Geographic Variation in Commercial Medical Care Expenditures: A Decomposition Between Price and Utilization

Abe Dunn, Adam Hale Shapiro, and Eli Liebman

July 15, 2011

Abstract

This study examines geographic variation in commercial medical care expenditures. Medical care expenditures are decomposed between service prices and service utilization. We find that both service prices and utilization contribute to overall differences in health spending across geographic markets. Our findings suggest that potential expenditure savings may be possible from more efficient utilization. However, the large variation in underlying service prices suggests that deviations in overall spending may persist, even if utilization differences across markets are diminished.

1 Introduction

There has been a considerable amount of research assessing the geographic variation of Medicare expenditures, but relatively little is known about geographic variation in commercial-market expenditures. This is a large hole in our understanding of the overall health care market. The private market includes around 174 million enrollees, compared to Medicare that has 44 million.[1] In addition, overall medical care expenditure from private insurers is 60 percent greater than the amount spent by Medicare.[2] As private markets are influenced by different economic forces than regulated markets, there is likely distinct variation in medical care expenditures to each. Unlike the Medicare markets where payments to providers are fixed by the Center for Medicare and Medicaid Services, prices in the private sector are set through negotiations between insurers and providers. Understanding how medical care expenditures differ across these two markets may ultimately provide important insights into how government regulation affects medical care spending and outcomes. Overall, a sound identification of geographic variation in total medical expenditures is one step in comprehending how to achieve lower health care costs. Indeed, there is most likely waste in medical care expenditures in the United States, since the country spends 50 percent more than most other developed nations on health care in terms of a fraction of its GDP, but does not produce measurably better health outcomes (See Garber and Skinner (2008), Anderson and Hussey (2001)).

In this study, we provide a number of contributions to the literature on medical care expenditures. First, we provide descriptive evidence of geographic variation in overall medical care expenditures in commercial markets across Metropolitan Statistical Areas (MSAs) in the United States. To do this, we construct a Medical Care Expenditure Index (MCE) that tracks the overall medical care expenditure of treating an episode of a disease. Specifically, we create an aggregate MCE index—a measure for a "typical" disease for an MSA—as well as a multitude of disease-specific MCE indexes which track the expenditure of a particular disease for an MSA. A major contribution of this study is that we decompose the MCE index between its two key components: a Service Price Index (SPI) and a Service Utilization Index (SUI). The SPI isolates the variation in underlying service prices (for example, the price of

[1] The 174 million in the private sector includes only those under 65 years of age (Health United States (2009)).

[2] Specifically, private market accounts for more than 35 percent of total medical care expenditures, while Medicare's share is 22 percent. The other sources of funding include Medicaid accounting for 17 percent, out-of-pocket costs accounting for 14 percent, and other sources accounting for the remaining 14 percent of spending. These figures are from the personal health expenditures reported in the National Health Expenditure Accounts for the year 2009. Also note that around 57 percent of the out-of-pocket costs are from individuals with private insurance (MEPS 2007 Data).

a visit to a doctor to manage a pregnancy), but holding service utilization constant (for example, fixing the number of visits to the doctor across markets for each pregnancy). By contrast, the SUI isolates the variation in medical care expenditures attributable to the quantity of services provided per episode of care. Specifically, the SUI holds the prices of the underlying services constant but allows the number of services to vary.

These measures are important for understanding expenditure differences and potential savings across markets. If significant variation in utilization is identified then market participants can seek out ways to control potentially wasteful spending. On the other hand, if variation in spending is primarily driven by price differences, then measures to curb utilization may have a limited impact on high spending areas, and policy makers should focus on factors that affect the price of services, such as the bargaining power of providers and factors affecting the cost of the underlying services (for example, regulatory barriers or wasteful administrative costs).

We find relatively limited variation in medical care expenditure across geographic markets (that is MSAs) in the aggregate. The coefficient of variation for the aggregate MCE index is 0.10 and appears to be similar in magnitude to price variation in non-health goods and services. Unlike Medicare markets where variation is primarily driven by differences in utilization across markets, it appears that variation in price is particularly important in commercial markets. Specifically, we find a coefficient of variation of 0.087 for the aggregate SPI and 0.059 for the aggregate SUI. These findings suggests that even if utilization is controlled for, differences in spending are likely to persist across areas.

Although the aggregate geographic variation in MCE, SPI and SUI are relatively low, the geographic variation when examining specific diseases is quite large. Looking at the variation observed at the disease level, we find the weighted average coefficients of variation for MCE, SPI, and SUI to be 0.21, 0.15 and 0.16, respectively; which is considerably greater than the variation in the aggregate indexes. Thus, it appears that averaging over diseases masks the underlying geographic variation in service utilization and spending across specific diseases. This finding leads us to believe that there are not necessarily high and low utilization areas, but rather they may be over-utilizing services for some diseases and under-utilizing for others. Therefore, there may be greater efficiency gains by looking at specific disease categories rather than focusing on the aggregate level. This is especially true if one factors in variation in service prices. For example, it appears that for some diseases the variation in service price is relatively large (for example, pregnancy), while for other diseases the variation in utilization appears large (for example, depression).

One of our main findings is that both the aggregate index and the disease-specific indexes suggest that variation

in service prices play a key role in explaining the differences in medical care expenditures across markets. To better assess the source of this variation, we divide the data among service categories: inpatient, outpatient, physician services, and pharmacies. We find large variation in the service prices for inpatient hospital, outpatient hospital, and physician services, with coefficients of variation of 0.15, 0.20, and 0.12, respectively. The prices for pharmacy products vary considerably less, with a coefficient of variation of 0.07. A likely explanation for this dichotomy is that the competition in pharmacy markets is very similar across the nation, since the same drugs are generally available across all markets in the United States so the competitive environment is quite similar across locations. This is not true for physician and hospital services, where both the competitive environment and the products themselves may vary greatly across markets.

To assess the economic importance of geographic variation in medical care expenditures, we calculate changes in the episode expenditures from conducting hypothetical shifts in service utilization and service prices.[3] First, we find that if MSAs in the high spending areas (that is, the top quartile) had similar utilization to those in low spending areas (that is, the bottom quartile), there would be a 7 percent reduction in per episode expenditures. In addition, we find that if prices in the low spending areas are applied to high spending areas, then expenditures in the high spending areas would fall by 13 percent. The effects of shifting utilization are considerably greater if we perform this exercise on a disease basis. Specifically, the expenditures per disease are 22 percent lower, on average, if we substitute the utilization levels from the high spending quartile to the levels observed in the low spending quartile for each disease.[4] Similarly large differences are shown by shifting the service prices from the low spending areas to high spending areas, where we find a reduction in the average episode expenditures of 18 percent. These results bolster our hypothesis that strategies to contain expenditures by eliminating unnecessary utilization may benefit from focusing on the optimal care on a disease basis, rather than attempting to replicate "low spending" markets. For example, there are some conditions that have particularly high variation in utilization that may benefit from efforts to curb over-utilization (for example, joint degeneration of the back), while other conditions have relatively low variation in utilization where the benefits are likely to be smaller (for example, asthma).

The analysis differs if we focus on variation between low utilization and high utilization areas. As we would

[3] These exercises are solely meant to highlight the importance of observed variation across markets because they ignore the behavioral response to market changes.

[4] Phelps (1995) discusses in greater detail the small area variation literature and how variation in utilization may be a sign of inefficiency and a loss in consumer surplus.

expect, we find that if MSAs in the high utilization areas (that is, the top quartile) had similar utilization to those in low utilization areas (that is, the bottom quartile), there would be a 12 percent reduction in spending per episode. However, if prices in the low utilization areas are applied to high utilization areas, then expenditure per episode in the high spending areas would *increase* by 9 percent. Thus, it appears that low utilization areas tend to be higher priced areas, so that the benefits of lower utilization are partly offset by higher prices. This negative relationship between service prices and service utilization is also confirmed by examining simple correlations between the SPI and SUI indexes, which show a negative and statistically significant correlation (both in the aggregate and at the disease level).

An additional contribution of this paper is that we assess how well these three measures (that is MCE, SPI, and SUI) are correlated with quality of treatment using guidelines from the National Committee for Quality Assurance (NCQA). Although evidence from Medicare data suggests that more spending is not necessarily associated with a greater quality of service, less research has studied the relationship between spending and quality in the commercial sector. Our results show little systematic relationship between these measures and the quality of care. The patterns are similar if we look at the relationship between quality and service price or the relationship between quality and service utilization. The limited association between each of these measures and indicators of quality suggests that savings are possible in commercial markets without an accompanied reduction in quality.

2 Literature Review

Geographic variation in health care utilization and expenditure is a growing area of study. The work by Dartmouth researchers has shown considerable variation in medical care expenditures and service utilization across areas of the United States. Results from studies using Medicare data show large geographic variation in spending and utilization that do not appear to be associated with better patient outcomes or quality. For example, for a national sample of Medicare enrollees, Baicker and Chandra (2004) find that spending and quality of care are actually inversely related.

While there is a large body of work studying variation across geographic markets for samples of Medicare beneficiaries, fewer studies have analyzed variation in medical care expenditure in the commercial sector.[5] Although it may

[5] The seminal work in the study of geographic differences in medical care expenditures was pioneered by John Wennberg, with a description of his earlier work and implications of geographic variation discussed in Wennberg (1984). A more recent review of the literature on geographic variation in health care spending is in Congressional Budget Office (CBO) (2008).

be tempting to draw conclusions from the significant analysis conducted in the Medicare market, evidence suggests that commercial and Medicare markets may be quite distinct, even within the same geographic area. In particular, Chernew et al. (2010) compare spending across Medicare and commercial markets and found only a small negative correlation. Although their study finds a significant positive correlation in utilization of 0.59 for inpatient days per capita, the link between the two markets remains unclear. Moreover, given the regulated structure of Medicare pricing, it appears that price variation seems to be relatively unimportant in Medicare markets (See Gottlieb et al. (2010)). However, this may not be the case in commercial markets where insurers and providers are free to negotiate over rates. Therefore, additional research may be necessary to understand the unique features of spending, pricing and utilization in commercial markets.

There are many approaches for analyzing geographic differences in spending and utilization across markets. Some research focuses on differences at the micro level, examining the use of specific procedures for certain diseases (for example, Chandra and Staiger (2007) look at different types of treatments for heart attack patients across markets), while other studies examine aggregate differences in overall medical care expenditures (for example, Cutler and Sheiner (1999), Fuchs, McClellan and Skinner (2004), and MedPac (2003)). This paper combines aspects of both these approaches because it focuses on aggregate medical care expenditures in a geographic area, but these measures are constructed at the most micro level possible. This approach may be useful for understanding aggregate differences in medical care expenditures, since providers differ by specialty (for example, cardiology or orthopedic doctors), type (for example, hospitals or physicians), and factors unique to a local market (for example, cultural or information differences (Chandra and Staiger (2007))).

Tracking expenditure at the episode level may be important given the potential for shifts in service between provider types. There are several documented shifts in treatment over time that have been studied in the literature, such as heart attacks (Cutler et al. (1998)), cataracts (Shapiro et al. (2001), and depression (Berndt et al. (2001)). General shifts in treatments across a broad range of conditions have also been documented by Aizcorbe and Nestoriak (2011). The existence of these observed shifts over time suggest that different allocations of services across geographic markets may also be important.

Similar to the work of Aizcorbe and Nestoriak (2011) we measure the medical care expenditure of an episode of care for hundreds of different disease categories. Aizcorbe and Nestoriak (2011) provide an innovative framework for studying medical care expenditure and differences in expenditure across a broad range of conditions. In this

study, we adopt much of their basic methodology. Similar to their work, we show how a simple service price index that holds service utilization fixed may be a misleading indicator for medical care expenditures, since differences in utilization may also be important in determining the overall expenditure of treatment. However, our analysis is distinct for three key reasons. First, rather than focus on the variation over time, we focus on the variation across geographic markets. Second, we decompose medical care expenditure into its service price as well as its service utilization components. This additional decomposition allows us to look at differences in the MCE index caused by differences in the quantity of services provided in addition to differences in service prices. Third, we analyze services at a more micro level—the level of the specific procedure. This allows us to capture greater heterogeneity in the types of services performed across markets. For example, rather than pricing a visit to a doctor, we focus on the price for a particular procedure and modifier code. This methodology is discussed in greater detail in the following sections.

3 Methodology of Index Construction

The MCE index is a measure of the medical care expenditures for the treatment of a disease episode, and is defined as the dollar amount of medical care used until treatment is completed.[6] Formally, denote the average expenditure per episode of treating disease d in area r as c_d^r. Denoting c_d^B as the average expenditure per episode across *all* areas, the MCE index for disease d is the ratio of the two measures:

$$MCE_d^r = \frac{c_d^r}{c_d^B} \tag{1}$$

Thus, if the MCE index is larger than one, it signifies that the expenditure for treating disease d is larger than average (or what we call the "base" area) and if the index is less than one it signifies that the expenditure is less than the average.

Our decomposition rests on the fact that the average expenditure, c_d^r, 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

[6] For 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.

(that is, patients): $c_d^r = \sum_s p_{d,s}^r Q_{d,s}^r / N_d^r$, where $Q_{d,s}^r$ is the quantity of services of type, s; $p_{d,s}^r$, is the service price; and N_d^r is the number of episodes treated.

Measuring service utilization is not a straightforward task since the definition of "service" is a bit ambiguous and there are a variety of ways that one could define it across various service types.[7] Ideally, we would like the definition of service to depend on how the price of the service is typically set and paid. For example, for visits to a physician's office, 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 "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.[8] For prescription drugs, we define the unit of service as a prescription filled, albeit this is a bit of 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 from inpatient or outpatient stays, the prices paid to facilities are often set based on the disease of the patient for a particular visit. Therefore, we define the unit of 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: $p_{d,s}^r = \frac{c_{d,s}^r}{Q_{d,s}^r}$ where $c_{d,s}^r$ is the average expenditure on disease d for service s in area r. 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^r be a vector of services utilized for the typical treatment of diseases in an area, $q_d^r = Q_d^r / N_d^r$, where the component of the utilization vector for service s is , $Q_{d,s}^r / N_d^r$. Also, let p_d^r be a vector of service prices, where the component of the

[7] The key service types are inpatient hospitals, outpatient hospital, general physician, physician specialist, and prescription drugs.

[8] This 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.

vector for service s is, $p^r_{d,s}$. The service price index (SPI) is then calculated as:

$$SPI^r_d = \frac{p^r_d \cdot q^B_d}{c^B_d}$$

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

$$SUI^r_d = \frac{p^B_d \cdot q^r_d}{c^B_d}$$

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:

$$MCE^r_d = SPI^r_d + SUI^r_d + (p^B_d \cdot q^r_d - c^B_d)(p^r_d \cdot q^B_d - c^B_d)/((c^B_d)^2) - 1$$

Here the MCE index is equal to the service price index, SPI^r_d, plus the service utilization index, SUI^r_d, plus a cross term, $(p^B_d \cdot q^r_d - c^B_d)(p^r_d \cdot q^B_d - c^B_d)/((c^B_d)^2)$, and subtracting 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 differences in utilization across markets, SUI^r_d is fixed near 1, then the MCE^r_d will entirely be determined by service prices. Similarly, if there are very few differences in services prices across markets, SPI^r_d, is near 1, and the MCE^r_d will entirely be determined by utilization. To aggregate across diseases, we weight by the expenditure share of that disease for the entire U.S.

4 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 concerning 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 2006 and 2007. We also limit our sample to enrollees with drug benefits because drug purchases will not be observed for individuals without drug coverage.[9] The MarketScan database tracks claims from all providers using a nationwide convenience sample of enrollees. Each enrollee has a unique identifier and can be linked to a particular MSA. All claims have been paid and adjudicated.[10]

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.[11] 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.

4.1 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.[12] 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 hospital, outpatient hospital, general physician, specialist

[9] In our selected sample that removes individuals without drug benefit information, we find about 62 percent of the data is from employer plans while the remaining 38 percent is from health insurance plans.

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

[11] The 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.

[12] For 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 and end in 2006 or 2007 are included in this study, while episodes that begin in 2007 and end in 2008 are not included.

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.[13]

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.[14]

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 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

[13] Although 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.

[14] As an alternative, we have also examined changing this definition to consider the facility price per inpatient day. The results to not change significantly based on these two alternative measures of utilization.

code represents the amount of the service used. If the expenditure on a prescription is greater than this amount, it suggests that prices are above average in an area.[15]

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.[16]

4.2 Selected Sample and Descriptive Statistics

When studying variation across MSAs, there is some concern that we have a large enough sample within each MSA so that an average over the population will be meaningful. To ensure that each MSA has a sufficient number of individuals, we select only those MSAs in the data that have an average of 20,000 enrollees per year over the 2006-2007 time period (that is 40,000 enrollee year observations). The sample size in each city is more than double the sample size of the commercially-insured sample from the Medical Expenditure Panel Survey, which is a national survey of health expenditures meant to be representative of the entire U.S. non-institutionalized population.[17] This first selection rule leaves a sample of 81 MSAs.[18] When analyzing diseases within this population, we wish to remove those diseases that appear relatively infrequently. The concern is that it may be challenging to obtain precise estimates

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

[16] The 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. We also obtain similar results when we define the amount done during an inpatient or outpatient hospital visit as a "visit" for that disease. That is, we treat the physician portion of the payment as part of the inpatient or outpatient visit and ignore the heterogeneity of procedures done.

[17] The commercially insured sample in the MEPS data is around 14,799 individual observations in each year. In this study we are using two years of data which includes more than 40,000 individual-year observations per MSA.

[18] These 81 MSAs account for 69 percent of the enrollment population available in the MarketScan data that are located in an identifiable MSA for these years of study.

For each MSA used in this analysis, at least three unique employers contribute to the data in each MSA and at least three unique carriers also contribute to this data.

of episode expenditures for infrequently observed diseases. We select those diseases for which we observe 40,000 episodes in the data.[19] Those diseases with 40,000 or more episodes account for 87 percent of overall expenditures and 96 percent of the episodes.

Population weights are applied to each MSA to adjust for differences in age and sex across populations, so the expenditure estimates may be comparable across markets. Specifically, enrollees in each MSA are assigned weights so the weighted population has an age and sex distribution that is identical to that of the US commercially insured population.[20]

Table 1 provides some basic descriptive statistics for the selected population and the overall disease expenditures. It shows the disease expenditures for the two-year period of 2006 and 2007 and is based on the weighted sample of enrollees, so these figures are representative of the 2007 U.S. population.[21] Based on the ETG groupings, the top 5 disease categories based on overall expenditure include ischemic heart disease, pregnancy, joint degeneration of the back, hypertension and diabetes. Although there are 207 diseases in the sample, these five diseases account for 19.8 percent of the expenditures in our sample. In general, most of the expenditures are accounted for by a limited number of diseases with the top 20 diseases listed here accounting for 47 percent of total expenditures from the selected diseases, so the aggregate price index will be heavily influenced by the high spending diseases. There is a wide range in the expenditure per episode across diseases. Hypertension costs just $733 per episode, while ischemic heart disease costs $5,783.[22]

[19] The results in this paper are not sensitive to a either the selection rule for the diseases or the MSAs. These robustness checks are outlined in greater detail in the appendix to this paper.

[20] Using the enrollment data in each MSA, weights are applied to different age and sex categories so that the total enrollment files match the population for commercially insured individuals in the U.S. for 2007. Information on the population is obtained from the Medical Expenditure Panel Survey.

MSA observations in 2006 and 2007 are each weighted to the national level population in 2007. That is, the sample in 2006 is weighted to the 2007 national population *and* the sample in 2006 is also weighted to the 2007 national population. After weighting the populations to the national level, the data is aggregated over the two years. This ensures that 2006 and 2007 receive equal weights in the price index, even if the enrollment within an MSA changes over these years.

We have conducted similar analysis looking at only 2006 and only 2007 year data. We obtain very similar results in each year.

[21] The national weights are applied to each city and the total expenditures and episodes are divided by the number of cities in our sample, 81, times the number of years of data, 2 (Thus we divide by 162 (=81*2)). Therefore, these figures actually overcount smaller MSAs included in the sample, relative to their share of the U.S. population.

[22] About 11 percent of expenditures are not assigned to any ETG disease category (that is screening for diseases and other records that cannot be assigned a category). Those claims that are not assigned disease categories are removed from our analysis.

Table 1. Total Average Annual Expenditures and Average Number of Episodes for 2006-07 - weighted to U.S. Totals for Commercial Insurance

Rank	Disease	Total Dollars (Billions) 2006-07	Episodes	Dollars Per Episode	Fraction of Spending
1	Ischemic heart disease	$46	15,861,278	$5,783	5.0%
2	Joint degeneration, localized - back	$38	30,094,861	$2,511	4.1%
3	Hypertension	$34	93,469,428	$733	3.7%
4	Pregnancy, with delivery	$33	6,724,011	$9,894	3.6%
5	Diabetes	$31	32,312,301	$1,911	3.4%
6	Malignant neoplasm of breast	$26	4,463,696	$11,593	2.8%
7	Routine exam	$24	230,082,880	$211	2.6%
8	Mood disorder, depressed	$22	33,712,821	$1,325	2.4%
9	Hyperlipidemia, other	$21	65,162,614	$649	2.3%
10	Joint degeneration, localized - neck	$20	18,954,534	$2,075	2.1%
11	Chronic sinusitis	$17	47,739,153	$699	1.8%
12	Asthma	$15	32,811,182	$936	1.7%
13	Joint degeneration, localized - knee & l	$15	10,724,987	$2,832	1.7%
14	Chronic renal failure	$15	2,266,644	$13,070	1.6%
15	Non-malignant neoplasm of female genital	$14	9,737,587	$2,893	1.5%
16	Joint derangement - knee & lower leg	$13	7,178,211	$3,651	1.4%
17	Inflammation of esophagus	$13	17,480,137	$1,483	1.4%
18	Cerebral vascular accident	$10	3,899,303	$5,058	1.1%
19	Cholelithiasis	$10	2,736,871	$7,198	1.1%
20	Tonsillitis, adenoiditis or pharyngitis	$10	83,965,236	$229	1.0%
	Other	$491	1,305,857,911	$752	53.4%
	Total	$919	2,055,235,646	$447.29	100.0%

In the analysis that follows each of the observed claims are aggregated to the MSA level, so the discussion of the variation in a disease price is the variation in the average disease price across MSAs. This is important because individual episode expenditures are likely to be idiosyncratic. However, this paper focuses on an average treatment expenditure within an MSA, often using hundreds of observed episodes of a disease per MSA.

5 Results

5.1 Aggregate Indexes

Table 2 below shows the results of the MCE indexes for each MSA in the data, although some MSAs are not shown due to confidentiality concerns. The MCE index ranges from a high of 1.29 in Milwaukee to a low of 0.81 in Youngstown, PA. The table also reports the SPI, that reflects differences in service prices, and the SUI, that reflects differences in service utilization. A glance at this table shows that the underlying cause for a high MCE may be due to either higher service prices (SPI), higher service utilization (SUI), or a combination of the two. For example, it appears that Milwaukee has an overall high MCE index primarily because it has a high SPI of 1.23, although the

SUI is close to 1, the national average. In contrast, Gary, IN has higher than average expenditures primarily because of service utilization, while the SPI is close to the national average. Other MSAs, such as Minneapolis, MN, have higher expenditures due to higher than average SPI and SUI.

It is also interesting to measure the overall variation in these indexes. Since the MCE index is normalized to have an average of 1.00, the standard deviation reported at the bottom of the table may be viewed as a coefficient of variation (COV), which is a typical measure of variation in studies looking at geographic differences in health care expenditures.

Overall the variation in each of the three indexes is low based on the literature that looks at variation in the use of a specific procedure across markets. The analysis here is not at the procedural level and therefore it may not be directly comparable, but it offers a useful benchmark. Phelps (1997) states that a coefficient of variation that is around 0.10 to 0.15 is considered to be very low based on variation in the use of procedures across areas. Each of the indexes is at the bottom of (or below) this range, with the COV of 0.10 for the MCE index and the other two indexes below this value. The variation in the SUI (0.06) is actually lower than the variation in the SPI (0.09).[23] More generally, the variation in these indexes is also low relative to price differences across regions for other goods and services.[24]

[23]The lower variation in the utilization index is not driven by selecting the coefficient of variation as the measure of dispersion. We also show differences in the 90th and 10th percentile, which also show that the utilization variation tends to be less.

[24]Using a regional price index for all goods from Aten and D'souza (2008) we find a coefficient of variation of 0.15 for a sample of 67 cities that match to our data. The CBO uses statistics from the BLS for the years 2004-2005 for a select sample of 24 cities and finds coefficients of variation of 0.12 for food, 0.143 for housing, and 0.143 for transportation.

Table 2. MSA Medical Care Price Indexes and Variation in Indexes - MCE, SPI and SUI

	MSA Name	Overall MCE Index	SPI Index	SUI Index
1	Milwaukee-Waukesha-West Allis, WI	1.287	1.228	1.031
2	Oakland-Fremont-Hayward, CA	1.231	1.263	0.972
4	Minneapolis-St. Paul-Bloomington, MN-WI	1.206	1.117	1.081
5	Fort Worth-Arlington, TX	1.174	1.121	1.039
7	Indianapolis, IN	1.158	1.105	1.027
8	Dallas-Plano-Irving, TX	1.151	1.142	1.009
9	Houston-Sugar Land-Baytown, TX	1.150	1.102	1.031
10	Gary, IN	1.136	0.971	1.147
12	Peoria, IL	1.123	1.141	1.006
13	Denver-Aurora, CO	1.104	1.030	1.062
14	Miami-Miami Beach-Kendall, FL	1.097	1.054	1.033
15	Portland-Vancouver-Beaverton, OR-WA	1.091	1.075	1.043
65	Louisville, KY-IN	0.924	0.952	0.959
66	Las Vegas-Paradise, NV	0.918	0.985	0.970
67	Pittsburgh, PA	0.906	0.837	1.111
69	Knoxville, TN	0.892	0.888	1.000
70	Memphis, TN-MS-AR	0.892	0.981	0.907
71	Nassau-Suffolk, NY	0.887	1.008	0.884
72	Providence-New Bedford-Fall River, RI-MA	0.885	0.879	1.017
73	Kingsport-Bristol-Bristol, TN-VA	0.884	0.843	1.050
74	Augusta-Richmond County, GA-SC	0.879	0.956	0.916
76	Detroit-Livonia-Dearborn, MI	0.860	0.885	0.983
80	Grand Rapids-Wyoming, MI	0.836	0.912	0.961
81	Youngstown-Warren-Boardman, OH-PA	0.805	0.823	0.986
	mean	1.000	1.008	0.996
	sd	0.099	0.087	0.059
	COV	0.099	0.087	0.059
	p10	0.884	0.912	0.917
	p90	1.150	1.117	1.063
	N	81	81	81

Although the variation appears to be quite low, much of the variation across markets appears to be smoothed out through this aggregation. While Table 2 shows indexes that aggregate over both the diseases and places of service, the next two sections take a more disaggregate view of these figures. The following sections look at variation in these indexes at the disease level as well as the service-type level.

5.2 Disease-Specific Indexes

The observed variation is much greater when looking at differences in the expenditures of specific diseases across MSAs. Table 3 reports the standard deviation of the MCE index, SPI and SUI for the top 20 diseases in the data (ranked by expenditures). The variation in the MCE index for the diseases listed in the table is, in general, much larger than the variation in the aggregate index (Table 2). Chronic Renal Failure has the largest variation in expenditure per episode across areas at 0.39, and Asthma has the lowest variation at 0.095.

The finding of higher variation in utilization at the disease level, relative to the variation using the aggregate indexes, is of particular interest. This finding suggests that although utilization patterns are different for each MSA,

MSAs are not systematically "high" utilization and "low" utilization areas for all diseases. Prior research suggests that much of the variation in utilization across medical care markets may be attributed to variation in practice styles and how information disseminates among physicians. For example, Wennberg (1984) reports huge variation in the probability of having tonsils removed across geographic markets. If factors influencing practice patterns are unique for each disease within an MSA, then averaging over the diseases may smooth the variation in utilization in the aggregate indexes.[25] For example, Gary, IN is ranked as the highest utilization city based on the overall SUI, but Gary ranks below the average based on utilization for the disease categories "Mood Disorder, Depressed" and "Pregnancy, with Delivery".

The underlying cause for the variation may be attributed to either utilization or price, which are shown using the standard deviation of the SPI and SUI. For some conditions it appears that price variation primarily affects the variation across areas, while for other conditions, the utilization variation appears to be more important. For example, the condition "Mood Disorder, Depression" has a relatively low price variation, but the utilization variation is greater. This could potentially be explained by talk therapy being more popular in some areas, while treatment with depression drugs may be more common in other areas; although the prices for each of these services may not vary by a large amount across areas. In contrast, a condition like pregnancy or the cost associated with routine examines have relatively little variation in utilization compared to variation in price across areas. One possibility is that treatments for these diseases are relatively clear, although the prices for the underlying services vary substantially across markets.

[25] Although one may be concerned that the analysis may be affected by outliers or small samples, we check for both of these. Specifically, we obtain similar results if we remove outliers for each disease. We also obtain a similarly larger COV at the disease level relative to the aggregate if we define the disease at the level of the Major Practice Category, which aggregates over many ETG disease categories. Although the COV shrinks when we aggregate to the MPC level, the variation we observe at this level, remains considerably larger than the aggregate SUI.

Table 3. Sources of Price Variation Across MSAs by Disease - MCE, SPI and SUI

	Description	COV of MCE Index	COV of SPI Index	COV of SUI Index
1	Ischemic heart disease	0.193	0.158	0.138
2	Joint degeneration, localized - back	0.207	0.130	0.155
3	Hypertension	0.104	0.091	0.109
4	Pregnancy, with delivery	0.161	0.149	0.052
5	Diabetes	0.126	0.060	0.109
6	Malignant neoplasm of breast	0.197	0.147	0.136
7	Routine exam	0.156	0.125	0.064
8	Mood disorder, depressed	0.199	0.068	0.168
9	Hyperlipidemia, other	0.102	0.078	0.104
10	Joint degeneration, localized - neck	0.206	0.121	0.189
11	Chronic sinusitis	0.181	0.092	0.137
12	Asthma	0.095	0.078	0.085
13	Joint degeneration, localized - knee & lower leg	0.198	0.142	0.151
14	Chronic renal failure	0.389	0.441	0.361
15	Non-malignant neoplasm of female genital	0.194	0.159	0.135
16	Joint derangement - knee & lower leg	0.215	0.184	0.161
17	Inflammation of esophagus	0.112	0.097	0.091
18	Cerebral vascular accident	0.260	0.217	0.174
19	Cholelithiasis	0.231	0.209	0.109
20	Tonsillitis, adenoiditis or pharyngitis	0.164	0.126	0.162
	Weighted Average	0.214	0.153	0.161

5.3 Service-Specific Indexes

Table 3 focuses on the variation in the diseases, but both Tables 2 and 3 suggest that price variation in the underlying services may play an important role in explaining differences in expenditures across markets. Given the importance of service prices in explaining differences in expenditures across areas, it may be useful to examine the underlying service types that might explain these differences. Table 4 below shows the key service types that are studied in the data along with the amount of expenditures associated with each type. Most of the spending is on doctor office visits (General MDs and Specialists), inpatient hospital visits, outpatient hospital visits and pharmacy. Only 13 percent of spending is on other services, such as, emergency and ambulatory care.

Table 4. Spending Share Across Services 2006-07

Place of Service	Average Total Spending (Billions)	Share of Spending
Inpatient Hospital	$189	20.6%
Outpatient Hospital	$208	22.7%
Office - General MD	$74	8.1%
Office - Specialist MD	$136	14.8%
Other (Emergency, Ambulatory Centers etc)	$120	13.1%
Pharmacy	$191	20.8%
Total	$919	100.0%

Similar to the analysis by disease, we will examine both the utilization and service price components of the expenditures. However, here we focus on the expenditures for specific services, instead of diseases. These indexes are constructed in an identical manner to the overall MCE, SPI and SUI measures, except we focus on only a single service category (that is, ignoring all other categories) and weighting each disease by its national expenditure share for that service type. For instance, to construct a SPI-service index for a specific type of service, the price of each service s for treating disease d, $p^r_{d,s}$, is weighted by the expenditure share of that service type across diseases.[26]

Table 5 shows the variation in the MCE-service, SPI-service, and SUI-service indexes for each of the main service types. The first column shows the variation in the MCE-service indexes. It appears that both Outpatient Hospital and Office-General MD vary by the most, with pharmacy varying the least. Interestingly, most of the variation in the MCE-service indexes stem from variation in the price of services. The middle column of the table shows the coefficient of variation in the service price index for each type of service. The table shows that prescription drug prices vary the least with a COV of 0.07 with outpatient hospital service prices varying the most with a COV of 0.20. The coefficient of variation for pharmaceutical prices is much lower than each of the other service types. One potential reason for the lower variance in price levels for pharmaceutical products is that competition among prescription drugs is likely to be very similar across markets, since the same drugs are typically available in each market. In contrast, the hospital and physician providers are offering services that are unique to each local market.

Table 5. Coefficient of Variation of Service Indexes Across Service Types

Service Category	COV MCE-service	COV SPI-service	COV SUI-service
Inpatient Hospital	0.175	0.152	0.069
Outpatient Hospital	0.247	0.198	0.069
Office - General MD	0.248	0.122	0.051
Office MD - Speciality	0.184	0.118	0.059
Pharmacy	0.078	0.069	0.056
Weighted Average	0.179	0.136	0.062

[26]For example, let the inpatient hospital expenditure share for disease d be denoted $s^{Service}_{d,Inpatient}$ where $\sum_d s^{Service}_{d,Inpatient} = 1$. Then the price index for the service category would be: $P_{Inpatient} = \sum_d p^r_{d,s} \cdot s^{Service}_{d,Inpatient}$. In contrast to the overall index that is weighted by the total expenditure share for each disease, this index is weighted by the expenditure share of a service. To normalize the prices we divide by the average price index for that service type across all MSAs.

Although the price variation appears to be relatively large for many of the services, there is relatively little variation in the utilization index across these service types. Similar to the aggregate SUI, the likely reason for the limited variation is that the index aggregates over disease types. In contrast, the price levels for the different services may be common across diseases for a specific MSA due to reasons related to bargaining leverage of physicians or hospitals (See Dunn and Shapiro (2011)). For example, a general MD that negotiates a higher price with an insurer (for example, 10 percent above Medicare rates) will receive a higher service price regardless of the disease of his patients. A more detailed listing of the MCE-service, SPI-service and SUI-service specific indexes is shown in the appendix of this paper.

6 The Economic Importance of Variation in Price and Utilization

6.1 MSA Variation

This section presents analysis demonstrating the economic importance of the variation in service price and utilization across markets. Specifically, we focus on the potential expenditure reduction if either utilization or prices were shifted from the levels observed in high expenditure MSAs to the levels observed in the low expenditure MSAs. The low and high expenditure indexes are determined by the overall MCE index. These exercises are solely meant to highlight the importance of observed variation across markets because they ignore the behavioral response to market changes.

To perform these exercises, we first rank the MSAs based on their overall MCE index and place each MSA into one of four "quartile-bins." Next, we calculate the average utilization for each service type and disease in each quartile-bin, $q_{d,s}^{quart^{MCE}}$ creating the vector $q_d^{quart^{MCE}}$ where the MCE superscript indicates that we are choosing quartile-bins based on the MCE index. We then take the average utilization in the low-spending quartiles, $q_{d,s}^{25^{MCE}}$, and apply these utilization levels to actual price levels:

$$qADJUST_d^{25^{MCE}} = \frac{p_d^r \cdot q_d^{25^{MCE}}}{c_d^B}.$$

We should note that this measure is computed in a similar fashion to how the SPI is calculated, except instead of fixing utilization to the average benchmark level, we use the utilization levels from the low (that is, 25th percentile) spending areas. One way to view $qADJUST_d^{25^{MCE}}$ is that it is a measure of the expenditure savings from lower utilization.

Similarly, to look at the reduction in expenditure from shifting service prices, we take the average price of each service for treating each disease in the low spending areas and apply these prices to actual utilization levels:

$$pADJUST_d^{25^{MCE}} = \frac{p_d^{25^{MCE}} \cdot q_d^r}{c_d^B},$$

which is analogous to the SUI calculation, except that prices are fixed to the price levels in low spending areas. The change in the MCE index represents the reduction in the overall MCE index due to either utilization shifting or prices shifting. The results of this analysis applied to those MSAs in the top quartile are shown in Table 6.1. The table shows a 6.5 percent change in episode expenditures from the utilization reduction, and a 12.7 percent change from the price reduction. So it appears that price is a more important factor than utilization when looking at overall health care expenditures in high spending areas compared to low spending areas. This is consistent with the observation that the variation in the SPI was greater than the variation in the SUI. However, the precise effect of these shifts depends on the specific MSA. In some MSAs the savings are much greater than in others (for example, the higher MCE areas tend to have larger savings). In addition, in some cases the price shift may actually lead to an increase in overall expenditures. For example, in markets such as Cambridge and Boston prices are higher due to utilization, not higher service prices. In this case, substituting prices from the low spending areas can actually increase the MCE index.

Table 6.1 Movement from Highest Spending Quartile to Average Levels in Low Spending Quartile

		Price Change		Utilization Change	
	MCE Index	MCE Adj.	Change	MCE Adj.	Change
Milwaukee-Waukesha-West Allis, WI	1.287	0.979	-0.239	1.182	-0.082
Oakland-Fremont-Hayward, CA	1.231	0.930	-0.244	1.213	-0.015
Minneapolis-St. Paul-Bloomington, MN-WI	1.206	1.032	-0.144	1.076	-0.107
Fort Worth-Arlington, TX	1.174	0.995	-0.152	1.080	-0.080
Indianapolis, IN	1.158	0.987	-0.148	1.064	-0.082
Dallas-Plano-Irving, TX	1.151	0.964	-0.163	1.098	-0.046
Houston-Sugar Land-Baytown, TX	1.150	0.983	-0.145	1.055	-0.082
Gary, IN	1.136	1.100	-0.032	0.934	-0.178
Peoria, IL	1.123	0.957	-0.148	1.094	-0.026
Denver-Aurora, CO	1.104	1.011	-0.084	0.991	-0.102
Miami-Miami Beach-Kendall, FL	1.097	0.989	-0.098	1.016	-0.074
Portland-Vancouver-Beaverton, OR-WA	1.091	0.996	-0.087	1.034	-0.053
San Diego-Carlsbad-San Marcos, CA	1.080	0.888	-0.178	1.098	0.016
Tacoma, WA	1.069	0.971	-0.092	0.994	-0.070
Boston-Quincy, MA	1.060	1.068	0.007	0.978	-0.077
Bridgeport-Stamford-Norwalk, CT	1.048	0.942	-0.101	1.015	-0.032
Cambridge-Newton-Framingham, MA	1.033	1.063	0.029	0.956	-0.075
Average Savings			-0.127		-0.065
Median Savings			-0.144		-0.076

Alternatively one can conduct a similar analysis as above, but instead we can sort the MSAs by the SUI instead of the MCE index. In this case, we substitute $q_d^{25^{MCE}}$ with $q_d^{25^{SUI}}$ and $p_d^{25^{MCE}}$ with $p_d^{25^{SUI}}$ where the superscript SUI indicates that we are ranking the MSAs by utilization. The analysis reported in Table 6.2 is identical to the

analysis in Table 6.1, but the quartiles of the data are divided based on the SUI index. That is, we examine the effects of taking the utilization and prices observed in the low utilization MSAs, and applying those amounts to the high utilization areas. As one would expect we find larger expenditure savings of around 12 percent from a shift in utilization. However, if the high utilization areas take on the low utilization price levels, the expenditures actually *increase* by 9 percent on average. Therefore, it appears that low utilization areas also tend to be higher price areas on average.

Table 6.2 Movement from High Utilization Quartile to Average Levels in Low Utilization Quartile

		Price Change		Utilization Change	
	MCE Index	MCE Adj.	Change	MCE Adj.	Change
Minneapolis-St. Paul-Bloomington, MN-WI	1.206	1.135	-0.058	1.041	-0.137
Fort Worth-Arlington, TX	1.174	1.088	-0.073	1.032	-0.120
Gary, IN	1.136	1.219	0.073	0.896	-0.211
Denver-Aurora, CO	1.104	1.107	0.003	0.955	-0.135
Miami-Miami Beach-Kendall, FL	1.097	1.079	-0.016	0.968	-0.118
Portland-Vancouver-Beaverton, OR-WA	1.091	1.107	0.014	1.014	-0.071
Boston-Quincy, MA	1.060	1.213	0.145	0.960	-0.094
Cambridge-Newton-Framingham, MA	1.033	1.203	0.164	0.941	-0.089
Philadelphia, PA	1.012	1.116	0.103	0.884	-0.126
Toledo, OH	1.009	1.210	0.200	0.837	-0.171
Wichita, KS	1.004	1.121	0.117	0.925	-0.079
St. Louis, MO-IL	0.979	1.103	0.127	0.861	-0.121
Phoenix-Mesa-Scottsdale, AZ	0.974	1.099	0.128	0.862	-0.115
Jacksonville, FL	0.964	1.082	0.123	0.856	-0.112
Cleveland-Elyria-Mentor, OH	0.963	1.122	0.166	0.850	-0.117
Pittsburgh, PA	0.906	1.188	0.311	0.788	-0.130
Kingsport-Bristol-Bristol, TN-VA	0.884	1.103	0.248	0.791	-0.105
Average Savings			0.089		-0.122
Median Savings			0.120		-0.119

6.2 Disease Variation

The calculation of the reduction in expenditures from adjusting utilization and prices may be much different than the amounts reported in Tables 6.1 and 6.2 if one looks at the expenditures from treating specific diseases across cities. For example, one location may have very low utilization for diabetes, but high utilization for pregnancy, which may be ignored by the overall utilization index. As an alternative measure of the potential savings, we focus on high expenditure and low expenditure areas on a disease by disease basis. That is, for each disease, we sort the data into low expenditure (low utilization) and high expenditure (high utilization) MSA quartile-bins. This exercise is identical in fashion to that shown above, but the MSA ranking into different quartiles is done for each disease separately, so that an MSA may belong in a high-quartile bin for one disease and a low-quartile bin for another. These steps are taken for every disease in the data and we calculate the weighted average savings using national expenditure shares of each disease.

Table 7.1 reports the results from the analysis where savings is computed for each disease by comparing prices

and utilization from high spending MSAs for each disease to the low spending areas for each diseases. As one might expect, those diseases with relatively large variation in price (as measured by the SPI), such as "Pregnancy, with Delivery", observe a greater reduction in expenditure from shifts in prices, relative to shifts in utilization. On the other hand, those diseases with greater relative variation in utilization, such as "Mood Disorder, Depressed", show a greater reduction in expenditure from shifts in utilization, relative to shifts in prices. The weighted average savings for the "typical" disease is shown at the bottom of Table 7.1. The typical reduction in expenditures from the shift in price is 18 percent and the reduction in expenditures is 22 percent based on quantity changes. That is, it appears that savings are quite large based on both price and utilization when one examines disease by disease expenditure differences. The reduction in expenditures reported here are in line with prior research. In particular, research examining the Medicare population suggests that savings may be as great as 20 to 30 percent if the high spending areas were brought in line with the low spending areas (See Skinner et al. (2005) and Fisher et al. (2003)).

Table 7.1. Percentage Change in Expenditures from Shifts in Utilization and Shifts in Price by Disease - From High Spending Quartiles to Low Spending Quartiles

	Description	%Chg from Shift in Price	%Chg from Shift in Utilization
1	Ischemic heart disease	-22.1%	-18.8%
2	Joint degeneration, localized - back	-17.3%	-21.2%
3	Hypertension	-2.8%	-13.9%
4	Pregnancy, with delivery	-28.7%	-3.9%
5	Diabetes	-4.8%	-19.0%
6	Malignant neoplasm of breast	-21.3%	-21.9%
7	Routine exam	-23.1%	-8.6%
8	Mood disorder, depressed	-9.5%	-29.0%
9	Hyperlipidemia, other	-5.5%	-12.9%
10	Joint degeneration, localized - neck	-8.6%	-30.4%
11	Chronic sinusitis	-12.6%	-26.6%
12	Asthma	-7.2%	-13.3%
13	Joint degeneration, localized - knee & lower leg	-12.8%	-24.1%
14	Chronic renal failure	-24.1%	-33.5%
15	Non-malignant neoplasm of female genital	-22.9%	-14.7%
16	Joint derangement - knee & lower leg	-25.2%	-14.2%
17	Inflammation of esophagus	-10.9%	-12.4%
18	Cerebral vascular accident	-24.7%	-19.5%
19	Cholelithiasis	-33.2%	-9.8%
20	Tonsillitis, adenoiditis or pharyngitis	-10.1%	-18.0%
	Weighted Average Savings (All Diseases)	-17.6%	-21.6%

In contrast to Table 7.1 that ranks MSAs into high and low expenditure areas for each disease, 7.2 categorizes MSAs into high and low utilization areas for each disease. Looking at the second column, one can see potentially large reductions in expenditures from shifting utilization levels from the low utilization areas to the high utilization areas. The weighted average savings in expenditures for the typical disease is 30 percent. However, looking at the

first column, when the prices from the low utilization areas are substituted into the high utilization areas, we see an *increase* in expenditures for most diseases. This suggests a negative correlation between price and utilization at the disease level. Simple correlations between the overall SUI and SPI or the SUI and SPI for each disease confirm a negative and statistically significant correlation between utilization and price. Specifically, the correlation between the log of the overall SUI and SPI is -0.18 which is significant at the 90 percent level. In addition, the correlation between the log of the SUI and the SPI at the disease level is -0.05 and statistically significant at the 1 percent level. If this represents a demand relationship, this could be an interesting finding since it would suggest that the full price (that is, the price paid by the insurer and the consumer) ultimately impact the demand for services and help shape utilization patterns across markets.

Table 7.2. Percentage Change in Expenditures from Shifts in Utilization and Shifts in Price by Disease - From High Utilization Quartiles to Low Utilization Quartiles

	Description	%Chg from Shift in Price	%Chg from Shift in Utilization
1	Ischemic heart disease	11.6%	-28.4%
2	Joint degeneration, localized - back	9.4%	-29.5%
3	Hypertension	13.9%	-21.8%
4	Pregnancy, with delivery	2.1%	-11.1%
5	Diabetes	1.4%	-20.8%
6	Malignant neoplasm of breast	4.7%	-28.5%
7	Routine exam	-7.5%	-12.9%
8	Mood disorder, depressed	-2.5%	-31.6%
9	Hyperlipidemia, other	11.4%	-19.2%
10	Joint degeneration, localized - neck	7.8%	-34.5%
11	Chronic sinusitis	-4.3%	-28.7%
12	Asthma	6.5%	-18.2%
13	Joint degeneration, localized - knee & lower leg	10.9%	-30.8%
14	Chronic renal failure	94.1%	-53.9%
15	Non-malignant neoplasm of female genital	18.3%	-26.3%
16	Joint derangement - knee & lower leg	25.7%	-29.8%
17	Inflammation of esophagus	9.9%	-18.6%
18	Cerebral vascular accident	9.5%	-30.6%
19	Cholelithiasis	12.8%	-22.6%
20	Tonsillitis, adenoiditis or pharyngitis	21.2%	-25.8%
	Weighted Average Savings (All Diseases)	10.4%	-29.6%

Overall, Tables 6.1, 6.2, 7.1 and 7.2 demonstrate that potential savings from either price or utilization shifts may be much greater if one focuses on changing pricing or utilization at the disease level. In addition, it appears that both differences in price and utilization are important determinants of spending differences across markets, which contrasts with the Medicare markets where utilization differences appear to be more important. Another interesting finding is that we observe a negative and significant relationship between the SUI and SPI measures. Additional research is necessary to determine if this captures a true demand relationship or just a spurious correlation.[27]

[27] Please see the Appendix for a list of robustness checks.

7 Medical Care Expenditures and Quality Measures

There is considerable variation in spending, service prices and service utilization, but it is unclear what this variation means for consumer surplus, since greater spending may be associated with high quality care. Although this relationship between spending and quality does not appear to be present in Medicare markets,[28] less research has been conducted in the commercial sector. Here we use a set of procedural quality measures constructed from the MarketScan data to examine whether there is an association between quality and the MCE index, SUI and SPI. The quality measures are constructed using Healthcare Effectiveness Data and Information Set (HEDIS) guidelines from the National Committee for Quality Assurance (NCQA). Additional details regarding the construction of the quality measures are included in the appendix.

Table 8 below shows the correlation between the log of quality measures and the different indexes. The results show some interesting patterns, but there does not appear to be a systematic relationship between quality measures and spending. For most measures of quality we find this correlation to be statistically insignificant, whether we look at the association with the SPI, the SUI, or the MCE index. However, there does appear to be a positive and significant relationship between measures of diabetes quality and utilization. An association with overall quality measure (that is, a composite of quality measures) and overall MCE, SUI and SPI are statistically insignificant for the MCE and SUI measures. There is a significant positive correlation between the overall index of quality and the overall SPI at the 10 percent level of significance.[29] Although the association is positive and significant for some measures (e.g. HbA1c tests for pediatric patients), it is negative and significant for other measures (for example, the fraction of people with hyperlipidemia taking drugs), suggesting no systematic pattern between prices, spending and utilization.

[28] For example, one study in this area by Fisher et al. (2003) shows that patients in high spending areas do not have better health outcomes or greater satisfaction scores.

[29] The regression is significant, but the explanatory power is tiny. A regression of the log(*overall quality*) on the overall SPI has an adjusted R-squared of 0.03.

Table 8. Correlation between Quality and Episode Indexes: Episode Cost, Price and Utilization

	Log(MCE)	Log(SPI)	Log(SUI)
Log Quality Measure - Persistence of Beta-Blocker Treatment after a Heart Attack			
Hypertension	-0.1089	0.1389	-0.1984*
p-value	0.3392	0.2221	0.0796
Ischemic Heart Disease	0.0515	-0.0086	0.0918
p-value	0.6522	0.94	0.4208
Log Quality Measure - HbA1c Test for Pediatric Patients			
Diabetes	0.2338**	0.1038	0.2392**
p-value	0.0369	0.3597	0.0326
Log Quality Measure - Complete Lipid Profile for Patients 18 years and older with Ischemic Vascular Disease			
Hyperlipidemia	-0.2417**	-0.2377**	-0.1291
p-value	0.0297	0.0326	0.2507
Log Quality Measure - Those with Back Pain Not Reporting an MRI within first 6 Months			
Joint Degeneration - Back	0.0104	0.1611	-0.0968
p-value	0.9269	0.1508	0.3899
Log Quality Measure - Composite (summation of four quality indexes above and two indexes of preventative care)			
Overall Indexes	0.1679	0.2041*	-0.012
p-value	0.1342	0.0676	0.9155

* 90% signficance level ** 95% signficance level

The lack of a consistent pattern between these measures and our quality variables confirms what others have found looking at Medicare markets that also show little relationship between overall spending and quality. This result is also supported by the recent work by Turbyville et al. (2011), which shows very little relationship between utilization and quality for commercial markets. The limited association between utilization and quality suggests that savings may be possible in commercial markets by shifting utilization, without an associated decline in the quality of services. However, there are a number of limitations to this analysis. First, we only look at simple correlations and there may be other explanatory factors affecting these relationships. Second, we have just a handful of quality measures that may not accurately reflect the true quality in the market. Third, these measures do not look at outcomes, which would be preferable measures of quality. Therefore, more work is necessary to determine whether there is a relationship between spending and quality in commercial markets.

8 Conclusion

Unlike Medicare markets where the service prices play a limited role in explaining variation in expenditures across markets, the variation in service prices in the commercial sector appear to be as important as utilization. Although there is variation in both service prices and service utilization measures across markets, aggregate differences in the typical expenditure for disease treatment are not that different from aggregate price indexes observed for other goods and services, with a coefficient of variation of around 0.10. However, this aggregate measure hides some of the underlying differences in spending across markets that are observed at the disease level. Focusing on variation on a disease-by-disease basis, we find that the coefficient of variation for the typical disease to be around 0.21.

The observed variation in spending across markets suggests the possibility of inefficiencies in some markets and potential savings. Several exercises are conducted to measure the economic importance of this variation. We find that the potential savings from controlling utilization on a disease-by-disease basis may have a substantial effect on overall expenditures. For example, those areas with the highest quartile of utilization for a disease may find savings of around 30 percent by adopting the utilization practices from the lower utilization areas. This implies large potential gains from controlling utilization, perhaps through bundled payments. However, lower utilization may not be caused only by differences in provider practice patterns across areas. Sensitivity to market price may also play an important role, since we find a negative correlation between utilization and prices across markets, so that lower utilization may partly be driven by higher service prices.

Although the observed variation in spending across markets suggests that there are potential inefficiencies, it is possible that greater utilization or service prices may be indicative of higher quality in commercial markets. However, in this study we find little correlation between measures of quality and each of the three indexes, which is consistent with the findings in the Medicare sector. This offers some evidence that spending may be reduced in commercial markets, either through reductions in price or utilization, without an accompanied loss in welfare.

There are some additional areas where more research may be beneficial. First, preliminary findings from this study show a negative and significant relationship between the SUI and the SPI measure. More work should be done to study if this negative correlation is spurious or whether it signifies an actual demand relationship. Second, it is unclear how closely these expenditures are actually linked to the amount consumers pay for their overall medical care. Profit margins and administrative costs of insurers may drive a wedge between the full price of services studied here (that is, the full amount paid to the provider by the insurer and the patient) and the out-of-pocket costs to

consumers (that is, the premium and the out-of-pocket costs). Third, the study between these indexes and quality may be greatly improved. For instance, future work may benefit from studying outcome measures of quality, rather than the process measures used in this paper. Fourth, this paper characterizes MCE, SPI, and SUI differences across markets, but does not attempt to explain underlying reasons for these differences. Future work may benefit by looking for explanatory causes for the observed variation across markets (e.g. Dunn and Shapiro (2011) and Dunn (2011)). While there are a number of areas to extend our research, we believe the framework presented here to define episode expenditures, service prices, and service utilization may be valuable for analyzing a variety of related topics in the future.

9 Bibliography

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] Anderson, Gerard and Peter Sotir Hussey. 2001 "Comparing Health System Performance in OECD Countries," *Health Affairs*.

[3] Baicker, Katherine, and Amitabh Chandra, (2004), "Medicare Spending, the Physician Workforce, and the Quality of Health Care Received by Medicare Beneficiaries", Health Affairs, Web Exclusive.

[4] Berndt, Ernst R., Anupa Birb, Susan H. Buschc, Richard G. Frank, , d 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

[5] Chandra, Amitabh and Douglas Staiger, (2007), "Testing a Roy Model with Productivity Spillovers: Evidence from the Treatment of Heart Attacks" *Journal of Political Economy*, 115(1)

[6] Congressional Budget Office (CBO), (2008), "Geographic Variation in Health Care Spending", The Congress of the United States, February

[7] Chernew, Michael, Lindsay Sabik, Amitabh Chandra, Teresa Gibson, and Joseph Newhouse, (2010), "Geographic Correlation Between Large-Firm Commercial Spending and Medicare Spending", *American Journal of Managed Care*, 16(2) pgs 131-138.

[8] 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.

[9] Cutler, David, and Louise Sheiner, (1999), "The Geography of Medicare", *American Economic Review: Papers and Proceedings* 89(2) pgs 228-233

[10] Dunn, Abe, Eli Liebman, Sarah Pack, and Adam Shapiro, 2010 "Medical Care Price Indexes for Patients with Employer-Provided Insurance: Nationally-Representative Estimates from MarketScan Data", Bureau of Economic Analysis, *Working Paper*

[11] Dunn, Abe. "The Demand for Medical Care: Measuring Price Sensitivity from Cross Regional Data" Working Paper

[12] Dunn, Abe and Shapiro, Adam. 2011. "Physician Competition and Medical Care Expenditures" Bureau of Economic Analysis, *Working Paper*

[13] Fisher, Elliot, David Wennberg, Therese Stukel, Daniel Gottlieb, F. Lucas, Etoile Pinder, (2003), "The Implications of Regional Variation in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care", Annals of Internal Medicine 138(4) pgs 273-287

[14] Fisher, Elliot, David Wennberg, Therese Stukel, Daniel Gottlieb, F. Lucas, Etoile Pinder, (2003), "The Implications of Regional Variation in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care", Annals of Internal Medicine 138(4) pgs 288-298

[15] Folland and Stano, (1989), "Sources of Small Area Variation in the Use of Medical Care", *Journal of Health Economics*, pgs 85-107

[16] Fuchs, Victor, Mark McClellan, and Jonathan Skinner, (2004), "Area Differences in Utilization of Medical Care and Mortality Among U.S. Elderly", Perspectives on the Economics of Aging, University of Chicago Press, Chapter 10

[17] Gottlieb, Daniel, Weiping Zhou, Yunjie Song, Kathryn Gilman Andrews, Jonathan S. Skinner, and Jason M. Sutherland, (2010), "Prices Don't Drive Regional Medicare Spending Variations", *Health Affairs*, 29(3) pgs 537–543.

[18] Medicare Payment Advisory Commission (MedPac), (2003), Report to the Congress: Variation and Innovation in Medicare, Chapter 1.

[19] Phelps, Charles E., (1995), "Welfare Loss from Variations: Further Considerations," *Journal of Health Economics* 14, pgs 253-260

[20] Phelps, Charles E., (1997), "Good Technologies Gone Bad: How and Why the Cost-Effectiveness of Medical Intervention Changes for Different Populations," Medical Decision Making, 17 pgs 107-117

[21] Shapiro I.L., Shapiro M., Wilcox D. (2001) "Measuring the Value of Cataract Surgery", in: D. M. Cutler and E. R. Berndt (Eds.), Medical Output and Productivity, University of Chicago Press Chicago. pp. 411-438.

[22] Turbyville, Sally, Meredith Rosenthal, L. Gregory Pawlson, and Sarah Scholle, (2011), "Health Plan Resource Use - Bringing Us Closer to Value-Based Decisions", *The American Journal of Managed Care*, 17(1) pgs e1-e7

[23] Wennberg, John, (1984), "Dealing with Medical Practice Variations: A Proposal for Action," *Health Affairs* 3, pgs 6-32

[24] Zhang, Yuting, Katherine Baicker, and Joseph Newhouse, (2010), "Geographic Variation in Medicare Drug Spending", *New England Journal of Medicine*, pgs 405-409

10 Appendix

10.1 Service-Specific Indexes

Table A5.1 Service MCE Across Service Types

	MSA	Overall MCE	Inpatient Hospital	Outpatient Hospital	Office - General MD	Office MD - Speciality	Pharmacy
1	Milwaukee-Waukesha-West Allis, WI	1.287	1.471	1.695	1.333	1.127	0.971
2	Oakland-Fremont-Hayward, CA	1.231	1.403	1.278	1.154	1.118	0.995
4	Minneapolis-St. Paul-Bloomington, MN-WI	1.206	1.478	1.073	1.498	1.126	1.005
5	Fort Worth-Arlington, TX	1.174	1.305	1.160	1.122	1.166	1.078
7	Indianapolis, IN	1.158	1.115	1.636	0.881	0.816	1.107
8	Dallas-Plano-Irving, TX	1.151	1.165	1.119	1.168	1.294	1.074
9	Houston-Sugar Land-Baytown, TX	1.150	1.187	1.222	0.992	1.128	1.073
10	Gary, IN	1.136	1.075	1.603	0.752	1.054	1.009
12	Peoria, IL	1.123	1.074	1.444	0.618	0.850	1.047
13	Denver-Aurora, CO	1.104	1.134	1.056	1.226	1.090	1.046
14	Miami-Miami Beach-Kendall, FL	1.097	1.133	1.069	1.264	1.254	0.863
15	Portland-Vancouver-Beaverton, OR-WA	1.091	1.143	1.163	1.494	1.023	0.872
65	Louisville, KY-IN	0.924	0.775	1.209	0.881	0.768	0.990
66	Las Vegas-Paradise, NV	0.918	1.077	0.582	1.133	0.983	0.875
67	Pittsburgh, PA	0.906	0.823	1.052	0.646	0.873	0.982
69	Knoxville, TN	0.892	0.699	0.842	0.890	1.067	1.058
70	Memphis, TN-MS-AR	0.892	0.810	0.701	0.960	1.059	0.842
71	Nassau-Suffolk, NY	0.887	0.928	0.510	0.759	1.338	0.863
72	Providence-New Bedford-Fall River, RI-MA	0.885	0.875	0.924	0.912	0.798	0.953
73	Kingsport-Bristol-Bristol, TN-VA	0.884	0.801	0.755	0.940	1.080	1.005
74	Augusta-Richmond County, GA-SC	0.879	0.877	0.949	0.659	0.952	0.789
76	Detroit-Livonia-Dearborn, MI	0.860	0.821	0.949	0.933	0.759	0.947
80	Grand Rapids-Wyoming, MI	0.836	0.762	0.897	1.306	0.594	0.959
81	Youngstown-Warren-Boardman, OH-PA	0.805	0.770	0.856	0.702	0.673	0.934
	standard deviation	0.099	0.175	0.247	0.248	0.184	0.078

Table A5.2 Service SPI Across Service Types

	MSA	Overall SPI	Inpatient Hospital	Outpatient Hospital	Office - General MD	Office MD - Speciality	Pharmacy
1	Oakland-Fremont-Hayward, CA	1.263	1.581	1.291	1.100	1.065	1.046
3	Milwaukee-Waukesha-West Allis, WI	1.228	1.219	1.333	1.331	1.318	1.035
4	San Diego-Carlsbad-San Marcos, CA	1.144	1.321	1.215	0.961	0.968	1.044
5	Dallas-Plano-Irving, TX	1.142	1.108	1.305	1.083	1.072	1.022
6	Peoria, IL	1.141	0.983	1.339	1.171	1.293	0.863
8	Fort Worth-Arlington, TX	1.121	1.145	1.254	1.056	1.069	1.011
9	Minneapolis-St. Paul-Bloomington, MN-WI	1.117	1.165	1.060	1.313	1.265	0.971
11	Indianapolis, IN	1.105	1.224	1.234	0.929	0.972	1.060
13	Houston-Sugar Land-Baytown, TX	1.102	1.073	1.155	0.993	0.994	1.052
14	Des Moines, IA	1.095	0.804	1.246	1.047	1.099	1.007
15	Charlotte-Gastonia-Concord, NC-SC	1.093	1.060	1.302	1.143	1.038	0.953
65	West Palm Beach-Boca Raton-Boynton Beach	0.931	0.880	0.936	0.816	0.894	0.989
67	St. Louis, MO-IL	0.925	0.777	0.884	0.939	0.932	1.072
69	Phoenix-Mesa-Scottsdale, AZ	0.925	0.960	0.848	0.920	0.914	1.041
70	Jacksonville, FL	0.920	0.876	0.905	0.929	0.969	0.991
72	Cleveland-Elyria-Mentor, OH	0.913	0.807	0.878	0.921	0.920	1.083
73	Grand Rapids-Wyoming, MI	0.912	0.826	0.624	1.049	1.014	1.030
76	Knoxville, TN	0.888	0.824	0.755	0.962	0.951	1.027
77	Detroit-Livonia-Dearborn, MI	0.885	0.764	0.732	0.969	0.972	1.059
78	Providence-New Bedford-Fall River, RI-MA	0.879	0.916	0.651	1.000	0.972	0.961
79	Kingsport-Bristol-Bristol, TN-VA	0.843	0.805	0.712	0.991	1.008	0.877
80	Pittsburgh, PA	0.837	0.816	0.630	0.905	0.914	1.047
81	Youngstown-Warren-Boardman, OH-PA	0.823	0.724	0.736	0.821	0.804	1.072
	standard deviation	0.087	0.152	0.198	0.122	0.118	0.069

Table A5.3 Service SUI Across Service Types

	MSA	Overall SUI	Inpatient Hospital	Outpatient Hospital	Office - General MD	Office MD - Speciality	Pharmacy
1	Gary, IN	1.147	1.149	1.169	1.129	1.157	1.137
3	Boston-Quincy, MA	1.122	1.089	1.143	1.115	1.124	1.132
4	Cambridge-Newton-Framingham, MA	1.115	1.071	1.133	1.122	1.140	1.132
5	Pittsburgh, PA	1.111	1.121	1.160	1.084	1.097	1.076
6	Minneapolis-St. Paul-Bloomington, MN-WI	1.081	1.139	1.089	1.024	1.048	1.077
9	Wichita, KS	1.063	1.131	1.056	1.032	1.015	1.098
11	Denver-Aurora, CO	1.062	1.078	1.057	1.050	1.074	1.072
12	Philadelphia, PA	1.060	1.062	1.061	1.063	1.052	1.093
13	Phoenix-Mesa-Scottsdale, AZ	1.052	1.115	1.058	1.048	1.057	1.013
14	St. Louis, MO-IL	1.051	1.072	1.060	1.048	1.045	1.049
15	Kingsport-Bristol-Bristol, TN-VA	1.050	1.050	1.066	1.047	1.050	1.047
67	Charlotte-Gastonia-Concord, NC-SC	0.939	0.932	0.930	0.953	0.936	0.985
69	San Antonio, TX	0.934	0.956	0.931	0.941	0.937	0.935
70	Raleigh-Cary, NC	0.933	0.955	0.914	0.937	0.933	0.964
71	San Diego-Carlsbad-San Marcos, CA	0.933	0.937	0.923	0.945	0.962	0.936
74	Augusta-Richmond County, GA-SC	0.916	0.928	0.914	0.938	0.914	0.922
76	Atlanta-Sandy Springs-Marietta, GA	0.909	0.902	0.896	0.937	0.922	0.934
77	New York-White Plains-Wayne, NY-NJ	0.907	0.874	0.877	0.939	0.916	0.960
78	Memphis, TN-MS-AR	0.907	0.907	0.923	0.943	0.928	0.867
79	Des Moines, IA	0.897	0.895	0.893	0.899	0.889	0.952
80	Nassau-Suffolk, NY	0.884	0.864	0.856	0.916	0.889	0.940
	standard deviation	0.059	0.069	0.069	0.051	0.059	0.056

10.2 Quality Measures

We construct six quality measures from the claims data using methods outlined by National Committee for Quality Assurance (NCQA). We focus on quality measures that may be constructed from administrative claims data. These quality measures are described in greater detail in NCQA Measure Technical Specifications.[30]

1. *Persistence of Beta-Blocker Treatment after a Heart Attack* - The percentage of patients 18 years of age and older during the measurement year who were hospitalized and discharged alive from July 1 of the year prior to the measurement year to June 30 of the measurement year with a diagnosis of acute myocardial infarction (AMI) and who received persistent beta-blocker treatment for six months after discharge (See Page 38).

2. *HbA1c Test for Pediatric Patients* - Percentage of pediatric patients with diabetes who had an HbA1c test in a 12-month measurement period (See Page 27).

3. *Complete Lipid Profile* - Percentage of patients 18 years and older with ischemic vascular disease who had a complete lipid profile (See Page 45).

4. *Use of Imaging Studies for Low Back Pain* - The percentage of patients with a primary diagnosis of low back pain who did not have an imaging study (plain X-ray, MRI, CT scan) within 28 days of the diagnosis (Page 12).

[30] http://www.ncqa.org/tabid/59/Default.aspx

5. *Cervical Cancer Screening* - The percentage of women 21–64 years of age who received one or more Pap tests to screen for cervical cancer. (within a 3 year period). (Page 94).

6. *Breast Cancer Screening* - The percentage of women 40–69 years of age who had a mammogram to screen for breast cancer. (within a 2 year period). (Page 92).

10.3 Robustness Checks

To check the robustness of the results presented in this paper, we generated the tables presented here under a number of alternative assumptions (Tables 2 through 7). The following is a list of robustness checks. Unless noted, no qualitative results changed.

1. Isolated 2006 and 2007 separately. The results are quite similar in each of the separate years. The key advantage from combining years is that we are able to use more observations from each MSA.

2. For selecting diseases to be included in the sample we alter the threshold for the number of episodes observed in the data. Recall that the threshold applied here was 40,000. Similar results are attained if the threshold falls to 20,000 (accounting for 95 percent of spending) or is increased to 200,000 (accounting for 66 percent of spending). The problem with a lower threshold is that there will not be a sufficient number of episodes to attain an accurate price. In contrast, the problem with too high a threshold is that it accounts for a more limited fraction of overall spending.

3. Removed the population weights.

4. Adjusted the sample of cities by changing the threshold from 20,000 to 30,000 (dropping 20 cities).

5. Dropped the top and bottom 2.5 percent of episodes based on the episode expenditure price. Similar results are also found when the top and bottom 5 percent of episodes are dropped.

6. Aggregated diseases to the Major Practice Category (MPC) level.