following year. We focus on population weighted spending estimates, so the total population and age distribution of the two samples is identical.²⁵ We find that the per capita spending for the exiting sample is 21 percent higher than the spending for the continuing sample. The allocation of spending also appears distinct. Specifically, for the exiting sample, a greater share of spending is allocated to potentially fatal diseases. For example, the exiting sample allocates 9.9 percent of spending to malignant cancers, while the continuing sample allocates 5.8 percent of spending. We also find more spending on severe conditions in the exiting sample with 14.9 percent of spending on severity 3 or severity 4 conditions, compared with 11.9 percent of spending on these conditions for the continuing sample.

Further investigation reveals that the difference between the exiting sample and the continuing sample may have major effects on expenditure growth, leading to a large

²⁵In this analysis we do not study 64 year olds, since 64 year olds typically enter the Medicare program and leave private insurance when they turn 65.

over-statement of the expenditure index and its components. Therefore, researchers studying expenditure growth should be mindful of this selection issue when using panel data. Despite these issues, there are potentially significant gains in our understanding of expenditure growth from exploiting the panel dimension of these data, but more research is necessary to better understand how to account for potential selection bias.

6 Conclusion

Researchers examining spending growth in the commercial sector often use convenience claims data, which may not be representative of the full commercially insured population. In this paper, we analyze the MarketScan commercial claims data and apply various weighting strategies to correct for the potential non-representative aspects of the data. In general, we find that spending growth is primarily driven by price growth and a growth in prevalence, with utilization per episode staying relatively flat. Although this main qualitative finding holds, even when no weights are applied; we find that the application of population weights to reflect the population distribution of the U.S., produces spending growth figures that are more aligned with other benchmark estimates of price and expenditure growth from national statistics. In general, the results in this paper complement those reported in Dunn, Liebman, and Shapiro (2012a) by showing how alternative weighting strategies impact key results and trends.

To further understand the components of spending growth and how they may be influenced by population weights, we look at growth rates for different subpopulations. In particular, we look at growth rates by age group and by geography. We find that a similar general pattern of spending growth holds across age groups, but we also find some interesting differences across age groups. Spending growth appears to be increasing most rapidly for the population below 25, primarily due to higher service price growth and utilization growth. Prevalence appears to be increasing most rapidly for the population over the age of 18. Looking at regional growth differences, we find that growth rates are slower in the South, due to both lower price and lower utilization growth.

Overall, we recommend applying population weights for studying expenditure growth in all circumstances when attempting to make national projections using a convenience

sample. However, another important consideration is the changing mix of data contributors, which could introduce a bias. Comparing estimates when the data contributors are fixed, to those estimates when the data contributors vary over time produce similar estimates; although prevalence growth rates tend to be lower and price growth trends tend to be higher for the fixed contributor sample. There are trade-offs with using either the full or the fixed sample. In this study and in our related studies, even if we focus on one set of estimates, we examine estimates from both samples and search for consistent patterns across each.

There are a couple of important areas for future research. First, it would be useful to look at other convenience samples to see if we observe similar patterns using alternative data sources. Second, there are interesting panel aspects of these claims data that could potentially be useful for obtaining more precise estimates, but researchers must first figure out how to deal with the selection issue caused by the most unhealthy people potentially exiting the sample through death. Third, the paper here is entirely descriptive of the trends, but does little to explain the observed trends. Future research may benefit from trying to understand the underlying health and economic factors that may cause these observed differences and changes over time.

References

- [1] Aizcorbe, Ana and Nicole Nestoriak, (2011), “Changing Mix of Medical Care Services: Stylized Facts and Implications for Price Indexes”, *Journal of Health Economics*, 30(3) pgs 568-574.
- [2] Berndt, Ernst, David Cutler, Richard Frank, Zvi Griliches, Joseph Newhouse, and Jack Triplett, (2000), “Medical Care Prices and Output”, *Handbook of Health Economics*, Chapter 3.
- [3] Berndt, Ernst, Anupa Birb, Susan H. Busch, Richard G. Frank, and Sharon-Lise T. Normande “The Medical Treatment of Depression, 1991–1996: Productive Inefficiency, Expected Outcome Variations, and Price Indexes”, *Journal of Health Economics*, 21(3) pages 373-396.

- [4] Bundorf, Kate, Anne Royalty, and Laurence Baker, (2009), “Health Care Cost Growth Among the Privately Insured”, *Health Affairs*, September/October.
- [5] Committee on National Statistics, (2011), *Accounting for Health and Health Care: Approaches to Measuring the Sources and Costs of Their Improvement*, The National Academies Press, Washington, D.C.
- [6] Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler. “Are Medical Prices Declining? Evidence from Heart Attack Treatments”, *Quarterly Journal of Economics* 113 (1998) pgs 991-1024.
- [7] Dunn, Abe, Eli Liebman, Sarah Pack, and Adam Shapiro, (2012a), “Medical Care Price Indexes for Patients with Employer-Provided Insurance: Nationally-Representative Estimates from MarketScan Data”, *Health Services Research*, Forthcoming.
- [8] Dunn, Abe, Eli Liebman, and Adam Shapiro, (2012a), “Decomposing Medical-Care Expenditure Growth”, Bureau of Economic Analysis, *Working Paper*.
- [9] Dunn, Abe, Eli Liebman, and Adam Shapiro, (2012b), “Implications of Utilization Shifts on Medical-Care Price Measurement”, Bureau of Economic Analysis, *Working Paper*.
- [10] Dunn, Abe, Eli Liebman, Adam Shapiro, and Rittmueller, (2012b), “Defining Disease Episodes and the Effects on the Components of Expenditure Growth”, Bureau of Economic Analysis, *Working Paper*.
- [11] Dunn, Abe, Adam Shapiro, and Eli Liebman, (2012), “Geographic Variation in Commercial Medical-Care Expenditures: A Framework for Decomposing Price and Utilization”, Bureau of Economic Analysis, *Working Paper*.
- [12] Roehrig, Charles, George Miller, Craig Lake, Jenny Bryant, (2009), “National Health Spending By Medical Condition, 1996-2005”, *Health Affairs*, February - Web Exclusive.

- [13] Roehrig, Charles and David Rousseau, (2011), “The Growth in Cost Per Case Explains Far More of US Health Spending Increases Than Rising Disease Prevalence”, *Health Affairs*, 30(9) pgs 1657-1663.
- [14] Matthew D. Shapiro and David W. Wilcox, (1996), “Mismeasurement in the Consumer Price Index: An Evaluation,” in Ben S. Bernanke and Julio J. Rotemberg, Editors, NBER Macroeconomics Annual 1996, Volume 11 MIT Press
- [15] Thorpe, Kenneth, Curtis Florence, and Peter Joski, (2004), “Which Medical Conditions Account for the Rise in Health Care Spending?”, *Health Affairs*, August - Web Exclusive.

7 Appendix

Table A1. Sources of Growth 2003-07 - Fixed Demographics, Full Sample, Regional Weights

	ECI	PREV	MCE	SPI	SUI
Infectious diseases	1.338	1.421	1.162	1.087	1.081
Endocrinology	1.366	1.305	1.068	1.152	0.937
Hematology	1.298	1.125	1.152	1.196	0.976
Psychiatry	1.235	1.141	1.083	1.129	0.994
Chemical dependency	1.572	1.500	1.079	1.110	1.018
Neurology	1.275	1.103	1.159	1.189	0.983
Ophthalmology	1.190	1.148	1.036	1.084	0.965
Cardiology	1.054	1.028	1.019	1.120	0.922
Otolaryngology	1.165	1.057	1.104	1.124	1.006
Pulmonology	1.117	1.004	1.122	1.169	0.963
Gastroenterology	1.260	1.124	1.135	1.140	1.000
Hepatology	1.053	1.018	1.033	1.098	0.951
Nephrology	1.209	1.423	0.864	0.851	1.025
Urology	1.189	1.113	1.083	1.112	0.983
Obstetrics	1.227	1.061	1.158	1.119	1.038
Gynecology	1.194	1.024	1.162	1.147	1.014
Dermatology	1.283	1.103	1.164	1.154	1.023
Orthopedics & rheumatology	1.290	1.145	1.129	1.121	1.026
Neonatology	1.277	1.135	1.129	1.122	1.002
Preventive & administrative	1.715	1.362	1.261	1.134	1.111
Late effects, environmental trauma & poisonings	1.222	0.969	1.268	1.230	1.035
Isolated signs & symptoms	1.124	1.018	1.104	1.106	1.010

Table A2. Distribution of Spending by Age Group - Average 2003 to 2007

	0 to17	18 to 24	25 to 34	35 to 44	45 to 54	55 to 64
Spending Per Capita	\$1428	\$1587	\$2424	\$2802	\$3998	\$5899
	<u>Spending Share by Disease for Each Age Group</u>					
Infectious diseases	2%	1%	1%	2%	2%	1%
Endocrinology	4%	4%	5%	7%	8%	9%
Hematology	3%	3%	2%	2%	2%	3%
Psychiatry	9%	8%	5%	5%	4%	2%
Chemical dependency	0%	1%	1%	1%	1%	0%
Neurology	6%	7%	6%	6%	6%	5%
Ophthalmology	2%	1%	1%	1%	1%	2%
Cardiology	4%	3%	4%	8%	14%	20%
Otolaryngology	16%	8%	6%	5%	4%	3%
Pulmonology	7%	3%	3%	3%	4%	5%
Gastroenterology	6%	8%	8%	9%	11%	10%
Hepatology	1%	2%	2%	3%	3%	2%
Nephrology	0%	1%	1%	1%	2%	2%
Urology	2%	3%	3%	3%	3%	5%
Obstetrics	1%	12%	22%	5%	0%	0%
Gynecology	1%	5%	8%	11%	8%	5%
Dermatology	8%	8%	5%	4%	4%	3%
Orthopedics & rheumatology	13%	16%	14%	18%	19%	18%
Neonatology	8%	0%	0%	0%	0%	0%
Preventive & administrative	7%	3%	3%	3%	2%	1%
Late effects, environmental trauma & poisonings	1%	1%	1%	1%	1%	0%
Isolated signs & symptoms	1%	1%	1%	1%	1%	0%
Total Share	100%	100%	100%	100%	100%	100%