NBER WORKING PAPER SERIES

ATTRIBUTING MEDICAL SPENDING TO CONDITIONS: A COMPARISON OF METHODS

David Cutler
Kaushik Ghosh
Irina Bondarenko
Kassandra Messer
Trivellore Raghunathan
Susan Stewart
Allison B. Rosen

Working Paper 25233
http://www.nber.org/papers/w25233

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2018

We are grateful to the National Institute on Aging (P01-AG031098 and R37AG047312) for research support and to Barry Bosworth and seminar participants for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w25233.ack

Attributing Medical Spending to Conditions: A Comparison of Methods
David Cutler, Kaushik Ghosh, Irina Bondarenko, Kassandra Messer, Trivellore Raghunathan, Susan Stewart, and Allison B. Rosen
NBER Working Paper No. 25233
November 2018
JEL No. I1

ABSTRACT

Partitioning medical spending into conditions is essential to understanding the cost burden of medical care. Two broad strategies have been used to measure disease-specific spending. The first attributes each medical claim to the condition listed as its cause. The second decomposes total spending for a person over a year to the cumulative set of conditions they have. Traditionally, this has been done through regression analysis. This paper makes two contributions. First, we develop a new method to attribute spending to conditions using propensity score models. Second, we compare the claims attribution approach to the regression approach and our propensity score stratification method in a common set of beneficiaries age 65 and over drawn from the 2009 Medicare Current Beneficiary Survey. Our estimates show that the three methods have important differences in spending allocation and that the propensity score model likely offers the best theoretical and empirical combination.

David Cutler
Department of Economics
Harvard University
1875 Cambridge Street
Cambridge, MA 02138
and NBER
dcutler@harvard.edu

Kaushik Ghosh
NBER
1050 Massachusetts Ave.
Cambridge, MA 02138
ghoshk@nber.org

Irina Bondarenko
University of Michigan
ibond@med.umich.edu

Kassandra Messer
University of Michigan
kasey@umich.edu

Trivellore Raghunathan
University of Michigan
teraghu@umich.edu

Susan Stewart
NBER
1050 Massachusetts Ave
Cambridge, MA 02138
sstewart@nber.org

Allison B. Rosen
Department of Quantitative Health Sciences
University of Massachusetts Medical School
368 Plantation Street, AS9-1083
Worcester, MA 01605
Allison.Rosen@umassmed.edu

I. Introduction

For many purposes, it is important to attribute medical spending to particular conditions. For example, researchers and policy makers often ask questions such as: Does spending align with the burden of disease, or are the two independent? How much of medical spending growth is associated with the rising prevalence of chronic disease? Each of these questions requires a condition-based look at spending. In this paper, we compare different methodologies for attributing spending to medical conditions.

Cost of illness studies have a long tradition in health economics, dating back to the 1960s (Scitovsky, 1967; Rice and Horowitz, 1967; Rice, 1967). The fundamental difficulty with attributing spending to particular conditions is comorbidities. If all people had only one medical condition at a time, it would be easy to measure condition-based spending. When people have multiple conditions, however, this becomes more difficult. If a person has a heart attack from which recovery is slow, is the extra spending during the post-acute period a result of an abnormally slow recovery, or might it result instead from pre-existing mental illness, which makes following a recommended medication and lifestyle pattern more difficult? One needs some type of disease attribution methodology to answer this question.

Traditionally, studies in this literature assign each claim to one or more conditions (e.g., Cooper and Rice, 1976; Berk, Paringer and Mushkin, 1978; Koopmanschap, 1995; Hodgson and Cohen, 1999; Leon and Neumann, 1999; Druss et al., 2001, 2002; Cohen and Krauss, 2003; Thorpe et al., 2004a, 2004b, 2005, 2006, 2007, 2010, 2013; Roehrig et al., 2009, 2011; and Starr et al., 2014). In the case of the heart attack example noted above, spending in the post-acute period would be allocated to either cardiovascular disease or mental illness based on which diagnosis the physician recorded as being the primary cause of the visit – or perhaps split

between multiple recorded diagnoses. The difficulty with this methodology is that physicians do not necessarily solve the attribution question well. If the patient is re-hospitalized because recovery from the heart attack is slow, the physician will (properly) record the new hospitalization as caused by cardiovascular disease, even though that is just the symptom of another problem.

Thus, there is a simultaneous history of researchers exploring other ways to measure cost for medical conditions. Typically, regression analysis is used to relate annual spending at the person level to a set of medical conditions that a person has in that calendar year (e.g., Dudley et al.,1993;Koopmanshap, 1998; Bloom et al., 2001; Sturm, 2002; Finkelstein et al., 2003a, 2003b, 2009; Akobundu et al., 2006; Trogdon et al., 2008; Honeycutt et al., 2009, 2013; Rosen and Cutler, 2009; Roberts et al. 2010; Gregori et al., 2011). The resulting coefficients are then used to attribute spending to conditions.

Three issues come up in this type of analysis. First, most of the models for spending are non-linear – often non-linear least squares or generalized linear models – and some method must be used to attribute spending to conditions in a way that neither overcounts nor undercounts total dollars. Trogdon et al. (2008) develop an attributable fraction methodology that does this, but it can be very computationally intensive for models with many conditions.

Second, these regressions typically have a large unexplained component – the constant term and other covariates. It is not clear what condition to assign that spending to. Third, regression models make parametric assumptions, and the results can be sensitive to violation of these assumptions. As is well known, the ordinary least squares(OLS) method is not well suited to handle expenditure data with common data problems like heteroskedasticity, heavy tails, and

large outliers (Manning et al. 1998, 2001, 2005), Basu and Manning 2009; Jones, 2000; Zaslavsky and Buntin, 2004; Cantoni and Ronchetti, 2006).

This paper has two goals. Our first goal is to develop a method of spending attribution to conditions that is more robust than claims and regression methods. We propose a two-step propensity score methodology to do this. The first step uses a propensity score method to compare people with a condition to observably similar people without that condition. The difference in spending between those individuals is an estimate of the condition-specific cost. In the second step, we propose a non-additive framework that models total spending as a non-linear function of attributed costs and a number of comorbidities. This allows us to incorporate the overall number of comorbidities into the estimation of the disease-specific spending. This method also attributes spending to all medical conditions without any residual.

We then compare three methods of allocating spending to conditions: a claims-attribution approach along the lines of the earlier literature; a traditional regression specification; and the propensity score methodology. The data that we use is the same for all three methods: the 2009 Medicare Current Beneficiary Survey (MCBS).

The results show significant differences between the approaches. The claims approach differs the most from the two other approaches in attributing more of the cost to acute medical conditions (e.g., a heart attack) relative to the regression or propensity score methods, which attribute more spending to comorbid conditions (e.g., mental illness, diabetes, hypertension, hyperlipidemia). In contrast to the regression approach, the propensity score approach benefits from not having a large unexplained component.

We also use a variety of statistical techniques to compare these models. In general, the propensity score model does better at matching the distribution of individual-level spending than

does the regression or claims-based model. In addition, the propensity score model has a lower out-of-sample mean squared error for predicted spending. For this reason, we believe that the propensity score model is likely the best way to measure condition-specific medical spending.

The paper is organized as follows. The next section discusses the data we employ. The third section presents the different methodologies for measuring condition-specific costs, using our specific data to highlight the issues in each case. The fourth section shows the results comparing spending using the different methodologies. The last section concludes.

II. Data

Our primary data source is the 2009 Cost and Use sample of the Medicare Current Beneficiary Survey (MCBS).[1] We restricted our sample to the population aged 65 and older since the MCBS is nationally representative for the older population. Importantly, the MCBS includes both community and institutionalized people. The MCBS has information on survey-reported events, supplemented by Medicare claims.

To define conditions, we started with the 259 Clinical Classification Software categories delineated by the Agency for Health Care Research and Quality. Not all of these conditions have a high prevalence in the elderly, and some combine categories that might usefully be disaggregated for this group. After combining and disaggregating, we determined *ex-ante* a set of 105 conditions: 98 diagnosed conditions; 3 undiagnosed conditions (high cholesterol, high blood pressure, and diabetes), and 4 cancer screening variables (colon cancer, cervical cancer, breast cancer and prostate cancer).

[1] The MCBS has two samples: a set of people who were enrolled for the entire year (the Access to Care sample) and a set of ever-enrolled beneficiaries (the Cost and Use sample). The latter differs from the former in including people who die during the year and new additions to the Medicare population. The Cost and Use sample adds up to national spending totals, so we use those data.

In any clinical taxonomy, there is always a residual category of conditions that are difficult to disaggregate. We defined a category of "other conditions" that includes signs and symptoms (for example, routine general medical examination at health care facility), residual conditions and unclassified conditions (for example, recurrent hypersomnia), and some E codes such as accidents).

In a preliminary analysis, several of the calibrated conditions had too low a prevalence to meaningfully estimate their cost in the elderly. Thus, for the purpose of cost estimation, we combined these 105 conditions into 78 conditions with a larger sample size. The appendix reports the prevalence of each condition (Raghunathan T, et al., 2017).

While MCBS has many strengths, it also has some limitations, for which we developed solutions. We sketch the issues and solutions here and provide more detail in the Appendix. First, MCBS has incomplete or no claims information for beneficiaries enrolled in Health Maintenance Organizations (HMOs). We reweighted the fee-for-service population to compensate for this exclusion, matching HMO enrollees to comparable fee-for-service beneficiaries. Second, we made an adjustment for the difference between survey spending and national spending estimates, along with the lines of Selden et al. (2001) and Sing et al. (2006). We began by removing expenditures from the NHEA for goods and services which were out of the scope of the survey data. We then reallocated spending across service categories so that total spending across service categories would match. To ensure that the reallocated spending matched national totals, we proportionately increased spending in each category so that total survey spending equaled the NHEA estimate of spending for the elderly. The aggregate adjustment was to increase survey spending by 11%, roughly evenly spread across service categories (Rosen A, et al., 2017).

Finally, the disease prevalence rates reported in MCBS for some conditions (for example, hypertension and hyperlipidemia) are lower than those observed in surveys such as the National Health and Nutritional Examination Survey (NHANES), which collects laboratory results. We used the NHANES to calibrate the prevalence of conditions in MCBS to more accurately reflect national levels. We term the resulting condition list the "calibrated medical conditions." The imputation procedure produced five imputed data sets. In our empirical analysis, we estimated the models for each of the five imputed data sets. We used proc mianalyze in SAS 9.4 to combine means and standard errors from the five imputed data sets using standard combination rules (Little and Rubin, 1987; Li, Raghunathan and Rubin,1991).

The distribution of per-person spending for the resulting data set is shown in the first part of figure 1 and the first row of table 1. Mean per capita spending on medical care for the elderly was $17,479 in 2009. The median is much lower: $7,281; as is well known, there is a long right tail in spending, which can be seen in figure 1 and in the last two columns of table 1.

III. Methods for Attributing Spending to Conditions

There are two fundamental methodologies in the literature for attributing spending to conditions: a claims-based method that assigns spending for particular claims to one or more conditions coded as the reason for the medical visit, and a regression method that uses total spending over a period of time and a set of medical conditions to decompose spending into attributable conditions. We discuss each method in turn, highlighting how we implement each.

Before we present our methodologies, we note one feature of the analysis. We seek to estimate the partial effect of each condition on spending, controlling for all other conditions that a person has. That differs from the total effect that spending on a disease might lead to,

including its effects on the prevalence of other conditions. As an example, suppose that physicians treat hypertension more aggressively than in the past. That will raise spending on hypertension. But it might lower overall spending if there are offsets in fewer heart attacks or strokes. Our estimates will give us the additional spending on hypertension since we hold constant the other cardiovascular conditions a person has. To estimate the total cost impact of spending more on any condition, we would need a set of equations for how each condition affects the incidence of others. We do not explore these inter-linkages in this paper.

A. CLAIMS-BASED METHOD

In the claims-based method, claims for which there is only one diagnosis are easy to assign to a disease. However, most medical claims have more than one condition associated with them. A typical claim in Medicare data contains a principal diagnosis code (ICD9-CM) and several other secondary diagnoses codes (sometimes up to 14). For example, a patient hospitalized for a heart attack will likely also be diagnosed with other comorbid conditions, such as congestive heart failure (CHF), diabetes, hypertension, or hyperlipidemia. The question to address is how to estimate the amount spent on treating each of the conditions.

In the first claims-based studies, developed in the 1960s, spending for each claim was allocated to the principal diagnosis. Using data from the 1960s, Rice (1967) estimated that the most costly conditions were diseases of the digestive system; mental, psychoneurotic and personality disorders; and diseases of the circulatory system. The most recent work following the claim-attribution methodology is by Thorpe et al. (2004a, 2004b, 2005, 2006, 2007, 2010, and 2013) and Roehrig et al. (2009, 2011) and uses data from the Medical Expenditure Panel Survey over the 1990s and 2000s. To address the problem of multiple conditions, Thorpe et al.

(2004a) estimate an "upper bound" (attributing total spending for each health care event to all the conditions reported), a "lower bound" (summing spending from each medical event for which only a single condition is reported) and a "best guess", where claims with more than one listed condition are divided using the relative spending of each condition in the single condition claims. Using the 259 disease classifications in the Clinical Classification System (CCS) (Elixhauser, Steiner, and Palmer, 2014) to define conditions, Thorpe et al. (2004a) estimate that the conditions accounting for the most spending are heart disease, pulmonary disease, and mental disorders. The finding that heart disease is the most costly condition has been mirrored in other recent studies (Roehrig et al., 2009).

Our version of the claims-based attribution method follows Thorpe's best guess, making adjustments as appropriate. To start, we note that this method is based on the observed claims for each person, not the calibrated claims that we derived– except as noted below. In each set of MCBS claims files (hospital inpatient, hospital outpatient, carrier, hospice, home health, skilled nursing facility and durable medical equipment), we first identified claims with only one listed condition. Table A6 in the Appendix gives detail as to whether each medical condition has claims that satisfy the "single condition claim" criteria. Out of the seven different types of claims, only carrier/physician and hospital outpatient claims have "single condition claims" – that is, a claim with single diagnosis code listed in it – for all 98 medical conditions (excluding the 3 undiagnosed conditions) and 4 cancer screening variables. In these data files, we estimated the average cost for single claims for all 98 medical conditions and 4 screening variables.

We used this information to apportion a cost to medical conditions for multiple condition claims in these two files. For example, if the average cost of condition 'a' in the single condition claims file is c_a, and the average cost of condition 'b' in the single condition claims file is c_b, a

claim that listed both conditions 'a' and 'b' would be allocated $c_a/(c_a+c_b)$ to condition 'a', and $c_b/(c_a+c_b)$ to condition 'b'.

Unfortunately, the "single-condition claim" criterion fails for inpatient, skilled nursing facility, home health, durable medical equipment, and hospice claims. In the inpatient file, only 13 (e.g., breast cancer, prostate cancer) out of 102 medical conditions (excluding 3 undiagnosed conditions) satisfy the "single condition claim" criteria. In skilled nursing facility file, only 36 (e.g., tuberculosis, HIV) out of 102 medical conditions satisfy the "single condition claim" criteria. In the home health, hospice, and durable medical equipment files, 20, 40 and 59 conditions respectively satisfy the "single condition claim" criteria. This is one of the biggest limitations of claims-based approach.

To proceed, we used a methodology somewhat akin to the multiple condition methodology presented above. We began with a sample of Medicare inpatient claims and assigned each claim to one of the 78 conditions discussed above based on the principal diagnosis of the claim. For each condition, we then averaged the DRG (diagnosis-related group) weight for all of the associated claims. This gave us a rough measure of severity for the condition. We used these relative severity weights to apportion the original claims to the multiple listed diagnoses. For example, suppose there was a claim for $500 in the inpatient hospital file and two conditions listed are diabetes and hypertension. Suppose further that in the full set of claims for which diabetes was the principal diagnosis, the average DRG weight was 1.52, and in the set of claims for which hypertension is the principal diagnosis, the average severity weight is 0.71. We would assign 1.52/(1.52+0.71)*$500 to diabetes and 0.71/(1.52+0.71)*$500 to hypertension.

The prescription drug data in MCBS have no diagnosis codes listed. Therefore, we identified all medical conditions the person was treated for in 2009 using the diagnosis codes

listed in their claims: inpatient, outpatient, skilled nursing facility, carrier, hospice, home health and durable medical equipment. Each condition was counted once and assigned a DRG weight based on inpatient admissions with that condition, as above. We then apportioned total prescription drug spending for the year based on the share of these DRG weights to each condition.

Finally, for 5.3% of beneficiaries there were dollar amounts in the personal summary file(s), but no claims.[2] For these beneficiaries, we used the calibrated claims that we derived using the NHANES instead of actual claims and assigned dollars to calibrated medical conditions on that basis. The result of this process was a complete decomposition of medical spending into 74 medical conditions and 4 cancer screening variables.

Table 2 reports the total attributed spending by multi-level clinical classification system (CCS) categories, and table A9 in the appendix reports spending for each of the 78 medical conditions. Diseases of the circulatory system were the most expensive category, costing $148.1 billion and accounting for 23 percent of all personal health care expenditure. Diseases of the nervous system and sense organs cost $64 billion, accounting for 10 percent of aggregate personal health care. The next three most costly disease categories were diseases of the musculoskeletal system and connective tissue ($62.2 billion, 10 percent); diseases of the respiratory system ($53.4 billion, 8 percent); and endocrine; nutritional; and metabolic diseases and immunity disorders ($49.2 billion, 8 percent). Together, these five disease groups accounted for almost 59% of all personal health care spending. At the level of specific conditions, the most expensive conditions were acute events with expensive and often prolonged hospitalizations:

[2] Person Summary or Service Summary files are created from the event-level files. The event-level files include self-reported events, some of which are not paid under Medicare FFS, so there will be no corresponding claims.

acute myocardial infarction, cardiac arrest, hematologic cancers, lung cancer, and acute renal failure.

One way to understand these results is to take the disease-specific spending and add it up to the individual level, using the conditions diagnosed for each patient. The result is predicted spending at the individual level, which can be compared to observed spending in the population. Figure 1 and table 1 report the distribution of the person-level expenditures from the claims model. By construction, mean person-level spending is the same as in the observed data. Median spending is significantly higher with the claims-based attribution: $15,006, compared to $7,281 in the observed data. Effectively, the cost of high spending outliers is attributed to some diseases, and these diseases are distributed throughout the population. Thus, the median individual is estimated to be spending more than is true in practice whereas the high end spends less. This is true to some extent for all of the methods we examine.

For people with no claims, spending is zero; figure 1 shows that the claims-based method effectively reproduces the share of people with zero spending.

B. REGRESSION APPROACH

The second method of attributing spending to medical conditions is to use regression analysis. Conceptually, the regression model relates total spending over the year to the full set of conditions that a person has. The coefficients can then be used to find spending for each condition. A typical equation is of the form:

$$\log(y_i + 1) = X_i\alpha + D_i\beta + \epsilon_i$$

A good deal of work has considered how to model medical spending. Issues such as the non-normal distribution of spending and heteroscedasticity are clearly important (Duan et al.

1983; Manning 1998; Jones, 2000; Manning et al., 2001, 2005; Zaslavsky and Buntin, 2004; Basu and Manning, 2009). The most common solutions to heavy tails are: (1) transformation to deal with skewness (e.g., ordinary least square on ln(spending+1)) and use of smearing factors for retransformation; and (2) different weighting approaches based on exponential conditional model (ECM) and generalized linear model (GLM) approaches.

Following Manning and Mullahy (2001), we explored several GLM estimation techniques. First, (1) we fit models that are variants of generalized linear models (GLM) for spending with a Gamma distribution and log link function. We also tried (2) a GLM model with log(y+1) as the dependent variable with a Gaussian distribution and identity link. Other models included (3) a cubic root model (cubic root of cost) and (4) a Box-Cox model.

Some studies make an additional adjustment for people with no medical spending, for example using a two-part model (Duan 1984, Mihaylova et al., 2011, Belotti et al., 2015): one equation for the probability of positive spending and the second for the amount spent. We estimate two-part models assuming the probability that a beneficiary has positive health care spending is a probit. For people with non-zero spending, an OLS or GLM regression is run with the same set of covariates as in the probit model (Belotti et al., 2015).

We tried several two-part models with the following specifications: (5) generalized linear models (GLM) for spending with a Gamma distribution and log link function; (6) log(spending) as the dependent variable with a Gaussian distribution and identity link, and finally (7) a cubic root model (cubic root of cost). We could not estimate the two-part Box-Cox model as the STATA software used in the estimation does not support Box-Cox (Belotti et al., 2015). In our data, only 2% of the sample has no spending in the year. This is not surprising

given the age of the population and its enrolment in Medicare. For this reason, we believe that the two-part models do not offer a huge improvement over one-part models.

A third concern that arises in the regression model is that conditions can be associated with decreased spending – that is, there is no constraint that the coefficients on the conditions are all positive. Generally, negative coefficients are ascribed to diagnostic mismeasurement (Hall and Highfill, 2013a, 2013b). When relatively healthy people see physicians, physicians still need some condition to code as the cause of the visit. In such circumstances, the physician may code a relatively benign condition (hypertension or high cholesterol, for example) that is common in the population and straightforward to justify. For patients with a variety of conditions, in contrast, physicians may focus on more acute conditions and not record risk factors that have little immediate impact on health. Effectively, this will lead to lower spending on less acute conditions than would be true in practice.

Some studies address this concern by forcing the coefficients to be positive. For example, one could model spending as:

$$\log(y_i + 1) = X_i\alpha + D_i\exp(\beta) + \epsilon_i$$

In this equation, the coefficient on spending (exp(β)) can never be negative. In practice, however, this equation does not always converge; the β's approach -∞, effectively wanting some conditions to not affect spending. Other studies address the issue by defining hierarchical categories to pull out difficult cases. For example, rather than coding for hypertension, the regression might include categories for 'hypertension without other conditions', 'hypertension with heart disease', and so on (Pope et al., 2004). Conditions are often defined so that spending on every condition is positive. For our purposes, we do not wish to change the set of conditions

across estimation methodologies. Thus, we let negative coefficients be negative and proceed as if that is the true impact of the condition.

To implement the regression models, we include as independent variables the 78 calibrated condition and screening indicators and other covariates that are expected to influence medical costs. These include: health status compared to one year ago (somewhat better/much better, about the same, somewhat worse/much worse); general health compared to others same age (excellent, very good, good, fair, poor); death in calendar year, any difficulty lifting/carrying 10 pounds; any difficulty in stooping/crouching/kneeling; any difficulty walking 1/4 mi. or 2-3 blocks; any difficulty dressing; any difficulty eating; private health insurance coverage for the calendar year; days in institution; count of inpatient nights; count of inpatient stays; current smoker; body mass image; age dummies; education dummies (<9 years, 9-11 grade (including 12 with no diploma); high school grad/GED or equivalent, some college or AA degree; college grad); ever served in armed forces; sex; marital status (married, widowed, divorced, separated, never married); poverty category (poor, near poor, low income, middle income, high income); race and ethnicity dummies (white, black, other; Hispanic ethnicity).

As noted in section 1, to form an accurate diagnosis profile for each individual, we imputed prevalence of diseases from NHANES using multiple imputations. This produces 5 imputed datasets, denoted D^m, for m=1,…,5. We then estimate five regressions, of the form:

$$\log(y_i + 1) = X_i\alpha^m + D_i^m\beta^m + \epsilon_i^m$$

where the regression is estimated using HMO-adjusted survey weights. We combine these estimates and calculate appropriate standard errors using standard techniques (Little and Rubin, 1987; Li, Raghunathan and Rubin,1991). Table A7 in the appendix shows the coefficient estimates from all five imputed data sets using the one-part model with the Gaussian distribution

and identity link. Although they are not the primary object of our analysis, we remark briefly on the estimates for the demographic characteristics. Generally, these are as expected. People who die spend more, as do people who have difficulties with ADLs or IADLs. Among the more surprising findings are that men spend less than women (conditional on their health conditions and health status), as do smokers. Importantly, these covariates suggest that there is more to spending than just the clinical factors.

From our estimates in appendix table A7, two of the condition groups have negative coefficients for all five of the replicates: Deep Vein Thrombosis (DVT) and Acute Renal Failure. These are relatively less common (6% and 9% respectively), but the claims-based approach attributed $2.2 billion and $7.7 billion to them respectively. We set spending to zero for those two conditions (as in Hall and Highfill, 2013a, 2013b). In some cases, coefficients were negative for only a selection of the five multiples. In this case, we treated those variables as missing and averaged the coefficients over the remaining replicates.

We did an extensive diagnosis of the residuals from all of the models to choose the most appropriate regression specification, shown in figures A3-A6 in the appendix. Figure A3 shows the kernel density plot comparison of the residuals (in blue) along with the normal density (in red). Figure A4 shows the standardized normal probability plots of the residuals. Figure A5 shows the quantile of residuals relative to the normal distribution. Figure A6 shows the scatter plot of the residuals relative to per person spending.

All of these residual plots use the in-sample estimates. An alternative way to measure fit is to use out-of-sample predictions. Table 3 shows the root mean square error (RMSE) based on out-of-sample predictions; we estimate the model in a random half of the data (N=3,100) and predict spending in the other half.

The out-of-sample predictions show that the one-part model with log(spending+1) as the dependent variable and with a Gaussian distribution and identity link had the lowest RMSE**.** Figure A3 shows that this model also fit well. The one and two part cubic-roots were second best. Surprisingly, the Gamma models with log link performed relatively poorly. These models are sensitive to high-spenders with big residuals – as typified by people in nursing homes.

The log specification implicitly assumes that each disease has a multiplicative effect on spending. An additional transformation is needed to turn this into dollars, and to ensure that the average dollars spent will match the known total. We follow a methodology described by Trogdon et al. (2007, 2008) which estimates expenditures associated with co-occurring diseases and reallocates these expenditures to individual diseases. For notational simplicity, we omit the m superscript denoting each multiply imputed dataset. In practice, we estimate the equations below for each data set and then average the estimates as noted above. We start by calculating the fraction of spending attributable to all of the individual's illnesses (AF_i):

$$\widehat{AF_i} = \frac{\exp\left(A_i\hat{\alpha} + D_i\hat{\beta}\right)\cdot\hat{S} - \exp(A_i\hat{\alpha})\cdot\hat{S}}{\exp\left(A_i\hat{\alpha} + D_i\hat{\beta}\right)\cdot\hat{S} - 1},$$

where $\hat{S} = \frac{1}{n}\sum_{i=1}^{n}\exp(\hat{\epsilon}_i)$ is a smearing factor to estimate the average error term. Then, we calculate the share $\left(S_{ij}\right)$ of spending by each individual that can be attributed to each diagnosed condition:

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

We then disaggregate the observed cost at the individual level by taking the product of the attributable fraction, the spending share for that condition, and the observed cost per person:

$\widehat{AC}_{ij}^{(T)} = \widehat{AF}_i \cdot \hat{S}_{ij} \cdot Cost_i$. To form an average cost per condition, we average spending over the subjects with the condition of interest.

$$\widehat{AC}_j^{(T)} = \frac{\sum_i \widehat{AC}_{ij}^{(T)} I\{i,j=1\}}{\sum_i I\{i,j=1\}}$$

The second column of table 2 shows spending by multi-level CCS category, and table A9 in the appendix shows the spending attributed to each of the 78 medical conditions. The regression estimates differ in many ways from the claims-based estimates. "Other conditions" are much more important in the regression approach, accounting for 19 percent of spending compared to 6 percent in the claims approach. Indeed, other spending is the leading cause of spending in the regression approach. Subsequent analysis suggests that this is largely due to high attributed spending for signs and symptoms. Circulatory disease is second in spending (19 percent).

The third most expensive condition is not an actual condition; it is the unexplained component – the constant term and other covariates. This picks up spending that the regression does not attribute to any condition. Unfortunately for the regression approach, this is very large – $58 billion, or 9 percent of total spending. The importance of this unattributed spending is one argument against the regression approach.

Endocrine, nutritional, and metabolic diseases and immunity disorders (8 percent) and diseases of the nervous system and sense organs (7 percent) round out the top five. We present more comparison below of how the regression and claims-based approaches estimate disease-specific spending.

To evaluate the distribution of individual spending resulting from this set of average costs, we calculated the predicted spending for each person as the sum of averages cost for each of the diagnosed conditions: $\hat{Y}_i^{(T)} = \sum_j \widehat{AC}_j^{(T)} D_{ij}$.

Figure 1 and table 1 report distribution of the estimates from the regression-based model. The regression approach has less variability in predicted spending than the claims-based (and propensity score) approaches. Effectively, the regression approach has a difficult time determining the cost of different conditions (as typified by the unexplained component), and thus finds less variation in predicted spending across individuals. A related byproduct of this is that the regression approach greatly understates the share of people with zero and very low spending. Figure 1 shows essentially no mass at very low spending with the regression approach.

C. PROPENSITY SCORE METHOD

Our proposed propensity score methodology consists of two steps: first, estimating the average difference in spending for individuals with and without each disease and second, incorporating the number of conditions and other measures of the amount of health care received to refine the estimates of cost attributed to each condition. We repeated each step using each multiply imputed dataset and average them as above. For notational ease, we omitted the m superscript in our presentation.

To estimate the average difference in spending AC_j for individuals with and without the condition of interest D_j, we began by estimating the probability of having condition D_j, conditional on the available covariates and all other conditions; this estimated probability is known as the propensity score. We computed the propensity score by fitting separate logistic regression models for each condition, where D_j is the outcome. The covariates used in the

propensity score models were the same as the covariates used in the regression approach. For each propensity score model, we normally included the other 77 calibrated health conditions as a part of the X vector, but excluded conditions that had a deterministic or extremely tight correlation with the condition of the interest. For example, in estimating the propensity of hypertension, we did not condition on undiagnosed hypertension (the two are mutually exclusive). Similarly, in estimating the probability of mammogram screening, we did not condition on breast cancer (all people with breast cancer have mammograms). Also, the propensity of gender-specific diseases such as breast and prostate cancer were estimated only for the relevant gender. We estimated the probability of death in a given year and then used the estimated probability of death as a covariate in the propensity score models for the 78 calibrated health conditions. We could not directly include death (as a binary variable) in the propensity score models, as there was small fraction of decedents (roughly 5%) in any calendar year, and including it in the logistic regression models with certain rare conditions as the outcome led to separation of cases and controls, causing convergence issues with the model estimates. We did not use spending as a covariate in any of our propensity score models.

After computing the propensity scores for a disease category, subjects were grouped into five homogeneous strata based on their propensity scores, using the quintiles of the propensity score as cutoffs to determine groups. We then estimated the mean difference in expenditures between cases and controls within each stratum, by simply taking the difference in average costs between those with and without condition D_j within each stratum. Finally, we combined the estimated difference in expenditures within each strata, weighting by the population size in each stratum, to obtain an estimate of healthcare expenditure attributable for that disease category. Using the propensity score to divide subjects into strata or subclasses, as described by

Rosenbaum and Rubin (1983), we could directly compare medical expenditures for individuals with D_j and those without D_j within groups containing same covariate profiles, enabling us to adjust for comorbidities and other covariates. Creating five strata using propensity score was shown to be an effective method for applying causal inference to observational studies, removing around 90% of the bias in the difference between the cases and controls caused by covariate imbalance (Rosenbaum and Rubin 1984). Thus, our estimates are for expenditures attributable to each condition, adjusted for all other conditions. This process of stratification and estimating the attributable spending was done separately for each of the 78 conditions.

We performed the Hosmer-Lemeshow goodness of fit test to assess our logistic regression models used in computing the propensity scores (Hosmer and Lemeshow 2000). Furthermore, we compared the distribution of covariates for those with the condition and those without within each propensity score stratum in order to ensure a balance of covariates within each stratum. We found good overlap in the covariate profiles between cases and controls to allow us to properly estimate the attributable costs for each condition.

From our propensity score models, we obtained estimates for the average difference in spending among individuals with and without each condition of interest; however, these estimates ignore several factors such as the number of diseases, severity, the number of days of hospitalization, etc. that can affect the actual cost for each individual. The adjustment model aims to model the observed cost as a function of the naïve cost (the sum of the costs of all the conditions as estimated in the first step) and other factors: the number of health conditions, history of hospitalization, institutionalization and death.

First, we compared actual medical expenditures at the individual level reported in the survey with predicted medical costs for each individual based our model estimates, obtained by

summing the relevant AC_j estimates for each subject. When examining the difference between observed and estimated costs, relationships became apparent between this difference and several variables, such as the number of health conditions, history of hospitalization and institutionalization, and death. The relationships can be seen by plotting the error term, measured as a difference between naïve predicted cost and true cost, against the aforementioned covariates, all of which are proxies for the volume or intensity of care. Figure A7 in the Appendix shows the different plots. Empirically, the difference is related to these characteristics. For example, the institutionalized population spends even more than predicted given its medical history.

We developed an adjustment model that predicts subject-specific observed cost as a product of the sum of the disease costs, AC_j , and a polynomial dependent on the number of the health conditions, history of hospitalization, institutionalization, and death within a given year:

$$E(Y_i) = C_i \sum_j (AC_j \ D_{ij}),$$

where D_{ij} is an indicator of disease j for subject i, and C_i is a polynomial

$$C_i = \alpha_0 + \sum_{l=1-9} \alpha_l K_{li},$$

where K_1,...K_9 are variables corresponding to the number of comorbidities, that number squared, an indicator for any hospitalization, number of nights in the hospital, number of hospital admissions, number of days institutionalized, survival in the indicated year, number of months survived for those who are deceased, and number of outpatient claims. Although we adjusted for death in given year in our propensity models, it was through estimating the probability of death, not directly conditioning on the indicator for death. Several of the other variables in the adjustment model were not included in the propensity score model due to either rarity or high

correlation with calibrated conditions, which could lead to issues in fitting the logistic regression model.

When fitting our model, we constrained the adjustment factors $(C_j\ )$ so that the weighted sum of predicted cost matched the weighted sum of observed cost. To do this we first fit the model for C_j with no constraints and obtained estimates for $\alpha_0,\ldots\ \alpha_9$. We then adjusted the intercept, α_0, by adding a term so that

$$\sum_i Y_i = \sum_i C_i \ \sum_j AC_j \ D_{ij}$$

This model yielded subject-specific adjustment factors $(C_i\ )$ that incorporated covariates related to the volume of medical care received. Constraining C_j ensured that the average predicted cost from our model would be equal to the observed costs, guarantying that predicted spending would match total spending in the population. Parameter estimates for the adjustment are displayed in table A8.

To integrate these adjustments to the estimated cost of specific health conditions, we first aggregated the adjustment factors C_i for each disease *j* by taking the average of the individuals with the condition *j*:

$$\delta_j \ = \frac{\sum_i C_i \ D_{ij}}{\sum_i D_{ij}}$$

Then, we applied this disease-specific adjustment factor δ_j to the average difference in spending for individuals with and without the condition of interest, by taking the product of $(AC_j\ )$ and δ_j : $AC_j^{(adj)} = AC_j \ \delta_j$. The results are estimates of spending at the condition level that are guaranteed to sum to total spending and that track the individual distributions of spending as well as possible. As above, we averaged across 5 multiple imputed data sets using

proc mianalyze in SAS 9.4. After fitting the adjusted model, we examined plots of the residuals on the variables in the adjustment model to check that we correctly modeled the relationship between cost and the number of health conditions, history of hospitalization and institutionalization, and death.

Table 2 reports the total attributed spending by multi-level CCS categories, and table A9 shows the specific conditions. Diseases of the circulatory system were the most expensive category, costing $144.8 billion and accounting for 23 percent of personal health spending. The next most expensive categories were "other conditions" (13 percent), diseases of the respiratory system (10 percent), endocrine, nutritional; and metabolic diseases and immunity disorders (8 percent), and diseases of the nervous system and sense organs (7 percent). Together, these six disease groups accounted for almost 61% of personal health spending. Two of the conditions (cancer screening; complications of pregnancy, childbirth, and the puerperium) are estimated to have negative spending. We remark on this below.

Figure 1 and table 1 report distributions of the estimates from the propensity score model. The propensity score approach shows a broader range of spending than the regression approach but not as broad as the claims-based approach. Effectively, our second stage makes spending at the individual level more dispersed than with the regression but does not allow as many very low and high spenders as with the claims.

IV. Comparison of Cost Attribution Results

The previous section showed that spending for different conditions differs across the methods. We start by showing more detail about these differences. Figure 2 shows two-way scatter plots of average attributed cost for the claims-based, regression-based, and propensity

score approaches. Panel A compares the claims-attribution approach and the regression-based approach. Panel B compares the claims attribution and propensity score approaches, and panel C compares the regression and propensity score methods.

The correlation between average disease-specific spending using the different approaches is reasonably high: 0.27 between the claims-based approach and the regression approach, 0.71 between the claim-based approach and propensity score approach, and 0.56 between the regression approach and the propensity score approach. The figures show that some large and significant outliers influence the correlations.

The average attributed cost from the claims-based method is higher for most of the conditions than for the other two approaches. This is not surprising; the claims approach estimates cost of the condition that resulted in the claim, whereas the propensity score and regression methods estimate cost attributed to the condition regardless of whether it was specified in the claim or remained latent. For example, the cost of a stroke that has a claim in a given year is likely to be higher than a past stroke whose residual effects are still present. A corollary of this is that the effective number of conditions for each claim is higher in the regression and propensity score approaches.

In general, acute medical conditions like acute myocardial infarction (AMI), cardiac arrest, hematologic cancer, and hip fracture have significantly higher average costs in the claims-based method than in the regression and propensity score methods.

Not surprisingly, the average attributed cost from the regression method and the propensity score method have much greater accord than either one has with the claim-based approach. Most of the conditions are around the 45-degree line, with few exceptions that are off-diagonal. The major differences are that conditions like lung cancer, hematologic cancer, and

schizophrenia have much higher average attributed cost in the propensity score method as compared to the regression-based method. On the other hand, hypertension and signs and symptoms have significantly higher attributed costs in the regression-based approach.

Intuition suggests that expensive, rare diseases are likely to generate high spenders. The regression method, by assuming parametric models and distributions, may smooth over these extreme values, whereas the non-parametric propensity score approach gives higher weight to these outliers. On the other hand, attributed cost estimation for very prevalent and less severe conditions such as hypertension and signs and symptoms may be influenced by collinearity with other conditions, which could be more of an issue for the regression approach.

In addition to average spending per condition, we also care about total spending on each condition – taking into account prevalence as well as cost per condition. Figure 3 shows the scatter plot of total attributed spending between the three approaches. To a great extent, the results in figure 3 mirror those in figure 2. The biggest outliers in the total spending figure are hypertension and signs and symptoms. The regression approach estimates the highest spending on these conditions, followed by the propensity score approach and then the claims-based approach. It is natural that hypertension is attributed fewer dollars by claims than by overall spending impact.

Figure 4 shows the ratio of total costs for the propensity score method to the claims-attribution method, using the total dollars displayed in Table 1. A ratio >1 implies that the propensity score method attributes more spending than claims, and a ratio <1 implies the opposite. Among the major identified diseases that are consistent with the ICD9-CM disease chapters, the biggest ratios (>1) are for diseases of the blood and blood-forming organs (largely anemia), injury and poisoning, and mental illness. Some of these are important comorbid

conditions. Anemia can generally be treated cheaply, but its presence indicates a more severe form of the disease. Thus, it is natural that people with anemia spend a lot more than people without. Similarly, mental illness has been shown in many studies to be a significant risk factor for spending (Finkelstein, 2003b). People who are mentally ill spend more for all comorbid conditions, even acute ones, than those without mental illness.

A flip occurs at diseases of circulatory system, where the claims approach attributes more dollars than the propensity score approach. Two ratios are negative for the conditions (cancer screening; and complications of menopause, pregnancy, childbirth, and the puerperium) because the PSM method attributed negative spending to those conditions whereas the claims approach attributed positive spending. In the elderly population, complications of menopause, pregnancy, childbirth, and the puerperium largely reflects menopause. Women coded for this may be relatively healthy – if there were other, more severe conditions, they would likely be coded. Screening may be negative for the same reason (because healthier people get screened) or because screening prevents more expensive diseases.

Using all three metrics, disease specific spending in the elderly population is very skewed. In the claims-based approach, the top 5 (10) conditions account for 20% (36%) of total dollars. The shares are 45% (61%) for the regression approach and 29% (46%) for the propensity score model.

Table A9 in the appendix shows the average and total cost associated with each condition using the propensity score method, alongside the other methods. The most expensive conditions in table A9 generally cost about $3,000 to $4,000 annually. For example, the cost of lung cancer ranges from $2,800 to $4,400. This is true using all methods and seems unusual given *ex-ante* expectations. For example, the cost of almost any chemotherapy regimen will exceed the few

thousand dollars that we estimate cancer to cost. The issue here is that not all of the prevalent cases are incident cases. Imagine that a person was diagnosed with lung cancer in 2008, receives the bulk of their care in that year and has a few visits for monitoring in 2009. We will record that person as (correctly) having lung cancer in 2009, even though the case is not new. Further, lung cancer spending will be relatively low in 2009 unless the person has a cancer recurrence. The net effect will be relatively low average spending per case.

If one were developing a model of lung cancer cost effectiveness, the spending we estimate would be of limited use. For such a model, one would want to know spending by the phase of cancer: the acute phase, maintenance phase, and (possibly) terminal phase. In studying the decomposition of total spending into conditions, however, the estimate we have is very relevant: it correctly indicates the average amount spent per person treated with cancer in the given year for a representative cross-section of the population, as well as the total dollars in the population.

V. Discussion

Our results show many similarities, but also important differences across the alternate methods of assigning costs to medical conditions. The obvious question to ask is, which one is best? There is no gold standard against which to compare the estimates from the different methods. Thus, we cannot give a definitive answer to the question. Still, some observations are possible.

Statistically, one way to evaluate the models is by comparing in-sample predicted and actual spending at the individual level. The root of the mean squared difference between actual and predicted spending across individuals is $2,679 for the claims-based method, $3,032 for the

regression-based method and $2,723 for the propensity score method. We also assessed the correlation between observed and predicted costs using each of the three methods. The correlation for the claims method is 0.63; the correlation for the regression approach is 0.61, and the correlation for the propensity score approach is 0.63.

Out-of-sample prediction is generally more informative than in-sample prediction, since allowing more degrees of freedom inherently reduces in-sample discrepancies between actual and predicted spending. In out-of-sample prediction, by contrast, allowing more degrees of freedom may overfit the data. To form out-of-sample estimates, we divided the data into two random sub-samples: one half for model-fitting and one half for prediction. We use the same observations for fitting and prediction in each of the models so that the difference results only from the ability to fit the underlying data. Table 3 shows the out-of-sample root mean square errors (RMSEs) for the various methods. The propensity score approach has the lowest RMSE, followed by the best regression model and last the claims-based model.

The ability to implement the models is also important. On this count, the claims-attribution methodology suffers from several problems. First, it is difficult to implement the methodology in a consistent manner. Because in most claims – especially the most expensive ones – involve more than one condition, the claims-based methodology will depend on the method used to parcel out costs into conditions. In our sample, 99.5% of dollars involve claims having multiple conditions listed, and 98.4% of beneficiaries have multiple conditions. This makes the claims-based estimation method extremely challenging. We made assumptions about how to divide claims into component conditions, but the assumptions do not have a strong theoretical rationale.

In Thorpe et al.'s (2004a) analysis, there was little difference between the claims-attribution model done different ways. However, Thorpe was using MEPS data for his analysis, and utilization events in MEPS are reported by individuals, rather than providers. Thus, underreporting of less severe conditions is more likely. Further, Thorpe et al. (2004a) were looking at the entire population rather than just the elderly; there is a much greater incidence of comorbidity in the elderly than in the non-elderly population.

Further, the claims-attribution methodology has difficulty with comorbid conditions that are not central to the primary reason for health care utilization. Relative to the regression and propensity score approaches, the claims-attribution methodology is low for anemia and mental illness, each of which is likely to increase medical spending across-the-board.

The regression and propensity score methods have a lot in common. Both methods facilitate cost attribution to any condition of interest: claim-based, calibrated health condition, self-report, or behavioral risk factors such as smoking. In addition, both methods are designed to adjust for other diseases and demographic covariates. The regression method is easier to implement because it does not require a new model for each health condition. If the data fits the parametric assumptions well and the set of health conditions is not highly correlated, then it produces unbiased and efficient estimates of the attributed costs.

However regression-based cost estimation has several limitations. First, it makes several parametric assumptions, which may not be satisfied. Second, there is a large residual spending amount that cannot be attributed to any disease. Third, some coefficients are estimated to be negative, and the approach we follow assigns zero spending to them. Finally, the variability in spending at the individual level implied by the regression approach is significantly smaller than the actual variation.

Relative to these issues, the propensity score approach has a number of strengths. It requires fewer parametric assumptions and is, therefore, more robust with respect to the high spenders and zero-spenders. Further, the second step of the propensity score approach allows us to account for the number of comorbidities and volume of medical care received. These features of the propensity score method permit us to relax the assumption that health care spending is additive.

In practice, the most important difference between the propensity score method and the regression method is the importance of unattributed spending. There is a base of $58 billion (9 percent of total spending) that the regression does not attribute to any condition. In the propensity score model, by contrast, all spending is automatically allocated to conditions. Across conditions, the major difference between the two approaches is the much greater total attributed to hypertension and signs and symptoms using the regression model.

While more research is certainly warranted, our tentative conclusion is that the propensity score model offers a good theoretical and empirical methodology to decompose total medical spending to conditions.

References

Akobundu E, Ju J, Blatt L, Mullins CD. Cost-of-Illness studies: a review of current methods. *Pharmacoeconomics*. 2006; 24(9):869–90

Basu A, Manning WG, "Issues for the next generation of health care cost analyses," *Medical Care*, 2009; 47(suppl 7):S109-S114.

Berk A, Paringer L, Mushkin SJ, "The Economic Cost of Illness Fiscal 1975," *Medical Care*, 1978; 16(9):785-790.

Belotti F, Deb P, Manning WG, Norton EC, "twopm: Two-part models," *Stata Journal*, 2015, 15(1):3-20.

Bloom BS, Bruno DJ, Maman DY, Jayadevappa R. Usefulness of U.S. cost-of-illness studies in health care decision making. *Pharmacoeconomics*. 2001; 19(2):207–213.

Cantoni E, Ronchetti E, "A Robust Approach for Skewed and Heavy-Tailed Outcomes in the Analysis of Health Care Expenditures," *Journal of Health Economics*, 2006; 25(2):198-213.

Cohen JW, Krauss NA, "Spending and Service Use among People with the Fifteen Most Costly Medical Conditions 1997," *Health Affairs*, 2003; 22(2):129–138.

Cooper BS, Rice DP, "The economic cost of illness revisited," *Social Security Bulletin*, 1976*;* 39:21-36.

Druss BG, Marcus SC, Olfson M, Tanielian T, Elinson L, Pincus HA ,"Comparing the national economic burden of five chronic conditions," *Health Affairs*, 2001; 20(6):233-242.

Druss BG, Marcus SC, Olfson M, Pincus HA, "The Most Expensive Medical Conditions in America," *Health Affairs*, 2002; 21(4):105-111.

Duan, N. (1983). Smearing Estimate: A Nonparametric Retransformation Method. Journal of the American Statistical Association, 78(383), 605

Duan, N., W. G. Manning, Jr., C. N. Morris, and J. P. Newhouse. 1984. Choosing between the sample-selection model and the multi-part model. Journal of Business and Economic Statistics 2: 283‑289.

Dudley RA, Harrell FE, Smith LR, Mark DB, et al., Comparison of analytic models for estimating the effect of clinical factors on the cost of coronary artery bypass graft surgery. *Journal of Clinical Epidemiology*, 1993; 46(3), 261-271.

Elixhauser A, Steiner C, Palmer L. "Clinical Classifications Software (CCS)," U.S. Agency for Healthcare Research and Quality, Available: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp.

Finkelstein EA, Fiebelkorn IC, Wang G, "National medical spending attributable to overweight and obesity: how much, and who's paying?" *Health Affairs*, 2003a, w3:219-226.

Finkelstein EA, Bray JW, Chen H, Larson, MJ, et al., Prevalence and costs of major depression among elderly claimants with diabetes. *Diabetes Care*, 2003b, 26(2), 415–420.

Finkelstein EA, Trogdon JG, Cohen JW, Dietz W, "Annual medical spending attributable to obesity: payer-and service-specific estimates," *Health Affairs*, 2009; 28(5):822-831.

Gregori, D., Petrinco, M., Bo, S., Desideri, A., Merletti, F., & Pagano, E. (2011). Regression models for analyzing costs and their determinants in health care: An introductory review. International Journal for Quality in Health Care, 23(3), 331-341

Hall A, Highfill T, "A Regression-Based Medical Care Expenditure Index for Medicare Beneficiaries," *BEA Working Paper*, 2013a.

Hall A, Highfill T, "Calculating Disease-Based Medical Care Expenditure Indexes for Medicare Beneficiaries: A Comparison of Method and Data Choices," NBER Working Paper No. 19720, 2013b.

Hodgson TA, Cohen AJ, "Medical expenditures for major diseases, 1995," *Health Care Financing Review*, 1999; 21(2):119-164.

Honeycutt AA, Segel JE, Hoerger TJ, Finkelstein EA, "Comparing cost-of-illness estimates from alternative approaches: an application to diabetes," *Health Services Research*, 2009; 44(1): 303-320.

Honeycutt, AA, Segel JE, Zhuo X, Hoerger TJ, Imai K, and Williams D, " Medical Costs of CKD in the Medicare Population. *Journal of the American Society of Nephrology*, 2013; 24(9), 1478-1483.

Jones, A, "Health Econometrics." In: Culyer A and Newhouse J (Eds.), *Handbook of Health Economics*, Elsevier, Amsterdam, 2000.

Koopmanschap MA, Rutten FFH, van Ineveld BM, and van Roijen L, "The Friction Cost Method for Measuring Indirect Costs of Disease," *Journal of Health Economics*, 1995, 14(2):171-189.

Koopmanshap M. Cost-of-Illness Studies. Useful for Health Policy? *Pharmacoeconomics*. 1998; 14:143–148

Leon J, Neumann PJ, "The cost of Alzheimer's disease in man-aged care: a cross-sectional study," *American Journal of Managed Care*, 1999; 5(7):867-877.

Li K, Raghunathan T, and Rubin D, "Large sample significance levels from multiply-imputed data using moment-based statistics and an F reference distribution," *Journal of the American Statistical Association*, 1991; 86: 1065–1073.

Little R, and Rubin D, *Statistical Analysis with Missing Data*. New York: John Wiley & Sons, 1987.

Manning WG, "The logged dependent variable, heteroscedasticity, and the retransformation problem," *Journal of Health Economics*, 1998; 17(3):283–295.

Manning WG, Mullahy J, "Estimating log models: to transform or not to transform?" *Journal of Health Economics*, 2001; 20(4):461-494.

Manning WG, Basu A, Mullahy J, "Generalized Modeling Approaches to Risk Adjustment of Skewed Outcomes Data," *Journal of Health Economics*, 2005; 24(3):465-488.

Mihaylova, B., A. Briggs, A. O' Hagan, and S. G. Thompson. 2011. Review of statistical methods for analysing healthcare resources and costs. Health Economics 20: 897‐916.

Pope GC, Kautter J, Ellis RP, Ash AS, Ayanian JZ, Iezzoni LI, Ingber MJ, Levy JM, Robst J, "Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model," *Health Care Financing Review*, 2004; 25(4):119-141.

Raghunathan T, et al., "Model Based Estimation of Prevalence Rates and Trends Using Imperfect Data Sources," *mimeo* (2017).

Roberts RR, Scott RD, Hota B, Kampe LM, Abbasi F, Schabowski S, Weinstein RA, et al., "Costs Attributable to Healthcare-Acquired Infection in Hospitalized Adults and a Comparison of Economic Methods," *Medical Care*, 2010; 48(11), 1026-1035.

Rosen AB, and Cutler DM, "Challenges in Building Disease-Based National Health Accounts," Medical Care, 2009; 47(Supplement), S7-13.

Rosen A, et al., "Strengthening National Data to Better Measure What We Are Buying in Health Care: Reconciling National Health Expenditures with Detailed Survey Data", NBER Working Paper No. 23290, Issued in March 2017

Rice DP, Horowitz LA, "Trends in medical care prices," *Social Security Bulletin*, 1967; 30:13-28.

Rice DP, "Estimating the Cost of Illness," *American Journal of Public Health*, 1967, 57(3):424-440.

Roehrig C, Miller G, Lake C, Bryant J, "National health spending by a medical condition, 1996–2005," *Health Affairs*, 2009; 28(2):w358–w367.

Roehrig CS, Rousseau DM, "The Growth in Cost Per Case Explains Far More of US Health Spending Increases Than Rising Disease Prevalence," *Health Affairs*, 2011; 30(9):1657-1663.

Rosenbaum, P. R. and Rubin, D. B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects". Biometrika, 70: 41–55.

Rosenbaum, Paul R., and Donald B. Rubin. 1984. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association 79(387): 516-524

SAS Institute Inc., SAS 9.4 Help and Documentation, Cary, NC: SAS Institute Inc., 2002-2004.

Scitovsky AA, "Changes in the costs of treatment of selected illnesses," 1951-1965. *American Economic Review*, 1967; 57(5):1182-1195.

Selden T, Levitt K, Cohen J, et al., "Reconciling Medical Expenditure Estimates from the MEPS and the NHA, 1996," *Health Care Financing Review*, 2001; 23(1):161-178.

Sing M, Banthin JS, Selden TM, et al., "Reconciling Medical Expenditure Estimates from the MEPS and the NHEA, 2002," *Health Care Financing Review*, 2006; 28(1):25-40.

Starr M, Dominiak L, Aizcorbe A, "Decomposing Growth In Spending Finds Annual Cost Of Treatment Contributed Most To Spending Growth, 1980-2006," *Health Affairs*, 2014; 33(5):823-831.

Sturm R, "The Effects of Obesity, Smoking, and Drinking on Medical Problems and Costs," *Health Affairs*, 2002; 21(2):245-253.

Thorpe KE, Florence CS, Joski P. "Which medical conditions account for the rise in health care spending?" *Health Affairs*, 2004a; 23:w437-w445.

Thorpe KE, Florence CS, Howard DH, Joski P. "The impact of obesity on rising medical spending." *Health Affairs*, 2004b; 24:w480-w486.

Thorpe KE, Florence CS, Howard DH, Joski P. "The rising prevalence of treated disease: effects on private health insurance spending." *Health Affairs*, 2005; 24:w317-w325.

Thorpe KE, Howard DH. "The rise in spending among Medicare beneficiaries: the role of chronic disease prevalence and changes in treatment intensity." *Health Affairs*, 2006; 25(5):w378-w388.

Thorpe KE, Howard DH, Galactionova K. "Differences in disease prevalence as a source of the U.S.-European health care spending gap." *Health Affairs*, 2007; 26(6):w678-w686.

Thorpe KE, Ogden LL, Galactionova K. "Chronic conditions account for rise in Medicare spending from 1987 to 2006." *Health Affairs*, 2010; 29(4):718-724.

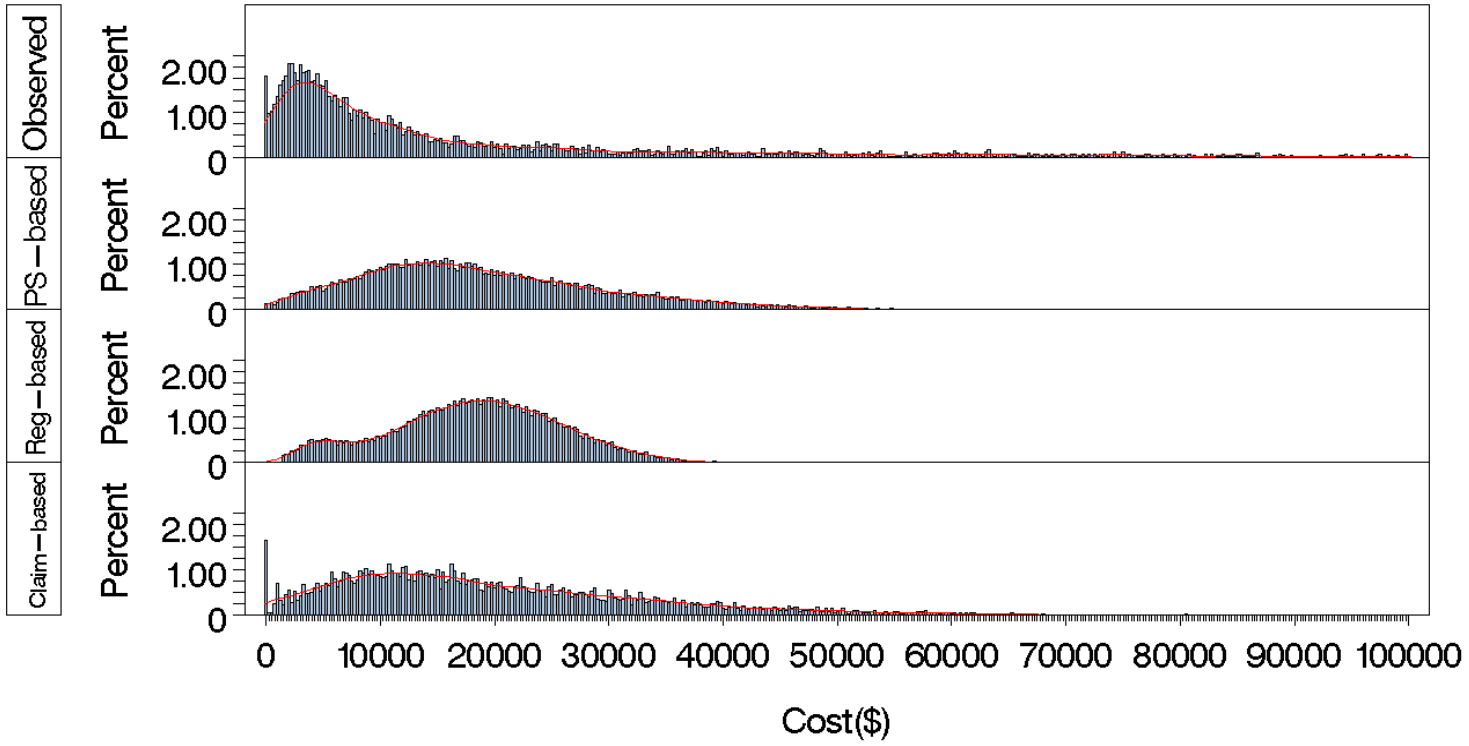
Thorpe KE. "Treated disease prevalence and spending per treated case drove most of the growth in health care spending in 1987–2009." *Health Affairs*, 2013; 32(5):851-858.

Trogdon JG, Finkelstein EA, Nwaise IA, Tangka FK, Orenstein D, "The economic burden of chronic cardiovascular disease for major insurers." *Health Promotion and Practice*, 2007; 8(3):234-242.

Trogdon J, Finkelstein EA, Hoerger TJ, "Use of Econometric Models to Estimate Expenditure Shares," *Health Services Research*, 2008; 43(4):1442-1452.

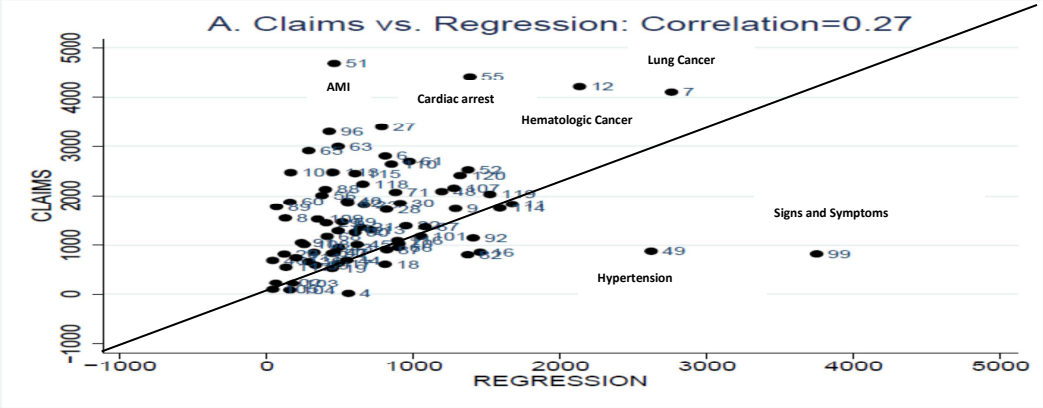
Zaslavsky AM, Buntin MB, "Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures," *Journal of Health Economics*, 2004; 23(3):525-554.

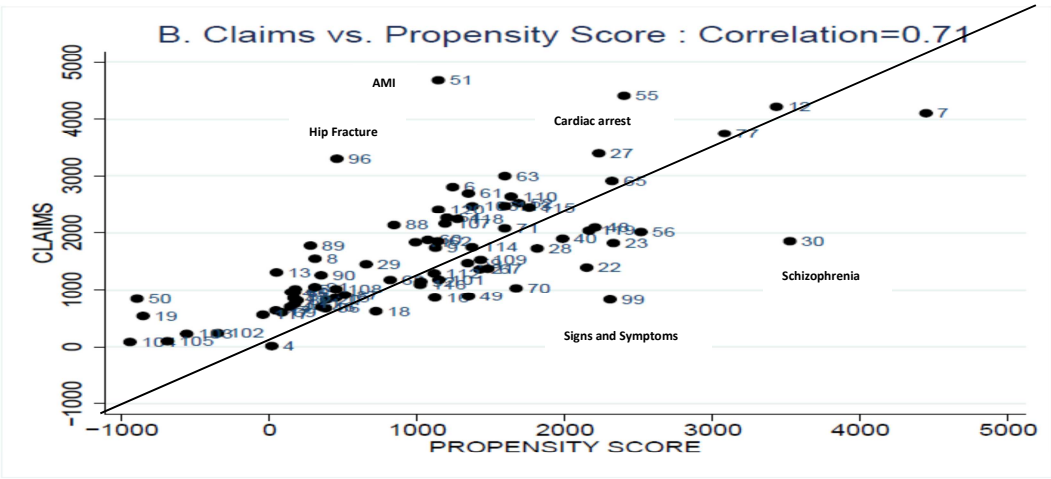
Figure 1: Distribution of Observed and Predicted Spending

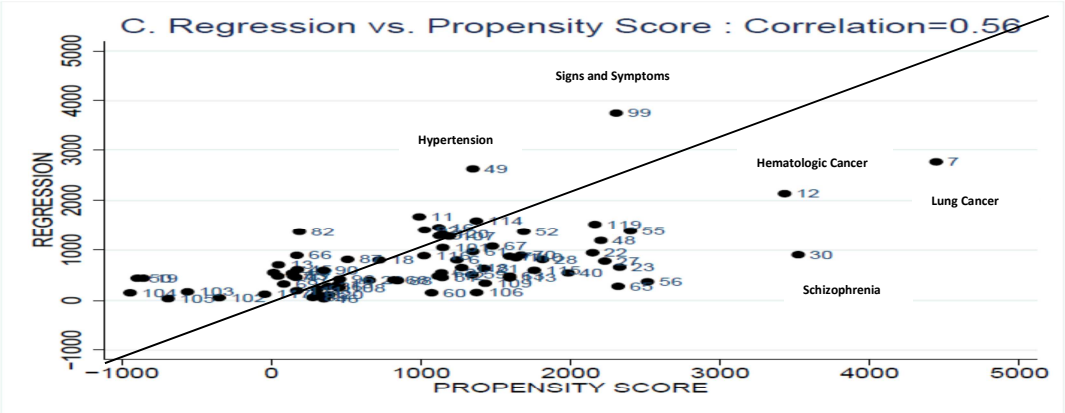


Note: The distribution of observed cost was cut off at $100,000, omitting about 2% of beneficiaries. There are no beneficiaries with a predicted spending above $80,500 in any of the models. Here, N=6,200 and weighted N=36,824,486. The regression and propensity score methods are based on calibrated claims. The claims approach is based on actual claims except for about 5% of beneficiaries with dollar amounts in the personal summary file(s) but no claims. For these beneficiaries, we used the claims imputed using the NHANES and assigned dollars to calibrated medical conditions on that basis. PS-based – Propensity Score based method. Reg-based: Regression based method. Claim-based: Claims based method.

Figure 2: Average Attributed Cost Using Three Attribution Methods

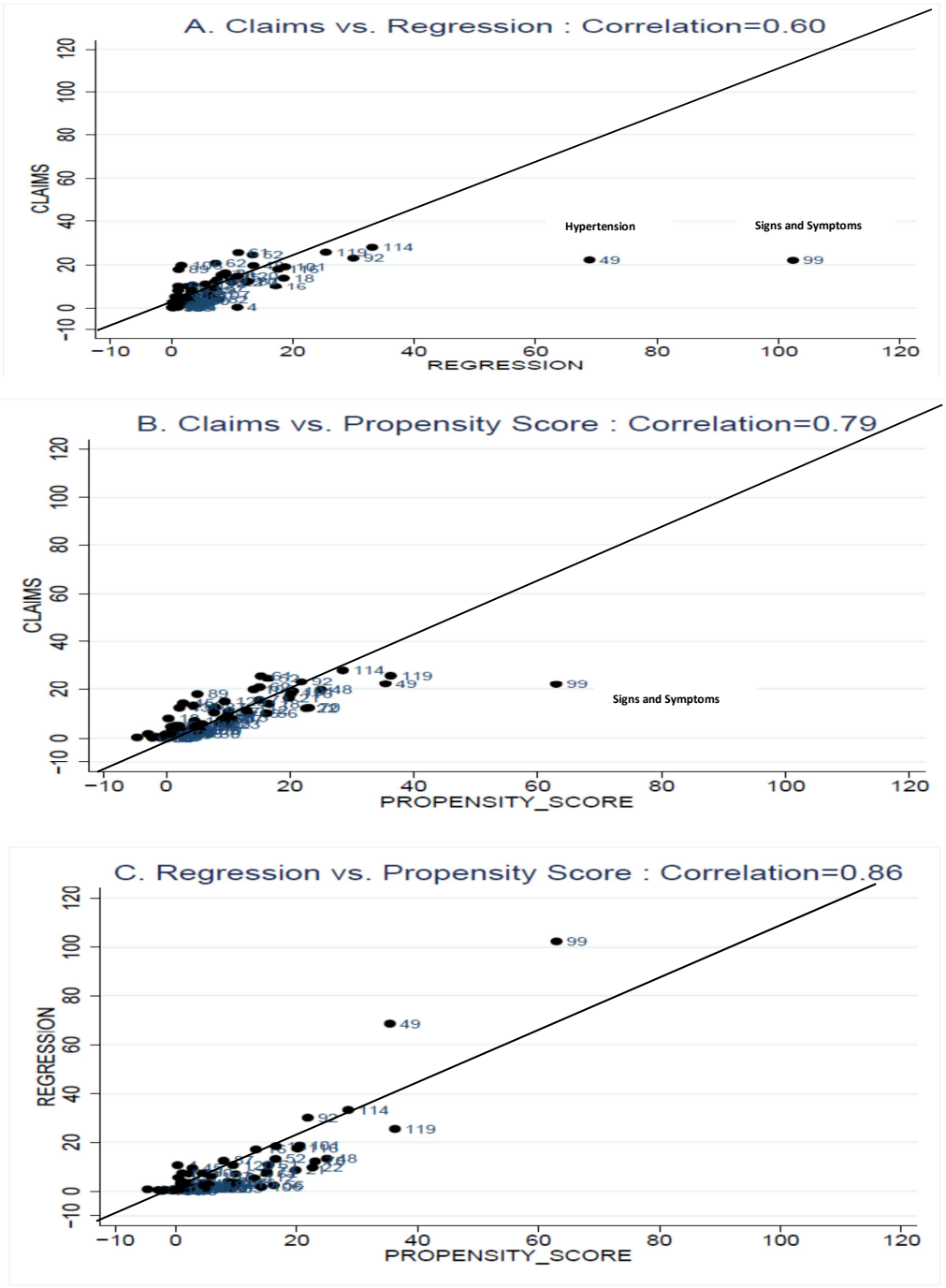






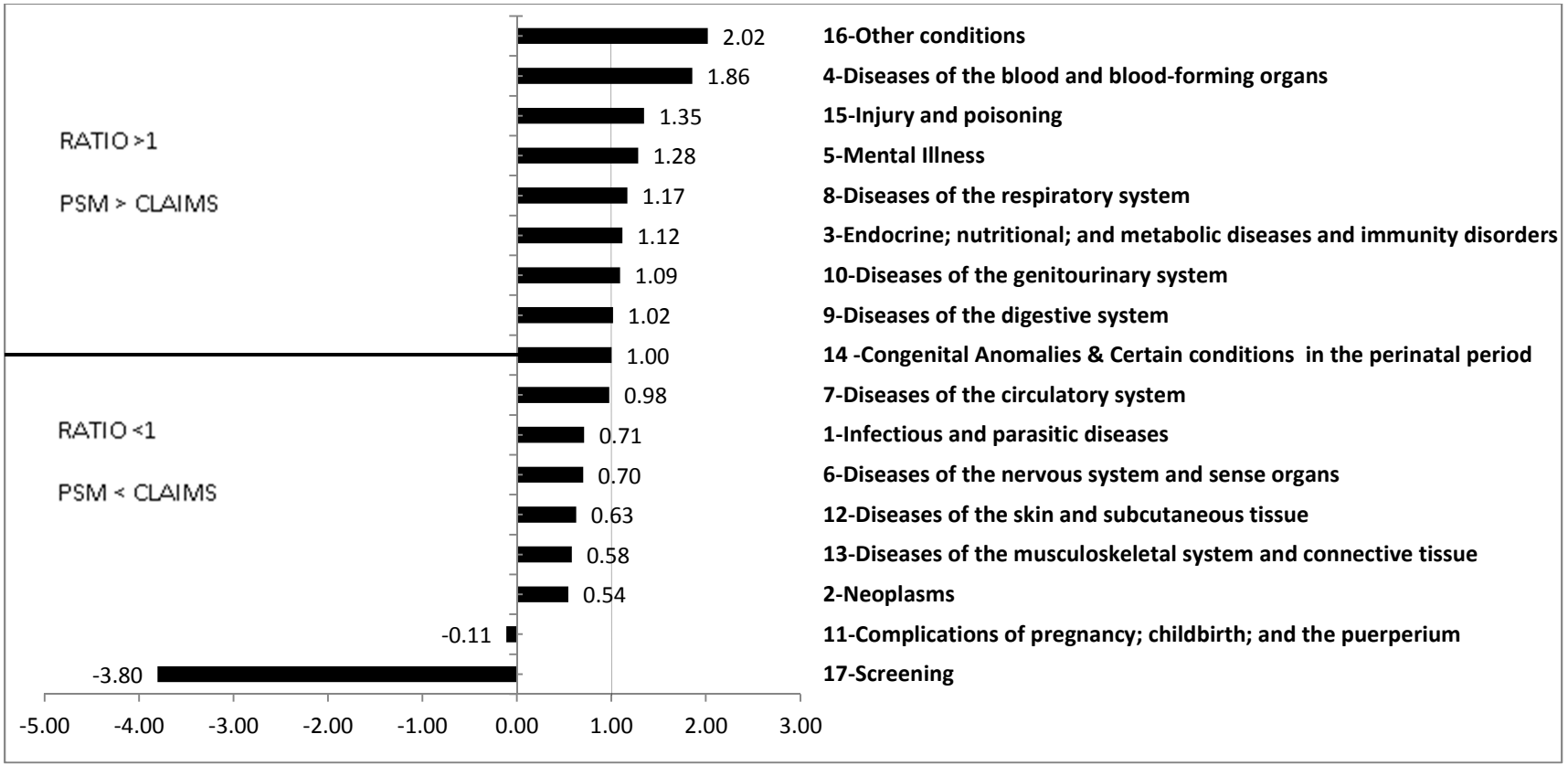
Note: Acute Renal Failure and Deep Vein Thrombosis (DVT) had negative coefficients in Regression Model and no cost attributed. As a result, they are not included in the correlation or in Panels A and C.

Figure 3: Total Attributed Cost (Billions of 2010 US $) Using Three Attribution Methods



Note: Acute Renal Failure and Deep Vein Thrombosis (DVT) had negative coefficients in the Regression Model and no cost attributed. As a result, they are not included in the correlation or in Panels A and C.

Figure 4: Ratio of Total Cost by Major CCS categories: Propensity Score vs. Claims



Note: Bars show the ratio of total costs from the propensity score method to those from the claims-attribution method. The ratios are based on total attributed dollars reported in Table 2. Ratio >1 indicates that the PSM method attributes more total spending than the claims method, and ratio <1 implies the opposite. Other Conditions (top bar) include Signs and Symptoms, Residual, unclassified and all other E codes. The propensity score method attributes negative spending to various cancer screenings and complications of pregnancy, making the ratios negative.

Table 1: Characteristics of person-level spending from different methods

	Dollar Amount (Standard Error)						N (Weighted N) of Beneficiaries with Cost>=	
	Mean	**25%ile**	**Median**	**75%ile**	**Min**	**Max**	**$100,000**	**$50,000**
Observed	$17,479 ($389)	$3,291 ($76)	$7,281 ($196)	$17,547 ($525)	$0	$441,857	225 (913,599)	795 (3,301,586)
Claims-based	$17,479 ($213)	$8,573 ($179)	$15,006 ($271)	$24,230 ($370)	$0	$80,500	0	153 (671,044)
Regression-based	$17,479 ($128)	$12,719 ($204)	$17,726 ($157)	$22,496 ($191)	$1,818	$40,394	0	0
Propensity score-based	$17,479 ($238)	$10,236 ($286)	$15,889 ($246)	$23,202 ($317)	$831	$63,242	0	36 (167,288)

Note: The table shows summary statistics for individual-level spending based on observed costs, costs using the claims-based method, costs using the regression-based method, and costs using the propensity score-based method.

Table 2: Estimated Total Cost in 2009 (Billions of 2010 US $) Using Different Methods

	Total Attributed Cost (Billions)			% of Total Spending		
Multi-level CCS	**Claims Method**	**Regression Method**	**Propensity Score Method**	**Claims Method**	**Regression Method**	**Propensity Score Method**
7-Diseases of the circulatory system	$148.1	$119.3	$144.8	23.0%	18.5%	22.5%
16-Other conditions	41.3	121.1	83.4	6.4%	18.8%	13.0%
8-Diseases of the respiratory system	53.4	35.5	62.4	8.3%	5.5%	9.7%
3-Endocrine; nutritional; & metabolic dz. & immunity disorders	49.2	48.9	54.9	7.6%	7.6%	8.5%
6-Diseases of the nervous system and sense organs	64.0	43.0	44.9	9.9%	6.7%	7.0%
10-Diseases of the genitourinary system	38.1	28.6	41.6	5.9%	4.4%	6.5%
5-Mental Illness	30.8	18.3	39.5	4.8%	2.8%	6.1%
15-Injury and poisoning	27.5	26.1	37.0	4.3%	4.1%	5.7%
13-Diseases of the musculoskeletal system and connective tissue	62.2	41.6	36.1	9.7%	6.5%	5.6%
4-Diseases of the blood and blood-forming organs	17.7	12.9	32.8	2.7%	2.0%	5.1%
9-Diseases of the digestive system	28.1	33.1	28.5	4.4%	5.1%	4.4%
2-Neoplasms	41.4	27.6	22.5	6.4%	4.3%	3.5%
1-Infectious and parasitic diseases	20.3	12.6	14.4	3.2%	2.0%	2.2%
12-Diseases of the skin and subcutaneous tissue	12.6	12.7	7.9	2.0%	2.0%	1.2%
14 -Congenital Anomalies & perinatal conditions	4.6	2.4	4.6	0.7%	0.4%	0.7%
11-Complications of pregnancy; childbirth; and the puerperium	1.4	0.5	-0.2	0.2%	0.1%	0.0%
17-Screening	3.0	1.9	-11.5	0.5%	0.3%	-1.8%
Other covariates (including intercept)		57.7			9.0%	
Total	$644	$644	$644	100%	100%	100%

Note: The table is sorted from highest to lowest attributed total cost from the propensity score method. The multi-level CCS category "16-Other conditions" includes Signs and Symptoms ($62.97 billion) and Residual, unclassified, all other E codes ($20.46 billion).

Table 3: Characteristics of Spending Using Out-of-Sample Predictions

Method	Root Mean Square Error of Prediction
Claims Based Model	$28,856
Regression Models	
One-part gamma –log link	$42,711
One-part-cubic-root-cost-Gaussian with identity link	$25,228
One-part-log(cost+1)-Gaussian-identity link	$24,410
One-part-Box-Cox model	$26,052
Two-part-gamma-log-link	$42,042
Two-part-cubic-root-cost-Gaussian-identity link	$25,065
Two-part-log(cost+1) –Gaussian-identity link	$31,858
Propensity Score Method	$22,621

Note: The data are divided into two sub-samples, one for model fitting and a second for out-of-sample prediction. The Root MSE is for the out-of-sample prediction.

APPENDIX

Attributing Medical Spending to Conditions: A Comparison of Methods

In this appendix, we present various technical aspects of the data assembly and analysis.

A.1 HMO enrollment adjustment

MCBS has incomplete or no claims information for beneficiaries enrolled in Health Maintenance Organizations (HMOs). In 2009, approximately one-quarter of elderly beneficiaries were enrolled in HMOs (Kaiser Family Foundation, 2014). To adjust for this, we developed a weighting adjustment similar to the non-response adjustments performed for unit non-response in national surveys (Little, 1986; Kreuter et al., 2010).

We began by defining two groups: those with complete Medicare enrollment and those enrolled in HMOs. We define complete Medicare enrollment as: (1) no participation in Medicare Advantage program for the year of study, and (2) enrollment in traditional fee-for-service Medicare parts A & B for the full 12-month study period, unless the participant died during the year.

We used a propensity score method to create the adjustment weights. We performed these adjustments separately for the community and institutionalized population. Since the majority of our sample (~92%) consisted of community residents, we present the results for the community population here. A logistic regression model was estimated using selected covariates

(demographics, health status, and socio-economic variables) to model traditional Medicare enrollment. Table A1 gives the list of covariates used for such adjustments. Using the predicted probability (p) of complete Medicare enrollment, the adjustment for HMO enrollees was calculated as 1/p. Model fit was assessed by a Hosmer-Lemeshow test.

Table A2 reports the Hosmer-Lomeshow goodness of fit statistics. In our estimation, there were 10 groups and hence we had 8 degrees of freedom. The corresponding Chi-squared values are also reported. The model passed the test suggesting a good fit.

To assess balance in the community population, the propensity of complete Medicare enrollment was estimated using Generalized Liner Models (GLM); F-ratios were reviewed for significance. In the institutionalized population, regression models were performed for each covariate to assess the association with the propensity for complete Medicare participation. Using the residuals from each model, we calculated effect size to assess balance. We then calculated the "final weight" as the product of the existing MCBS survey weight and the Medicare HMO adjustment weight.

Figure A1a gives the distribution of propensity scores for complete Medicare enrollment for the 2009 Medicare community population. The overlap between the complete Medicare and HMO populations was high. Table A3 shows the comparison between the complete Medicare and HMO-adjusted elderly populations. In 2009, around 24.3 million elderly beneficiaries were enrolled in traditional fee-for-service Medicare, and 12.5 million were in HMOs. The distributions of age, sex, race, education, marital status and health status were mostly similar between the complete Medicare enrollment and HMO-adjusted populations.

A.2 Survey spending adjustment

A second adjustment was made so that total medical spending in the MCBS matched what was reported to be national spending on the elderly. This adjustment was performed in several steps. First, we removed expenditures from the National Health Expenditure Accounts (NHEA) for goods and services which are out of scope of the MCBS survey: other non-durable medical equipment (2.7%), other personal healthcare (2.6%), graduate medical education and disproportionate share medical payments to hospitals (1%), hospital non–patient revenue such as in the gift shop and for parking (3.2%), and spending by foreign visitors (0.12%). In total, this accounted for 9.6% percent of NHEA spending (Rosen et al., 2017, NHEA, 2014).

Second, we redefined some categories of medical services in the NHEA and MCBS, shifting expenditures as appropriate, to create consistent categories between the two sources. Table A4 shows the adjustments we made and the dollar amounts moved. The total portion of money shifted across categories was 4.2%.

Third, we proportionately increased spending in the MCBS by the factors necessary to have total survey spending equal the remaining portion of the NHEA total in each service-by-payer category. Figure A2 gives the adjustment factors by each service category. Overall, the NHEA-adjusted spending was 11 percent higher than the total spending reported in MCBS. The adjustment was largest for home health (43%), but was generally small for the other types of services.

A.3 Condition Definitions and Prevalence

We developed a classification schema for medical conditions building upon the Agency for Healthcare Research and Quality's (AHRQ) Clinical Classification Software (CCS), which

aggregates the 14,000+ ICD-9-CM diagnosis codes and 3,900+ ICD-9-CM procedure codes into a smaller number of clinically meaningful, mutually exclusive categories (Elixhauser, Steiner, and Palmer, 2014). For brevity, we do not describe the full process of forming the categories here; interested readers are referred to Raghunathan et al. (2001, 2017). Creation of the mutually exclusive, collectively exhaustive categories required the clinical expertise of physicians, and extensive data management and analytic investigation, which resulted in 101 medical condition categories and 4 cancer screening categories (for breast, prostate, colon and cervical cancer).

Our physician working group determined that a few conditions identified in larger CCS categories should be stand-alone disease categories because of their clinical significance in the elderly (mostly mental health). These were grouped into independent disease categories. For example, while the CCS has a single "mood disorders" category, we separated this into two separate groups – depression and bipolar disorder.

Prevalence rates for some conditions in the MCBS were below those based on self-reports and physical assessment in national surveys. This was generally true for chronic diseases that are not serious enough to warrant a medical visit on their own, or at least not every year: hypertension and high cholesterol, for example. By definition, undiagnosed conditions are also not in MCBS. For such conditions, we used self-reports and diagnosed condition rates in NHANES to estimate 'calibrated' health conditions that more accurately reflect national prevalence rates. We term the results Calibrated Claims (CCL).

The imputation method proceeded in several steps. We chose to impute the community and institutionalized populations separately given the differences in these populations. We began by appending data from MCBS (2009) and NHANES (2009-2010). Each person was placed into one of three groups: having the condition in the self-report (NHANES) or claims

(MCBS); not having the condition if the NHANES self-report indicated the beneficiary did not have the health condition and there was no claim for the condition; and missing if there was no claim for the health condition in the MCBS. We then had a standard missing data problem for which we a used sequential regression multivariate imputation procedure.

For conditions present in NHANES, let $D_{(-j)}$ denote the collection of disease indicators for all diseases except disease j. We constructed a propensity score for having disease j based on fitting a logistic regression model to the other conditions and exogenous covariates, X, and predicting with (X, $D_{(-j)}$) strata based on the propensity scores. Within each propensity score class, we estimated the prevalence rate using the self-report, S_j, and the claims C_j . If the prevalence rate based on the claims was greater than or equal to that based on the self-report, then we set all missing D_j to 0. That is, no additional imputation was necessary and all those with no claims were considered not to have that health condition. If the self-report prevalence rate was greater than the prevalence rate based on the claims, we randomly set some missing D_j to 1 so that the prevalence rates after the imputation matched the self-report prevalence rates. We used five Bernoulli draws within each propensity score class to achieve this calibration, resulting in five imputed data sets.

Note that medical expenditure and health conditions without self-report are missing in the NHANES portion of the appended data. To be fully conditional, these missing values were imputed in the NHANES. These two steps – the disease imputations into MCBS and the medical spending/health condition imputations into NHANES – were iterated across all diseases several times until the multiply imputed prevalence rates stabilized.

The regression relationship between the multiply imputed D_j and claims-based C_j for conditions available in NHANES may be viewed as a measurement error model and this

relationship is then used to calibrate other health conditions not present in NHANES. In this step, we chose the most similar prevalent condition for the imputation.

The NHANES is a sample of the community dwelling population only. Thus, the claims imputation for the institutionalized sample required some differences. For this population, the calibrated non-institutionalized MCBS data was considered as the 'donor' survey in imputing condition prevalence in the institutionalized population. For each claim, subjects were matched according to the estimated propensity of being institutionalized given the self-report and demographic information, and the remaining claims. To estimate this propensity, logistic regression was utilized with a forward selection procedure on the principal components of the set of variables of interest. Principal component analysis was used in an effort to explain as much of the variation in propensity scores as possible while avoiding a complete separation of data points given the small number of people who are institutionalized. Assuming that the probability of being calibrated is the same conditional on institutionalization status, calibrated conditions are drawn for the institutionalized population matching the distribution for the community population.

The calibration process produces five imputed data sets for both community and institutionalized populations. We use all five imputed data sets in our analysis using appropriate survey weight and sample design adjustments. Importantly, prevalence based on the calibrated conditions indicates diagnosis or treatment either currently or at any time in the past.

Because some of our 101 calibrated condition categories have relatively low prevalence in the elderly, even after calibration, we collapsed our initial set of 101 medical conditions to 74 conditions with generally higher prevalence. Table A5 shows prevalence rates for the 74 calibrated medical conditions and 4 cancer screening variables. We have also aggregated these

74 conditions into 17 multi-level categories (analogous to ICD-9-CM chapters), including one for any cancer screening, which is how we group the table.

A.4 Analysis of different models

Figure A3 shows the kernel density plot comparison of the residuals (in blue) along with the normal density (in red). Figures A4–A6 show additional residual diagnostic plots, including standardized normal probability plots of the residuals, quantile plots of the residuals, and scatter plots of the residuals relative to person-level costs. The out-of-sample predictions show that the one-part model with log(cost+1) as the dependent variable and with a Gaussian distribution and identity link had the lowest RMSE. Figure A3 shows that this model also fit well. The one and two part cubic-roots were second best. The Gamma models with log link performed relatively poorly than expected. These models are sensitive to high-spenders with big residuals – as typified by people in nursing homes.

To evaluate the PSM estimates, we examined the relationship between predicted costs at the individual level – estimated as the sum of the relevant AC_j estimates for each subject – and the observed cost. We plotted the error term, measured as the difference between naïve predicted costs and true costs, against the number of health conditions, history of hospitalization and institutionalization, and death. All these variables can be considered proxies for the volume or intensity of care.

Figure A7 shows multiple plots of the difference between naïve predicted and observed costs using the propensity-score models. Figure A7A relates the spending error to the number of calibrated conditions. When the number of comorbidities is low – roughly 3 or fewer – there is no systematic difference between predicted cost and observed cost: the error is about zero.

However, as the number of comorbidities increases, the naïve cost progressively overestimates the observed cost. We have much fewer data for people with a very high number of conditions (30 or more), but the estimates appear to come closer together after that point.

Figure A7B considers whether this pattern is different for people living in the community versus institutions. Negative residuals are much more prevalent in the institutionalized population, indicating that naïve costs underestimate observed costs for institutionalized beneficiaries – a natural finding given the high cost of institutionalization. However, the dependence on number of comorbidities remains similar in both groups.

Figure A7C shows how number of hospitalizations affects costs, given the number of health conditions. The incidence of hospitalization is associated with a higher number of comorbidities. A subgroup of beneficiaries without hospitalizations shows a roughly linear dependence of costs on the number of calibrated conditions. For the those who were hospitalized at least for one night, the dependence is nonlinear, and the trend is similar to one observed in Figure A7A.

Finally, Figure A7D shows how survivor status affects costs. Death slightly reduces a positive bias of naïve cost. All told, Figure A7 suggests a need for a non-linear adjustment to the disease cost estimates. We have adjusted for the non-linearity on the propensity score models.

References

Elixhauser A, Steiner C, Palmer L. "Clinical Classifications Software (CCS)," U.S. Agency for Healthcare Research and Quality, Available: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp.

Kaiser Family Foundation report on Medicare Advantage enrollment. http://kff.org/medicare/fact-sheet/medicare-advantage-fact-sheet/, May 2014, Internet accessed: 2015

Kreuter F, Olson K, Wagner J, Yan T, Ezzati-Rice TB, Casas-Codero C, Lemay M, Pitches A, Groves RM, Raghunathan T, "Using Proxy Measures and Other Correlates of Survey Outcomes to Adjust for Nonresponse: Examples from Multiple Surveys," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2010; 173(2):389-407.

Little RJ. "Survey non-response adjustments," *International Statistical Review*, 1986; 54(2):139-157.

Meara, E., White, C., & Cutler, D. M. (2004). Trends In Medical Spending By Age, 1963-2000. Health Affairs, 23(4), 176-183

National Health Expenditure Data: http://www.cms.gov/Research-Statistics-Data-and Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html: Internet accessed, 2014

Raghunathan, TE, Lepkowski JM, Hoewyk JV, Solenberger P, "A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models," *Survey Methodology*, 2001; 27(1):85-95.

Raghunathan T, et al., "Model Based Estimation of Prevalence Rates and Trends Using Imperfect Data Sources," *mimeo* (2017).

Rosen A, et al., "Strengthening National Data to Better Measure What We Are Buying in Health Care: Reconciling National Health Expenditures with Detailed Survey Data", NBER Working Paper No. 23290, Issued in March 2017

Rosenbaum, P. R. and Rubin, D. B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects". Biometrika, 70: 41–55.

Rosenbaum, Paul R., and Donald B. Rubin. 1984. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association 79(387): 516-524

Figure A1: Propensity Score for Complete Medicare Enrollment, 2009 Community Sample

Figure A1a

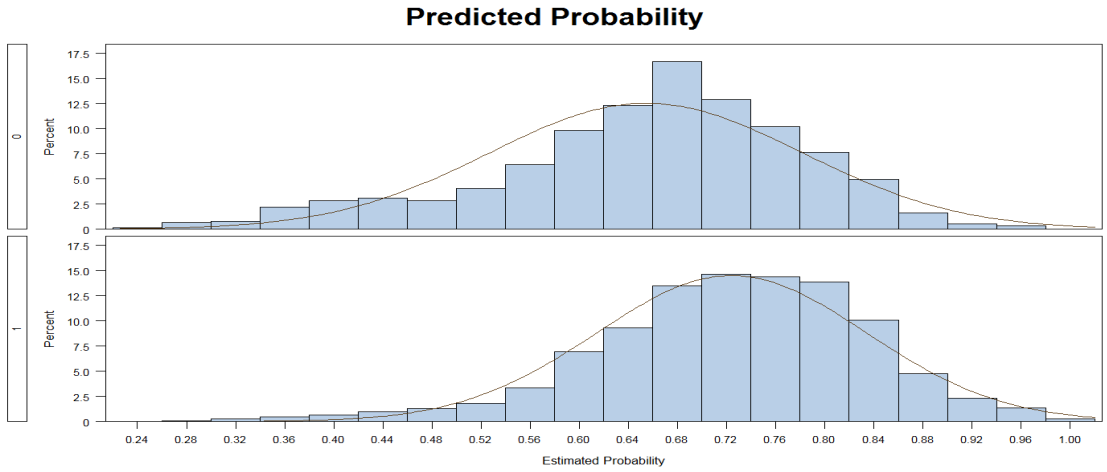
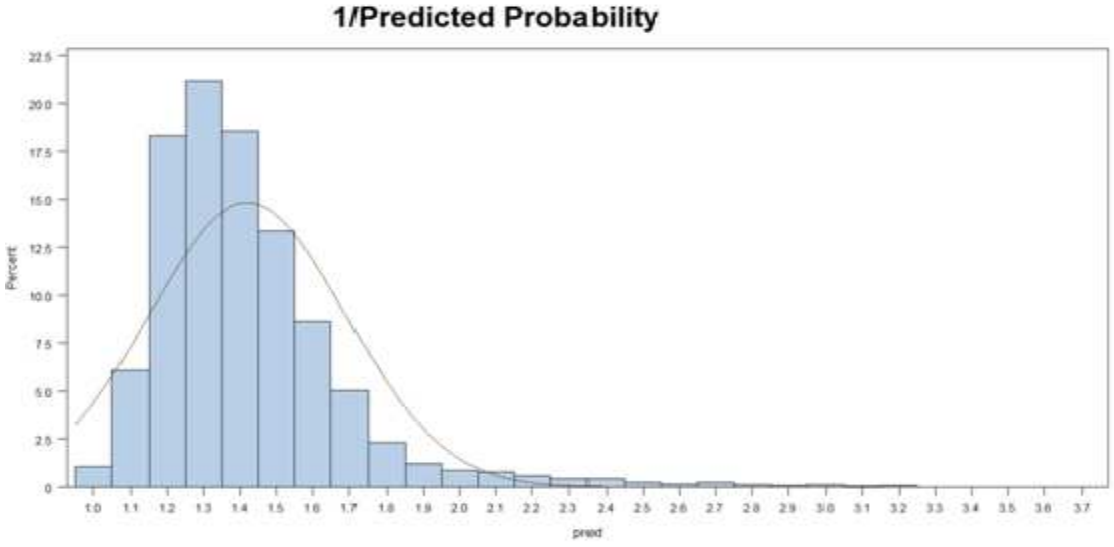


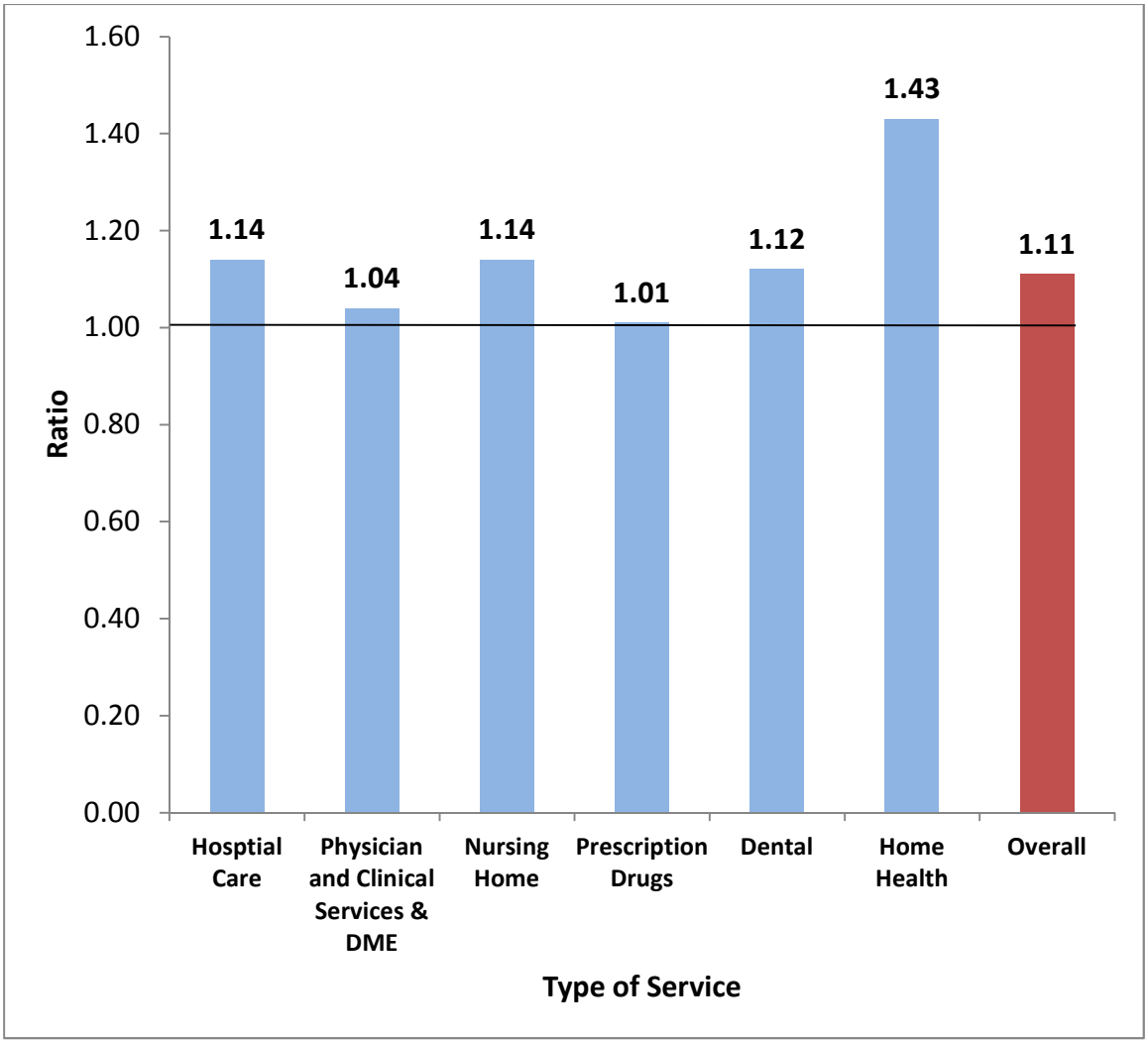
Figure A1b

Complete =1



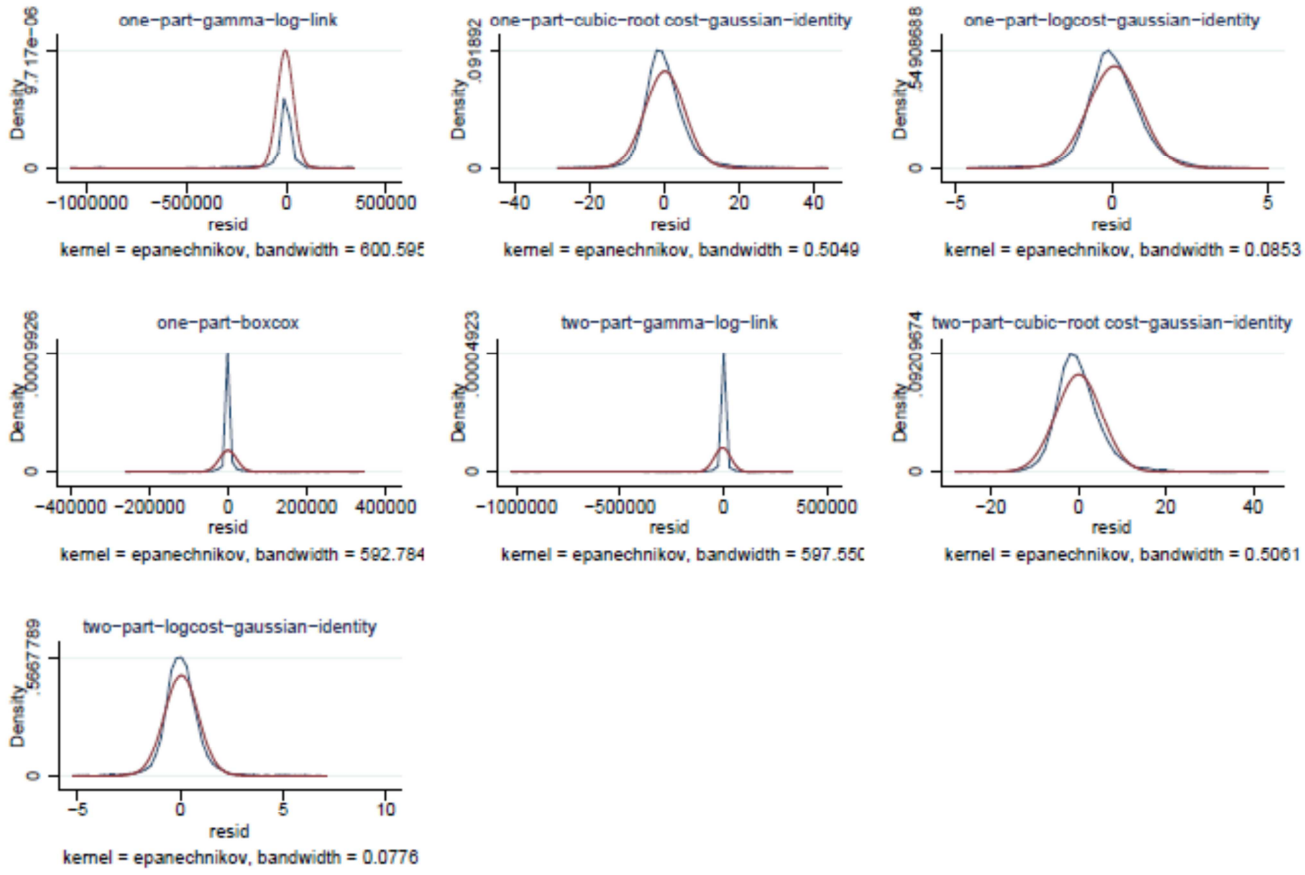
Note: Predictors in the propensity score model include demographics, health status (including ADLs and IADLs), and socioeconomic variables. Here, 1 is complete Medicare enrollment (full-year enrolment in traditional fee-for-service Parts A and B, and 0 includes at least some HMO enrollment. Figure 1b implies that adjustment factors are in a tight range.

Figure A2: Ratio of total spending: Adjusted NHEA and MCBS by Service Categories



Note: Figure 2 gives the adjustment factor for different types of services. Overall, the adjusted National Health Expenditure Accounts (NHEA) spending is 11 percent higher than the total spending reported in 2009 MCBS. We use these adjustment factors by service categories to adjust costs for MCBS-reported services to the national level.

Figure A3: Kernel density plot of residuals from Regression Models



Note: blue line:Kernel density estimate & red line: Normal density

Figure A4: Standardized normal probability plot of residuals

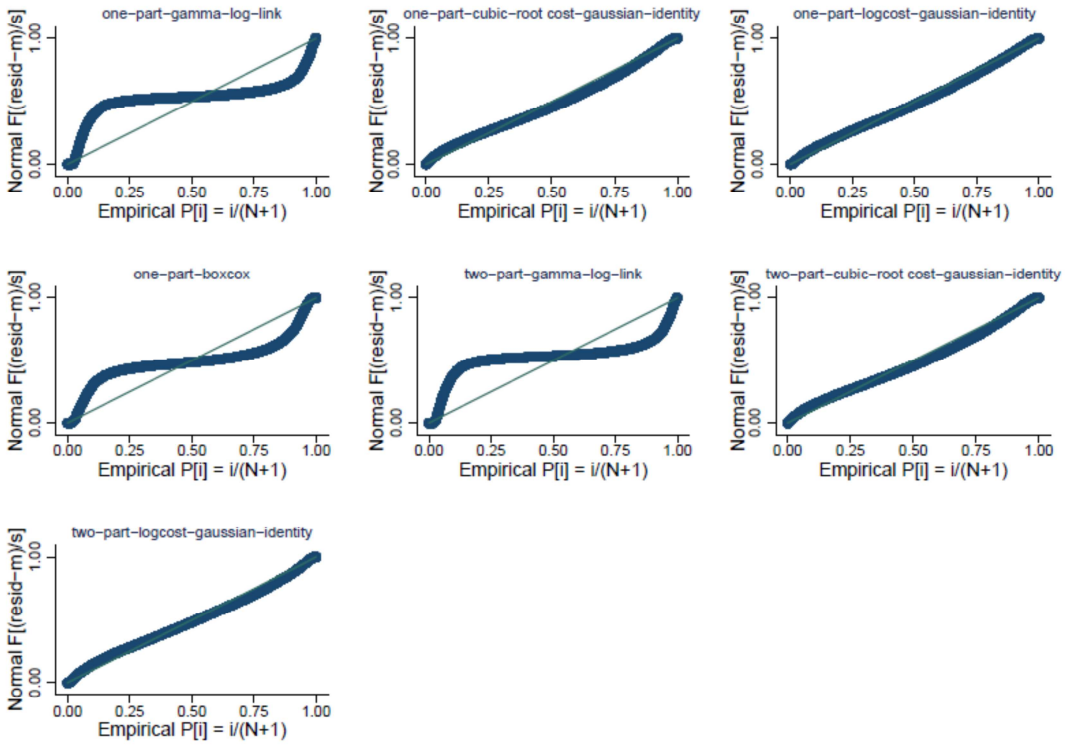


Figure A5: The quantile of residual against the quintiles of the normal distribution

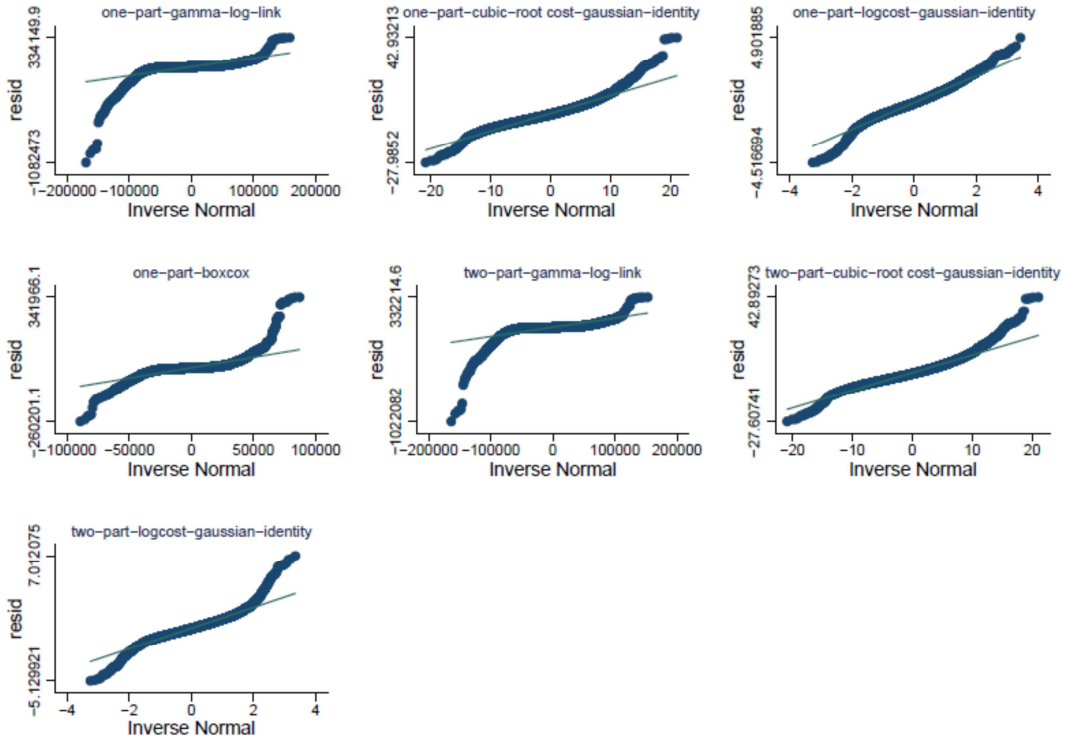


Figure A6: Scatter plot of Residuals

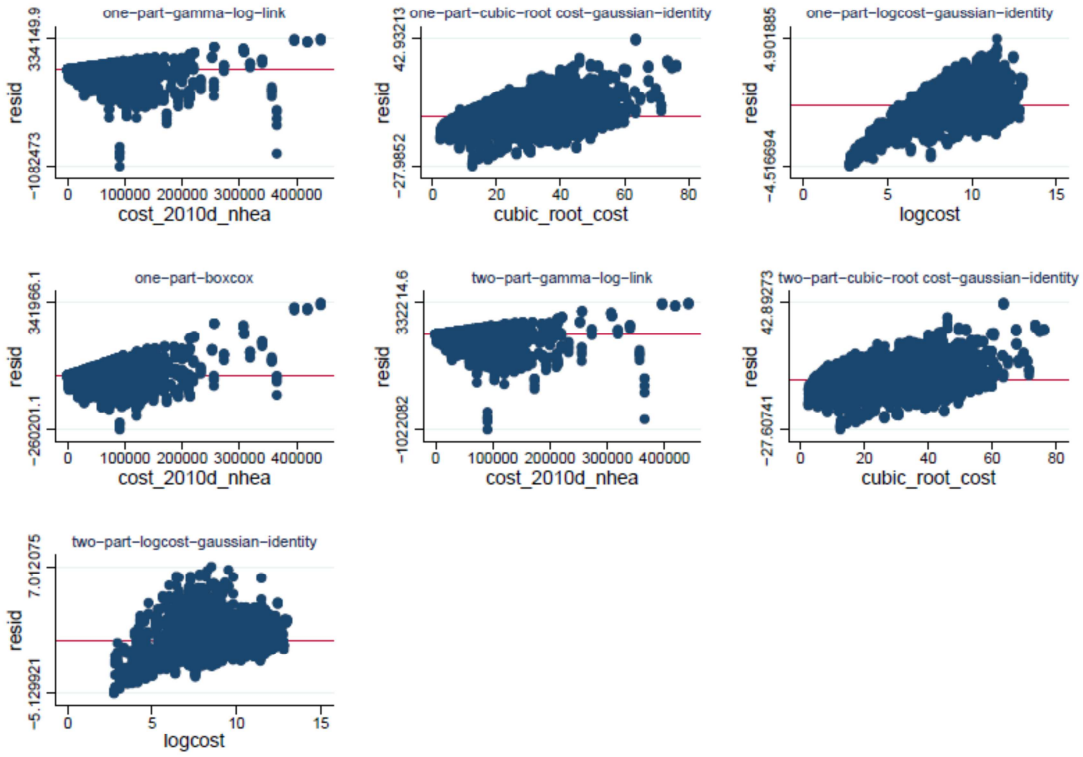


Figure A7: Difference between Naïve and Predicted Cost

Figure A7A: All individuals

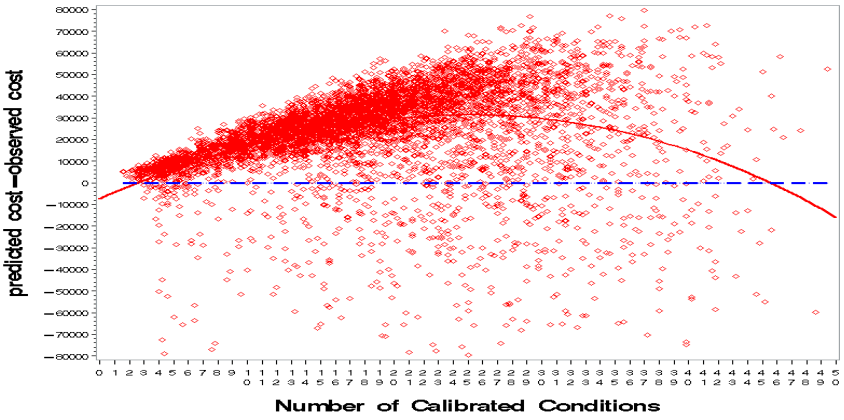


Figure A7B: Separated by Community and Institutional Living Status

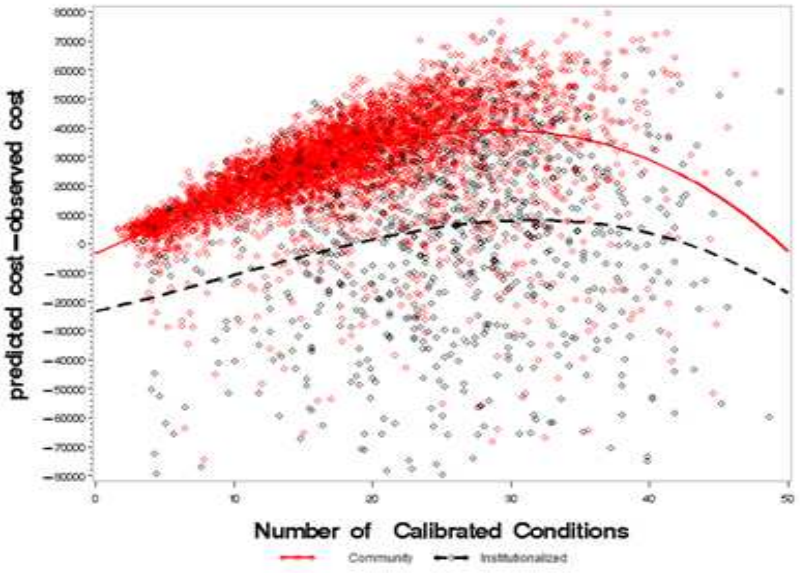


Figure A7 C: Separated by hospitalization experience

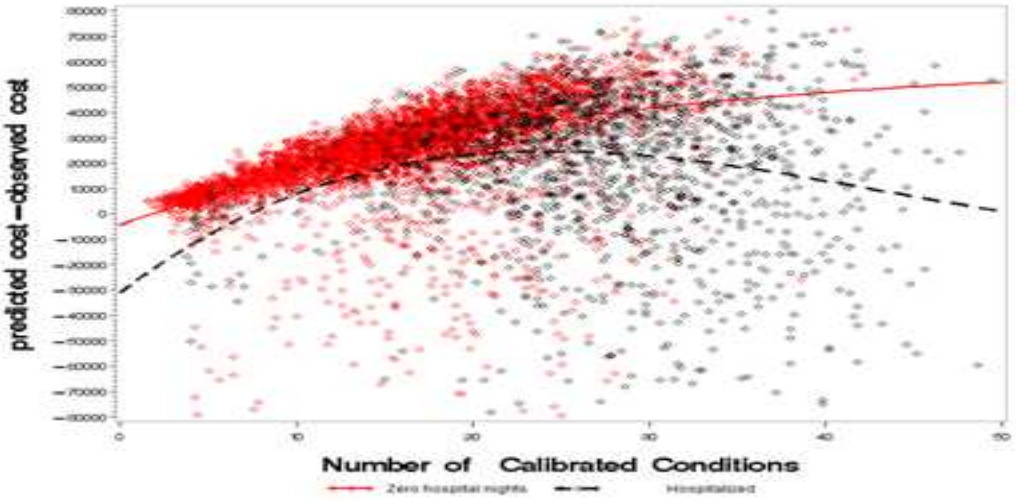
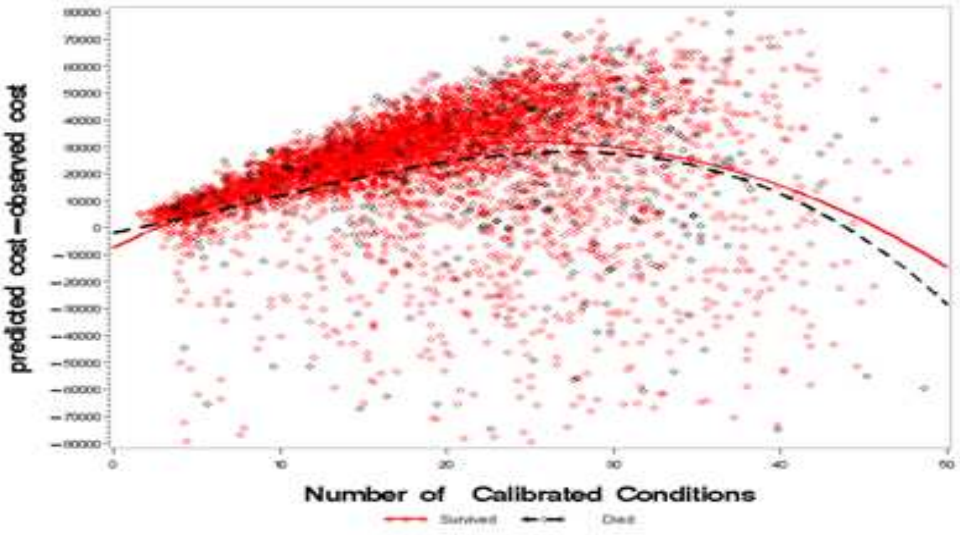


Figure A7D: Separated by survivorship



Note: Figures show the error term, measured as the difference between naïve predicted cost and true cost, against the number of health conditions, history of hospitalization and institutionalization and death. The results suggest a need for the non-linear adjustment.

Table A1: List of covariates used in HMO adjustment

[1] Age
[2] Age squared
[3] Asthma/emphysema
[4] Blood cholesterol checked
[5] Blood pressure checked-categorical
[6] Routine place receive care*Employment status
[7] Health compared to 1 yr ago-categorical
[8] Served in armed forces
[9] Died in study year
[10] Difficulty lifting/carrying 10 pounds-categorical
[11] Difficulty stooping/crouching/kneeling-categorical
[12] Difficulty walking 1/4 mi
[13] Education-categorical
[14] Ever smoke
[15] Flu shot in last year
[16] Employment status-have job
[17] Routine place receive care
[18] Self-reported health status-categorical
[19] Hearing
[20] Wear hearing aid
[21] Height (cm)-continuous
[22] Had hysterectomy
[23] Number of days in institution-squared
[24] Number of days in institution-continuous
[25] Inpatient nights-continuous
[26] Inpatient stays-continuous
[27] Male
[28] Mammogram/breast x-ray in last year
[29] Marital status category 2*Hispanic Race
[30] Marital status category 4*Black Race
[31] Marital status category 5*Hispanic Race
[32] Marital status
[33] Number of people in household
[34] Pap smear in last year
[35] Inpatient stays-squared
[36] Pneumonia vaccination
[37] Routine place receive care*Poverty status
[38] Poverty Status-categorical
[39] Any difficulty dressing
[40] Any difficulty eating
[41] PSA test in last year
[42] Race
[43] Poverty status category 5*Black Race
[44] Served in armed forces*Black Race
[45] Employment status*Hispanic Race
[46] Inpatient stays*Hispanic Race
[47] Smoke now
[48] Weight (kg)-continuous

Table A2: Hosmer-Lemeshow Goodness-of-Fit Test

	Hosmer and Lemeshow Goodness-of-Fit Test		
Community Calibrated	**Chi-Square**	**DF**	**Pr > ChiSq**
1	7.40	8	0.49
2	6.08	8	0.64
3	9.51	8	0.30
4	8.44	8	0.39
5	14.39	8	0.07

This table shows the Hosmer-Lemeshow Goodness of Fit test for the imputation used in the HMO adjustment.

Table A3: Complete Medicare and HMO Adjusted Samples

	Complete Medicare	HMO Adjusted
Variables	**N=6,200 (Weighted N=24,283,071)**	**N=6,200 (Weighted N=36,824,486)**
Gender	**Percent**	**Percent**
Men	42.14%	43.43%
Women	57.86	56.57
Age		
65-69	25.45	26.78
70-74	24.10	24.15
75-79	19.81	19.17
80-84	15.54	14.71
≥85	15.09	15.19
Race		
White	83.42	80.01
Black	6.86	8.15
Other	11.30	11.83
Education		
<=High School	52.86	53.94
Some College	26.09	26.17
College and above	21.05	19.89
Married	52.55	53.26
Health Status		
Excellent	17.07	16.45
Very good	30.74	29.81
Good	32.50	32.38
Fair	14.95	16.12
Poor	4.74	5.24

Note: Percentages and averages are weighted using sample weights. "Complete Medicare" population is defined as follows: (1) no participation in a Medicare Advantage for the year of study, and (2) enrollment in Medicare parts A & B for the full 12-month study period unless the participant died during the year.

Table A4. Adjustments to the National Health Expenditure Accounts: Exclusions and Transfers for Elderly

Health Care Service or Type of Expenditure	Amount in millions
Exclusions for Out-of-Scope Services or Expenditure	
Other Non-Durable Medical Equipment [a]	$19,327
Other Personal Health Care [a,d]	$18,685
Graduate Medical Education and Disproportionate Share Payments [b]	$6,998
Non-Patient Revenue [a]	$22,497
Exclusions for Out-of-Scope Populations	
Foreign Visitors [b]	$700
Total Exclusions	**$68,208**
Transfers between Service Categories	
Hospital-Based Personal Health Care [b]	$693
Hospital-Based Home Health Care [c]	$6,927
Hospital-Based Nursing Home Care [c]	$5,672
DME provided by Physicians	$477
Rx supplied in Hospitals [b]	$1,187
Rx supplied by Physicians [b]	$1,815
Other Professional Services provided in Physician Offices [a]	$13,372
Total Transfers	**$30,143**

a We follow Meara, White and Cutler (2004) and Sing et al. (2006) in this adjustment.
b We follow Sing et al. (2006) in this adjustment.
c We follow Meara, White and Cutler (2004) in this adjustment.
d We exclude all expenditure on Other Health, Residential and Personal Health Care plus from the Hospital services estimated to be hospital-based Other PHC services**.**

Table A5: List of Conditions in each multiple CCS categories and prevalence rates

Multi-level CCS / Condition label(ID)	Prevalence (Calibrated Conditions)	Combined Prevalence (Multi-level CCS)
1-Infectious and parasitic diseases		**65%**
Immunizations and screening for infectious disease (4)	53%	
Tuberculosis, STD, non-HIV, HIV, Other Infectious disease (106)	28%	
2-Neoplasms		**50%**
Cervical Cancer and Other Cancer (120)	22%	
Benign Neoplasm (13)	22%	
Skin Cancer (8)	14%	
Breast Cancer (9)	6%	
Prostate Cancer (11)	6%	
Colon Cancer (6)	3%	
Lung Cancer (7)	2%	
Hematologic Cancers (12)	2%	
3-Endocrine; nutritional; and metabolic diseases and immunity disorders		**83%**
Hyperlipidemia (18)	62%	
Other Endocrine Diseases (21)	38%	
Diabetes Mellitus (16)	32%	
Thyroid Disorders (15)	28%	
Gout and other crystal arthropathies (20)	8%	
Undiagnosed Diabetes Mellitus(17)	2%	
Undiagnosed Hyperlipidemia (19)	2%	
4-Diseases of the blood and blood-forming organs		**35%**
Anemias (22)	29%	
Other Hematologic Disease (23)	12%	
5-Mental Illness		**46%**
ETOH Abuse, Illicit Drug Use, Tobacco Use (107)	13%	
Depression (28)	13%	
Anxiety and Posttraumatic Stress Disorder (108)	13%	
Dementia (27)	12%	
Attention Deficit Hyperactivity Disorder ADD-ADHD, Mental Retardation (HCC term), Other Mental Health Disorders (109)	10%	
Schizophrenia (30)	5%	
Bipolar disorder (29)	3%	

Table A5 (Contd.): List of Conditions in each multiple CCS categories and prevalence rates

Multi-level CCS /Condition label (ID)	Prevalence (Calibrated Conditions)	Combined Prevalence (Multi-level CCS)
6-Diseases of the nervous system and sense organs		**82%**
Eye Disorders(45)	43%	
Cataract(43)	40%	
Other Disease of the Central Nervous System (48)	31%	
Glaucoma(44)	20%	
Other Ear Disorders(47)	16%	
Vestibular Disorders(46)	14%	
Headaches and Migraine(111)	12%	
Parkinson's Disease, Multiple Sclerosis, Paralysis(110)	9%	
Seizure Disorders(40)	6%	
Otitis Media(36)	6%	
7-Diseases of the circulatory system		**90%**
Hypertension(49)	71%	
Other Vascular Diseases(62)	36%	
Atrial Fibrillation and Flutter, Other Arrhythmias(112)	32%	
Other Cardiovascular Diseases(61)	31%	
Coronary Atherosclerosis and other heart disease(52)	27%	
Peripheral Vascular Disease(60)	20%	
Cerebrovascular disease(59)	18%	
Congestive Heart Failure(56)	18%	
Acute myocardial infarction(51)	11%	
Acute hemorrhagic stroke, Acute ischemic stroke(113)	9%	
Deep Vein Thrombosis or DVT(64)	6%	
Pulmonary Embolism(63)	4%	
Undiagnosed Hypertension(50)	2%	
Cardiac Arrest(55)	3%	
8-Diseases of the respiratory system		**66%**
Respiratory symptoms(70)	38%	
Acute respiratory infection(69)	27%	
Other Respiratory Diseases(71)	26%	
Chronic Obstructive Pulmonary Disease (aka Emphysema)(67)	18%	
Asthma(68)	14%	
Pneumonia (65)	11%	
Influenza(66)	3%	

Table A5 (Contd.): List of Conditions in each multiple CCS categories and prevalence rates

Multi-level CCS /Condition label (ID)	Prevalence (Calibrated Conditions)	Combined Prevalence (Multi-level CCS)
9-Diseases of the digestive system		**57%**
Reflux/Ulcer Disease, Biliary Tract Disease, Liver Disease, Gastrointestinal Bleeding, Other Gastrointestinal Disorders(114)	57%	
10-Diseases of the genitourinary system		**64%**
UTI , Urinary Incontinence, Other Genitourinary Diseases(116)	54%	
Acute Renal Failure(77)	9%	
Chronic Renal Failure, End-stage Renal Disease (ESRD)(115)	15%	
Hyperplasia of the Prostate(82)	15%	
11-Complications of menopause; pregnancy; childbirth		**10%**
Menopause, Pregnancy and Childbirth, Contraception and Procreation(117)	10%	
12-Diseases of the skin and subcutaneous tissue		**42%**
Dermatologic Diseases(87)	42%	
13-Diseases of the musculoskeletal system and connective tissue		**82%**
Other Rheumatic Diseases(92)	58%	
Osteoarthritis(89)	49%	
Back Pain(90)	34%	
Osteoporosis(91)	18%	
Rheumatoid Arthritis(88)	9%	
14-Congenital anomalies & Certain conditions originating in the perinatal period		**10%**
Congenital Disorders, Newborn conditions(118)	10%	
15-Injury and poisoning		**47%**
Trauma, Fractures, Poisoning and other injury, Motor vehicle accident(119)	46%	
Hip Fracture(96)	4%	

Table A5 (Contd.): List of Conditions in each multiple CCS categories and prevalence rates

Multi-level CCS /Condition label (ID)	Prevalence (Calibrated Conditions)	Combined Prevalence (Multi-level CCS)
16-Other conditions		**83%**
Signs and Symptoms(99)	74%	
Residual, unclassified, all other E codes(101)	48%	
17-Screening		**41%**
Screening: Breast Cancer (102)	22%	
Screening: Prostate Cancer(104)	13%	
Screening: Cervical Cancer(105)	9%	
Screening: Colon Cancer(103)	8%	

Note: CCS refers to 259 Clinical Classification Software categories delineated by the Agency for Health Care Research and Quality. Prevalence rates are for health conditions that were 'calibrated' using information from NHANES to refine the estimates of medical conditions in MCBS. First, missing data on demographic, socio-economic and self-reported medical conditions in MCBS and NHANES were imputed using a sequential regression multiple–imputation procedure. We then used a multi-step calibration processes to impute a set of 101 health indicators and 4 screening variables that we call "calibrated medical conditions." This process included these steps: (1) calibration of conditions that have Self-Report (SR) available in NHANES; (2) calibration of conditions with no Self-Report (SR) in NHANES and (3) calibration of the institutionalized population for all conditions. This calibration process produced five imputed data sets. The average prevalence rates reported here were calculated using all five imputed data sets using MIANALYZE SAS 9.4. Table A3 is sorted from highest to lowest prevalence within each broad disease category.

Appendix Table A6: Claims reporting single condition in different types of services in MCBS (2009)

	Services With Single Condition Claims In MCBS							
Conditions	**Carrier**	**Outpatient**	**Inpatient**	**SNF**	**HHA**	**DME**	**Hospice**	**Rx**
Immunizations and screening for infectious disease	Yes	Yes	No	No	No	No	No	N/A
Tuberculosis	Yes	Yes	No	Yes	Yes	Yes	Yes	N/A
STD, non-HIV	Yes	Yes	No	Yes	Yes	Yes	Yes	N/A
HIV	Yes	Yes	No	Yes	Yes	Yes	Yes	N/A
Other infectious disease	Yes	Yes	No	Yes	Yes	Yes	Yes	N/A
Cervical Cancer	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Other cancer	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Colon cancer	Yes	Yes	No	No	No	Yes	Yes	N/A
Lung Cancer	Yes	Yes	No	No	No	Yes	Yes	N/A
Skin Cancer	Yes	Yes	No	No	No	No	No	N/A
Breast Cancer	Yes	Yes	Yes	No	No	Yes	Yes	N/A
Prostate Cancer	Yes	Yes	Yes	No	No	Yes	Yes	N/A
Hematologic Cancers	Yes	Yes	No	No	No	No	Yes	N/A
Benign Neoplasm	Yes	Yes	No	No	No	No	No	N/A
Thyroid Disorders	Yes	Yes	No	No	No	No	No	N/A
Diabetes Mellitus	Yes	Yes	No	No	No	Yes	No	N/A
Undiagnosed Diabetes Mellitus	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Hyperlipidemia	Yes	Yes	No	No	No	No	No	N/A
Undiagnosed Hyperlipidemia	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Gout and other crystal arthropathies	Yes	Yes	No	No	No	No	No	N/A
Other Endocrine Diseases	Yes	Yes	No	No	No	Yes	Yes	N/A
Anemias	Yes	Yes	No	No	Yes	Yes	No	N/A
Other Hematologic Disease	Yes	Yes	No	No	No	Yes	No	N/A
ETOH Abuse	Yes	Yes	No	No	No	No	No	N/A
Illicit Drug Use	Yes	Yes	No	No	No	No	No	N/A
Tobacco Use	Yes	Yes	No	No	No	No	No	N/A
Dementia	Yes	Yes	Yes	No	No	Yes	Yes	N/A
Depression	Yes	Yes	No	No	No	No	No	N/A
Bipolar Disorder	Yes	Yes	No	No	No	No	No	N/A
Schizophrenia	Yes	Yes	No	No	No	No	No	N/A
Anxiety	Yes	Yes	No	No	No	No	No	N/A
Posttraumatic Stress Disorder (PTSD)	Yes	Yes	No	No	No	No	No	N/A

Appendix Table A6 (Contd.): Claims reporting single condition in different types of services in MCBS (2009)

Conditions	Carrier	Outpatient	Inpatient	SNF	HHA	DME	Hospice	Rx
Mental Retardation (HCC term)	Yes	Yes	No	Yes	No	No	No	N/A
Other Mental Health Disorders	Yes	Yes	No	Yes	No	No	No	N/A
Otitis Media	Yes	Yes	No	No	No	No	No	N/A
Vestibular Disorders	Yes	Yes	No	No	No	No	No	N/A
Other Ear Disorders	Yes	Yes	No	No	No	No	No	N/A
Parkinson's Disease	Yes	Yes	N0	No	No	Yes	Yes	N/A
Multiple Sclerosis	Yes	Yes	N0	No	No	Yes	Yes	N/A
Paralysis	Yes	Yes	No	No	No	Yes	Yes	N/A
Seizure Disorders	Yes	Yes	No	No	No	No	No	N/A
Headaches	Yes	Yes	No	No	No	No	No	N/A
Migraine	Yes	Yes	No	No	No	No	No	N/A
Cataract	Yes	Yes	No	No	No	Yes	No	N/A
Glaucoma	Yes	Yes	No	No	No	No	No	N/A
Eye Disorders	Yes	Yes	No	No	No	Yes	No	N/A
Other Disease of the Central Nervous System (CNS)	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Hypertension	Yes	Yes	No	Yes	Yes	No	No	N/A
Undiagnosed Hypertension	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Acute myocardial infarction (AMI)	Yes	Yes	No	No	No	Yes	No	N/A
Coronary Atherosclerosis and other heart disease	Yes	Yes	No	Yes	Yes	Yes	Yes	N/A
Atrial Fibrillation and Flutter	Yes	Yes	No	No	No	No	Yes	N/A
Other Arrhythmias	Yes	Yes	No	No	No	Yes	Yes	N/A
Cardiac Arrest (includes VF)	Yes	Yes	No	No	No	No	No	N/A
Congestive Heart Failure	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Acute hemorrhagic stroke	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Ischemic stroke	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Cerebrovascular Disease	Yes	Yes	No	Yes	Yes	Yes	Yes	N/A
Peripheral Vascular Disease	Yes	Yes	Yes	Yes	No	Yes	Yes	N/A
Other Cardiovascular Diseases	Yes	Yes	No	No	No	Yes	Yes	N/A
Other Vascular Diseases	Yes	Yes	No	No	No	No	No	N/A
Pulmonary embolism	Yes	Yes	No	Yes	No	No	Yes	N/A
DVT	Yes	Yes	No	Yes	No	No	No	N/A
Pneumonia (non-TB, non-STD)	Yes	Yes	No	No	Yes	Yes	Yes	N/A
Influenza	Yes	Yes	No	No	No	No	No	N/A
Chronic Obstructive Pulmonary	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Asthma	Yes	Yes	Yes	No	No	Yes	No	N/A
Acute respiratory infection	Yes	Yes	No	Yes	Yes	Yes	No	N/A
Respiratory symptoms	Yes	Yes	No	No	No	Yes	Yes	N/A
Other Respiratory Diseases	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Reflux/Ulcer Disease	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A

Appendix Table A6 (Contd.): Claims reporting single condition in different types of services in MCBS (2009)

Conditions	Carrier	Outpatient	Inpatient	SNF	HHA	DME	Hospice	Rx
Biliary Tract Disease	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Liver Disease	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Gastrointestinal Bleeding	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Other Gastrointestinal Disorders	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Acute Renal Failure	Yes	Yes	No	No	No	No	Yes	N/A
Chronic Renal Failure	Yes	Yes	No	No	No	Yes	Yes	N/A
Endstage Renal Disease (ESRD)	Yes	Yes	No	No	No	Yes	Yes	N/A
UTI	Yes	Yes	No	Yes	No	Yes	No	N/A
Urinary Incontinence	Yes	Yes	No	Yes	No	Yes	No	N/A
Other Genitourinary Diseases	Yes	Yes	No	Yes	No	Yes	No	N/A
Hyperplasia of the Prostate	Yes	Yes	No	No	No	Yes	No	N/A
Pregnancy and Childbirth	Yes	Yes	No	No	No	No	No	N/A
Menopause	Yes	Yes	No	No	No	No	No	N/A
Contraception and Procreation	Yes	Yes	No	No	No	No	No	N/A
Dermatologic Diseases	Yes	Yes	Yes	No	Yes	Yes	No	N/A
Rheumatoid Arthritis	Yes	Yes	No	No	No	Yes	No	N/A
Osteoarthritis	Yes	Yes	No	Yes	No	Yes	No	N/A
Back Pain	Yes	Yes	Yes	No	No	Yes	No	N/A
Osteoporosis	Yes	Yes	No	No	No	Yes	No	N/A
Other Rheumatic Diseases	Yes	Yes	No	Yes	Yes	Yes	No	N/A
Congenital Disorders	Yes	Yes	No	No	No	Yes	No	N/A
Newborn conditions	Yes	Yes	No	No	No	Yes	No	N/A
Trauma	Yes	Yes	No	No	No	No	No	N/A
Fractures	Yes	Yes	No	No	No	No	No	N/A
Poisoning and other injury	Yes	Yes	No	No	No	No	No	N/A
Motor vehicle accident	Yes	Yes	Yes	Yes	Yes	Yes	No	N/A
Hip Fracture	Yes	Yes	No	Yes	Yes	Yes	No	N/A
Signs and Symptoms	Yes	Yes	No	Yes	Yes	Yes	No	N/A
Residual, unclassified, all other E codes	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Attention Deficit Hyperactivity Disorder	Yes	Yes	No	Yes	No	Yes	Yes	N/A
Screening: Breast Cancer	Yes	Yes	No	No	No	No	No	N/A
Screening: Colon Cancer	Yes	Yes	No	No	No	No	No	N/A
Screening: Prostate Cancer	Yes	Yes	No	No	No	No	No	N/A
Screening: Cervical Cancer	Yes	Yes	No	No	No	No	No	N/A
Conditions satisfying single condition claim	102	102	13	36	20	59	40	0

Note: This table reports the full set of 101 medical conditions and 4 screening variables. We have combined these 101 medical conditions to 74 conditions. SNF: Skilled Nursing Facility, DME: Durable Medical Equipment; HHA: Home Health Agency.

Appendix Table A7: Dependent variable: log(spending) : GLM (Gaussian distribution with identity link)

Coefficient/Standard Error (in parenthesis)

Disease Category	Imputation 1	Imputation 2	Imputation 3	Imputation 4	Imputation 5
Residual, unclassified, all other E codes	0.112 (0.033)	0.121 (0.033)	0.125 (0.032)	0.167 (0.032)	0.127 (0.036)
Screening: Breast Cancer	-0.014 (0.041)	-0.047 (0.044)	0.009 (0.042)	0.059 (0.043)	-0.014 (0.047)
Screening: Colon Cancer	-0.022 (0.05)	0.099 (0.049)	0.002 (0.056)	0.021 (0.052)	0.051 (0.054)
Screening: Prostate Cancer	-0.015 (0.056)	0.059 (0.052)	-0.027 (0.055)	0.038 (0.05)	0.059 (0.057)
Screening: Cervical Cancer	-0.036 (0.066)	0.025 (0.056)	-0.032 (0.051)	-0.133 (0.075)	0.024 (0.053)
Tuberculosis, STD, non-HIV, HIV, Other Infectious disease	0.026 (0.031)	0.018 (0.026)	0.008 (0.044)	0.033 (0.03)	-0.027 (0.027)
ETOH Abuse, Illicit Drug Use, Tobacco Use	0.187 (0.056)	0.213 (0.049)	0.145 (0.056)	0.066 (0.05)	0.149 (0.049)
Anxiety, Posttraumatic Stress Disorder (PTSD)	0.057 (0.036)	0.077 (0.046)	-0.014 (0.036)	0.010 (0.05)	0.013 (0.044)
ADD-ADHD, Mental Retardation (HCC term), Other Mental Health Disorders	-0.031 (0.072)	0.110 (0.041)	0.004 (0.054)	0.050 (0.049)	-0.003 (0.049)
Prostate Cancer	0.215 (0.061)	0.242 (0.067)	0.256 (0.065)	0.187 (0.063)	0.233 (0.067)
Parkinson's Disease, Multiple Sclerosis, Paralysis	0.072 (0.065)	0.073 (0.058)	0.150 (0.048)	-0.019 (0.078)	0.126 (0.055)
Headaches, Migraine	0.003 (0.033)	-0.112 (0.056)	0.068 (0.034)	0.066 (0.039)	-0.004 (0.045)
Atrial Fibrillation and Flutter, Other Arrhythmias	0.057 (0.033)	0.060 (0.034)	0.074 (0.033)	-0.002 (0.033)	0.099 (0.03)
Acute hemorrhagic stroke, Acute hemorrhagic stroke	-0.039 (0.058)	0.128 (0.045)	0.036 (0.044)	0.037 (0.062)	-0.084 (0.057)

Appendix Table A7: Dependent variable : log(spending) : GLM (Gaussian distribution with identity link)

	Coefficient/Standard Error (in parenthesis)				
Disease Category	**Imputation 1**	**Imputation 2**	**Imputation 3**	**Imputation 4**	**Imputation 5**
Reflux/Ulcer Disease, Biliary Tract Disease, Liver Disease, Gastrointestinal Bleeding, Other Gastrointestinal Disorders	0.238 (0.036)	0.202 (0.036)	0.224 (0.036)	0.198 (0.04)	0.136 (0.037)
Chronic Renal Failure, Endstage Renal Disease (ESRD)	0.051 (0.051)	0.089 (0.037)	0.043 (0.04)	0.051 (0.049)	0.064 (0.061)
UTI, Urinary Incontinence, Other Genitourinary Diseases	0.142 (0.037)	0.127 (0.044)	0.103 (0.033)	0.129 (0.036)	0.089 (0.045)
Menopause, Pregnancy, and Childbirth, Contraception and Procreation	0.036 (0.049)	0.007 (0.051)	-0.006 (0.04)	0.017 (0.052)	0.070 (0.046)
Congenital Disorders, Newborn conditions	0.069 (0.052)	0.096 (0.039)	0.040 (0.063)	0.114 (0.047)	0.070 (0.052)
Trauma, Fractures, Poisoning and other injury, Motor vehicle accident	0.171 (0.034)	0.184 (0.035)	0.219 (0.03)	0.127 (0.043)	0.192 (0.034)
Hematologic Cancers	0.249 (0.113)	0.212 (0.08)	0.253 (0.081)	0.210 (0.09)	0.161 (0.126)
Cervical Cancer, Other Cancer	0.173 (0.033)	0.163 (0.04)	0.144 (0.038)	0.203 (0.03)	0.184 (0.032)
Benign Neoplasm	0.062 (0.044)	0.134 (0.037)	0.111 (0.045)	0.161 (0.033)	0.134 (0.036)
Thyroid Disorders	0.037 (0.03)	0.060 (0.028)	0.053 (0.029)	0.045 (0.031)	0.032 (0.033)
Diabetes Mellitus	0.186 (0.029)	0.176 (0.03)	0.173 (0.03)	0.160 (0.029)	0.204 (0.028)
Undiagnosed Diabetes Mellitus	0.224 (0.122)	-0.164 (0.155)	-0.123 (0.201)	-0.586 (0.302)	0.143 (0.103)
Hyperlipidemia	0.122 (0.035)	0.147 (0.032)	0.130 (0.042)	0.142 (0.036)	0.102 (0.041)
Undiagnosed Hyperlipidemia	-0.157 (0.218)	0.170 (0.168)	-0.382 (0.289)	-0.161 (0.256)	0.223 (0.187)

Appendix Table A7: Dependent variable : log(spending) : GLM (Gaussian distribution with identity link)

	Coefficient/Standard Error (in parenthesis)				
Disease Category	**Imputation 1**	**Imputation 2**	**Imputation 3**	**Imputation 4**	**Imputation 5**
Gout and other crystal arthropathies	0.007	0.072	-0.001	-0.025	-0.086
	(0.047)	(0.048)	(0.049)	(0.06)	(0.07)
Other Endocrine Diseases	0.087	0.019	0.069	0.092	0.096
	(0.038)	(0.04)	(0.032)	(0.038)	(0.032)
Anemias	0.101	0.065	0.098	0.096	0.116
	(0.039)	(0.041)	(0.025)	(0.036)	(0.032)
Other Hematologic Disease	0.120	0.001	0.065	0.075	0.054
	(0.037)	(0.056)	(0.05)	(0.047)	(0.042)
Dementia	0.108	-0.119	0.067	-0.010	0.127
	(0.051)	(0.094)	(0.047)	(0.06)	(0.035)
Depression	0.021	0.088	0.101	0.068	0.093
	(0.045)	(0.041)	(0.035)	(0.04)	(0.036)
Bipolar Disorder	0.062	0.017	-0.039	0.224	-0.014
	(0.079)	(0.07)	(0.081)	(0.083)	(0.118)
Schizophrenia	0.037	0.075	0.118	0.093	0.057
	(0.085)	(0.061)	(0.065)	(0.062)	(0.057)
Otitis Media	0.099	-0.053	-0.007	0.060	0.055
	(0.054)	(0.058)	(0.06)	(0.053)	(0.058)
Immunizations and screening for infectious disease	0.111	0.095	0.070	0.111	0.087
	(0.032)	(0.029)	(0.039)	(0.035)	(0.038)
Seizure Disorders	0.002	-0.037	0.012	0.151	0.096
	(0.058)	(0.066)	(0.076)	(0.055)	(0.066)
Cataract	0.034	0.109	0.087	0.103	0.090
	(0.042)	(0.028)	(0.033)	(0.029)	(0.033)
Glaucoma	0.087	0.113	0.114	0.046	0.098
	(0.031)	(0.025)	(0.028)	(0.038)	(0.029)
Eye Disorders	0.123	0.112	0.086	0.056	0.115
	(0.036)	(0.027)	(0.03)	(0.038)	(0.031)
Vestibular Disorders	-0.025	-0.017	-0.022	0.027	-0.019
	(0.036)	(0.052)	(0.03)	(0.037)	(0.045)
Other Ear Disorders	0.079	0.065	0.042	0.062	0.096
	(0.037)	(0.038)	(0.04)	(0.032)	(0.033)

Appendix Table A7: Dependent variable : log(spending) : GLM (Gaussian distribution with identity link)

	Coefficient/Standard Error (in parenthesis)				
Disease Category	**Imputation 1**	**Imputation 2**	**Imputation 3**	**Imputation 4**	**Imputation 5**
Other Disease of the Central Nervous System (CNS)	0.092 (0.039)	0.169 (0.031)	0.118 (0.036)	0.090 (0.039)	0.130 (0.036)
Hypertension	0.346 (0.051)	0.330 (0.048)	0.311 (0.049)	0.334 (0.059)	0.385 (0.058)
Undiagnosed Hypertension	0.191 (0.156)	-0.003 (0.195)	-0.233 (0.196)	0.291 (0.242)	0.090 (0.194)
Acute myocardial infarction (AMI)	0.014 (0.061)	0.136 (0.039)	0.056 (0.044)	0.053 (0.065)	0.019 (0.056)
Coronary Atherosclerosis and other heart disease	0.152 (0.033)	0.130 (0.032)	0.133 (0.03)	0.169 (0.034)	0.146 (0.034)
Cardiac Arrest (includes VF)	0.171 (0.076)	0.133 (0.089)	0.130 (0.066)	0.147 (0.063)	0.057 (0.059)
Congestive Heart Failure	0.048 (0.033)	-0.026 (0.047)	0.044 (0.037)	0.032 (0.041)	0.040 (0.04)
Cerebrovascular Disease	0.100 (0.03)	0.058 (0.034)	0.032 (0.053)	0.022 (0.038)	0.056 (0.033)
Colon cancer	0.117 (0.076)	0.064 (0.109)	0.066 (0.092)	0.106 (0.071)	0.155 (0.055)
Peripheral Vascular Disease	0.010 (0.039)	0.018 (0.038)	0.034 (0.041)	0.021 (0.04)	-0.016 (0.048)
Other Cardiovascular Diseases	0.069 (0.029)	0.119 (0.03)	0.129 (0.032)	0.135 (0.03)	0.095 (0.035)
Other Vascular Diseases	0.069 (0.032)	0.091 (0.035)	0.061 (0.029)	0.048 (0.029)	0.068 (0.026)
Pulmonary embolism	0.133 (0.062)	-0.099 (0.117)	0.004 (0.085)	0.055 (0.08)	0.071 (0.076)
Deep Vein Thrombosis (DVT)	-0.036 (0.047)	-0.193 (0.074)	-0.060 (0.064)	-0.044 (0.06)	-0.121 (0.101)
Pneumonia (non-TB, non-STD)	0.011 (0.043)	0.043 (0.042)	-0.027 (0.046)	-0.008 (0.044)	0.063 (0.044)
Influenza	0.221 (0.06)	0.044 (0.076)	0.203 (0.077)	0.196 (0.06)	-0.091 (0.113)

Appendix Table A7: Dependent variable : log(spending) : GLM (Gaussian distribution with identity link)

Coefficient/Standard Error (in parenthesis)

Disease Category	Imputation 1	Imputation 2	Imputation 3	Imputation 4	Imputation 5
Chronic Obstructive Pulmonary Disease (aka Emphysema)	0.090 (0.033)	0.116 (0.038)	0.109 (0.037)	0.117 (0.036)	0.133 (0.038)
Asthma	0.108 (0.035)	0.090 (0.042)	-0.002 (0.061)	0.063 (0.045)	0.017 (0.061)
Acute respiratory infection	0.049 (0.027)	0.065 (0.029)	0.066 (0.036)	0.048 (0.031)	0.026 (0.029)
Lung Cancer	0.252 (0.08)	0.240 (0.07)	0.248 (0.081)	0.226 (0.091)	0.195 (0.086)
Respiratory symptoms	0.104 (0.03)	0.085 (0.032)	0.105 (0.031)	0.105 (0.034)	0.103 (0.036)
Other Respiratory Diseases	0.101 (0.024)	0.099 (0.031)	0.100 (0.036)	0.066 (0.036)	0.122 (0.032)
Acute Renal Failure	-0.056 (0.062)	-0.008 (0.043)	-0.007 (0.062)	-0.083 (0.078)	-0.061 (0.076)
Skin Cancer	0.021 (0.04)	-0.005 (0.056)	0.064 (0.032)	0.018 (0.043)	0.008 (0.04)
Hyperplasia of the Prostate	0.257 (0.058)	0.226 (0.059)	0.230 (0.053)	0.177 (0.068)	0.150 (0.065)
Dermatologic Diseases	0.106 (0.032)	0.103 (0.033)	0.076 (0.032)	0.109 (0.03)	0.150 (0.027)
Rheumatoid Arthritis	0.083 (0.061)	-0.091 (0.074)	0.050 (0.057)	0.111 (0.067)	0.021 (0.069)
Osteoarthritis	0.028 (0.033)	-0.002 (0.045)	-0.003 (0.034)	0.005 (0.043)	0.016 (0.043)
Breast Cancer	0.132 (0.066)	0.192 (0.063)	0.228 (0.052)	0.200 (0.062)	0.131 (0.062)
Back Pain	0.078 (0.031)	0.059 (0.04)	0.088 (0.03)	0.135 (0.036)	0.078 (0.029)
Osteoporosis	0.039 (0.039)	-0.032 (0.046)	0.052 (0.033)	-0.028 (0.043)	0.080 (0.03)

Appendix Table A7: Dependent variable : log(spending) : GLM (Gaussian distribution with identity link)

Coefficient/Standard Error (in parenthesis)

Disease Category	Imputation 1	Imputation 2	Imputation 3	Imputation 4	Imputation 5
Other Rheumatic Diseases	0.169 (0.04)	0.184 (0.039)	0.161 (0.038)	0.213 (0.044)	0.176 (0.035)
Hip Fracture	-0.262 (0.122)	0.085 (0.064)	-0.042 (0.074)	0.091 (0.075)	0.049 (0.053)
Signs and Symptoms	0.475 (0.052)	0.438 (0.05)	0.498 (0.053)	0.401 (0.052)	0.468 (0.051)
OTHER COVARIATES					
Health compared to one year ago (about the same)	-0.184 (0.055)	-0.200 (0.055)	-0.178 (0.055)	-0.183 (0.056)	-0.200 (0.054)
Health compared to one year ago (somewhat worse/much worse)	-0.100 (0.053)	-0.118 (0.056)	-0.091 (0.053)	-0.103 (0.054)	-0.089 (0.054)
Ever served in armed forces	0.102 (0.07)	0.129 (0.071)	0.088 (0.069)	0.114 (0.069)	0.137 (0.069)
Death in year	0.141 (0.100)	0.216 (0.106)	0.153 (0.104)	0.187 (0.105)	0.163 (0.101)
Any Difficulty lifting/carrying 10 pounds.	0.028 (0.017)	0.023 (0.018)	0.019 (0.017)	0.014 (0.017)	0.017 (0.018)
Any difficulty stooping/crouching/kneeling	0.039 (0.019)	0.039 (0.019)	0.041 (0.019)	0.037 (0.021)	0.027 (0.021)
Any difficulty walking 1/4 mi. or 2-3 blocks.	0.203 (0.043)	0.197 (0.041)	0.201 (0.041)	0.219 (0.041)	0.204 (0.042)
Education-9-11 grade (includes 12 with no diploma)	0.036 (0.094)	0.050 (0.098)	0.044 (0.094)	0.044 (0.095)	0.067 (0.096)
Education- High school grad/GED or equivalent	0.116 (0.085)	0.102 (0.088)	0.119 (0.085)	0.100 (0.09)	0.119 (0.088)
Education- some college of AA degree	0.218 (0.098)	0.211 (0.099)	0.223 (0.098)	0.206 (0.102)	0.231 (0.104)
Education- college grad or more	0.243 (0.087)	0.213 (0.089)	0.242 (0.088)	0.215 (0.095)	0.230 (0.095)
Ever smoked cigarettes/cigars/tobacco	0.008 (0.039)	-0.023 (0.04)	0.010 (0.038)	0.002 (0.04)	-0.004 (0.042)

Appendix Table A7: Dependent variable : log(spending) : GLM (Gaussian distribution with identity link)

Coefficient/Standard Error (in parenthesis)

Disease Category	Imputation 1	Imputation 2	Imputation 3	Imputation 4	Imputation 5
General health compared to others same age (Very good)	0.089 (0.069)	0.125 (0.07)	0.110 (0.066)	0.115 (0.072)	0.109 (0.073)
General health compared to others same age (Good)	0.217 (0.074)	0.269 (0.072)	0.241 (0.07)	0.255 (0.075)	0.247 (0.074)
General health compared to others same age (Fair)	0.319 (0.087)	0.373 (0.082)	0.337 (0.088)	0.362 (0.085)	0.333 (0.086)
General health compared to others same age (Poor)	0.431 (0.1)	0.449 (0.094)	0.427 (0.1)	0.496 (0.103)	0.451 (0.101)
Days in institution	0.004 (0.000)	0.004 (0.000)	0.004 (0.000)	0.004 (0.000)	0.004 (0.000)
Count of inpatient nights	0.005 (0.003)	0.006 (0.003)	0.006 (0.003)	0.005 (0.003)	0.005 (0.003)
Count of inpatient stays	0.327 (0.025)	0.333 (0.024)	0.324 (0.023)	0.351 (0.027)	0.330 (0.024)
Male	-0.288 (0.108)	-0.314 (0.119)	-0.262 (0.108)	-0.263 (0.107)	-0.264 (0.112)
Marital Status-Widowed	-0.011 (0.036)	-0.017 (0.037)	0.002 (0.036)	0.000 (0.037)	-0.004 (0.037)
Marital Status-Divorced or Separated	-0.098 (0.082)	-0.087 (0.084)	-0.072 (0.083)	-0.084 (0.08)	-0.090 (0.084)
Marital Status- Never married	-0.394 (0.161)	-0.421 (0.162)	-0.388 (0.161)	-0.405 (0.16)	-0.431 (0.159)
Poverty category (near poor)	-0.069 (0.092)	-0.080 (0.091)	-0.071 (0.094)	-0.070 (0.095)	-0.107 (0.096)
Poverty category (low income)	-0.044 (0.077)	-0.036 (0.083)	-0.038 (0.083)	-0.054 (0.086)	-0.035 (0.083)
Poverty category (middle income)	0.050 (0.074)	0.046 (0.073)	0.050 (0.078)	0.030 (0.08)	0.032 (0.077)
Poverty category (high income)	0.140 (0.071)	0.162 (0.072)	0.155 (0.079)	0.146 (0.079)	0.151 (0.077)

Appendix Table A7: Dependent variable : log(spending) : GLM (Gaussian distribution with identity link)

	Coefficient/Standard Error (in parenthesis)				
Disease Category	**Imputation 1**	**Imputation 2**	**Imputation 3**	**Imputation 4**	**Imputation 5**
Any difficulty dressing	0.065 (0.063)	0.104 (0.06)	0.049 (0.058)	0.075 (0.06)	0.076 (0.059)
Any difficulty eating	0.083 (0.064)	0.082 (0.067)	0.124 (0.071)	0.132 (0.069)	0.123 (0.064)
Private health insurance coverage for the year	-0.023 (0.059)	-0.030 (0.059)	-0.037 (0.061)	-0.020 (0.059)	-0.033 (0.059)
Race- Black	-0.042 (0.085)	-0.065 (0.085)	-0.069 (0.081)	-0.053 (0.086)	-0.061 (0.087)
Race- Hispanic	0.030 (0.096)	0.006 (0.106)	-0.004 (0.099)	0.011 (0.105)	0.031 (0.099)
Race- Other	0.007 (0.096)	0.036 (0.09)	0.014 (0.092)	0.053 (0.098)	0.036 (0.093)
Current smoker	-0.221 (0.111)	-0.227 (0.114)	-0.206 (0.106)	-0.224 (0.108)	-0.197 (0.107)
Age of the beneficiary	-0.001 (0.003)	0.000 (0.003)	-0.001 (0.003)	0.000 (0.003)	-0.002 (0.003)
Body Mass Index	0.003 (0.003)	0.003 (0.004)	0.003 (0.003)	0.002 (0.003)	0.004 (0.003)
Intercept	6.407 (0.31)	6.396 (0.293)	6.440 (0.319)	6.330 (0.311)	6.436 (0.314)

Note: The cost model regressed the logarithm of spending (plus $1) on the 78 condition and screening indicators and 22 demographic and other covariates. Lifestyle covariates include information on inpatient activity and institutionalization, as well as difficulty with common life-skills such as eating, walking, stooping, lifting and dressing. In addition to the common demographic characteristics such as age, race, and gender, the regression incorporates self-reported indications of health comparisons to one year prior and to others of the same age as well as smoking history, poverty category, educational attainment, and service in the armed forces. Two of the condition groups have negative coefficients for all five of the multiply imputed data sets: DVT and Acute Renal Failure. As a result we are unable to attribute spending to these conditions. In the situation where coefficients were negative for only a selection of the five multiples they were treated as missing in the subsequent steps of the analysis.

Table A8: Adjusted Observed Cost Model Estimates

Parameters	Coefficient/Standard Error (in parenthesis)
Intercept	0.3956 (0.0462)
Number of comorbidities	-0.0042 (0.0015)
Number of comorbidities squared	0.0001 (0.0001)
Any hospitalization	-0.1663 (0.0166)
Number of nights in hospital	0.0143 (0.0004)
Number of hospital admissions	0.0246 (0.0068)
Number of days in an institution	0.0021 (0.0000)
Patient survived the calendar year	0.0828 (0.0332)
Number of months survived in the calendar year (if deceased)	0.0071 (0.0042)
Number of outpatient claims	0.0058 (0.0007)

Table A9: Attributed Cost by Medical Conditions using different methods

		Average Cost			Totals Cost (in Billions)		
ID	**Multiple-CCS/Conditions**	**Claims**	**Regression**	**Propensity Score**	**Claims**	**Regression**	**Propensity Score**
1-Infectious and parasitic diseases							
4	Immunizations and screening for infectious disease	$22	$556	$15	$0.40	$10.89	$0.30
106	Tuberculosis, STD, non-HIV, HIV, Other Infectious disease	2,469	163	1,372	19.88	1.68	14.14
2-Neoplasms							
6	Colon cancer	2,802	807	1,240	1.89	0.92	1.41
7	Lung Cancer	4,102	2,758	4,444	2.30	1.56	2.52
8	Skin Cancer	1,547	126	308	5.17	0.65	1.58
9	Breast Cancer	1,737	1,286	1,123	2.86	2.59	2.28
11	Prostate Cancer	1,832	1,671	988	3.58	3.43	2.05
12	Hematologic Cancers	4,214	2,131	3,431	2.60	1.73	2.82
13	Benign Neoplasm	1,309	710	47	8.04	5.76	0.38
120	Cervical Cancer and Other Cancers	2,410	1,317	1,143	15.00	10.90	9.46
3-Endocrine; nutritional; and metabolic diseases and immunity disorders							
15	Thyroid Disorders	859	322	448	7.10	3.29	4.59
16	Diabetes Mellitus	862	1,454	1,119	10.19	17.22	13.25
17	Undiagnosed Diabetes Mellitus	637	485	43	0.01	0.33	0.03
18	Hyperlipidemia	623	806	720	14.04	18.51	16.57
19	Undiagnosed Hyperlipidemia	540	445	-856	0.07	0.25	-0.51
20	Gout and other crystal arthropathies	825	119	384	1.28	0.34	1.11
21	Other Endocrine Diseases	1,362	639	1,422	16.47	8.92	19.81

Table A9 (Contd.): Attributed Cost by Medical Conditions using different methods

ID	Multiple-CCS/ Conditions	Average Cost			Totals Cost (in Billions)		
		Claims	Regression	Propensity Score	Claims	Regression	Propensity Score
4-Diseases of the blood and blood-forming organs							
22	Anemia	$1,393	$948	$2,147	$12.22	$9.99	$22.62
23	Other Hematologic Disease	1,818	662	2,328	5.44	2.87	10.15
5-Mental Illness							
107	ETOH Abuse, Illicit Drug and Tobacco	2,156	1,275	1,188	5.77	6.22	5.85
27	Dementia	3,388	783	2,229	10.93	3.53	10.11
28	Depression	1,724	815	1,812	5.85	3.90	8.65
29	Bipolar Disorder	1,451	405	654	0.34	0.43	0.62
30	Schizophrenia	1,850	909	3,522	1.81	1.78	6.94
108	Anxiety, Posttraumatic Stress Disorder	1,017	254	450	3.01	1.22	2.20
109	Attention Deficit Hyperactivity Disorder ADD-ADHD., Mental Retardation (HCC term), Other Mental Health Disorders	1,526	346	1,429	3.05	1.25	5.12
6-Diseases of the nervous system and sense organs							
36	Otitis Media	673	284	379	0.77	0.68	0.90
46	Vestibular Disorders	699	40	351	2.58	0.21	1.80
47	Other Ear Disorders	854	461	165	3.77	2.78	1.00
110	Parkinson's Disease, MS, Paralysis	2,638	849	1,635	3.29	2.88	5.48
40	Seizure Disorders	1,893	549	1,984	1.61	1.16	4.07
111	Headaches, Migraine	753	199	169	1.85	0.93	0.77
43	Cataract	960	490	146	12.40	7.21	2.15
44	Glaucoma	701	544	141	3.70	3.95	1.02
45	Eye Disorders	1,017	615	174	14.25	9.69	2.75
48	Disease of the Central Nervous System	2,089	1,192	2,203	19.76	13.53	24.99

Table A9 (Contd.): Attributed Cost by Medical Conditions using different methods

		Average Cost			Total Cost (in Billions)		
ID	**Multiple-CCS/Calibrated Conditions**	**Claims**	**Regression**	**Propensity Score**	**Claims**	**Regression**	**Propensity Score**
7-Diseases of the circulatory system							
49	Hypertension	$883	$2,619	$1,347	$22.33	$68.86	$35.42
50	Undiagnosed Hypertension	840	445	-897	0.18	0.37	-0.81
51	Acute myocardial infarction (AMI)	4,680	459	1,142	4.63	1.98	4.71
52	Coronary Athero. and other heart disease	2,524	1,371	1,687	24.64	13.40	16.49
112	Atrial Fib. and Flutter, Other Arrhythmias	1,293	486	1,114	11.41	5.63	12.97
55	Cardiac Arrest (includes VF)	4,407	1,384	2,400	1.64	1.50	2.63
56	Congestive Heart Failure	2,009	375	2,514	10.18	2.41	16.20
113	Acute hemorrhagic or Ischemic Stroke	2,473	452	1,593	5.09	1.46	5.06
59	Cerebrovascular Disease	1,470	514	1,341	7.60	3.46	9.01
60	Peripheral Vascular Disease	1,873	158	1,071	10.42	1.15	7.75
61	Other Cardiovascular Diseases	2,692	971	1,346	25.66	11.05	15.32
62	Other Vascular Diseases	1,848	552	1,137	20.81	7.30	15.07
63	Pulmonary embolism	2,993	486	1,593	1.30	0.70	2.31
64	Deep Vein Thrombosis (DVT)	2,259	-----	1,200	2.22	-----	2.68
8-Diseases of the respiratory system							
65	Pneumonia (non-TB, non-STD)	2,909	285	2,318	8.01	1.14	9.31
66	Influenza	958	897	170	0.29	1.13	0.21
67	Chronic Obstructive Pulmonary Disease	1,373	1,079	1,477	9.16	7.22	9.88
68	Asthma	1,178	411	817	3.37	1.99	4.07
69	Acute respiratory infection	604	331	83	4.67	3.29	0.82
70	Respiratory symptoms	1,033	901	1,669	12.39	12.45	23.05
71	Other Respiratory Diseases	2,073	877	1,592	15.55	8.27	15.05

Table A9 (Contd.): Attributed Cost by Medical Conditions using different methods

ID	Multiple-CCS/Calibrated Conditions	Average Cost			Total Cost (in billions)		
		Claims	Regression	Propensity Score	Claims	Regression	Propensity Score
9-Diseases of the digestive system							
114	Reflux/Ulcer Disease, Biliary Tract Disease, Liver Disease, Gastrointestinal Bleeding, Other Gastrointestinal Disorders	$1,747	$1,588	$1,370	$28.06	$33.09	$28.52
10-Diseases of the genitourinary system							
77	Acute Renal Failure	3,738	----	3,080	7.74	-----	10.65
115	Chronic Renal Failure & ESRD	2,449	601	1,757	8.39	3.32	9.74
116	UTI , Urinary Incontinence, Other Genitourinary Diseases	1,096	891	1,020	18.12	17.60	20.14
82	Hyperplasia of the Prostate	813	1,369	187	3.85	7.65	1.05
11-Complications of menopause, pregnancy, childbirth, and the puerperium							
117	Menopause, Pregnancy and Childbirth, Contraception and Procreation	563	130	-45	1.43	0.46	-0.16
12-Diseases of the skin and subcutaneous tissue							
87	Dermatologic Diseases	912	817	509	12.61	12.70	7.91
13-Diseases of the musculoskeletal system and connective tissue							
88	Rheumatoid Arthritis	2,130	399	843	2.66	1.26	2.71
89	Osteoarthritis	1,774	66	277	17.96	1.22	5.05
90	Back Pain	1,260	601	349	13.32	7.58	4.40
91	Osteoporosis	1,053	236	307	5.11	1.57	2.08
92	Other Rheumatic Diseases	1,149	1,406	1,024	23.14	29.98	21.83

Table A9 (Contd.): Attributed Cost by Medical Conditions using different methods

ID	Multiple-CCS/Calibrated Conditions	Average Cost			Totals Cost (in Billions)		
		Claims	Regression	Propensity Score	Claims	Regression	Propensity Score
14 -Congenital Anomalies & Certain conditions originating in the perinatal period							
118	Congenital Disorders, Newborn conditions	$2,237	$654	1,272	$4.63	$2.38	$4.64
15-Injury and poisoning							
119	Trauma, Fractures, Poisoning and other injury, Motor vehicle accident	2,033	1,520	2,163	25.83	25.46	36.25
96	Hip Fracture	3,293	425	456	1.63	0.66	0.73
16-Other conditions							
99	Signs and Symptoms	829	3,747	2,305	22.13	102.37	62.97
101	Residual, unclassified, all other E codes	1,182	1,051	1,146	19.20	18.76	20.46
17-Cancer Screening							
102	Screening: Breast Cancer	243	61	-350	1.79	0.50	-2.89
103	Screening: Colon Cancer	231	177	-561	0.67	0.52	-1.67
104	Screening: Prostate Cancer	90	157	-945	0.34	0.77	-4.67
105	Screening: Cervical Cancer	103	41	-691	0.21	0.13	-2.23
	Other covariates (including intercept)	N/A	1,566	N/A	N/A	57.68	N/A

Note: Cost attribution in the "claims based approach" is based on health conditions reported in 2009 MCBS. Cost attribution in Regression and Propensity Score Method is based on calibrated health conditions (refer to data section for details on calibrated health conditions). Regression coefficients are negative in all five imputed data set for Acute Renal Failure and Deep Vein Thrombosis (DVT). No spending is attributed to these conditions in regression based approach. N/A = Not applicable.