MEDICARE EXPENDITURE GROWTH AND ITS HEALTH RETURNS ACROSS COHORTS

By YI-SHENG CHAO

A dissertation submitted to the

School of Public Health

University of Medicine and Dentistry of New Jersey and the

Graduate School – New Brunswick

Rutgers, the State University of New Jersey

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

UMDNJ – School of Public Health

Awarded jointly by these intuitions and

Written under the direction of

Professor Alan C. Monheit

And approved by

____

____

____

____

Piscataway/New Brunswick, New Jersey

October, 2012

Abstract of the dissertation

Medicare expenditure growth and its health returns across cohorts

By Yi-Sheng Chao

Dissertation Director:

Alan Monheit

There are several key findings in the following chapters. In Chapter 2, the individual characteristics associated with higher spending growth over the period 1996 to 2008 were identified based on analyses with pooled cross-sectional data from the Medical Expenditure Panel Survey. The key factors that were associated with the adjusted growth rates higher than the actual Medicare spending growth rate (5.8% annually) from 1996 to 2008 include races other than the whites and blacks, Hispanic origin, high income, residence in the West, and very good health status.

Findings from Chapter 3 reveal that enrollment in HMOs under Medicare is not random, but is systematically related to characteristics of Medicare enrollees. Using the longitudinal Health and Retirement Study, there were factors associated with a higher likelihood of becoming enrolled in Medicare Advantage/Part C.

Chapter 4 examines the relationship between health care spending and returns to health with regard to five dimensions of health: mortality, hypertension, arthritis, self-assessed health status and mental health status (Center for Epidemiologic Studies Depression scale, CESD scale). The pre-Medicare characteristics were used to predict the change in these five dimensions of health after four years of Medicare coverage. The results reveal that increases in Medicare total and out-of-pocket spending were associated with poor health outcomes. Total spending was associated with a higher likelihood of death, worse self-rated health status (five categories) and mental health status. The increase in out-of-pocket health expenditure was associated with a higher chance of getting a worse category in self-rated health status after controlling for health status and other characteristics before being enrolled in Medicare. The findings of these studies suggest that policies to constrain Medicare spending should recognize and target the multiple factors contributing to Medicare expenditure growth and the dubious returns to health, as well as target integrated care for Medicare enrollees and incentives for individuals to prevent the onset of chronic health conditions.

Acknowledgements

I would like to thank my advisor, Dr. Alan Monheit, for providing instruction for my PhD study and giving me insightful suggestions to improve my dissertation. Dr. Irina Grafova provided encouraging comments that improve the quality of this dissertation, as Dr. Joel Cohen and Dr. Jeannette Rogowski also provided constructive comments. They are the members of my PhD committee and the most valuable resources to accomplish this dissertation. I also appreciate Rizie Kumar, Margaret Mitchell, and Vanerette Cramer for their help in my school life. Finally, I am very lucky to have the support from Chao-Jung Wu, Jeremie Chao, Fu-Juan Hsieh, Shu-Wen Chao, Guei-Ling Chao, Tom Hsieh, and An-Sheng Chao. Without their support, I would not be able to go through all the challenges I received during this study. This study is partly supported by the Taiwan Government Scholarship for Overseas Study.

Table of Contents

List of tables

List of figures

List of figures

Chapter 1: Introduction to Medicare

The purpose of Medicare's creation in 1965 was to provide health insurance for the elderly and thus to improve their access to health services. The two Parts, A and B, of Medicare were first put into law in 1966, and since that time, the benefit structure of Medicare has been revised and expanded (Kaiser Family Foundation 2010a). Part A was established for hospital care and Part B, Supplemental Health Insurance that covered physician care, was to be purchased voluntarily by enrollees (CMS 2010a). After three decades, Part C, first called Medicare+Choice in 1997 and currently Medicare Advantage, was designed to provide enrollees with a choice of private health plans in place of traditional Medicare (Kaiser Family Foundation 2010b). The adoption of Part D prescription drug coverage in 2003 was a relatively recent change to the structure of Medicare (Kaiser Family Foundation 2010b).

Principally, applicants or their spouses aged 65 years or older become eligible after working for more than 10 years in Medicare-covered employment. However, those aged 65 years or more who did not contribute for 10 years could be enrolled after paying a certain amount of money (Morrisey 2007; CMS 2010b). Additionally, non-elderly persons with end-stage renal disease (ESRD) are also eligible as are those who qualify for Social Security Disability Insurance (SSDI) (CMS 2010b). Overall, Medicare provides nearly universal coverage for the population aged 65 years and over (Birnbaum and Patchias 2008). However, it would be inaccurate to treat Medicare as a purely universal health care coverage as in other countries because it requires payroll tax contributions to be eligible and barriers to Medicare coverage exist[1] (Birnbaum and Patchias 2008).

The problem of Medicare expenditure growth

There are serious fiscal concerns about Medicare. Medicare expenditures have increased several-fold since its creation (Palmer and Saving 2006). The projections show that the trend in overall health care costs, which is in part due to the growth in Medicare spending, would grow to 41% of GDP (Gross Domestic Product) by 2060 if there were no effective and large-scale policy intervention (CBO 2008). Moreover, this problem is becoming a budget issue for everyone because Medicare expenditures alone currently represent 12% of the federal government budget (Kaiser Foundation 2010a) and have contributed to the national deficit (CBO 2008). Medicare reform is a necessary policy intervention to control the national budget deficit based on the Medicare budget growth estimated by Congressional Budget Office (CBO) (Winfree 2011).

Areas of inquiry

Sources of aggregate spending increases

There have been a significant number of studies focusing on the factors contributing to the general health expenditure inflation in the US. Some studies analyzed the trend through a supply-side perspective by analyzing factors such as medical care providers, institutions, and technology (Farrell et al. 2008). Other studies have applied a more general approach by taking both medical care providers' and users' contributions

[1] The barriers include the payment for less than ten years of Medicare tax contribution.

into account (CBO 2008; Zuckerman and McFeeters 2006; Ginsburg 2008; Farrell et al. 2008). However, it is not clear how these factors contributed to the growth in *Medicare* spending.

In the meantime, demand-side factors, contributed by patients and health care enrollees that have affected the increase in Medicare spending over time have not been well studied. Cross-sectional analyses examining the demand side showed that individual decisions in purchasing and maintaining third-party coverage to supplement Medicare contributed to the overall health care spending because of reduced beneficiary cost-sharing (CBO 2008; Kaiser Family Foundation 2010a). The first goal of this thesis is to identify the types of enrollee characteristics that have contributed to the increase in Medicare spending over time.

Health plan selection and Medicare spending

Medicare enrollees have the option to select Medicare Advantage plans in place of traditional Medicare or to purchase Medigap policies to supplement their traditional Medicare coverage. In addition, Medicare beneficiaries may also hold employer-sponsored retirement coverage, Medicaid coverage, and other supplemental policies. Enrollment and spending across the types of coverage might not only reflect differences in beneficiary health status, but also reflect different cost sharing packages by plan type. Both factors could add to Medicare spending and thereby threaten the financial stability of Medicare (CBO 2008).

More specifically, Medicare expenditure growth and related Medicare coverage selection have been studied in the literature. In Biles, Dallek and Nicholas (2004), enrollees in Medicare Advantage plans chose to obtain more generous coverage. Because Medicare Advantage failed to adjust for high-cost enrollees' health status, the sizable premium increase and the declining enrollment in Medicare Advantage by 26% from 1998 to 2003 were observed at the same time. The Medicare Advantage Prescription Drug Plan was also threatened by the rapid growth of total and out-of-pocket spending that subsequently caused the enrollee to select plans with less cost sharing. (Biles, Dallek and Nicholas 2004) In summary, the literature provides some information about the possible sources of coverage selection that caused plan attrition and provides evidence of individual selection to alternative programs under Medicare, but how differences in enrollee health coverage type (especially Medicare Advantage, Medicare+Choice, and others) influenced health spending relative to Medicare has not been fully explored. Hence, the second goal of this research is to evaluate how individuals eligible to enroll in Medicare selected among available plan choices and how significantly the sustained coverage choice affected their use of health care.

Health outcomes and Medicare spending

Lastly, researchers have analyzed the general contribution of health care in the US and found that the mortality decline and the monetary value of the deaths averted were worthwhile compared with the total national health expenditures from 1960s to 2000 (Cutler, Rosen and Vijin 2006). Other researchers also pointed out that the returns from Medicare were worthwhile because the health coverage benefited those of lower education attainment. (Bhattacharya and Lakdawalla 2002; McClellan and Skinner 1999) The relationship between spending and health returns studied by Cutler, Rosen, and Vijin (2006) was that health spending improved health and increased life expectancy.

However, this is only one dimension of health improvement and other aspects of health improvement, such as the decrease of disease incidence and better control of chronic conditions, require further investigation. This research would evaluate the effects of health spending by Medicare enrollees on their health and disease status and could improve the understanding of effects on health. It will do so by adding other specified health conditions (including diabetes, arthritis, activities of daily life, and others) into the model and by quantifying the health impact from health plans. A study with a broader perspective on health outcomes associated with Medicare spending is necessary to evaluate the expenditure impact on individual health status change.

Given the above discussion, this proposal was thus developed based on ***three main questions***: (1) what characteristics of Medicare enrollees drove the increase in Medicare expenditures historically; (2) how has the variation in individual health coverage among Medicare enrollees changed the amount of total and out-of-pocket spending over time; and (3) what are the health returns for Medicare enrollees covered for several years compared to total and out-of-pocket spending incurred.

Chapter 2: Sources of differential growth of Medicare expenditures

To study the sources of health expenditure growth among the Medicare enrollees, data from the individual-level Medical Expenditure Panel Survey (MEPS) is used to estimate the contribution of individual characteristics, such as gender, race/ethnicity, geographic region, health status, health coverage, income, and others to Medicare spending. The MEPS is particularly well suited for this analysis since it contains detailed information on person-level health spending and sources of payment, and on characteristics such as those noted above. Some notable factors have been identified as affecting health spending, including poverty, poor health, disability and chronic conditions (Kaiser Family Foundation 2010). However, the question of how much of Medicare expenditures over time were accounted for by such factors has not been fully investigated. The analysis in this chapter will examine the following null hypothesis:

Total Medicare expenditures over time did not differ according to enrollees' characteristics, including individual health status, demographic characteristics, socioeconomic status, and type of health care coverage.

To illustrate the contribution of each factor, there are descriptive and econometric analyses in this chapter.

Method

Descriptive analysis:

In the first step, a thorough descriptive analysis of the expenditures incurred by Medicare beneficiaries over the period 1996 to 2008 provides a preliminary overview of the expenditure trend in Medicare. The descriptive study in this chapter tabulates health spending in each year for well-defined groups (defined on the basis of characteristics such as age, gender, presence of chronic health conditions, health insurance status, and poverty status). Aggregating total Medicare spending over the study period and aggregating spending for each group of interest, this descriptive study examines how the growth of Medicare spending for specific groups appeared to be contributing disproportionately to aggregate spending growth. See Appendix A for the results of tis descriptive study of Medicare spending.

For this descriptive analysis, MEPS sampling weights and survey design features are applied to tabulate population mean expenditures and their standard errors in different years by types of insurance plans, age groups, race, ethnicity, education attainment, income, health conditions (including heart disease, mental health, and hypertension), and functional status. The absolute amounts of spending are listed for each year. Apart from determining the growth of aggregate spending by specific groups over the study period, this descriptive analysis also shows how Medicare expenditures varied over time based on the above-mentioned characteristics of individuals and how Medicare expenditures grew for enrollees in different types of health plans. However, to better express the relationship between factors and actual amount of spending, the second part of the analysis applies econometric models of health spending to control for other potential confounders that could have affected spending for specific groups and thus, to quantify the growth rates of specific health plan types, health status measures, and demographic factors with other relevant factors held constant.

Econometric analysis

Model specification

In applying an empirical model of health spending, it is important to note that the health expenditure distribution displays considerable skewness and the presence of observations containing zero or little expenditure (Buntin & Zaslavsky 2004). There are two important issues to consider. The first is whether to use a one- or two-part model to estimate individual health spending. A one-part model does not differentiate between the decision to incur health spending and the level of spending incurred, while the two-part model makes this distinction. The second issue is whether the health spending data were to be transformed according to the data's inherent distribution and variance structure (Manning and Mullahy 2001). Because of these two issues, the models considered in this chapter include a one-part OLS (ordinary least square) model; a two-part model (a logit model for the likelihood of service use and an OLS model for the level of spending, given health care use); and a one or two-part GLM (general linear models with log link) with different variance structures (constant variance, proportional-to-mean variance, and variance proportional to squared mean).

By comparing the models applicable to the health spending of Medicare beneficiaries from the Medicare Current Beneficiary Survey (MCBS), Buntin and Zaslavsky (2004) concluded that the one-part GLM (log link) should be tried first based on considerations of efficiency and fit, as the two-part OLS models might have imprecise estimates due to a data transformation that was designed to adjust for heteroscedasticity and its variance structure[2]. If the probability of health care consumption is part of the research question, or if the decision process is based upon the decision to incur spending and then, upon how much to spend, a two-part GLM (log link) model should be used to estimate the probability of incurring health spending and the amount of health spending conditional on incurring health expenditures in separate equations.

Because the data used in this analysis (from the Medical Expenditure Panel Survey) differed from that used by Buntin & Zaslavsky (2004) and because a different set of variables were used in this chapter, these models, including one- and two-part OLS models, and one- and two-part GLMs, are implemented and compared for model fit. With the results from these health-spending models, the model that best fit actual spending is used for the estimation of the historical health-spending trend. See Appendix B to this chapter for specific details regarding model selection.

Functional forms

The two-part expenditure model was discussed first. In the first part of this expenditure model, the probability of consuming health care services (i.e., incurring health spending) is estimated based on the individual characteristics in the dataset. The second part of the model estimates the actual amount or transformed amount of health spending (transformed from a logarithmic to a natural scale), conditional on some spending. The product [$HE(y|x)$] of the predicted probability of health care spending [Pr(y

[2] Transforming data into log scale can adjust for the heteroskedasticity issue efficiently, but it might not necessarily decrease the bias originating from this variance structure.

> 0)] (where y is health spending) and the predicted amount of spending conditional on incurring expenditures [$E(y|y>0)$] give us the total predicted health spending for each individual in the dataset.

$$HE(y|x) = \Pr(y > 0) \times E(y|y > 0) \tag{2.1}$$

In a logit function, the first part of the model to predict the probability of incurring health spending is expressed in a functional form as follows:

$$\Pr(y_{it} > 0) = \frac{e^{\beta_0+\beta_1 X_{it}+\beta_2 T+\beta_3 (TX_i)+\varepsilon}}{1+e^{\beta_0+\beta_1 X_{it}+\beta_2 T+\beta_3 (TX_i)+\varepsilon}} \tag{2.2}$$

The probability of incurring health spending or not, [Pr $(y_{it} > 0)$], is modeled as a dichotomous outcome taking the value of 0 for not incurring health spending and 1 for incurring health care spending, and was estimated based on the characteristics of the enrollees and other covariates. X_{it} denotes a vector of the individual characteristics for person (i) observed in each year (t). The individual characteristics that were essential to health spending modeling and available from MEPS datasets include age, sex, race[3] and ethnicity, regions of residence, income, years of education, self-assessed health status, mental health status, regions of residence, functional status (ADL, IADL, difficulties in mobility and others), and insurance coverage. In a reduced model, only basic characteristics (age, sex, race and ethnicity, income, years of education, and self-assessed health status) are used to serve as a comparable model for the full model. The other extended model adds chronic conditions to be compared with the full model, although the observation time is limited, from 2000 to 2008.

[3] Race and ethnicity were re-categorized in the MEPS over time. See Appendix B for details.

Time fixed effects for each year in the MEPS datasets from 1996 (from 2000 in the extended model) to 2008 are represented by the variable, T. The interaction term, $[X_i T]$, indicates which individual factors significantly influences the probability of incurring health spending based on different years.

The second part of two-part expenditure model involves estimating health care spending conditional on incurring health expenditures. Here, I draw upon the modeling approaches discussed in Buntin and Zaslavsky (2004). In their analysis, Buntin and Zaslavsky found the logarithm transformation appropriate for the expenditure data, but other considerations were necessary to adjust for potential heteroskedasticity associated with this transformation. Similar to Buntin and Zaslavsky (2004), this step tests multiple expenditure models mentioned above by examining the model fit of log-transformed and non-transformed models with mean square error (MSE), and mean absolute prediction error (MAPE).

Because the coefficients across regression models (the one- or two-part models noted earlier) differ from each other, the test of model fit is used to select the "best" regression results to adopt for interpretation. Since the dataset tested by Buntin and Zaslavsky (2004), from the Medicare Current Beneficiary Survey, differs from the MEPS datasets used in this chapter, this chapter tests which regression model fit the data best. Moreover, Manning and Mullahy (2001) evaluated spending data with a large share of zero consumption and positively skewed distribution of health spending. In their study, the choice of estimation model and the effects of data retransformation were evaluated based on the estimator precision and bias by fitting the medical visit statistics in the 1992 National Health Interview Survey (NHIS). In their study, the best variance function was first evaluated with Park test (Park 1966) and then they adjusted the variance function after assessing the value of λ_1 in the regression equation, $[\ln[(y_i - \hat{y}_i)^2] = \lambda_0 + \lambda_i \ln(\hat{y}_i) + v_i]$. The main purpose of this equation is to understand the relationship between the variance function and the mean. This relationship of the variance to the mean could be constant (λ_1=0, homoscedastic distribution), proportional to the mean (λ_1=1, Poisson variance function), or proportional to the mean squared (λ_1=2, gamma variance function) (Buntin and Zaslavsky 2004).

Based on the above discussion, this part of this analysis estimates conditional expenditure models. A general model specification could be expressed as follows:

$$y_{it} = \beta_0 + \beta_1 X_i + \beta_2 T + \beta_3 (X_i T) + \varepsilon \qquad (2.3)$$

The values of individual health expenditures among those with any consumption (y_{it}) are estimated based on the vector of individual characteristics [X_i] and time fixed effects (T) to capture any time-specific trends in Medicare spending. Most importantly, the interaction term, $[X_i T]$, will indicate how much individual factors influenced the amount of health spending in different years.

Individual characteristics are defined differently in different studies. In Mullahy (1998), the individual characteristics include age, gender, education, race, marital status and health status. In other studies, activities of daily living (ADL), instrumental activities of daily living (IADL), chronic conditions (stroke, heart disease, diabetes and others) (Buntin and Zaslavsky 2004), poverty status (Cook, McGuire, Meara and Zaslavsky 2009), and residential characteristics (Hill and Miller 2010) are used. MEPS datasets include all of these variables as well as other person-level characteristics that have been used in studies on health spending.

Finally, predicted health care spending for each individual in the sample can be obtained. In the two-part model, this entails estimating for *each observation in the sample,* the predicted probability of using health care and the predicted amount of health spending (from the conditional expenditure model) and taking their product. Using a one-part model, predicted health spending for each observation in the sample can be obtained directly from predictions based on this model. Predictions regarding health care spending are assessed by these two types of expenditure models. However, for purposes of this analysis, we use a slightly different approach described below.

After estimating individual health spending over time by the regression models, the remaining question is how to identify the differential effects of individual characteristics on health spending in each year. This highlights the importance of the regression coefficients of the interaction terms [X_iT]. In a summary table of regression coefficients, the interaction terms are evaluated for their significance in each year to understand the differential effect of specific individual characteristics over time on the amount of health spending, controlling for other factors.

Because of the model specification and the use of interactions, there is a group of people serving as reference group: non-Hispanic white married males age 65 or older living in Northeast region without any income and education in excellent health and mental health status in 1996.

Steps to quantify the adjusted health spending growth in different population groups

The first step to examine the contribution of a particular subgroup to health spending is to use the estimated regression model to predict the contribution of the specific subgroup to health spending in different years. The individual coefficients from the regression model provide estimates of the "average" contribution of each characteristic over the study period.[4] Using regression coefficients from the interaction terms (between characteristics and years) provides estimates of the contribution of specific individual characteristics in each year of the study period, relative to the coefficients in 1996. As described below, base year estimates are based on the intercept term and the coefficients of the explanatory variables and subsequent year estimates are based on these factors, plus the time (year dummy) variable, plus the interaction of time and the variable of interest. The mean values of continuous variables (age, income in dollars, and years of education) and the average proportions of categorical variables (gender, regions of residence and others) are also obtained for each population characteristic subgroups from 1996 to 2008 for use in the estimation of subgroup spending. In the extended model, the prevalence of chronic conditions is documented from 2000 to 2008 and used for spending estimation.

---

[4] For Ordinary Least Square (OLS) models, the regression coefficients could be taken as the average contribution to total annual health spending (the dependent variable in the model). For the Generalized Linear Models (GLM with log link), the average contribution could be obtained from the marginal-effect estimations by taking derivatives from the predicted equations.

Aggregate population health spending and rates of growth based on specific subgroup characteristics

The estimated aggregate health spending for subgroups (individuals with a specific characteristic) is obtained as follows.[5] First, the subset of individuals with the subgroup characteristic is selected. Next, using the estimated pooled regression model, the regression coefficients of all associated characteristics and the constant are combined to predict the average Medicare spending for individuals of this characteristic subset. This is obtained by using the regression coefficient for the subgroup characteristic of interest (e.g., the coefficient on "black" for the subset of blacks) and the mean values (or proportions) of specific subgroup characteristics (e.g., mean age for the black subset). Then the average prediction is multiplied by the size of the population with this characteristic (e.g., the number of blacks in the specific year for which the estimate is being produced). This generates aggregate population spending estimates in 1996 (2000 for the extended model with chronic conditions). Estimates for subsequent years (i.e., 2008) are obtained by including the contribution of the year dummy variables and the interaction between subgroup characteristic and subsequent year dummy variable. Aggregate spending is obtained using the 2008 subgroup population size and mean values for the individual characteristics. The baseline and 2008 spending estimates are used to

---

[5] The estimation of aggregate predicted Medicare spending is based on the equation: $\sum_{i=1}^{n} \hat{y}$ or $\sum_{i=1}^{n}(\beta_0 + \beta_i x_i + \varepsilon)$, estimated from the pooled cross-section regression. Here, n denotes the number of observations in a subgroup with a specific grouping characteristic. We aim to predict $\hat{y}$ (the estimated spending for observations within this group, $\beta_0 + \beta_i x_i + \epsilon$) in the prediction formula for each characteristic in different years (for 1996, 2000, and 2008) using estimates from the pooled regression. The mean values of other associated subgroup characteristic (age, income and years of education) are also imported into this equation to produce an average spending estimation for this subgroup. The other associated independent variables are categorical and the coefficients of these variables are multiplied by proportions of subgroup observations with these specific characteristics. For example, the average amount incurred by blacks in the high-income group was calculated by multiplying the proportion of blacks in this group by their coefficient on individual spending. The contribution of the blacks in high-income group is then added to the average spending estimation.

obtain annual growth rates[6] from 1996 (2000 for the extended model) to 2008 and the summary measures for the growth in spending. The important concept in this step is that the other subgroup characteristics that might be correlated with the specific characteristic of interest are taken into consideration and their effects on spending are included in this estimation.

Data

For the analyses of this chapter, I use data from the Medical Expenditure Panel Survey (MEPS) designed and conducted by the Agency for Healthcare Research and Quality (AHRQ). MEPS is a nationally representative dataset and the respondents consist of a nationally representative sample of the civilian non-institutionalized population in the US (AHRQ 2011a). The MEPS is based on a complex stratified survey design with appropriate weights assigned to individuals in the sample in order to produce national estimates of outcomes of interest. When using MEPS data, standard errors of the estimates in descriptive and regression analyses must be adjusted for the complex and clustered sampling design. MEPS household component (HC) datasets contain a rich set of information on individuals and families as well as their health plans that directly and indirectly influence decisions about health consumption (AHRQ 2011b).

Since 1996, AHRQ has created and disseminated household component datasets each year. The latest HC dataset at the time of this study is for 2008. (AHRQ 2011c) To study the historical trend of health care spending conditional on individual characteristics

---

[6] The annual growth rate was calculated by the equation: $\ln\left(\frac{\hat{y}_{2008}}{\hat{y}_{1996}}\right)/D$. The notations, $\hat{y}_{1996}$, $\hat{y}_{2008}$ and D, represented the aggregated predicted spending in 1996, 2008 and the number of years in this period. In the extended model, year 2000 was the base year and the percentage change of health spending was divided by 8 years.

in this chapter, these annual datasets were combined to make estimates of spending across years. AHRQ also provides a file of longitudinally consistent strata and PSUs for use in examining trends over time.

Results

The first results are based on the descriptive analysis and focused on the aggregate Medicare health spending change from 1996 to 2008. In the second part, after reviewing these descriptive findings and their implications, the econometric analysis was applied to assess the role of key population attributes holding other factors constant and thus to assess whether the findings from the descriptive analysis remained valid.

Descriptive analysis

Aggregate Medicare health spending change from 1996 to 2008

This section investigated the aggregate health spending for different groups of individuals based on the personal characteristics of Medicare beneficiaries that did not change over time, including race/ethnicity, gender, birth year, and educational attainment. In absolute terms, health expenditures by different individual characteristics were listed in Table 2.1. Aggregated Medicare spending (for those age 65 and over) grew from $183 billion in 1996 to $366 billion in 2008 and out-of-pocket health spending grew from $27.8 billion in 1996 to $53.2 billion in 2008.

For specific groups, there were differential rates of growths in health care spending between 1996 and 2008, several of which exceeded the average annual growth rate in aggregate Medicare spending (5.8%). The subgroups include females, persons of ages 65-74 years, black and other races, Hispanics, persons with more than eight years of education, those with near poor, middle or high incomes, residence in the West, other marital status, very good health status, and very good mental health status.

For chronic conditions reported after 2000, certain conditions were associated with percentage changes in spending that were higher than that of total aggregate Medicare spending in the same period (180% from 2000 to 2008, 7.4% growth annually). These conditions included diabetes, stroke, emphysema, hypertension, coronary heart disease, other heart disease, and joint pain (Table 2.1).

Medicare enrollment and characteristics of Medicare enrollees from 1996 to 2008

Of all Medicare enrollees age 65 and over, the population size increased from 34 million to 39 million and the percentage that had any health spending remained stable, 95.9% in 1996 and 97.0% in 2008.

The rate of increase was higher after 2000, from 35 million in 2000 to 39 million in 2008. Females were the dominant population (more than 56% from 1996 to 2008) (Table 2.2). Whites were the racial majority (more than 80% from 1996 to 2008) but their share of the Medicare population was declining. There was an upward trend for age, education, and income, as the percentages of top tiers of these characteristics were growing over time. The geographic distribution of the Medicare enrollees indicated that the largest percentage was living in the South (more than 33%) and the least in the Northeast (in 2008) or West (in 1996 and 2000). The distribution of health status changed over time with those in excellent health declining as a proportion of enrollees (18.5% to 14.1%). The number of Medicare enrollees with very good and good health status was increasing, while the proportions of those with excellent or fair or poor health status decreased in this population. The distribution in self-reported mental health occurred with a decreasing proportion of enrollees reporting excellent mental health status from 1996 to 2008. Similar to self-assessed health status, the proportions of persons in of excellent or fair/poor mental health status were declining and the proportions in very good and good mental health were decreasing from 1996 to 2008.

The percentages of Medicare enrollees having activity limitations declined by three percentage points (from 23.2% to 20.0%), while those with any limitation increased in prevalence from 55.2% in 1996 to 57.8% in 2008. However, the proportions did not decline steadily. After 2000, there were small percentage-point increases for limitations in ADL, IADL, activity and cognition. For insurance coverage, the percentages of dual eligibilities and private coverage for Medicare enrollees declined from 11.7% to 9.6% and from 67.1% to 47.3% (1996 to 2008) respectively.

Except for asthma whose prevalence decreased by less than half a percentage point, the chronic conditions documented after 2000 in MEPS became more prevalent. The most prevalent conditions were joint pain[7] (51.9% in 2000 and 53.4% in 2008) and hypertension (49.9% in 2000 and 67.2% in 2008). Even for the less prevalent conditions, such as emphysema and stroke, the prevalence increased at a faster rate than the growth of Medicare population.

[7] Joint pain was defined as the appearance of the symptoms including pain, swelling or stiffness around a joint in the last 12 months for those ages 18 years and over (MEPS 2011). This was not equivalent to the diagnosis of arthritis. Another reason to use joint pain instead of the diagnosis of arthritis was because the diagnosis of arthritis was introduced to MEPS questionnaires in 2001, one year after joint pain was asked among MEPS participants.

Econometric analysis - Health Spending Modeling

Sources of historical health spending growth among Medicare enrollees

Absolute amount of Medicare spending change from 1996 to 2008

In Table 2.3 (second row), the actual total amount of Medicare spending incurred by those age 65 years and over was $183 billion in 1996 and $366 billon (nominal dollars). Medicare spending for these enrollees doubled from 1996 to 2008, yielding an average growth rate of 5.8% per year over this period (See the descriptive analysis).

Comparison between the actual and adjusted spending in different Medicare groups from 1996 to 2008

The advantage of the econometric analysis is to adjust for the other factors that influenced the levels of spending in each group in the descriptive analysis. The aggregate spending in each population subgroup among the Medicare population age 65 years and over was calculated by multiplying the prevalence of this characteristic and the mean spending in this grouping characteristic predicted by one-part Poisson GLM[8]. In Table 2.3, the aggregate Medicare spending levels[9] in 1996 and 2008 are listed to capture the aggregate spending growth over this period. In addition, the table displays the rates of growth in spending for different Medicare subgroups.

---

[8] The one-part spending model based on Poisson GLM (log link) estimates was chosen for its superiority in fitting the predicted spending with the actual amounts, compared with other one- or two-part OLS models and GLM. See Appendix B for the details in model selection.

[9] The aggregate spending in one grouping characteristic among Medicare population age 65 years and over was calculated by multiplying the prevalence of this characteristic and the mean spending in this grouping characteristic predicted by one-part Poisson GLM.

Annual growth rates of Medicare spending due to different characteristics from 1996 to 2008

There were several important findings based on the computed growth rates. First, there were groups with estimated spending growth rates higher than the actual or adjusted amounts in Table 2.3. Of all subgroups, the other races (8.4%), high-income individuals (6.6%), and very good health status (6.9%) had growth rates higher than the overall growth rates (5.8%) from 1996 to 2008 after taking the average levels of other characteristics into account.

Comparison between regression coefficients of individual characteristics and aggregate subgroup growth rates

After the other characteristics were adjusted, there were characteristics associated with higher amounts of individual health spending increase from 1996 to 2008, including females, races other than the whites, Hispanic origin, high income, regions (South and West), health status, and mental health status.[10] In these subgroups, the aggregate Medicare spending increase (Table 2.2) could be partly attributed to the individual spending increase in this period.

However, there were groups with higher growth rates but the grouping characteristics were associated with lower individual Medicare spending from 1996 to 2008 (including those related to age or higher education attainment, and those married). The growth in these groups should be attributed to the correlated increase in the growth-

[10] Their coefficients in 2008 were higher than those in 1996, leading to positive individual spending growth in this period. See Appendix B and Table B.7 for details.

related characteristics[11]. For example, income was associated with higher individual spending in 2008 than in 1996 and the average income in many groups increased. The changes in the associated characteristics (such as income growth) lead to high aggregate spending growth from 1996 to 2008, overwhelmed the effects of the spending-diminishing grouping characteristics and induced growth rates higher than that of the total actual Medicare spending.

Second, groups for which had negative computed growth rates of adjusted spending had decreasing shares of population, especially for those with lower education attainment (zero to eight years of education), requiring IADL assistance, activity limitation, cognitive limitation, Medicaid coverage and private coverage. The decrease in the prevalence of these characteristics was associated with the negative growth rates of aggregate adjusted spending. They were also associated with negative individual spending growth in 2008 that provided another source for the negative aggregate spending growth for these groups.

However, the negative coefficient of being divorced in 2008 was larger than its positive effect on Medicare spending in 1996.[12] This contributed to the negative growth in the adjusted spending among those divorced even with an increase in population size (Table 2.2). The contribution of the associated changes in the other individual characteristics to these negative spending growths was not clear.

[11] Growth-related characteristics had higher coefficients in 2008 than those in 1996. See Table B.7 for details.

[12] Its coefficient in 2008 contributed to a negative coefficient whose absolute value was larger than the value of the positive coefficient in 1996. See Table B.3 for details.

Finally, there were other grouping characteristics were also important for Medicare spending growth. For example, whites had a growth rate that was 0.3 percentage points less than overall adjusted spending. Because of the large share of contribution from the whites, the absolute increase of Medicare spending in other racial groups was much less than that for whites in this period. Although growth-leading characteristics could be identified easily, the lesson was that these factors might not have the largest impact on spending growth over time.

To conclude, the illustration in Figure 2.1 helps to show the two major factors determining aggregate spending in different groups, population sizes and the spending levels in groups. As noted, population sizes for groups of Medicare enrollees changed differentially across characteristics. The spending levels in different groups were the sum of the coefficients of the specific grouping characteristics and the coefficients of the other associated factors (independent variables) for each of the characteristic subgroups.

Medicare spending change with the adjustment in chronic conditions from 2000 to 2008 (the extended model)

In Table 2.4 (first row), the actual total amount of Medicare spending incurred by those age 65 years and over was \$203 billion in 2000 and \$366 billon (nominal dollars). The annual growth rate was 7.4% in this period.

Growth rates of the Medicare spending in different groups from 2000 to 2008 (the extended model with chronic conditions)

The extended model that used year 2000 as reference with control for chronic conditions did not produce the same estimates as the full model in Table 2.3. In Appendix

B, the detailed results in the full and extended models were listed. The sample sizes in these two models were not the same. In addition to including more independent variables in the extended model, the number of original and weighted eligible observations in the extended model was less than for the full model.[13]

The annual population growth rate of the Medicare enrollees aged 65 years and over (1.4%) was smaller than the rate of growth in overall Medicare spending (7.4%) from 1996 to 2008 (Table 2.4). The result suggested several implications. First, there were groups with negative growth rates, including those with less education attainment, private insurance, the diagnosis of emphysema, and current smoking. The reduction in Medicare spending could be partly attributed to decreasing shares of population, except for those diagnosed with emphysema and other heart diseases that had higher shares of population and negative interaction effects with year 2008.

Second, there were groups with growth rates higher than that of the actual spending (7.4%), including other races (28.0%), Medicare enrollees diagnosed with coronary heart disease (8.4%) and other heart diseases (7.5%). Other races and patients with coronary heart disease were associated with increases in regression coefficients of individual spending from 2000 to 2008. However, individuals with other heart diseases who had negative change in regression coefficients from 2000 to 2008[14] had a higher growth rate than overall growth, because their share of population almost doubled in this period (14.8% and 28.0% in 2000 and 2008 respectively).

---

[13] See Table B.7 and B.8 for the comparison.

[14] The patients with other heart diseases were associated with negative growth of mean adjusted individual spending, but the population spending grew. See Table B.9 for details in the negative coefficients in 2008.

Finally, with respect to income, the high-income group (5.1%) had the highest growth rate, as the near-poor and negative/poor groups had the second and third highest growth rates of 3.6% and 2.4% respectively. The other groups, low and middle income, had the lowest growth rates, 0.7% and 0.1% respectively. Because those in the middle of the income distribution had the lowest spending growth rates over time, it is very likely that there were many other factors influencing spending for those at both extremes of the income distribution.

Discussion

Limitations

Sources of estimation imprecision

Merged datasets

There were threats to the validity and precision of estimation. First, the analysis dataset was created by merging multiple datasets that differed in terms of variable definitions. The definition of race/ethnicity, insurance categories, and health conditions changed over time and there was a major overhaul in the questionnaire of 2000. The revised questionnaire could use self-administered questionnaires to obtain information on chronic conditions and other factors. However, the inclusion of these new variables after 2000 would exclude the time frame from 1996 to 1999. This particular tradeoff between time span and individual-level details was the first difficult choice to make.

Single model for long observation periods

The second challenge was to model health expenditure over years, from 1996 to 2008 with a single equation. The mean square error (MSE) and mean absolute prediction error (MAPE) were much larger than those observed in Buntin and Zaslavsky (2004), which analyzed single-year data from Medicare Current Beneficiary Survey (MCBS) in 1996. This problem might be exacerbated because inflation was not accounted for in this study[15] but was in part captured by year-specific main effects (which also capture factors such as technology and policy interventions which vary over time).

Sample size and the minorities

The third limit was sample sizes, which necessitated that subgroups with too few observations had to be merged. For example, four racial categories, American Indians, Asians, Pacific Islanders and multiple races reported (available only after 2002), were put in the same category, other races. Two of the worst categories of mental health, fair and poor, were merged. Although merging in some variables might help to increase the sample size of specific groups, the danger was to combine different groups together. The standard error might not decrease with a larger sample size but in fact increase due to more diverse values in this newly merged subgroup. Even though MEPS contained good precision of expenditure information and a rich set of variables over time; it was still limited when we focused on the minorities.

For the extended model, the threat of small sample sizes was more significant because of a shorter observation time and additional sources of non-response for the

[15] The focus of this study was on actual Medicare spending unadjusted for inflation to capture the current year value of resources used to support health care for Medicare enrollees.

questions on chronic conditions in the self-administrated questionnaires after 2000. The estimation in the extended model shared higher rates of spending growth (other races) than those in the full model. Caution should be taken to compare the results in the extended model with those in the full model.

Unknown precision of estimates in aggregate spending

The last limit was that the precision (variance and standard errors) of projected aggregate total Medicare spending was unknown. As many of the main effects and interactions terms were not significant in the individual levels. A major concern was that the projected spending might have a wide confidence interval that included the null value and produced unreliable population projection. Although some researchers did not consider the statistical insignificance in some estimates as an issue to construct long-term health expenditure simulations (Alemayehu and Warner 2004), a possible solution would be to conduct a research that observed the total health spending in different communities and compare the total health spending over time. As the population compositions were controlled, this type of community-level research had a larger potential to increase the precision of population expenditure estimates.

Poliy implications

This section summarizes findings and draws implications from the study's results. A major contribution of the study in Chapter 2 was that groups of high spending or characteristics of high spenders were identified among Medicare enrollees. Only after these sources of high spending were identified, we could make concrete policy implications and consider approaches to control health spending among Medicare enrollees. Some individual characteristics, such as gender and race/ethnicity, were found to be associated with high spending due to interaction with other factors or being proxy of other unobservable high spending factors. Such characteristics might provide useful information that could be used to more efficiently target high cost enrollees.

Dimensionality of cost-containment policy

The results in Chapter 2 showed that there were different dimensions at work on the problem of spending growth in Medicare. The first dimension was to prevent health events as a way to reduce the growth in spending. Many of the characteristics in the high-spending Medicare enrollees were related to a worse health status or physical limitation.[16] Moreover, the extended model showed that heart (angina and heart attack) and chronic (diabetes) conditions were related to higher degree of health cost inflation.[17] The effects of these conditions also persisted and significantly influenced the spending growth from 2000 to 2008 (Table 2.4). In Stampfer, Hu, et al. (2000), the value of cardiovascular disease prevention was emphasized and its potential to curb health care spending was highlighted in their study. On the other hand, treating patients with better-integrated health care could be another answer to these high-spending factors and conditions, as suggested in Fisher (2008). The consideration over prevention of health conditions or efficient care for these chronic patients would be the first dimension for the policymakers.

Second, another dimension was the contrast between the interventions in population or in personal levels. This approach proposed by Geoffrey Rose indicated that

[16] See Table B.7 for these high-spending characteristics.

[17] See the regression coefficients in Table B.9.

health policy could focus on changing the distribution of certain characteristics or influencing a specific group within the population (Rose 1985; Rose 2001; Doyle, Furey et al. 2006). The regression coefficients showed the effects of individual characteristics on the magnitude and direction of individual spending from 1996 to 2008. However, these individual growth-related characteristics were not necessarily the population groups that contributed the most to overall Medicare spending. The leading causes of health spending growth for individuals were health status other than "excellent", female, residence in the West, "fair" or "poor" mental health status, and needing any help in ADL[18]. However, the leading causes of growth for aggregate Medicare spending included other races, Hispanic origin, high income, residence in the West, and "very good" health status.[19] Only residence in the West was the common factor for both individual and population Medicare spending growth.

A population approach to contain Medicare spending will use measures to influence the Medicare spending universally, as an individual approach will simply target high spending groups or patients. This contrast provided policy options for the public, to implement new health programs whether for populations or for personal characteristics. If particular population were targeted, some general measures, such as cost sharing ratios (especially for the services with higher price elasticity) and regulations on visits to specialties (for example, adding gatekeepers to reduce the visits to medical specialists),

---

[18] The leading causes were the categorical individual characteristics that had the highest change in the coefficients of individual characteristics from 1996 to 2008. The coefficients were derived from the regression coefficients in one-part Poisson GLM (log link) in Table B.10.

[19] These Medicare population subgroups were estimated to have adjusted spending growth rates higher than the annual actual Medicare spending growth (5.8%) from 1996 to 2008 in Table 2.3.

could be used to modify the financial incentives, such as cost sharing or copayments, and discourage high spending growth for a population.

If specific individual characteristics or health conditions were targeted, the way health care was delivered should be changed to select these individuals and deliver integrated health services for them. The recent trend in "medical home" type health care could be an example for this (Fisher 2008).

The third dimension was to choose or implement policies aiming at one-time spending saving or spending growth. In the one-part spending model and the two-part spending model[20], the regression coefficients of individual characteristics in 1996 differed and were not the same as the interaction terms between characteristics and year dummies. This meant that those characteristics associated with higher spending in 1996 might not lead to a larger share of Medicare spending growth from 1996 to 2008.

Two directions for health policy planning were found: reducing initial treatment cost and intensity for specific treatments (i.e., the spending in the base year, 1996, in this Chapter), or controlling the growth rate of specific treatment (i.e., the interaction terms between characteristics and year 2008). For example, one health care plan could provide the female enrollees less costly care, but this change in health plan might not influence future health spending growth. The scales of Medicare costs and growth might be tamed at the same time, but to achieve this might require major reforms in the payment to the providers and long observation periods mentioned in White (2008).

[20] See Table B.7 for the regression coefficients in the one-part GLM model.

Chapter 3: HMO coverage selection in Medicare and accumulated total and out-of-pocket health spending

Introduction

The second question noted in the Introduction was how the variation in health coverage of Medicare enrollees affects total and out-of-pocket medical spending and the extra spending induced by the privately purchased supplemental coverage, especially for health maintenance organizations (HMOs) mentioned in Luft (1981), Miller and Luft (1994) and Greenwald, Levy and Ingber (2000). Based on the literature on health insurance, it has been well established that the level of cost sharing influences total health care spending and out-of-pocket payments (Morrisey 2007; Newhouse and the Insurance Experiment Group 1993). What has been a concern is that the extra coverage among Medicare enrollees beyond basic Medicare (especially that found in Medicare Advantage or Medicare+Choice program) might lead to moral hazard and yield an increase in overall Medicare expenditure (CBO 2008). The challenge in assessing the contribution of differences in coverage status *per se* has been to assess and control for the possibility of biased selection (either adverse or favorable selection) into Medicare Advantage/Part C, compared to traditional Medicare plans. Failure to control for adverse selection, i.e., the endogeneity of different types of Medicare enrollment, can yield an upward bias in estimates of the contribution of different coverage types to Medicare spending and a upward bias in out-of-pocket spending that is associated with such total spending compared to what might be observed for an average health risk. This will occur when persons with unobserved adverse health status select into Medicare Advantage/Part C compared to those in better health who obtain traditional Medicare and no supplemental coverage. Alternatively, if there is favorable selection into HMO coverage, estimates of the impact of such health plans on total spending may be downward biased as will estimates of out-of-pocket spending compared to what might be expected on average.

On this basis, the types of coverage (traditional Medicare or Medicare Advantage/Part C) held by Medicare enrollees may be determined by their underlying health status. In the classic literature discussing insurance, the two-party Rothschild-Stiglitz model (Rothschild and Stiglitz 1976) illustrated the adverse selection problem. High-risk individuals might prefer more generous plans and low-risks might prefer less generous plans, so that high-risk individuals may have the ability to incur greater health care spending. The separating equilibrium might stand if the low-risk individuals could find the more restrictive plans preferable (Monheit, Cantor, Koller and Fox 2004; Lo Sasso and Buchmueller 2004). In theory, Medicare enrollees will have different spending patterns based on their level of coverage, especially the cost sharing of Medicare Advantage/Part C relative to that of traditional Medicare in Luft 1981; Miller and Luft 1994; Greenwald, Levy and Ingber 2000. These researches found that biased selection – specifically, favorable selection by health plans – existed in Medicare Advantage/Part C enrollment. The favorable selection was caused by the enrollment of relatively healthy individuals in Medicare managed care plans.

The use of MEPS datasets that provided cross-sectional observation of the non-institutionalized population in the US limits the ability of researchers to observe the health status before and after plan selection. Although MEPS datasets could be linked to form two-year panels, the length of time is not sufficient to observe the health status and other characteristics before and after enrolling in Medicare. Controlling for the health status and individual characteristics before plan enrollment can help to adjust for factors that were not observed in cross-sectional datasets and contributed to individual decisions regarding their selection of health plans. Therefore, MEPS datasets could not be used to investigate the endogeneity between health plans and the health status before plan selection.

To examine whether enrollment in Medicare Advantage/Part C relative to traditional Medicare programs reflects biased selection, this chapter takes advantage of the longitudinal study design in the Health and Retirement Study (HRS). Because the health status before Medicare that might be related to the plan selection was recorded in HRS, the HRS permitted assessment of whether health-related characteristics are associated with enrollment in such plans. In this study, the HRS enrollees were selected only if they recorded individual characteristics and provided their health status evaluation for the two-year period right before being covered by Medicare. To investigate the existence and the magnitude of biased selection to HMO coverage under Medicare, the probability of enrollment in each type of coverage (Medicare Advantage/Part C or traditional Medicare) was estimated based on the controls for individual characteristics, such as prior and current health status, current income, education, age and race/ethnicity.

Therefore the null hypothesis for this Chapter was as follows:

The probability that individuals will choose a health maintenance organization (HMO) to replace traditional Medicare was not systematically influenced by individual characteristics and pre-Medicare health status. Compared to individuals with traditional

Medicare, having HMO coverage (Medicare Advantage/Part C) did not affect out-of-pocket and total Medicare spending.

Method

Model specification

To test the hypothesis, the analysis must address the endogeneity problem that arises because of unobserved individual health factors that can influence both coverage type and the amount of health spending (Wilcox-Gok and Rubin 1990; Cameron, Trivedi and Milne 1988). To solve the endogeneity problem, there were two approaches taken in the literature: instrument-variable regression (for example, in Desmond, Rice, and Fox 2006; Cabral and Mahoney, 2010) and propensity score matching (for example, in White and Seagrave 2005; Stuart, Doshi, Briesacher, Wrobel, and Baysac 2004; Van Houtven, Jeffreys and Coffman 2008; Ellis et al. 2003). Some articles have applied both approaches in specific populations (Stukel, Fisher, Wennberg, Alter, Gottlieb and Vermeulen 2007). However, meeting the requirements for the use of instrumental variables can be challenging. In particular, the selected instrument can only be correlated with the endogenous variable of interest and not with the second-stage outcome of interest or with the equation's error term. Moreover, the instrument must exhibit adequate statistical power in the first-stage equation (Staiger and Stock 1997). Finding an instrument that meets these conditions can be difficult, especially since variables that affect health plan choice were also likely to affect spending.

The present chapter uses two estimation methods: Propensity score matching and a GLM regression model. The propensity score method works by giving an individual observation a summarized value based on observable factors (Rosenbaum and Rubin 1983; D'Agostino 1998). Individuals with identical propensity scores enrolling in different health insurance plans (Medicare Advantage/Part C and traditional Medicare in this chapter) were similar in their propensity to choose one of the health plans and taken as comparable in order to distinguish the effects of health plan implementation. This method adjusted for endogeneity in health plan purchasing because the propensity score aimed to balance the likelihood of receiving health plans that was determined by other factors. These factors, especially health status, induced biased selection by drawing individuals of relatively better health status into less generous, low-cost plans. This selection was influenced by pre-Medicare health and affected the health spending of individuals (the outcome of this Chapter). Once the control group (defined as having Medicare plans only) was matched with the treated individuals (selecting alternative HMO coverage, Medicare Advantage/Part C) of the same or similar propensity score, the effects of treatment (having Medicare Advantage/Part C) could be estimated by calculating the differences between similar individuals from different health plans.

Estimating the probability and propensity score to enroll in HMOs under Medicare

To begin the propensity score matching, the first step was estimating the likelihood (propensity score) of individuals (Caliendo and Kopeinig 2008) to participate in Medicare Advantage/Part C, compared to the chance of joining traditional Medicare. Because the selection into Medicare Advantage/Part C, rather than traditional Medicare, was the interest in this chapter, the likelihood was estimated with binominal logistic regression after controlling for exogenous individual characteristics, especially pre-Medicare health status that was theoretically a motivation for adverse selection. As such, these estimating equations should be considered reduced-form rather than structural equations. The probability of Medicare beneficiaries to be enrolled in HMOs was estimated compared to the chance of being covered by traditional Medicare. Two commonly used models to predict this probability, probit and logit, and the predictions in both models tend to be similar (Wooldridge 2002). Following the methodology in Caliendo and Kopeinig (2008), logit regression was used in order to obtain the results that could be used not only to make statistical inferences, but also to get the predicted probability of enrolling in each type of coverage. The statistical inferences from the regression helped to answer the first part of the hypothesis and suggested whether any biased selection existed based on pre-Medicare health status.

Matching methods to estimate the average treatment effect on the treated

In the second step, the predicted probability of enrolling in Medicare Advantage/Part C was used as the propensity score (between zero and one) to match individuals or to group individuals of similar scoring and to estimate the average effects of HMO plans (by comparing the mean values in the treated and control groups) on the amount of total health spending for those covered by Medicare from age 65 to 68 years (three to four years of Medicare coverage). The issue was what type of matching methods to use for the HRS enrollees and choosing the matching algorithms depends on the trade-off between efficiency and bias in Table 3.1 (Caliendo and Kopeinig 2008).

Trade-offs between matching algorithms

In Table 3.1, different matching algorithms estimated the average treatment effect on the treated (ATT) using different control and treated observations. The spending differences between the treated and control groups could be taken as an estimation of

ATT because the selection bias was theoretically removed and these two groups became comparable after matching. In the first matching method listed in Table 3.1, nearest neighbor (NN) matching, the controls were compared with treated observations with closest propensity score (PS). The possible risk with this method is poor matches. (Heinrich, Maffioli and Vázquez 2010) The second matching method, radius matching, selected those with similar propensity scores in the same caliper and the treated and controls were matched within calipers. (Heinrich, Maffioli and Vázquez 2010) The third matching method, kernel matching, weighted the closeness and produced a lower variance. (Heinrich, Maffioli and Vázquez 2010) There was no single superior matching algorithm for all conditions and each model was tested and compared (Caliendo and Kopeinig 2008). A general approach was to test each matching algorithm and compare the robustness of the matching algorithm according to the result similarity in each matching algorithm (Heinrich, Maffioli and Vázquez 2010).

GLM regression model (GLM with a log link)

In addition to propensity score matching, we use a GLM regression with a log link to assess the effect of Medicare Advantage/Part C coverage on health care expenditure relative to traditional Medicare coverage. This regression used one-part GLM (log link) to control for the same individual characteristics and health status that were used in the logit model to generate propensity score. As the HMO coverage was taken as one of the independent variables that influenced the amount of health spending, the selection to Medicare Advantage/Part C was not explicitly adjusted.

The coefficients estimated by this GLM (log link) could provide a reference value for how much Medicare Advantage/Part C influenced the total and out-of-pocket health spending in the first three to four years under Medicare. Contrary to Chapter 1, a one-part expenditure model was adopted in this chapter for two major reasons. First, the percentage of zero total spending in the HRS data set was 1.3%. The small percentage did not follow one of the rationales to adopt two-part expenditure model, a large number of zero spending in the data set. Buntin and Zaslavsky (2004) suggested that two-part models did not significantly improve the precision of estimators relative to one-part models due to a small share of zero spending (< 9%). The second reason was that the use of two-part model could further limit the sample size in the second part, estimation of health spending among those incurring positive health spending.

Data and empirical specification

Data

MEPS datasets in Chapter 2 are cross-sectional and do not contain historical data about an individual's health status or other characteristics before and after Medicare enrollment. This chapter needs a dataset that allows researchers to adjust for individual characteristics before Medicare enrollment. Since the Health and Retirement Study (HRS) has a longer observation period prior to Medicare enrollment and follow-up there after, this data set is used for the next two objectives (in Chapter 3 and 4) as the second nationally representative dataset. The HRS consists of interviews that were implemented every two years from 1992,[21] with the latest cohort enrolled in 2010, and provides information about health status, health care consumption and expenditures from individuals age 50 and over (RAND Center for the Study of Aging 2010). Because of the

[21] The interviews were not implemented exactly every two years before 1996. For details, see RAND Center for the Study of Aging (2010).

availability of information before and after individuals' health plan selection, the information from HRS is appropriate for this study that required individual longitudinal information to address the health plan selection issue (RAND Center for the Study of Aging 2010).

According to the RAND HRS data documentation, version K, (RAND Center for the Study of Aging 2010), it is feasible to identify HRS participants with traditional Medicare (the reference group), Medicare Advantage/Part C plans (HMO coverage), dual eligibility with Medicare and Medicaid, and Medicare with employer-sponsored or retirement plans for the HRS participants age 65 and over. Although the types of coverage under Medicare are identifiable, the outcome, self-reported total health spending, is not available in all years. Information on total health spending, including the medical bills paid by the third parties, was not asked after 2002. By contrast, self-reported out-of-pocket medical expenses are available from 1992 to 2008 (RAND Center for the Study of Aging 2010). Since the data on total medical expenses is limited, this restricts sample size. Goldman, Zissimopoulos, et al. (2011) found that out-of-pocket spending in HRS was consistent with the statistics in MEPS and MCBS (Medicare Current Beneficiary Survey), but the total health spending surveyed before 2002 in HRS was overstated, compared to the results in MEPS and MCBS. Recognizing that total spending may be somewhat overstated, for some models available data on total spending was treated as an outcome and out-of-pocket spending was the outcome measure in other regression models.

Empirical specification

As previously noted, some factors such as health status, gender and income influenced both health plan selection and future health spending, so that propensity score matching was used here to address the endogeneity problem. The functional form used to estimate propensity score for each individual was a reduced-form equation that predicted the likelihood of enrolling in Medicare Advantage/Part C coverage instead of the baseline plan (*HC*=Medicare health coverage; traditional Medicare, *HC*=0; Medicare Advantage/Part C, *HC*=1). The health plans before individuals being enrolled in Medicare were the pre-Medicare characteristics (individual characteristics observed prior to being enrolled in Medicare, the independent variables in this model).

As the alternative plans was coded differently (Medicare Advantage/Part C, *HC*=1), the probability of enrolling in and maintaining this plan was analyzed by using the following equation.

$$\Pr(HC_{it} = 1) = \frac{e^{\beta_0+\beta_1 X_i+\beta_2 Cohort_i+\beta_3 HC_{it}+\beta_4 T+\varepsilon}}{1 + e^{\beta_0+\beta_1 X_i+\beta_2 Cohort_i+\beta_3 HC_{it}+\beta_4 T+\varepsilon}} \qquad (3.1)$$

The probability of enrolling in and maintaining health coverage (Medicare Advantage/Part C or traditional Medicare) continuously from age 65 to 69 [Pr(HC_{it}), compared to traditional Medicare with or without supplemental private coverage, *i*: individual characteristics; *t*: at age 69] depended on other pre-Medicare individual characteristics (X_i), including prior health coverage, age, gender, race/ethnicity, regions of residence and others, at age 64, which birth cohort they belong to ($Cohort_i$, coded as the birth years in which they were born), and the year in which they were interviewed (*T*) at age 64 years. After estimating the probability of selecting HMO coverage compared to those enrolled in traditional Medicare, each individual can be assigned a predicted probability of selecting Medicare Advantage/Part C, [HC_{it}=1 in Pr ($\widehat{HC}_{it}$)], similar to the procedures in Wilcox-Gok and Rubin (1990) and Cameron, Trivedi and Milne (1988).

After the predicted propensity to enroll in an alternative plan (Medicare Advantage/Part C relative to traditional Medicare) was assigned, matching algorithms were applied to group individuals for comparison and to estimate how this selection influenced the overall and out-of-pocket health expenditure. The individual health spending was estimated compared to their matched peers in other types of coverage with controls of individual characteristics and other potential confounders.

Sample selection and exclusion

To conduct the hypothesis testing in this chapter, adequate sample selection from HRS enrolles is the most important for hypothesis testing. In the HRS dataset, there are total 30,548 observations from 1992 to 2008. The following four major steps were uded to target the adequate HRS participants.

First, the individuals who were only observaed for their characteristics after age 65 years were excluded, because their health status before Medicare (one of the pre-Medicare characteristics) was not revealed and their selection behavior was unknown. Second, the individuals aged less than 65 years covered with Medicare for reasons other than age eligibility were not taken into consideration because they were more likely to have specific diseases, such as end-stage renal disease, and maintain Medicare coverage after age 65. Therefore, they were treated as outliers that might bias the selection regression away from null. Those who were not excluded for these two reasons were kept in the sample. Their insurance coverage and characteristics were recorded to understand the mechanism of HMO selection under Medicare. Third, individuals had to be interviewed twice (more than three years of observation) after obtaining and maintaining Medicare coverage. By the same token, individuals enrolled after 2004 were excluded because they were only observed from 2004 to 2008, less than the desired length in this study (first interview for pre-Medicare characteristics and two other consequetive interviews for information on the Medicare coverage from age 65 to age 68 years, a total of three inteviews or six years). Those who were not excluded based on these criteria were kept for the logit regression (to generate propensity score for matching) and health expenditure modeling with GLM (log link). The actual sample sizes were listed in the following sections.

Propensity score matching for total and out-of-pocket health spending

Statistical package and programs for propensity score matching

Propensity scores for different matching algorithms were generated by the user-defined program, *psmatch2* (Leuven and Sianesi 2003) written for STATA 9 or later versions (STATA Corp, College Station, Texas). This program provided a range of matching algorithms (listed in Table 3.1) to obtain mean differences and the matching results between the treated and control groups. After the mean differences were estimated, it is feasible to use the bootstrapping method to simulate multiple sets of matching and estimate the z statistics and p values (Leuven and Sianesi 2003). The process and results of selecting the best performing matching algorithms were written in detail in Appendix C.

The out-of-pocket and total health expenditure of each enrollee during the first four years of Medicare coverage (equivalent to enrollees' age 65 to 68 years) were summed and the spending of each type of coverage was estimated using individual characteristics, birth cohort identity, regions, and health plans estimated from previous section. Although the HRS was designed as a longitudinal study to follow up enrollees for more than four years, its longer observation time might lead to more sample attrition that threatened sample size and validity. In this Chapter, four years after Medicare coverage was assumed to be an adequate time period to observe and distinguish the differential effects on spending due to differences in Medicare enrollment status from age 65 to 68 (four years of Medicare coverage).

Number of eligible Medicare enrollees

The criteria to select eligible Medicare enrollees in propensity score matching limited the number of observations. Of all 30,584 HRS participants, the number of HRS observations with pre-Medicare information (characteristics before Medicare coverage at age 64 years) was a third of the number of total participants, 10,740. The requirement of enrolling those without Medicare coverage before age 65 years further restricted the sample size to 8,831 individuals. After excluding observations with missing data[22] and health plan transition[23] after 65 years of age during Medicare, the exclusion of those that were covered with Medicare due to reasons other than age eligibility further limited the

---

[22] A casewise deletion was applied for the observations lacking any information in the independent variables (pre-Medicare characteristics only) of the logit model (selection into Medicare Advantage/Part C).

[23] Health plan transition meant that the respondents had different health coverage in the first and second interviews after being covered by Medicare. Because health plan transition was not a focus of this study, only those with the same health coverage were kept in the analysis.

sample size. There were 4,949 participants selected for qualifying these criteria and maintaining the same health policy during the desired observation period (three to four years of Medicare coverage). Finally, the lack of information on health expenditures and other variables further limited the sample size to 1,438 and 3,580 for use in total and out-of-pocket spending models, respectively.

This number of observation was still more than the minimal number calculated according to the methods proposed by Long (1997) and Pamel (2000), as Peduzzi, Cancato, et al. (1996) had simulations that indicated that ten or more observations in a single variable would not lead to problematic logit results.

Results

The results are written in different sections. The first section shows the results of propensity score generation. The regression coefficients in the logit model are listed to discuss the significant factors of these observable individual characteristics. These significant factors could be seen as evidence that biased selection into Medicare Advantage/Part C existed. Then the estimated effects and the summary statistics in matching algorithms are compared to choose a matching result that is the least vulnerable to endogeneity or hidden bias. By listing the results of different matching methods, the theoretical effects of matching algorithms in Table 3.1 could be compared with actual estimations. Third, the most robust matching estimations for total and out-of-pocket health expenditures are chosen to investigate the balancing of the individual characteristics before and after matching. The results in variable balancing may help to indicate some possible sources of endogeneity of health plan selection. Finally, the results of propensity score matching are compared with the effects of HMO coverage on total and out-of-pocket health expenditure in the one-part GLM. The treatment, enrollment in Medicare Advantage/Part C or not, and the variables used in the logit model that predicts the propensity score are put in the model as independent variables.

Propensity score matching

Demographic characteristics

In Table 3.2, two groups of eligible enrollees (in total and out-of-pocket spending models) were compared.[24] The outcome (total and out-of-pocket health spending) and the treatment (traditional Medicare or Medicare Advantage/Part C) were the information observed after enrollees were enrolled in first three to four years of Medicare coverage (equivalent to their age 65 to 68 years). Except for the Medicare interview years, the other variables were individual pre-Medicare characteristics. The mean total spending for the selected individuals was $26,350.7 (SD = 64,906.7) and their out-of-pocket spending was $5,762.3 (SD = 12,452.1). In the out-of-pocket spending model, the mean spending was $6,514.0 (SD = 19,348.0)[25]. The percentages of Medicare Advantage/Part C coverage were 29.7% in total spending model and 26.6% in out-of-pocket spending model.

[24] There were 1,820 individuals whose total and out-of-pocket health expenditures were available among all eligible HRS enrollees. Their information on spending and characteristics were included in both columns.

[25] Because only 1820 of 4126 eligible observations in out-of-pocket spending model had information on total spending, this was not reported.

Comparing their pre-Medicare characteristics, most of them were 64-year-old[26], white, married and in good health status (self-assessed health status and the Center for Epidemiologic Studies Depression, CESD, score). The gender distribution was not the same (54.2% and 56.2% female for total and out-of-pocket expenditure). The proportion of those living in the South (39.8% and 40.6% for total and out-of-pocket health expenditure respectively) was higher than those in the other regions.

Because the information on total health spending was only collected before 2002, total expenditure sample represented an older birth cohort and their mean values of education attainment and income (nominal dollars) were less than those for the out-of-pocket health-spending model. However, the standard errors of income and educational attainment in both groups were also large.

On average, there were few mental problems (CESD scale)[27] for the individuals in the total spending model (1.27, SD = 1.79) and out-of-pocket spending models (1.31, SD = 1.85). In the total spending model, most individuals were observed to have no difficulty in ADL (Activities of Daily Living)[28] or IADL (Instrumental Activities of Daily Living)[29]

---

[26] Because the differences in the timing of Medicare enrollment, some individuals were covered after or before the Medicare age criteria, 65 years, but most of them were 64-year-old while being enrolled.

[27] In the HRS codebook, CESD scores were defined as follows (RAND Center for the Study of Aging 2010). "The CESD score is the sum of six "negative" indicators minus two "positive" indicators. The negative indicators measure whether the respondent experienced the following sentiments all or most of the time: depression, everything is an effort, sleep is restless, felt alone, felt sad, and could not get going. The positive indicators measure whether the respondent felt happy and enjoyed life, all or most of the time." There were no CESD scores reported before 1994.

[28] ADL included five tasks in HRS: bathing, eating, dressing, walking across a room, and getting in or out of bed (RAND Center for the Study of Aging 2010).

[29] IADL included the following tasks: using a telephone, taking medication, and handling money (RAND Center for the Study of Aging 2010). However, there were different IADL definition in 1992 and 1993 and these definitions were not used for this longitudinal RAND dataset. Therefore, individuals enrolled in the HRS before 1994 were not used for this logit regression because of the lack of adequate IADL information.

or mobility[30] (89.5%, 95.3%, and 61.8% respectively), as the percentages were 90.3%, 96.2%, and 58.6% respectively in the out-of-pocket spending model.

The pre-Medicare insurance status included government-provided plans, Medicaid and Champus/VA, covering 3.2% and 4.8% of the eligible individuals, as the role of private insurance remained substantial before Medicare coverage. The self-purchased private plans covered 28.0% and 30.2% of the eligible individuals in total and out-of-pocket spending models. The private plans purchased by their spouses covered 13.8% and 15.2% of the individuals in the total and out-of-pocket spending models.

The first years of HRS interview after Medicare coverage were from 1994 to 2003 (total health expenditure group) or to 2007 (out-of-pocket health expenditure group). Most people were born from 1929 to 1935 (total expenditure) or to 1941 (out-of-pocket expenditure).

Prediction of propensity score

Because of different sample sizes used to estimate total and out-of-pocket health expenditures, the predicted propensity score also differed for these two models. In Table 3.3, the sample sizes were 1,841 and 4,126 for total and out-of-pocket health expenditure in the logit regression models. These two logit models were both statistically significant ($p<0.05$).

The propensity score (the estimated probability of enrolling in Medicare Advantage/Part C rather than the traditional fee-for-service Medicare plan) predicted by

---

[30] Mobility was defined by five tasks: walking several blocks, walking one block, walking across the room, climbing several flights of stairs and climbing one flight of stairs (RAND Center for the Study of Aging 2010). These questions were not asked before 1994.

the logit models in the total spending model was on average 0.40 and 0.25 for individuals with traditional Medicare and Medicare Advantage/Part C respectively, as the mean propensity score in OOP spending model was 0.36 and 0.23 respectively. The higher propensity score in total spending model indicated those in the total spending model had an average higher probability of enrolling in HMOs than those in the out-of-pocket spending model.

Factors associated with the selection into HMOs under Medicare

Pre-Medicare characteristics and selection into HMOs

Table 3.3 shows that being black, Hispanic origin, regions of residence, pre-Medicare insurance plans (Champus/VA and private coverage) were significant factors ($p < 0.05$) for enrollees to select Medicare Advantage/Part C once they became Medicare-eligible at age (65 years) in both models. In the out-of-pocket spending model, four difficulties in mobility and widowhood were significantly associated with a smaller propensity score to be enrolled in Medicare Advantage/Part C ($p < 0.05$). Years of education and income (log scale) were not significant for this selection ($p<0.05$ in both models) when individuals became eligible for Medicare coverage at age 65 years.

Geographic locations and regions for the HMO selection

The regions of residence were common significant predictors for these two logit models, the residence in the Midwest and South ($p<0.05$) that was negatively associated with the likelihood of selecting Medicare Advantage/Part C, compared to the residence in the Northeast (Table 3.3).

The residence in the West was associated with a higher likelihood of selecting Medicare Advantage/Part C ($p<0.01$), compared to the residence in the Northeast in both models. The sample sizes in the other regions in both models were not sufficient (less than ten) to estimate the probability of being enrolled in Medicare Advantage/Part C.

Health status, mental health status, functional status and the selection in Medicare Advantage/Part C

In both models, self-rated health status, mental health status (CESD scale), limitations in ADL and mobility were not significantly associated with the selection in Medicare Advantage/Part C in both models.

However, four difficulties in mobility were negatively associated with the likelihood of selecting Medicare Advantage/Part C in out-of-pocket model ($p<0.05$). If four difficulties in mobility was a proxy for a larger likelihood of health events among Medicare enrollees, the results might show that a biased selection to Medicare Advantage/Part C existed in the out-of-pocket model.

Pre-Medicare insurance coverage and the selection in Medicare Advantage/Part C

There were three major pre-Medicare insurance plans (the health plans individuals had before they were eligible for Medicare at age 65 years) documented in HRS datasets, Medicaid, private coverage[31], and Champus/VA. Medicaid was not significantly associated with the selection in Medicare Advantage/Part C in both models in Table 3.3. However, pre-Medicare Champus/VA coverage was negatively associated with the likelihood of selecting Medicare Advantage/Part C ($p<0.01$) in both models. The pre-

[31] Private coverage included the coverage purchased by the respondents themselves or by their spouses.

Medicare private coverage purchased by enrollees themselves or their spouses were positively associated with the likelihood of Medicare Advantage/Part C coverage ($p < 0.05$).

Effects of Medicare Advantage/Part C on expenditure predicted by propensity score matching and GLM

GLM (Gamma family with log link) estimates

The cost-saving effects of HMO coverage under Medicare on the amount of total health care spending in the first three to four years of coverage was $1,515.5 (SE = 2,743.9, P = 0.58) in total-spending model and $1,772.5 (SE = 445.7, $p < 0.01$) less than those in the traditional fee-for-service Medicare plan after individuals' pre-Medicare characteristics were controlled for (first part of data in Table 3.4). The GLM prediction[32] indicated an insignificant effect on total health spending and a significant effect on out-of-pocket health spending in the first three to four years of Medicare coverage. However, the GLM estimates did not consider the process of selection into Medicare Advantage/Part C coverage and the results should be interpreted with caution.

Propensity score matching estimates

Similar to the prediction in GLM, the effect of HMO coverage in total spending model was not significant, $2,651.0 less than the traditional Medicare plan (SE = 3,761.2 in kernel matching with caliper [0.1] or kernel matching with bandwidth as 0.6,

[32] See Appendix C for the modified Park test results and the regression (GLM) coefficients.

bootstrapped p = 0.46) in the second part of Table 3.4[33]. Its cost-saving effect on out-of-pocket spending was significant, \$1,411.5 less than the traditional Medicare plan (SE = 620.8 in kernel matching with bandwidth as 0.1, bootstrapped p = 0.03).

However, the unmatched differences in the mean Medicare spending between those enrolled in Medicare Advantage/Part C (treated group) and traditional Medicare (control group) in the first three to four years of coverage (equivalent to the Medicare enrollees' age 65 to 68 years) were not significant in the total spending models ($p = 0.78$). The difference in total spending (\$2,411.0 less than the traditional Medicare plan, SE = 3,312.4) was smaller than the estimated magnitude in propensity score matching (\$2,651.0, p=0.46). The unmatched out-of-pocket spending difference (\$1,244.0 less than traditional Medicare plan, SE = 620.83, $p < 0.01$) was smaller than the absolute magnitude predicted by propensity score matching (\$1,411.5 less, SE = 681.4, $p < 0.01$). This piece of evidence showed that part of the out-of-pocket spending difference between traditional Medicare and Medicare Advantage/Part C might be induced by selection into Medicare Advantage/Part C[34], rather than purely by the variation of cost-sharing and benefit packages in different plans.

---

[33] The results and selection of propensity score matching were written in detail in Appendix C. The propensity score matching estimated the average treatment effect on the treated by taking the differences of the mean Medicare spending in the treatment and control groups.

[34] Because the observed spending difference was smaller than the estimated effect of HMO coverage (from propensity score matching), the selection into HMOs by those with a larger likelihood of spending might contribute to the difference that was smaller than expected.

Discussion

Limitations in this chapter

Propensity score matching is a technique that compares the treated and control individuals of similar characteristics to remove the threat of endogeneity and obtain an average treatment effect on the treated (ATT). However, the missing data and changes in questionnaires lead to a smaller sample size for propensity score matching, as the HRS ceased to gather or impute information on total health spending after 2002. The cessation in data gathering in total spending further limited the sample size to less than two thousand for the propensity score estimation in the total health expenditure model. However, the sample size was still more than the minimal number calculated according to the methods proposed by Long (1997) and Pamel (2000) and Peduzzi, Cancato, et al. (1996), who had simulation results that indicated ten or more observations in a single variable would not lead to problematic logit results.

Although this small sample size did not cause important threat to the validity of this analysis, it might be the main reason why there was no statistically significant difference (ATT) found in total health expenditure. On the other hand, the continued collection of the information on out-of-pocket health expenditure helped to maintain most of the observations selected from the reconstructed cohorts.

Obtaining and maintaining (for three to four years) HMO enrollment

The logit model in Table 3.3 showed which individual characteristics might be associated with the decision to enroll in HMOs under Medicare. The significant factors included being black, Hispanic origin, regions, four difficulties in mobility (out-of-pocket spending model), widowhood (out-of-pocket spending model) and pre-Medicare health coverage.[35] This result was surprising because in the literature health status was the major driver for biased selection (Morrisey 2007). One reason why this differed from the classic literature might be that the previous studies focused on a younger generation, most of who had no or few health issues and the demand for health care was relatively low (Chernew and Cutler et al. 2005). It might be possible that only those who perceived higher rate of health capital attrition[36] in the younger generation would seek for more generous health coverage and health status was a proxy for health capital attrition.

For the studies focusing on the same age range, findings also differed. Although one study on the supplementary health coverage, Medigap, found adverse selection (Ettner 1997), other studies on the Medicare population found favorable selections to Medicare Advantage/Part C with relatively healthy demanding more coverage (Mello, Stearns, et al. 2003; Shen, Hendricks, et al. 2005; Cox and Hogan 1997; Greenwald, Levy, & Ingber 2000). However, Feldman, Dowd, et al. (2003) pointed out that there was an adverse selection to Medicare Advantage/Part C plans and the main factor was the generous drug plan that attracted those with worse health status, based on their structural model and data from MCBS. They also observed favorable selection in dental coverage at the same time, with those in relatively good dental health purchasing dental supplemental coverage (Feldman, Dowd, et al. 2003).

[35] This assumed these factors were significant in spite of future total or out-of-pocket health spending individuals had expected because the total health-spending group was merely a subsample of out-of-pocket spending group.

[36] See Grossman (1972) and Grossman (1999) for a detailed definition of health capital and its attrition.

In a natural experiment in Minnesota, Zhang, Kane, et al. (2008) found that favorable selection into HMOs was found in the dual eligible (Medicare and Medicaid) population, especially those living in the nursing home and communities.

To conclude, these studies indicated that selection existed in different types of services in HMO coverage and populations. The direction and magnitude of biased selection varied across these studies. In this chapter, the logit model suggested that those with four difficulties in mobility on IADL were less likely to seek Medicare Advantage/Part C coverage. Other characteristics, including region of residence and pre-Medicare Chapmus/VA coverage[37] also influenced individual decisions to enroll in Medicare Advantage/Part C, relative to traditional Medicare. The selection in Medicare Advantage/Part C was biased due to these factors, which were not exactly the same as the characteristics reported in the literature.

The effect of HMO on total and out-of-pocket health expenditures among Medicare enrollees

After applying propensity score matching to deal with potential bias from endogeneity, the effects of Medicare Advantage/Part C on total health spending were not statistically significant in all matching algorithms used in this chapter. In the out-of-pocket health-spending model, HMO coverage was estimated to be associated with an average $1,411.5 less out-of-pocket health spending for individuals of similar health status and characteristics in the first four years of Medicare coverage.

[37] The Champus/VA coverage directly served as the secondary payer for enrollees' Medicare coverage and entitled for the Part A premium waiver so that the cost sharing for the Champus/VA eligible individuals was lower and the there was no strong incentive for eligible individuals to switch to the Medicare Advantage/Part C plan (U. S. Department of Veterans Affairs 2010).

The variable balance (similar distributions of variables between the treated and control groups) in OOP spending model showed that the factors that were associated with biased selection into Medicare Advantage/Part C were similarly distributed in the treatment group and their matched neighbors in the control group. This balanced distribution of individual characteristics that influenced individual decision to enroll in Medicare Advantage/Part C and contributed to the endogeneity issue helped to ensure the treated cases and matched control neighbors comparable. In theory, the endogeneity problem of plan selection should be removed by propensity score matching.

The sensitivity analysis, Rosenbaum bounds, identified Kernel matching algorithm (bandwidth 0.1) as the most robust and the least vulnerable from hidden bias for OOP spending model.[38] Although different magnitudes were found, the cost-saving effect on OOP spending was also captured by the Gamma GLM ($1,772.5 less OOP spending in Table 3.4).

Other details in health expenditure GLM

The results of Gamma GLM (OOP health spending) showed that there were significant effects from characteristics other than Medicare Advantage/Part C. These significant factors directly influencing the amount of OOP health spending included Medicare Advantage/Part C, being black, Hispanic origin, living in the South, education attainment, health status, three difficulties in ADL, one difficulty in mobility, being divorced, pre-Medicare health plans (Medicaid coverage and Champus/VA coverage) and the years when they spent.

---

[38] See Appendix C for details in the sensitivity analysis and selection process.

These significant variables for the amount of out-of-pocket spending did not have imbalance after matching (the individual characteristics were similar between the treated and control groups after matching). Some of them also influenced the probability of choosing Medicare Advantage/Part C, including being black, region and pre-Medicare health coverage. These characteristics seemed to play an important role in determining the level of health spending, at least in the first three to four years of Medicare coverage. Further studies with larger sample sizes might be necessary to understand their effects and the mechanism to induce different levels of health spending.

Difference in the predicted HMO effect on OOP spending between propensity score matching and GLM

In Stukel, Fisher, Wennberg, Alter, Gottlieb and Vermeulen (2007), endogeneity was dealt with propensity score matching and instrumental variable regression. After using geographic region as the instrument, their results in the two-step regression were similar to those in propensity score matching. This suggested that the regression models without instrumental variables in this chapter might not be optimal to control endogeneity. The estimates in the regression remained vulnerable to endogeneity.

Moreover, the biased estimates in regression seemed to be due to the lack of controlling for the selection to HMO coverage. Because four difficulties in mobility were positively related to gaining HMO coverage and Medicare Advantage/Part C was negatively associated with out-of-pocket spending, these two directions combined created a plausible explanation for the biased estimates in regressions. However, the limitation in sample sizes and difficulty in finding an instrument made further research necessary to illustrate the detailed relationship from health status to actual spending.

Chapter 4: Long-term health returns among the Medicare enrollees

Introduction

In literature discussing the returns from health expenditures, returns were typically defined as the extension of life expectancy in different eras or the estimated value of the deaths averted (for example, in Cutler, Rosen and Vijin 2006; Cutler and Richardson 1999; Cutler and McClellan 2001). These studies made some important assumptions, including that the changes in life expectancy were the result of the application of advances in health technology that were associated with increased health spending. However, the most recent studies indicated that health changes may not be immediately observed and could accrue over time, and encompass other health dimensions besides mortality. For example, the studies on the effect of the prescription drug caps which reduced spending showed that these policies resulted in higher incidence of poorly-controlled chronic health conditions that increased emergency room use and subsequent health spending much later than previous estimations (Hsu, Price, Huang, Brand, Fung, Hui, et al. 2006 and Newhouse 2006).

The advantage of the longitudinal design of the datasets such as the HRS is its ability to track long-term health changes for specific individuals and their health spending over the same time periods. HRS has longer lengths of panel observation compared to MEPS and the expenditure data will be compared to MEPS expenditure data. The longer HRS panel could give us a better insight toward health returns from such health care spending.

The null hypothesis for this analysis was as follows:

Among Medicare enrollees, the health returns (mortality, changes in health status, disease incidence, and mental health) did not change according to differences in the total and out-of-pocket health spending among enrollees.

Method

Data Source and Sample

Data

This chapter needed a dataset that allows researchers to adjust for the individual status before Medicare enrollment. Since the Health and Retirement Study (HRS) had a longer observation period prior to Medicare enrollment and follow-up there after, this data set was used for the research objectives in Chapter 3. The HRS consisted of interviews that were implemented every two years from 1992 with the latest cohort enrolled in 2010 and provided information about health status, health care consumption and expenditures from individuals age 50 and over (RAND Center for the Study of Aging, 2010). Because of the availability of information before and after individuals' health plan selection, the information from HRS is appropriate for this study that requires individual longitudinal information to address the health plan selection issue.

According to the RAND HRS data documentation, version K, (RAND Center for the Study of Aging 2010), it is feasible to identify HRS participants with traditional Medicare with or without private supplemental coverage (the reference group), Medicare Advantage/Part C, dual eligibility with Medicare and Medicaid, and Medicare with employer-sponsored or retirement plans for the HRS participants age 65 and over. Although the types of coverage under Medicare were identifiable, the outcome, self-reported total health spending, was not available in all years. Information on total health spending, including the medical bills paid by the third parties, was not asked after 2002. By contrast, self-reported out-of-pocket medical expenses were available from 1992 to 2008 (RAND Center for the Study of Aging 2010). Since the data on total medical expenses was limited, this restricted sample size. Goldman, Zissimopoulos, et al. (2011) found that out-of-pocket spending in HRS was consistent with the statistics in MEPS and MCBS (Medicare Current Beneficiary Survey), but the total health spending surveyed before 2002 in HRS was overstated, compared to the results in MEPS and MCBS. Recognizing that total spending may be somewhat overstated, for some models available data on total spending were treated as an outcome and out-of-pocket spending were the outcome measure in other regression models.

Sample selection and exclusion

To conduct the hypothesis testing in this chapter, adequate sample selection from HRS enrollees was the most important for hypothesis testing. In the HRS dataset, there were total 30548 observations from 1992 to 2008. There were four major steps to target the adequate HRS participants.

First, the individuals whose characteristics were observaed only after age 65 years were excluded, because their health status before Medicare (pre-Medicare health status) was not revealed and their selection behavior was unknown. Second, the individuals aged less than 65 years covered with Medicare for reasons other than age were not taken into consideration because they were more likely to have specific diseases, such as end-stage renal disease or were disabled, and maintain Medicare coverage after age 65. Therefore they were treated as outliers that might bias the selection regression away from null. Those who were not excluded for these two reasons were kept in the sample. Their insurance coverage and characteristics were recorded to understand the mechanism of HMO selection. Third, individuals had to be interviewed twice (more than three years of observation) after obtaining and maintaining Medicare coverage. By the same token, individuals enrolled after 2004 were excluded because they were only observed from 2004 to 2008, less than the desired length in this study (2-year observation before age 65 years and 4 years from age 65 to 69 years). Those who were not excluded based on these criteria were kept for the logit regression (assessing the change in the probability of obtaining returns to the five health dimensions) and health expenditure modeling with GLM (log link) (quantification of the financial impact of death events).

Health indicators in HRS

Health returns can be defined very differently. In the HRS codebook, major health indicators were grouped into health status and its change, health conditions (hypertension, heart disease, stroke, psychiatric problems and arthritis), activities of daily living (ADLs), other functional limits, health behaviors, physicians' diagnosis, and mental health [Center for Epidemiologic Studies Depression (CESD) scale] (RAND Center for the Study of Aging, 2010). Mortality was also recorded with an exit survey (RAND Center for the Study of Aging, 2010). Among these health indicators, health status change, cardiovascular diseases (incidence and associated spending) and mortality were frequently used to assess the return from health spending (Cutler and Richardson 1999 and Cutler and McClellan 2001).

Estimating the relationship between the probability of health return change and health spending (total and out-of-pocket)

Statistical package and programs for assessing health returns

The statistical package used to execute the logit and ordered logit regression is STATA 11 (STATA Corp, College Station, Texas). The health returns with dichotomous outcomes, including mortality, hypertension incidence and arthritis incidence, are analyzed with logit regression models to adjust for its non-linear relationship between the outcome and the independent variables. The outcomes with multiple categories, including health status and mental health rating, are assessed with ordered logit regression model not only to fit multiple categories of these outcomes, but also adjust for the non-linear relationship between outcomes and independent variables.

The out-of-pocket and total health expenditure of each enrollee during the first four years of Medicare coverage were summed and the spending associated with different health conditions was estimated using individual characteristics, birth cohorts, regions, and health plans estimated from previous section. Although the HRS was designed as a longitudinal study to follow up enrollees for more than two years, its longer observation time might lead to more sample attrition that threatened sample size and validity. In this Chapter, four years after Medicare coverage was assumed to be an adequate time period to observe and distinguish the differential effects of spending due to differences in Medicare enrollment status from age 65 to 68 years (written as 65-68 in the equations or texts).

Number of eligible Medicare enrollees

Of all 30,584 HRS participants, the number of HRS observations with pre-Medicare information (individual information before Medicare coverage) was a third of the number of total participants, 10,740. The requirement of enrolling those without Medicare coverage before age 65 years further restricted the sample size to 8,831 individuals. After excluding observations with missing data[39] and health plan transition[40] during Medicare, the exclusion of those that were covered with Medicare due to reasons other than age eligibility limited the sample size. There were 4,949 participants selected for providing information about the selected independent variables, qualifying these criteria and maintaining the same health policy during the desired observation period (three to four years after Medicare coverage). Finally, the lack of information in health expenditure and other variable further limited the sample size to 1,752 and 4,032 in total and out-of-pocket spending models.

After being selected based on the criteria, there were 1,752 and 4,032 individuals that had information about total and out-of-pocket health expenditure respectively, because of the plan attrition and missing data issues[41] after pre-Medicare interviews (the latest HRS interview before Medicare coverage). This number of observation was still more than the minimal number calculated according to the methods proposed by Long (1997) and Pamel (2000), as Peduzzi, Cancato, et al. (1996) had simulation that indicated

[39] A casewise deletion was applied for the observations lacking any information in the independent variables of the logit model (HMO selection).

[40] Health plan transition meant that the respondents had different health coverage in the first and second interviews after being covered by Medicare. Because health plan transition was not a focus of this study, only those with the same health coverage were kept in the analysis.

[41] Because difficulty in mobility, HMO coverage and cognitive limitation were not collected before 1994, 1994 and 1996 respectively, this was part of the missing data problem.

that ten or more observations in a single variable would not lead to problematic logit results.

Health spending definition

In this chapter, spending on health care was taken as a form of investment on health capital. Because this study aimed to study the long-term relationship between health spending and the changes in health outcomes, the total or out-of-pocket spending on health care in the first three to four years of Medicare coverage (equivalent to 65 to 68 years of age of Medicare enrollees) was summed to determine its influence on health outcomes. The advantage of HRS is to provide longitudinal information for specific individuals and the health spending in the first three to four years of Medicare coverage was summed. However, because of the data imprecision problem, total health spending was collected until 2002. The sample sizes in the total spending models were smaller than those in the out-of-pocket spending models. The notation 65-68 in the equations was used to indicate the accumulated sum of total or out-of-pocket health spending in the first three to four years of Medicare coverage (equivalent to Medicare enrollees' 65 to 68 years of age).

Health Outcomes Definition

Mortality

Information on mortality (1 = yes for being recorded dead, 0 = no) was collected from multiple sources: the exit survey for the deceased members within households and National Death Index (NDI) (RAND Center for the Study of Aging 2010). These sources provided detailed dates of death that help researchers to estimate the survival time after particular events. In this study, the length of survival was measured by months.

Hypertension

HRS participants were asked whether they were diagnosed with high blood pressure or hypertension (1 = yes for being diagnosed with this condition, 0 = no) by a medical doctor or general practitioner (RAND Center for the Study of Aging 2010). Because the probability of disease incidence was the study focus, those already having hypertension before Medicare coverage were excluded. The logit model predicted the chance of having hypertension after three to four years of Medicare coverage among those without hypertension before Medicare coverage.

Arthritis

This is a dichotomous variable (1 = yes for being diagnosed with arthritis, 0 = no). The HRS participants were asked whether a doctor had ever told them that they had arthritis or rheumatism (RAND Center for the Study of Aging 2010). Because the logit model was used to investigate the probability of having this condition after three to four years of Medicare coverage, those had arthritis or rheumatism before getting Medicare coverage were excluded.

Endogeneity consideration about the relationship between disease incidence and health spending

However, the goal of this study is to understand the long-term effect of health spending on disease incidence (at least three years in this chapter) and there is a possibility that the Medicare enrollees become afflicted with this chronic condition soon after being covered by Medicare (for example, within two years of Medicare coverage). Including the individuals with early-onset conditions will create endogeneity problem for the analysis that focuses on the long-term relationship. This endogeneity is caused because these early-onset conditions will not only indicate the occurrence of this condition after this period of time (dependent variable of this regression model), but also increase the amount of health spending incurred for disease treatment (independent variable of this regression model).

To partially remove this endogeneity problem, this section of analysis will first exclude individuals with existing and early-onset chronic conditions, especially within the first one or two years, to capture the long-term relationship between the disease and health spending. Moreover, the incidence of early-onset chronic conditions will be analyzed with the other logit model to understand the characteristics that lead to higher propensity of being diagnosed with these conditions soon after being enrolled in Medicare.

Therefore, there will be two logit regression models for chronic conditions (hypertension and arthritis in this chapter). The first model will assess the long-term relationship between spending and the incidence after three to four years of Medicare coverage after excluding individuals with existing and early-onset (occurred within first two years of Medicare coverage) chronic conditions. The second one will determine which characteristics are associated higher likelihood of having an early-onset condition within the first two years of Medicare coverage.

Self-assessed health status

There were five categories for self-rated health status in the HRS questionnaires, including excellent (the reference value, coded as 0), very good (coded as 1), good (coded as 2), fair (coded as 3), and poor (coded as 4) health status (RAND Center for the Study of Aging 2010). The HRS participants were asked to provide a general assessment of their health status and chose one of these five categories in each interview. Additionally, the health status before Medicare (pre-Medicare health status at 64 years of age) was used to predict the health returns after three to four years of Medicare coverage (usually from 65 to 68 years of age). By controlling for the pre-Medicare health status, the likelihood of having one of these five self-assessed health status at age 68 years of age was estimated.

Mental health status (CESD scale)

The CESD (Center for Epidemiologic Studies Depression) scale summed eight indicators[42] and the interviewees were assigned a score ranging from zero (0 = reference value, the mental status without any problem) to eight (8 = the mental status with problem in all indicators). Higher score suggested a respondent's more negative feeling regarding his/her mental health status (RAND Center for the Studies of Aging 2010). The scale was used as a categorical health outcome for the ordered logit model. However, this variable was not collected before 1993 and those who were not interviewed for this question were excluded.

---

[42] In the HRS codebook, CESD scores were defined as follows (RAND Center for the Study of Aging 2010). "The CESD score is the sum of six "negative" indicators minus two "positive" indicators. The negative indicators measure whether the respondent experienced the following sentiments all or most of the time: depression, everything is an effort, sleep is restless, felt alone, felt sad, and could not get going. The positive indicators measure whether the respondent felt happy and enjoyed life, all or most of the time." There were no CESD scores reported before 1994.

Functional forms

The effects of total or out-of-pocket health spending on dichotomous health outcomes were estimated by using binary logit regression models, and health outcomes with multiple categories were assessed with ordered logit regression models. To have a broader understanding, multiple measures of health outcomes, especially two of the most incident diseases among the elderly, hypertension and arthritis, were chosen to reveal differential effects of prior total and out-of-pocket (OOP) health spending (all health care expenditure incurred after individuals were covered in the first three to four years of Medicare coverage [equivalent to Medicare enrollees' age from 65 to 68 years], denoted by ***Spending*** in the following equations) on individual health. The models were specified as:

$$\begin{aligned}\ln\,(odds\ of\ Mortality_{i(68)}) &= \beta_0 + \beta_1 X_{ij} + \beta_2 Cohort_i + \beta_3 HC_{ki(65-68)} + \beta_4 Spending_{i(65-68)} \\ &+ \beta_5 HealthStatus_{i(64)} + \beta_6 MentalHealth_{i(64)} + \varepsilon \qquad (4.1)\end{aligned}$$

$$\begin{aligned}\ln\,(odds\ of\ Hypertension_{i(68)}) &= \beta_0 + \beta_1 X_{ij} + \beta_2 Cohort_i + \beta_3 HC_{ki(65-68)} + \beta_4 Spending_{i(65-68)} \\ &+ \beta_5 HealthStatus_{i(64)} + \beta_6 MentalHealth_{i(64)} + \varepsilon \qquad (4.2)\end{aligned}$$

$$\begin{aligned}\ln\,(odds\ of\ Arthritis_{i(68)}) &= \beta_0 + \beta_1 X_{ij} + \beta_2 Cohort_i + \beta_3 HC_{ki(65-68)} + \beta_4 Spending_{i(65-68)} \\ &+ \beta_5 HealthStatus_{i(64)} + \beta_6 MentalHealth_{i(64)} + \varepsilon \qquad (4.3)\end{aligned}$$

$$\ln\left(odds\ of\ HealthStatus_{i(68)}\right) = \beta_0 + \beta_1 X_{ij} + \beta_2 Cohort_i + \beta_3 HC_{ki(65-68)} + \beta_4 Spending_{i(65-68)} + \beta_5 HealthStatus_{i(64)} + \beta_6 MentalHealth_{i(64)} + \varepsilon \tag{4.4}$$

$$\ln\left(odds\ of\ MentalHealth_{i(68)}\right) = \beta_0 + \beta_1 X_{ij} + \beta_2 Cohort_i + \beta_3 HC_{ki(65-68)} + \beta_4 Spending_{i(65-68)} + \beta_5 HealthStatus_{i(64)} + \beta_6 MentalHealth_{i(64)} + \varepsilon \tag{4.5}$$

The HRS includes measures at age 68 years of individual deaths, $Mortality_{i(68)}$, the occurrence of hypertension, $Hypertension_{(68)}$, the occurrence of arthritis or rheumatism, $Arthritis_{(68)}$, individual health status, $HealthStatus_{i(68)}$, and individual health coverage from age 65 to 68 years, $HC_{(65\text{-}68)}$. Individuals' characteristics before being covered by Medicare, X_{ij}, included not only their demographic, socioeconomic, health, and functional characteristics, but also their pre-Medicare health coverage. The pre-Medicare health coverage documented in HRS datasets included Medicaid, Champus/VA coverage, and private plans purchased by the enrollees themselves or their spouses, as those covered by Medicare before age 65 years were excluded.

For the health coverage after age 65 years, Medicare Advantage/Part C (coded as 1) is the only health plan alternative to traditional Medicare coverage (coded as 0) because of the limited sample size and data gathering in HRS datasets that could not reflect the diverse choices of Medicare supplement private coverage (for example, more than nine types of Medigap plans for the Medicare enrollees (Centers for Medicare & Medicaid Services 2012)). The other advantage of grouping the Medicare coverage into these two major categories is its similarity to the approach in Centers for Medicare & Medicaid Services (2012) that guides the public to choose their own Medicare plans.

The likelihood of mortality (1 for death; 0 for survival), hypertension (1 for being diagnosed; 0 for none), and arthritis (1 for being diagnosed; 0 for none) were estimated using binary logit models; health status (categorical variable) was estimated using ordered logit regression models to control for other covariates and to evaluate the effects of spending on health care and health status. Mental health change in CESD (Center for Epidemiologic Studies Depression) scale that was rated by eight indicators with scores ranging from 0 to 8 (RAND Center for the Study of Aging, 2010) was assessed by ordered logit regression to control for other factors.

Quantification of the financial impact of death

Assumptions

The financial impact of death could be estimated by comparing the difference in the level of total or out-of-pocket health spending, if death events met following assumptions. First, the length of time was long enough to reveal the financial impact of death-associated health spending. In literature, end-of-life care could be defined as the care delivered within 6 months (Kaul, McAlister, et al. 2011; Unroe, Greiner, et al. 2011; Gibson 2011) to 12 months before death (Lubitz and Riley 1993). Those surviving and deceased were at least observed for 40 months (three to four years of observation) in this chapter. The first assumption was met.

Second, the information on death and death date was accurate and verified. Because the death events were also verified by the National Death Index (NDI), the quality of the data on death was reliable. Third, the unobserved endogeneity that caused both more health spending and higher likelihood of death should be controlled. In this chapter, the longitudinal study design and statistical modeling (GLM) might help to solve the endogeneity issue.

Quantification of the financial impact of death events with the Gamma GLM (log link)

To quantify the financial impact of death events on total and out-of-pocket health spending within three to four years of Medicare coverage, regression models that control for other observable factors serve for this purpose.[43] The dependent variable is total or out-of-pocket health spending spent within first four years of Medicare coverage. This regression used one-part GLM (log link) to control for the same individual characteristics and health status that were used in the logit model to generate propensity score. Besides the HMO coverage added in Chapter 3, this model introduced chronic conditions (hypertension and arthritis before Medicare coverage) and mortality within three to four years of Medicare coverage as independent variables in the equation. The coefficients estimated by this GLM (log link) could provide a reference value for how much these chronic conditions and death events influenced the total and out-of-pocket health spending in the first three to four under Medicare.

Contrary to Chapter 1, a one-part expenditure model was adopted in this chapter for two major reasons. First, the percentage of zero total spending in the HRS data set was 1.3%. The small percentage did not follow one of the rationales to adopt two-part

---

[43] The functional form of this spending model is as follows: $HealthSpending_{i(65-68)} = \beta_0 + \beta_1 X_{ij} + \beta_2 Cohort_i + \beta_3 HC_{ki(65-68)} + \beta_4 Condition_{i(64)} + \beta_5 Death_{i(65-68)} + \varepsilon$. The pre-Medicare characteristics ($X_{ij},\ Cohort_i,\ Region_i,\ and\ Condition_{i(64)}$) and individual characteristics under Medicare coverage ($HC_{ki(65-68)}\ and\ Death_{i(65-68)}$) were used to predict the amount of total and out-of-pocket health spending.

expenditure model, a large number of zero spending in the data set. Buntin and Zaslavsky (2004) suggested that two-part models did not significantly improve the precision of estimators relative to one-part models due to a small share of zero spending (< 9%). The second reason was that the use of two-part model could further limit the sample size in the second part, estimation of health spending among those incurring positive health spending.

Hence, this section of research would take advantage of the Gamma GLM that well modeled the variance structure of HRS spending data (variance proportional to the square of the mean values, $\lambda = 2$) in Chapter 3. The financial impact of death-associated health spending was estimated by the Gamma GLM spending model and compared with the spending estimates without mortality in Chapter 3.

Results

Health dimension one: mortality

Length of observation and mortality

In Table 4.1, the observed lengths of time for enrollees in Medicare Advantage/Part C or traditional Medicare were listed. Because of missing data, total number of observation was larger than the number of those with information in Medicare coverage. These deceased Medicare enrollees were observed for at least 37 months.

In Table 4.2, the characteristics of those surviving and deceased in the first three to four years of Medicare coverage were listed. The characteristics of those deceased and survived in the first three to four years of Medicare coverage were not exactly the same.

The proportions of persons having chronic conditions (hypertension and arthritis) were significantly higher among the deceased ($p = 0.02$ and 0.03 respectively). The percentage of female enrollees was significantly lower among the deceased ($p < 0.001$). The level of total spending was significantly higher among the deceased ($p < 0.001$), who did not have significantly different level of out-of-pocket spending ($p = 0.12$). However, there were only 1174 observations valid for the total health expenditure model, compared to 3279 valid for out-of-pocket spending analysis.

The other major discrepancies between those who survived and the deceased included higher percentage of blacks ($p < 0.01$), different regions of residence ($p = 0.04$), less education attainment and income for the deceased (both $p < 0.01$), different health status distribution and functional limitations ($p <0.001$), more mental problems among the deceased ($P<0.001$), different insurance coverage and earlier birth years among the deceased ($p<0.001$).

Survival probability and curve

In Figure 4.1, the probability of death was summarized and the shortest length of observation was 40 months after indicating the pre-Medicare status. The proportion of death increased after 45-month observation and the increase in mortality slowed after 50-month observation. However, this figure did not adjust for the loss of follow-up and right censoring in the HRS.

Kaplan-Meier survival curve

To adjust for the censoring in the surviving individuals, a formal survival curve, Kaplan-Meier survival curve, was estimated in Figure 4.2. After taking the censoring and sample attrition in the surviving individuals into consideration, the curve showed a gradual decline of survivorship until 60-month observation. However, the survival estimates adjusted the censoring issue and sample attrition without controlling for other factors. The effects of health spending on mortality remained unknown.

Logit regression: the probability of mortality and the level of health spending

Model summary

In Table 4.3, the individual characteristics and health expenditure were used as the independent variables to predict the probability of death after the first three to four years of Medicare coverage. The number of observations for total health expenditure model (1,752) was less than that for the OOP spending model (4,032). However, both models were statistically significant ($p<0.001$) and the OOP health spending had higher pseudo R^2 (0.22) than that of the total health spending model (pseudo $R^2 = 0.14$).

Association between individual characteristics and mortality

The tables in the following sections list the coefficients of all logit models and these coefficients were the log-odds of individual characteristics. The effect of health expenditure differed in terms of payment types. The total health expenditure was significantly associated with higher likelihood of mortality with three to four years of Medicare coverage ($p < 0.01$), but this positive association was not significant for OOP spending ($p = 0.17$). Other commonly significant factors included being female, health status, and difficulty in mobility. Females had a significant survival advantage during the first three to four years of Medicare coverage ($p < 0.01$ in both models). Health status worse than very good (good, fair and poor) was associated with higher mortality rate (p <

0.01 in total and OOP spending models). However, only two difficulties in mobility was observed to be associated with higher mortality in both models ($p < 0.01$), as one, three and five difficulties were significantly related to higher mortality in total spending model ($p < 0.05$).

There were other characteristics significant in single models. Medicare Advantage/Part C was associated with a lower likelihood of death in the first three to four years of Medicare coverage. Widowhood was positively associated with higher likelihood of death ($p < 0.01$ in total spending models). Birth years were negatively associated with the probability of death ($p < 0.05$ in out-of-pocket spending model). This suggested that the younger cohorts had a relatively lower likelihood of death than older cohorts. However, these two chronic conditions, hypertension and arthritis, were not significantly associated with mortality.

Health dimension two: hypertension

Long-term relationship between health spending and the incidence of hypertension

The results of logit model to predict the incidence of hypertension were listed in Table 4.4. The total-health-spending model had only 898 observations and this logit model was not significant at all ($p = 0.92$). The OOP spending model with 1905 enrollees was statistically significant ($p=0.002$). Although the pseudo R^2 of the OOP spending model was not large (0.05), there were some notable significant characteristics. First, self-rated health status ($p=0.045$, <0.001, and 0.01 for very good, good, and fair health status respectively) was associated with a higher likelihood of having hypertension and the effect was incremental with a worse health category, except for the poor health category (p=0.98).

The other significant characteristics were pre-Medicare insurance coverage (health plans before being covered by Medicare). Medicaid and Champus/VA coverage before age 65 years were significantly associated with a higher likelihood of being diagnosed with hypertension after three to four years of Medicare coverage ($p = 0.012$ and 0.048 respectively). On the contrary, individuals having obtained private plans themselves before Medicare coverage were less likely to be diagnosed with hypertension (0.22). Most important of all, total or out-of-pocket health expenditure were not related to a higher likelihood of hypertension incidence in the first three to four years of Medicare coverage.

Characteristics associated with diagnosis of hypertension within first two years of Medicare coverage

In Table 4.5, the total and out-of-pocket spending models for the association between individual characteristics and the probability of being diagnosed with (early-onset) hypertension in the first two years of Medicare coverage are listed. In the total spending model, there were 1455 and 2279 observations in total and out-of-pocket models. These two models were significant with small pseudo R^2 (0.06 and 0.04 for total and out-of-pocket spending respectively).

In both models, only being separated or divorced was associated with less likelihood of being diagnosed with hypertension ($p < 0.05$ in both models). One and three difficulties in mobility were significantly related to higher incidence of hypertension in total spending model ($p < 0.01$), as good health status was associated in the out-of-pocket spending model ($p = 0048$). Other characteristics, such as income, education and health spending, were not significant factors that determined the hypertension incidence in the first two years of Medicare coverage.

Health dimension three: arthritis

The total and OOP health spending models both failed to be significant in predicting the arthritis incidence after three to four years of Medicare coverage ($p = 0.83$ and 0.87 respectively) (Table 4.6). This might be due to the small sample sizes in total and OOP spending models (610 and 1380 respectively). The characteristic associated with higher likelihood of having arthritis included total health spending within three to four years of Medicare coverage ($p < 0.01$ in total spending model). This spending estimate also included part of the spending incurred after individuals were diagnosed with arthritis because this HRS dataset did not distinguish the spending incurred before or after certain conditions. The amount of spending on arthritis treatment is very likely to cause endogeneity problem, but this issue was minimized by excluding individuals with early-onset arthritis. Other characteristics, such as income, education, and insurance coverage, were not found to be significant in arthritis incidence after being coverage by Medicare for three to four years.

Characteristics associated with diagnosis of arthritis within first two years of Medicare coverage

The total and OOP health spending models were not statistically significant ($p =$ 0.45 and 0.62 respectively) (Table 4.7) with small sample sizes in total and OOP spending models (1067 and 1719 respectively). The characteristics that were associated with higher likelihood of being diagnosed with arthritis in the first two years of Medicare coverage included being black ($p < 0.01$ in total spending model) and very good health status ($p = 0.04$ in out-of-pocket spending model). Other characteristics, such as income, education, health spending and insurance coverage, were not found to be significant in arthritis incidence within two years of Medicare coverage.

Health dimension four: self-assessed health status

With more observations for the total and OOP health spending (1,731 and 4,029 respectively), both models were statistically significant ($p <0.001$ for both) in Table 4.8 and the pseudo R^2 (0.22 and 0.21 respectively) indicated that these two models had predictive power better than the disease incidence models (for hypertension and arthritis).

In detail, total and OOP health expenditures were statistically significant and positively associated with getting a worse category of health status ($p<0.001$ for both). In addition, being female was significantly associated with lower likelihood of having worse health status. Health status worse than excellent ($p < 0.01$ for very good, good, fair and poor health status before Medicare coverage), CESD scale ($p < 0.01$ for both models), difficulties in ADL ($p = 0.03$ and 0.02 for two and three difficulties in out-of-pocket spending model), difficulties in mobility ($p < 0.05$ for one to four difficulties), widowhood ($p < 0.01$ in total spending model) were positively associated with a worse health status, as education attainment ($p < 0.01$ in out-of-pocket model) was associated with a higher likelihood of getting a better health category.

The insurance coverage before Medicare had diverse effects on the health status change under Medicare. Medicaid enrollees were more likely to have worse health status ($p < 0.01$ in out-of-pocket spending model), as private coverage purchased from enrollees themselves ($p < 0.05$ in both models) and from their spouse ($p = 0.04$ in total spending model) were associated with better health status.

One of the chronic conditions, hypertension, was associated with worse health status change ($p < 0.01$ in out-of-pocket spending model) after three to four years of Medicare coverage. The other condition, arthritis, was not significant in the health status change.

Finally, there were four cutpoints in these two ordered logit models. The first cutpoint estimated that individuals with any higher values of log-odds would have health status other than excellent status. The second cutpoint indicated that those with any higher values of log-odds would have good, fair or poor health status and those with any lower values would have excellent or very good health status. The third cutpoint suggested that enrollees with any higher values of log-odds would have fair or poor health status and those with any lower values would have excellent or very good or good status. The fourth indicated those with any higher values of log-odds would have poor health status. If the characteristics had positive coefficients in log-odds, these were associated with more likelihood of having a larger sum of log-odds and exceeding a cutpoint of a worse health status after three to four years of Medicare coverage. The interpretation in the arthritis incidence models is similar, except for more cutpoints.

Health dimension five: mental health (CESD scale).

Both models had fewer observations (1,658 and 3,922 for total and out-of-pocket spending models respectively) in Table 4.9 than those in the self-rated health status. However, these models remained statistically significant ($p < 0.001$ for both) with relatively lower levels of pseudo R^2 (0.11 for both).

Other races, health status (except for "very good"), CESD scale and difficulty in mobility (two to four difficulties) were significantly associated with a higher likelihood of having worse mental health status (CESD scale) after three to four years of Medicare coverage in total and out-of-pocket spending models ($p < 0.05$).

In total spending model, health spending was associated with worse mental health ($p = 0.01$). In out-of-pocket spending model, being female ($p = 0.01$) was found to be significantly associated with worse mental health status. However, years of education were significantly associated with better mental health status ($p<0.01$).

Summary of the returns to five dimensions of health

Total or out-of-pocket health spending

The odds ratios for more dollars spent showed its association with higher mortality (OR[44]=1.0044 per thousand dollars in total spending model, $p < 0.01$), worse health status (OR = 1.0056 and 1.0154 per thousand dollars in total and out-of-pocket spending models respectively, $p<0.01$) and more mental problem on the CESD scale (OR = 1.0018 per thousand dollars in total spending model, $p = 0.01$). This suggested that

---

[44] The odds ratios were derived from the logit regression coefficients (log-odds). The full list of these derived odds ratios were not shown.

health spending or more investment on health care could be associated with a higher likelihood of death (total spending), a worse health category (total and OOP spending) and a worse mental health condition (total spending model).

The other way to understand the association between mortality and spending is to estimate the associated spending increase among those deceased in this period. The Gamma GLM computed a marginal effect of $17,669 increase in total spending for death events ($p < 0.01$), along with $3,186 and $997 increase for the diagnosis of hypertension and arthritis ($p = 0.23$ and 0.51 respectively).[45] In out-of-pocket spending model, hypertension, arthritis and death events were associated with $1,632, $785 and $1,974 more spending in the first three to four years of Medicare coverage ($p < 0.01$).

The effects of health plans before and after Medicare coverage on Medicare enrollees

HMO coverage under Medicare was not significantly associated with the change in different dimensions of health except for mortality in OOP spending model (OR = 0.63, $P<0.01$). It was associated with a 37% reduction in the mortality odds in the OOP spending model for Medicare enrollees observed in this period.

Moreover, there were some lasting effects of pre-Medicare insurance plans recorded in HRS datasets. There were 171 Medicaid enrollees before being covered by Medicare and 129 (75.4%) of them became dual eligible, as 126 (3.5%) of the 3583 individuals without Medicaid coverage before age 65 years became dual eligible. Medicaid coverage before Medicare was associated with higher probability of being

[45] This marginal effect was derived from the regression coefficient in total spending model in Table C.5 (left column) in Appendix C. The diagnosis of these two chronic conditions was two of the pre-Medicare characteristics. The spending incurred by these two conditions diagnosed after Medicare coverage was not included.

diagnosed with hypertension (OR = 2.49, $p = 0.01$) and getting worse health status (OR = 1.72, $p < 0.01$) in out-of-pocket spending model.

Before age 65 years, there were 207 individuals being covered by Champus/VA and 131 (63.3%) remained enrolled after being covered by Medicare. This pre-Medicare coverage was associated with higher likelihood of being diagnosed with hypertension in out-of-pocket spending model after three to four years of Medicare coverage (OR = 1.76, $p = 0.048$).

There were 1741 individuals purchasing private plans before covered by Medicare and 1034 (59.4%) of them maintained this private coverage, as 122 (6.4%) of the 1895 individuals without private coverage from themselves purchased private plans after being covered by Medicare. This pre-Medicare private plans purchased by the HRS respondents were associated with better health status change after three to four years of Medicare coverage (OR = 0.73 and 0.86, $p < 0.01$ and 0.04 in total and out-of-pocket spending models respectively).

There were 808 individuals covered by the private plans purchased by their spouses before age 65 years and 417 (51.6%) of them retained this type of coverage as they became qualified for Medicare. However, 124 (4.3%) of those without any private coverage from their spouses (2851) became covered with Medicare and the private plan purchased by their spouses. This pre-Medicare private coverage from their spouses were related to lower likelihood of being diagnosed with hypertension (OR = 0.55, $p = 0.02$ in the out-of-pocket spending model) and having worse health status (OR = 0.75, $p = 0.04$ in total spending model) after three to four years of Medicare coverage.

Socioeconomic status and individual characteristics

Being female was found to be significantly associated with multiple health dimensions. The directions of effects differed in these five dimension. Although the females' negative association with mortality (OR=0.39 and 0.49 in total and OOP spending model respectively, P<0.01) and health status deterioration (OR = 0.72 and 0.85 in total and OOP spending models respectively, p < 0.01) was found, their positive associations with a worse mental health status (OR=1.19 in out-of-pocket spending models, p = 0.01) was also noticed after three to four-year Medicare coverage.

For races and ethnicity, being black was not significantly associated with worse health after three to four years of Medicare coverage. Other races were more likely to have mental health problem on the CESD scale (OR = 3.08 and 1.68 in total and OOP spending models respectively, p < 0.01) Hispanic origin was not significantly associated with these five dimensions.

As regards socioeconomic status, pre-Medicare annual income in nominal dollars was not associated with any protective effects on these five chosen health dimensions. These five health dimensions were observed after income before Medicare was recorded. This ensured that income was not influenced by the short-term health shock that both reduced productivity and increased health spending. Income before Medicare, one of the socioeconomic status indicators, did not have benefits for the health status after three to four years of Medicare coverage.

In contrast, years of education had protective effects on different health dimensions, including education's effects on self-rated health status (OR = 0.95 in OOP spending models, $p < 0.01$) and mental health (OR = 0.96 in OOP spending model, $p < 0.01$). The odds ratios of education were less than one, showing its protective effects and a 5% decrease in the odds of worsening health status (OOP spending model) and a 4% reduction in the odds of mental health deterioration (OOP spending model).

After controlling for the socioeconomic and individual characteristics, widowhood, one of the marital status categories, had negative effects on mortality probability (OR = 1.73 in the total spending model, $p = 0.01$) and health status (OR=1.56 in the total spending model, $p<0.01$). Despite of the significant effects of widowhood, the other categories of marital status, including being divorced or separated and never married, did not have significant associations on these dimensions of health in the first three to four years of Medicare coverage.

Effects of pre-Medicare self-assessed health status on these dimensions

Original health status before individuals obtained Medicare coverage for their age eligibility (pre-Medicare health status) was statistically significant for the deterioration of health or attrition of health capital under Medicare. The self-assessed health status was categorized as excellent, very good, good, fair and poor. The worse health status, relative to excellent health status, that individuals had before Medicare, the more likely they had their health status deteriorating after the first three to four year of Medicare coverage.

Pre-Medicare health status on the probability of mortality

In total spending models, good, fair and poor health statuses (OR=2.81, 3.38, and 4.98, $p<0.01$) were associated higher chance of mortality after three to four years of Medicare coverage, compared to those with excellent health status before Medicare. The magnitude of this effect was proportional to the attrition of pre-Medicare health capital. The coefficient of very good health status before Medicare was not statistically significant.

This effect was also observed in the out-of-pocket spending model, as good, fair and poor health statuses (OR = 3.28, 4.25 and 5.62, $p < 0.01$) were also associated with higher likelihood of mortality after three to four years of Medicare coverage.

Original health status on the incidence of hypertension and arthritis

Self-assessed health status was significantly associated with hypertension incidence (OR=1.68, 2.49 and 2.25 for very good, good and fair health status respectively) in the OOP spending models.

Effects of pre-Medicare health status on health status after three to four years of Medicare coverage

The very good, good, fair and poor health statuses (OR = 3.56, 14.12, 44.68 and 104.31 respectively, $p<0.01$) before Medicare coverage were associated with higher chance of health status deterioration after three to four years of Medicare coverage in total-spending model. Similarly, the very good, good, fair and poor health statuses (OR = 4.06, 14.12, 44.68 and 104.31 respectively, $p<0.01$) before Medicare were associated with health status worsening after three to four years of Medicare coverage in OOP spending model.

Effects of original health status on mental health status (CESD scale) after three to four years of Medicare coverage

There was no significant effect of very good health status on CESD scale in total or OOP spending models. However, good, fair and poor health statuses (OR=1.72, 2.03 and 2.54 respectively in total spending model and 1.69, 2.22 and 2.92 in out-of-pocket spending model, $p<0.01$) were respectively associated with more problem in mental health on the CESD scale after three to four years of Medicare coverage.

Effects of pre-Medicare mental health status (CESD scale)

The CESD scale (mental health rating) was not significant in the association with mortality probability and disease incidence (hypertension and arthritis). One unit increase in the CESD scale was associated with health status deterioration (OR = 1.11 and 1.08 in total and out-of-pocket spending models respectively, $p < 0.01$) and more mental health problem on CESD scale (OR = 1.56 and 1.54 in total and OOP spending models respectively, $p < 0.01$) after being covered by Medicare for three to four years.

Effects of difficulty in mobility before Medicare coverage

The difficulties in mobility did not have significant effect on disease incidence models (for hypertension and arthritis). Difficulties in mobility were associated with mortality (OR= 1.52, 1.99, 1.99 and 4.12 for one, two, three and five difficulties in total spending model and 1.91 for two difficulties in out-of-pocket spending model, $p < 0.05$), compared with those without any difficulty.

For self-assessed health status, one to four difficulties were related to health status decline in total spending (OR = 1.51, 2.04, 1.92 and 2.85 respectively, $p < 0.01$) and out-of-pocket spending model (OR = 1.35, 1.70, 1.43 and 1.95 respectively, $p < 0.05$). Five difficulties in mobility were not significant for this association.

The pre-Medicare difficulties in mobility were associated with mental health problem on the CESD scale in total spending (OR = 1.84, 2.01 and 2.23 for two, three and four difficulties respectively, $p < 0.01$) and out-of-pocket spending model (OR = 1.31, 1.75, 1.60 and 1.96 for one to four difficulties respectively, $p < 0.01$).

Chronic conditions and health returns

Two chronic conditions were added to understand their effects on these five selected health dimensions. Arthritis was not significant in all models, while it was omitted in the arthritis incidence models because individuals with arthritis before Medicare coverage were excluded. Hypertension was significantly associated with higher chance of health status deterioration in out-of-pocket spending model (OR = 1.18, $p < 0.01$), as it remained insignificant in the other models.

Discussion: returns to different dimensions of health

Health is a vague term and is defined differently. To study the returns to health, the outcome measures must be specified. In this chapter, five common and important health indicators were chosen, death, hypertension, arthritis, self-assessed health status and mental health status (CESD scale). Logit or ordered logit models were used to assess whether health spending was associated with better health outcome and which characteristics were associated with the change in these health dimensions.

Total health spending at the first three to four years of Medicare coverage was associated with slightly higher chance of death in total spending model ($p = 0.01$), getting worse health status ($p < 0.01$) and worse mental health on the CESD scale ($p = 0.01$). Although this increased odds was not empirically meaningful, the accumulated spending (for example, $18,000 more among those deceased) could lead to more than 8% increase in the probability of mortality and getting a worse health status in the first three to four years of Medicare coverage.

Out-of-pocket health expenditure in the same period was also associated with a slightly but not empirically meaningful higher likelihood of getting worse health status ($p < 0.01$). The effects of health spending on other dimensions of health were not statistically significant. Because those deceased in the first three to four years of Medicare coverage had shorter time to consume health care than those surviving (Table 4.1), the association between death and total or OOP health spending were likely to be underestimated.

The other important issue was that individual characteristics behaved differently in these five chosen health dimensions. The significant factors for other health dimensions varied. For example, gender and health status had pervasive effects on all health dimensions, except for the insignificant models (hypertension total spending model and arthritis models). The effects of race were quite limited in the first three to four years of Medicare coverage, although the other race was found to have significant association with mental health problem in this chapter and other health conditions in AHRQ (2011c). The role of socioeconomic status was less potent. Education was only significant for better health status (out-of-pocket spending model) and mental health (out-of-pocket spending model), as pre-Medicare income was not significant for these five dimensions. Mental health (CESD scale) and difficulty in mobility were significant for worse health status and a larger CESD score. These findings demonstrate that health was multi-dimensional and individual characteristics varied in their influence on these five health dimensions.

Limitations

The sample size in this chapter was again limited by the available observations with spending information (1,752 and 4,032 for total and OOP spending in mortality regression) and many observations were excluded because some of the information on individual characteristics (difficulty in ADL, IADL and mobility, CESD scale) was not collected in the first two years of HRS. The exclusion of existing cases of hypertension and arthritis further reduced the sample size to less than 900 and 2000 for total and out-of-pocket spending models. This was the main reason why the logit models for arthritis (total and out-of-pocket spending) and hypertension (total spending only) were not statistically significant.

The other challenge was to establish a clear causal relationship between health expenditure and returns to health. This study was designed to use health spending as the main treatment that these observations received after being covered by Medicare. By using logit or ordered logit model, the original health status (before Medicare) and other individual characteristics were controlled and the temporal relationship between spending and death was clear. A general statement was that each thousand-dollar increase in total health expenditure was associated with a small increase in the likelihood of death in the first three to four years of Medicare coverage (also 65 to 68 years of age of Medicare enrollees who were qualified for their age). The temporality from pre-Medicare health status to death was clear, but the interaction between death and spending was not.

Although spending on health care occurred before death and other health events, other causal links and third mechanisms that influence how higher spending lead to death and other conditions analyzed in this study were not clear. One clear message in this chapter was the association between spending levels and the occurrence of health conditions.

Chapter 5: Discussion of findings

Research motivation

This dissertation aims to investigate which demand-side factors were associated with the high rates of growth in Medicare spending, what effects the HMO coverage had on spending and how much returns to health the Medicare enrollees actually gained over time. The first question is important because the high rate of growth in Medicare spending creates financial and budget pressures on the nation and may reversely makes this system unsustainable. Other researchers already studied the supply-side factors or took mixed approaches to understand the growth of Medicare spending growth (Farrell et al. 2008; Zuckerman and McFeeters 2006; Ginsburg 2008). This dissertation can supplement with information on the patients' characteristics and associated spending over time.

However, the validity of the first research question was threatened by the renowned biased selection to HMOs (Luft 1981; Miller and Luft 1994). The threat of this selection issue should be greatly reduced through appropriate statistical methods, such as instrumental variables and propensity score matching (Stukel, Fisher, Wennberg, Alter, Gottlieb and Vermeulen 2007). Propensity score matching was used in this chapter because it did not require variables with the statistical properties necessary for instrumental variables. With the help of propensity score matching, the influence of biased selection into Medicare Advantage/Part C should have been minimized and the estimates of the average spending differences relative to traditional Medicare should be less vulnerable for this issue.

Moreover, the health returns from the spending on Medicare coverage were not certain. Although other researches reviewed the health returns from health expenditures in history and found positive returns (Cutler, Rosen and Vijin 2006; Cutler and Richardson 1999), it was not sure whether the high rate of growth in Medicare was associated with better health outcomes among enrollees. Through using longitudinal datasets and controlling for individual characteristics and health status before Medicare, the association between health spending and its returns to health could be better understood for policy discussion.

Analysis and findings

In Chapter 2, the well-designed and implemented survey that focused on the use of health care and expenditures, the Medical Expenditure Panel Survey (MEPS), is the database of choice to investigate the relationship between demand-side factors and Medicare spending. After fitting multiple regression models, one-part Poisson GLM (log link) had the least number of model-fit problems and was used to predict individual Medicare spending from 1996 to 2008 (or 2000 to 2008 in the extended model). The multiplication of adjusted individual spending and population size better estimated the trend of spending growth in different groups. In Table 2.3, the leading groups with adjusted growth rates higher than overall Medicare spending growth from 1996 to 2008 (5.8% annually) includes races other than the whites and blacks, Hispanics, high-income groups, residents in the West, and individuals with very good health status. These groups are different from those found in the descriptive analysis in Table 2.1. This analysis helps to formulate Medicare policies that focus on the leading groups of Medicare spending. The analysis also provided policy implications in different dimensions. With them, the public can think about which groups associated with high growth should be the targets of cost-containment policies and what consequences may occur if implemented.

In Chapter 3, the aim is to investigate the biased selection issue that threatened the validity of the analysis with cross-sectional datasets. In the literature, biased selection to HMO coverage existed in various populations and Medicare enrollees. To deal with this issue, advanced statistical tools (propensity score matching in this chapter) and longitudinal datasets (Health and Retirement Study, HRS, in this chapter) are necessary to understand the impact of HMO coverage on the levels of spending observed in the cross-sectional datasets. By modeling binominal outcomes (selection into Medicare Advantage/Part C or traditional Medicare), the estimates from propensity score matching suggest an insignificant effect on total spending and a cost saving effect on out-of-pocket spending ($1,411.5) from 1992 to 2008, as the regression model finds a similar estimate, $1,772.5 saving in out-of-pocket spending. This analysis helps to confirm that there are individual characteristics associated with the selection into Medicare Advantage/Part C and the saving effect on out-of-pocket spending exists.

In Chapter 4, this dissertation estimated the health returns from the rapidly growing Medicare spending with HRS datasets. Controlling for individual characteristics before Medicare coverage, this chapter reviewed five dimensions of health, mortality, hypertension incidence, arthritis incidence, self-rated health status and mental health (Center for Epidemiologic Studies Depression, CESD, scale). The logit or ordered logit models shows that Medicare spending was associated with higher mortality (total spending in Table 4.3), worse health status (total and out-of-pocket spending in Table 4.8) and more problems in mental health (total spending in Table 4.9). Although a causal link could not be fully confirmed in these seemingly negative returns on Medicare spending, the level of current Medicare spending may not benefit population health of Medicare enrollees. However, more research on how and why Medicare spending was not associated with better health status would be necessary.

Limitations

Sample size

Although MEPS data sets contained a rich set of variables and ample observations in each year, the research method used in this dissertation might require more than the data could provide. The limitation in sample size required some categories to be merged with each other. For races, the racial groups other than the white and black were merged as the other races. Marital status was also simplified into four categories, married, widowed, divorced and others. Two of the mental health categories with the least number of observations, fair and poor, were combined. However, most of the interaction terms between individual characteristics and years in the logit model, one- and two-part expenditure models remained insignificant.

This issue also prevailed in HRS datasets. Although it is a longitudinal study from 1992, selecting individuals with information before and after Medicare coverage largely limited the number of eligible observations. The missing data of each variable in Chapter 3 and 4 reduced the sample size to less than two thousand Medicare enrollees in total spending model. This lack of sufficient number of observations lead to insignificant results in all matching algorithms for total health spending. In Chapter 4, the logit models for hypertension (total spending model) and arthritis (both total and OOP spending models) were not significant at all (Table 4.4 to 4.7). This sample size issue could be partially relieved in the future, as these surveys, MEPS and HRS, were still gathering information and following up individuals.

Different data sets

To take advantage of the characteristics of different data sets, the trade-off was losing the consistency of information between sources. In MEPS and HRS, they had different ways of categorizing health plans, marital status, mental health status, ADL, IADL, and other mobility limitation. This limited the comparability between Chapter 2 (using MEPS datasets) and the other chapters (using HRS datasets).

Moreover, the data precision from MEPS and HRS also differed. The total health expenditure reporting in HRS was discontinued in 2002 for the lack of precision and the frequent use of imputation. Because total health spending served as one of the outcomes in Chapter 3, the issue became more acute and the out-of-pocket spending was considered as another proxy for the intensity of health care use.

Poliy implications

Besides the potential targets of cost-saving policies identified in Chapter 2, there were other policy implications for the results in Chapter 3 and 4. In these two chapters, the policy implication from the investigation with HRS was to consider the impact of different types of Medicare enrollment on Medicare spending after controlling for purposeful selection into types of Medicare enrollment. Using the longitudinal HRS data provided a distinct advantage over the shorter time frame available in the MEPS. The contrast between Chapter 2 and Chapter 4 highlighted this issue that some factors might reduce spending in the short run but increase the expenditure in later years. Some policies, such as changes in cost sharing and imposing target intervention groups, could be formed to address this issue.

In Chapter 4, health returns, such as the decrease in disease incidence probability and better control of chronic conditions could be assessed based on the individual health spending incurred. In contrast to the first two chapters, the results in Chapter 4 did not yield evidence suggesting that health spending helped to reduce likelihood of being afflicted with some health conditions or provided with payoff in improved health. On this basis, policy makers might want to encourage greater investment or efforts targeted to preventing such illnesses while reducing spending on conditions for which there was little evidence of preferable outcomes. As regards the latter point, as more total spending was not found to reduce the likelihood of obtaining serious health conditions, policymakers could respond by considering lower-cost interventions or trying to constrain spending in other ways. These findings from these three chapters could contribute to the points that targeted at current policy debate on the contribution of Medicare to the health of elderly Americans and the utility of current spending levels.

Externality of private insurance or uninsurance to Medicare health plans

The inequality in the quality of care before Medicare was one of the determinants to population health (AHRQ 2011c), along with other social, genetic and behavioral determinants mentioned in WHO (2012). Medicare relieves part of the quality of care problem by providing universal coverage, but mostly for the elderly. As this study showed that there were high-spending characteristics in Medicare enrollees, Medicare passively cover these high-spending health status and functional limitations that occurred before individuals were enrolled.

Levy and Meltzer (2004) reviewed different types of health studies and commented that quasi-experiments and the RAND Health Insurance Experiment (HIE) consistently yielded the evidence of the benefits of health insurance on health. This implied that it might be economically beneficial to provide health coverage to prevent high-spending characteristics from happening before individuals being covered by Medicare. If these high-spending characteristics were not prevented in advance, Medicare had to deal with the consequences of poor health and the externality of inferior health status resulting from insufficient pre-Medicare health coverage or inadequate health behaviors.

Moreover, the research on the pent-up spending after being enrolled in Medicare showed that individuals with insufficient health coverage before Medicare coverage might delay their care, especially physician care, (Chen et al. 2004) and incur more Medicare spending or delayed conditions (Schimmel 2006). The evidence suggested that the early intervention for individuals who were waiting for Medicare eligibility with insufficient coverage might improve their health status and reduce the likelihood of major conditions that might be averted with adequate treatment. The economical efficiency could be improved if Medicare took action to reduce the incidence of high-spending characteristics among those whose ages were approaching the age eligibility of Medicare. The implication is that Medicare could provide these individuals or heath care providers with incentives that aim to promote health and reduce conditions before Medicare coverage. This type of Medicare policies would be particularly useful if preventable high-spending characteristics were targeted, such as chronic conditions in Chapter 2 and functional limitations in Chapter 3 and 4.

Biased selection and adjustment for the spending estimation

In Table 3.3, there were individual characteristics that increased the likelihood of joining HMOs, including being black, Hispanic origin, region (West), and pre-Medicare private coverage. The other variables that decreased the probability of enrolling in HMOs included regions (Midwest and South), and enrolling in Champu/VA. This result was not exactly the same as other similar research (Olin and Lavis 1998). As the likelihood of joining HMOs was associated with these factors, they increased or decreased the chance of adopting the HMO cost saving effects.

This also suggested that the HMO selection might strengthen the cost saving effect of living in the West (showed in Table B.7). The cross-sectional observation based on MEPS estimates in Chapter 2 should be interpreted with these adjustments.

Returns from health care spending

After evaluating returns to multiple health dimensions, there seemed to be enough evidence to claim that the returns from the current level of health care expedniture did not compensate for its costs. In Table 4.3 to 4.9, the intensity of health care in terms of dollar values were associated with higher mortality rate (total spending model), worse health statuses (total- and OOP-spending models) and more mental problem (OOP spending model) for similar individuals. This piece of evidence was not the first discovery to indicate the negative rate of returns for the investment on health care. There was evidence showing that current medical care was on the flat of the curve with little marginal benefits for health (Schoder and Zweifel 2011).

However, there were limitations to this study to make a formal conclusion that the intensity of health care was related to diminishing health returns. First, the types of health care services were not specified in the analysis and the returns from types of health services varied (Cutler and Rosen 2006). The differences of returns to health suggested that one single dimension of health care intensity measure (total and out-of-pocket spending in this study) might not be enough to draw a conclusion for the use of health care resources. Second, the diversity of health care also limited the strength of this finding. As medical care could be categorized as preventive, curative and plliative treatment (USAID 2009), the health returns were expected to differ for a variety of patients and some patients receiving palliative care were commonly treated for needs other than the demand in medical care. It would be arbitary to assume that spending in health care aimed to get direct returns from it. The unobserved indirect benefits in certain types of medical care, such a palliative care for the terminal cancer patients, might matter more for the elderly and their family.

Third, the strong link between health care and medical investment might be due to third factors that contributed an uncontrolled upward bias. An ideal study design would be to choose two groups of similar individuals to be randomly assigned to treatment or control groups. In this study, regression models were used to adjust for other characteristics and make individuals in different spending intensities comparable. However, endogeneity or the effect from the unobservables was not completely removed. This chapter used temporal relationship between pre-Medicare characteristics and health dimensions to ensure the causes, higher intensity of health care, occurrd before outcomes. this was only one of the essential components of causation (Rothman and Greenland 2005; Renton 1994). Unable to perform this random assignment and establish a biological pathway, this study was not free from biases or unobservable endogeneity and could merely build an association between health care spending and worse health outcomes, including a higher likelihood of death (total-health-expenditure mode) or a worse category of health status.

Cost-effectiveness analysis: the value worth the spending?

The negative association between health spending and health outcome (mortality in total spending model and health status in total and OOP spending models) suggested that there was a need to analyze the cost-effectiveness in medical care. After establishing a value rating mechanism, an essential step would be to make patients aware and able to pay for the value.

The insurers could assess the effectiveness of each treatment and consider what to provide to their enrollees. There were many ways to assess and judge clinical values of comparable treatments, for example, Health Technology Assessment mentioned in Sheldon (1992). The physicians should be able to obtain and provide high-priority services to their patients so that patients could pay for the value they get from health care [for example, value-based cost sharing in Kleinke (2004)]. However, one problem was that this process required an integrated health care systems and well-coordinated medical service delivery to help the informed patients to pay for the values.

The other issue was that we had to define the length of observation and the dimensions of health that we would like to use for the cost-effectiveness study. In this dissertation, the length of observation was determined based on the data availability and the concern on sample attrition. Besides missing data and loss of follow-up, the datasets from HRS also suffered from major changes in the questionnaires and many of the observations were excluded for these changes. To better formulate an analysis of the relationship between health care intensity and outcomes, compromises between services of evaluation and length of observation should be made to construct analysis that helps to understand the detailed mechanism from health capital investment to outcomes.

As more issues regarding Medicare enrollees' spending patterns were raised in this discussion, this study simply served as a beginning of solving the problems in health care systems and a footnote of the inquiry toward Medicare policies.

Appendix

Appendix A: Medicare enrollees and changes in individual spending

Population profile and individual spending

Change in individual health spending among Medicare enrollees from 1996 to 2008

The mean spending in different groups from 1996 to 2008

In Table A.1, the amount of mean individual total health spending among Medicare enrollees age 65 years and over was listed in different categories. As the mean total health spending grew from $5,423 in 1996 to $9,303 in 2008 (4.5% annually), the average ratio of out-of-pocket (OOP) payment to total spending remained low and decreased from 0.34 in 1996 to 0.24 in 2008.[46] The growth rate of OOP payments was smaller than that of total health payments.

In the same table, there were high-spending groups identified. Some subgroups had lower mean values in 1996, especially those age 65-74 years, other races, those with less than eight years of education, those with middle incomes, those with residence in the South and the West, those married, those with excellent and very good health status, and those with excellent and very good mental status than the mean Medicare spending ($5,423 in 1996). Those of functional or cognitive limitations, such ADL and IADL limitations, had average spending twice than the Medicare average in 1996.

[46] The ratio was the mean value of individual ratios of out-of-pocket spending to total health spending among the Medicare enrollees age 65 years and over.

Growth of mean Medicare spending in different groups from 1996 to 2008

There were different patterns of growth, 71.5% and 64% increase (4.50% and 4.12% annually) from 1996 to 2008 for total and OOP spending respectively. The leading high-growth subgroups included those with ages 65-74 years, other races, those with less than eight years of education, those with middle incomes, those living in the West, and those married. However, those with chronic conditions, except for hypertension patients, did not have growth rates higher than the average Medicare patients, 58.3% increase (5.6% annually) from 2000 to 2008. The current smokers did not have higher spending level in 1996 or higher growth rate from 2000 to 2008.

Comparison between individual and aggregate Medicare spending growth from 1996 to 2008

The comparison between aggregate (Table 2.1) and individual Medicare spending (Table A.1) revealed some important distinctions. First, the growth of total Medicare expenditure, 200% (5.8% annually) from 1996 to 2008, was higher than the per-capita growth (4.5% annually). Second, the ratio of OOP payment to the total aggregate health expenditure[47] did not drop and remained below 20% from 1996 to 2008. Third, the relatively lower growth for some groups, such as stroke patients, did not have a large weight on the overall Medicare spending. Instead, the health spending from those of higher income and more education attainment contributed to a larger share of total health expenditure and its growth.

[47] Ratio=(aggregate out-of-pocket spending)/(aggregate total health spending), not the mean ratios among Medicare enrollees.

Health spending distribution among Medicare enrollees

From 1996 to 2008, the individual health spending distribution was illustrated in Figure A.1 (a). The distribution of total health spending was positively skewed and a large number of people were included in the category of spending less than one hundred dollars (nominal values) from 1996 to 2008. The mean values of the health spending in 1996 ($5,423) differed from the amount in 2008 ($9,303). In Figure A.1 (b), there were more Medicare enrollees had health spending less than one hundred in 1996 and more individuals spent less than $1,800 (nominal dollars) than the distribution in 2008. The skewness in 1996 and 2008 differed and the range of spending were from zero to $218,700 in 1996 and to $189,900 in 2008 respectively.

In Figure A.2, the number of population was log-transformed[48] except for those with zero spending. The range of transformed Medicare expenditure was smaller than untransformed range. The shape of this distribution became less skewed in both tails. However, a formal test of heteroscedasticity was not feasible for data with complex survey design in MEPS and the skewness could not be quantified.

Changes in health status and chronic conditions among the Medicare enrollees age 65 years and over

As population aging was noticed, the prevalence of chronic conditions decreased for less than one percentage point (asthma) or increased for one to twenty percentage points (the other chronic conditions in Table 2.2). Hypertension and joint pain were the most prevalent and more than half of the aged Medicare enrollees were found to have

[48] The logarithm base was *e* (2.718).

these two conditions. The spending growth rates of the mean Medicare expenditure in the hypertension patients (59.8% from 2000 to 2008, 5.9% annually) were slightly higher than those in all Medicare enrollees (58.3%, 5.6% annually) or in individuals with joint pain (56.5%, 5.6% annually).

The selected preventive health behavior, smoking, showed a decreasing share of smokers. The growth rates of mean Medicare spending (70.2% for smoking, 6.7% annually) were higher than that in average Medicare spending (5.6% annually).

Appendix B: model selection process for Chapter 2

Data management

Data linkage

The version of linkage file used to combine the annual HC files is called *h036b08* (AHRQ 2010) in Figure B.1. It took several steps to integrate these annual files into a multi-year collection of individual spending pattern from nationally representative population for specific year. First, the relevant variables had to be selected from a complete list of observed variables in the annual files. Then, these variables had to be assigned consistently with new variable names. Because most variables were named with year-specific labels, the differences in year-to-year variable name lead to difficulties in formulating a correct regression. As noted in the MEPS document, *h036b08* (AHRQ 2010), the observations in each year were preserved without year-specific tags thereby permitting empirical analyses. Second, the pooled observations from these annual household component (HC) files were appended to each other.

The final step was to match renewed sampling units and strata in the linkage file. Two key variables, *dupersid* and *panel*, were used to perform many-to-many dataset merging. The merging process introduced new structures of sampling units and strata to each participant in order to formulate a dataset with historical information.

MEPS questionnaire change and variable selection

Although these annual surveys were conducted by similar methods, the questionnaires were revised with several major changes. First, the categories in races and ethnicity were changed. The observed racial variable, *racex*, was defined differently for the surveys conducted before and after 2002. The different racial indicators were harmonized to perform a uniform and consistent statistical analysis.

Second, the self administered questionnaires (SAQ) and parent administered questionnaires (PAQ) were introduced to MEPS in 2000 (AHRQ 2003). The use of SAQ and PAQ expanded the scope of this survey and the participants could provide information about their health conditions (for example, joint pain[49] and hypertension in this study), diseases (for example, diabetes, myocardial infarction, angina, stroke and other heart disease), health behaviors (for example, smoking status), health care quality and others (AHRQ 2003).

Third, there was a major discrepancy between the definition and identification of health insurance coverage for the surveys before 1999 and after 2000. In 1996, the users of health maintenance organization (HMO) and managed care (MC) were indentified and each type of health care delivery was further identified as public or private. From 1997 to 1999, information regarding the provider of health care was not revealed in the household component (HC) datasets. The surveys after 2000 identified HMO and MC access from Medicaid and private plans. Information regarding prescription drug coverage was also included in the surveys after 2000 to enable researchers to assess the impact of prescription drug coverage on drug expenditures. Therefore, the information on HMO or MC for the HC datasets before 2000 needs to be introduced from other MEPS datasets.

---

[49] To ask whether the survey respondent had joint pain was not intended to reveal the diagnosis of arthritis (AHRQ 2011b). The diagnosis of arthritis by health professionals was introduced to the SAQ section in 2001, one year later than the introduction of joint pain. Hence, joint pain was used to keep the 2000 MEPS dataset in this study, rather than arthritis.

Finally, other minor changes included the ceasation of collecting the information on characteristics , such as alternative care, coverage source, percentage paid by health plans, and dental checks, in the MEPS HC surveys conducted after 2000. This lack of information limited the range of variables used in this dissertation.

Considerations regarding the complex survey design of MEPS

To conduct a nationally representative survey with optimal sample size, surveys with multiple strata could be adopted and MEPS took the same approach. However, the interpretation of such survey results required the researchers to use specialized statistical procedures to adjust for the complex survey design. While STATA 11 (STATA Corp, College Station, TX) provides the capacity to handle the complex survey design[50], researchers could simply indentify the variables that denoted strata (*stra9608*) and sampling units (*psu9608*) in MEPS datasets (AHRQ 2011b). The application of person-level survey weights for observations in each survey year was required to obtain nationally representative estimates of the individual behaviors and outcomes of interest.

Model selection process

Candidate expenditure regression models

There were eleven expenditure models to be compared in Chapter 2. There were two major types of models: ordinary least square (OLS) and generalized linear models (GLM). For OLS models, the log transformation and the use of smearing factors were important comparisons for the other one- and two-part models. For GLM, the log link

[50] To program the survey design, a statement was made to identify the personal weight and primary sampling units. Then the *svy* syntax was added in front of the coding that aimed to adjust for the survey design.

functions were used in all one- and two-part models. However, the family of GLM should be chosen based on the variance function between variances and predicted means of the dependent variables (variance constant to, proportional to, or proportional to the square of means).

Determination of variance structure - Modified Park test for GLM

Homoskedasticity is an important assumption for regression analysis. The regression estimates could be biased if this heterogeneity issue was not adjusted (Manning 1998). However, the complex survey design in MEPS and two-part expenditure models prevented researchers from conducting a straightforward heteroscedasticity test for OLS models (White test) mentioned in Deb, Manning et al. (2011). To obtain a preliminary estimation of the residual distribution, modified Park test was still applicable to quantify the heterogeneity issue and find an optimal family for GLM.

Modified Park test was to regress the square of residuals (variance) on the predicted values of Gamma GLM regression (Deb, Manning et al. 2011; Manning and Mullahy 2001)[51]. For the one-part model, spending information from all Medicare beneficiaries aged 65 years and over was tested with modified Park tests. Those with any spending were examined with modified Park tests for two-part expenditure model. The regression coefficient of the modified Park test reflected the relation between the residual squared (variance) and the means. If the regression coefficient was zero, one or two, the

---

[51] The test followed the equation: $var(y|x) = \alpha[E(y|x)]^{\lambda}$ (Manning and Mullahy 2001). The value of λ determined the relationship between the variance and the mean predicted by different GLM. The detailed STATA coding sample could be found in Deb (2011). The value of λ was not necessarily integers and could be negative or positive.

square of residuals was constant to, proportional to the predicted values of the Gamma GLM, or proportional to the value squared (Buntin and Zaslavsky 2004; Matsaganis, Mitrakos, et al. 2008). In Table B.1, the regression coefficient for the one-part GLM was 1.77 and the adjusted Wald test indicated smaller *f* statistics if the regression coefficient assumed to be 2. This suggested Gamma GLM might fit the variance structure of the actual spending. In the two-part GLM, the regression coefficient was close to zero and suggested the Gaussian GLM might be a more appropriate GLM family in this context.

Comparison between actual health spending and predictions from models

Average annual per-capita cost (AAPCC) prediction

In Buntin and Zalavsky (2004), the average annual per-capita cost (AAPCC) was first estimated by the region[52], gender, and age to get an average value of health spending in the region where individuals resided. The AAPCC estimation in this chapter served as a comparison that represented the mean health spending in local environments and removed individual variation in health spending. The survey year was also added as dummy variables to estimate AAPCC of each year in this chapter.

In Table B.2, the actual health spending and the expenditure predictions from each regression model were summarized. The health expenditure was summarized based on individual characteristics, including health status, activities of daily living (ADL)[53],

---

[52] The regional variables used in McBride, Penrod, et al. (1997) and Buntin and Zaslavsky (2004) were "counties" where individuals resided. The regional variables in MEPS were "regions" in the US. Because of the differences in how geographic locations were defined and the datasets used in different studies, the AAPCC estimates in this dissertation might not be fully comparable to other studies.

[53] For those who were identified to need ADL help or supervision due to an impairment or physical or mental health problem, they were then asked "whether they were expected to need help or supervision with these activities for at least three more months" (AHRQ 2011b).

instrumental activities of daily living (IADL)[54], and other health related indicators. Among these models, one-part Gaussian GLM failed to predict the expenditure for all observations or specific groups. Most of the models performed quite well and were close to the actual mean values. However, one- and two-part log-transformed OLS models seemed to underestimate health expenditure in all groups.

Moreover, these models predicted expenditure in each year (not shown) and the finding was similar to the finding in Table B.2. The one-part Gaussian GLM failed to fit the actual spending trend across years and the one- or two-part transformed OLS models consistently underestimated the health expenditure.

Error structure of prediction models

In Table B.3, the mean standard errors (MSE) and mean absolute prediction errors (MAPE) were listed for each model. Except for one-part Gaussian GLM, the values of different models in these two columns were quite similar. One-part transformed OLS model had the prediction with the lowest mean absolute prediction error (MAPE) and one-part Gamma GLM had the lowest mean square error (MSE). However, this table was not a formal model fit test and there was no test for the differences between MAPE or MSE under complex survey design in MEPS.

Model fitting with regression coefficient and pseudo R^2

Conventionally, R^2 is a convenient summary for OLS models and gives a general understanding how much of the variation in the dependent variable in the model can be explained by the observed variation in independent variables. However, two-part models

---

[54] The respondent was asked “if they received help or supervision with IADLs such as using the telephone, paying bills, taking medications, preparing light meals, doing laundry, or going shopping” (AHRQ 2011b).

in this chapter limited the use of R^2 (Buntin and Zaslavsky 2004; Matsaganis, Mitrakos, et al. 2008). The complex survey design in MEPS imposed the same limitation and did not support the calculation of R^2.[55]

To obtain a summary statistics for each model, the predicted health spending from each model was used to fit the actual spending and to produce the pseudo R^2 that indicated how much of the variation in the predicted model could be explained by the variation in the actual spending data. In Table B.4, the total health expenditure of each observation was regressed based on the predicted values from each model. These OLS regression models[56] with the survey design adjustment produced regression coefficients, standard errors of regression coefficients, *p* values, and R^2 that indicated how well the predicted values matched the actual spending. In Table B.4, the predicted values from two-part Gaussian GLM had the highest R^2, 0.17, but its 95% confidence interval of regression coefficient did not contain one, which meant a perfect prediction of the actual spending.

There were five models that had the 95% confidence interval of regression coefficient including one: one- and two-part non-transformed OLS model, one-part Poisson GLM, two-part Gaussian GLM and two-part Poisson GLM. The R^2 of these four models were larger than 0.15. All these models seemed to perform with a certain degree of precision except for one-part Gaussian GLM that had very low R^2, less than 0.001, and its regression coefficient, less than 0.01, did not suggest good prediction.

---

[55]Although Zheng and Agresti (2000) proposed an alternative formula to calculate a summary value (similar to R^2 in OLS) for GLM, their formula was not applicable for two-part models in this chapter.

[56] This OLS model was specified as
$y = \beta_0 + \beta_1\hat{y} + \epsilon;\ y: health\ spending;\ \hat{y}: predicted\ health\ spending.$

Model fitting based on the average annual per-capita cost (AAPCC) adjustment

This section examined the health spending based on AAPCC amount that represented the health spending adjusted for local contexts. In Figure B.2, the total health expenditure was adjusted by taking the ratio to individual AAPCC amount. The ratio of spending to AAPCC could be seen as the individual variation in health spending after the effects of region, age and gender were considered. (Buntin and Zaslavsky 2004) Moreover, the AAPCC values in this chapter were also adjusted for the year when the Medicare enrollees were observed.

The ratio of predicted spending to AAPCC was calculated and the ratios of deciles, from 10th percentile to 100th percentile, were illustrated in Figure B.2 (similar to the procedures in Buntin and Zaslavsky (2004) and Matsaganis, Mitrakos, et al. (2008)). At each decile of the ratio, the actual and predicted spending ratios were compared. From the first to eighth deciles, the actual and predicted spending did not vary greatly. However, the ratio of actual spending increased greatly after ninth and tenth deciles and two-part Gaussian GLM seemed to pick up some of the high-spending individuals. Other models behaved similarly, except one-part Gaussian GLM that failed to match other curves.

However, as we changed the scale in Figure B.2 (b), the curve of actual expenditure was not adjacent to other curves until 70th percentile of the ratio of health expenditure to AAPCC. Except for one- and two-part transformed OLS model, one-part Gaussian GLM and AAPCC estimation, all estimators produced similar results and the estimation curves seemed to match well in the figure.

In Figure B.3, the ratio of predicted health expenditure to AAPCC was plotted against the ratio of total spending to AAPCC in order to assess how these models predicted the spending pattern. In this figure, deciles of the ratios of predicted values to AAPCC were plotted against the ratios of total health expenditure to AAPCC, but one-part Gaussian GLM was eliminated because of its lack of precision in previous figures. If the predicted values perfectly estimated the actual values, the line would match the 45-degree line and the value in each decile would form a straight line against the ratios of total health expenditure to AAPCC.

The first comment to be made from Figure B.3 would be that these models lined up with the 45-degree line and their precision to predict the high-spending group differed. The points of tenth decile in each model extended to different ratios. Notably, one- and two-part Poisson GLM could estimate some of the high-ratio individuals, as two-part Gaussian GLM had the highest amount in the 10th decile in Figure B.2. Second, these models performed similarly while predicting the individuals of lower AAPCC ratio. In Figure B.3, most of the model predictions were within the AAPCC ratio range less than ten (Y axis). Third, there was negative prediction from one- and two-part OLS model and this problem was not found in other models. Finally, there were models that provided twisting prediction curves in Figure B.3 (b), which might suggest that these models did not capture some important expenditure characteristics.

Model check with formal model fit tests

Comparison in goodness of fit

There were several ways to compare the goodness of fit across models, including the modified Hosmer-Lemeshow test[57], correlation test[58], Pregibon's Link test[59] (Manning, Busa and Mullahy 2005; Deb, Glick, Doshi and Polsky 2004; Basu and Manning 2009; Jones 2010) and Ramsey's RESET test[60] (Jones 2010; Ramsey 1996; Deb, Manning and Norton 2011) listed in Table B.5. Correlation test examined the "systemic bias in fit on raw scale" (Glick, Doshi and Polsky 2004). Modified Hosmer-Lemeshow test, Pregibon's Link test and Ramsey's RESET test examined the model fit (Deb, Manning and Norton 2011).

---

57 This tested was performed after predictions available from all models. The predicted values were sorted and separated into 10 groups. The test was to regress the residual on the mean values of these 10 groups. See Deb, Manning, et al. (2011) for detail. Because modified Hosmer-Lemeshow test required sorting percentiles and STATA did not allow labeling deciles of weighted data under complex survey design, the deciles were labeled based on the unweighted data without survey design adjustment. However, these deciles were adjusted for complex survey design in the regression function. This differed from the studies using database without the necessity of design effect adjustment.

[58] This correlation test was to assess the correlation between the residuals and the predicted values in each model. Because of the complex survey design, this correlation was tested by regressing the residuals on the predicted values or by regressing the predicted values on the residuals. Either one of these two regressions had a large R^2 was used and its p value less than 0.05 was seen as problematic in "systemic bias in fit on raw scale" (Glick and Doshi 2007). See Glick and Doshi (2007) for detail in STATA programming and Sribney (2005) for detailed methodology to conduct the correlation test for survey-design datasets.

[59] The functional form of Link test was "$y = \delta_0 + \delta_1(x\hat{\beta}) + \delta_2(x\hat{\beta})^2 + v$" (Deb, Manning, et al. 2011), whereas the *y* denoted the dependent variable, total annual health expenditure in this study, and $(x\hat{\beta})$ denoted the predicted values in different models. The goal was to use adjusted Wald test to assess the *p* value of the null hypothesis, "$\hat{\delta}_2 = 0$". If the p value was greater than 0.05, the Link test suggested that there was no problem in model fit. Although STATA provided the syntax "*linktest*" to directly perform this model fit test, the survey design in MEPS datasets did not allow using this syntax. See Deb, Manning, et al. (2011) for detail.

[60] The functional form of Ramsey's test was "$y = \delta_0 + \delta_1(x\hat{\beta}) + \delta_2(x\hat{\beta})^2 + \delta_3(x\hat{\beta})^3 + \delta_4(x\hat{\beta})^4 + v$", whereas the y denoted the dependent variable, total annual health expenditure in this study, and $(x\hat{\beta})$ denoted the predicted values in different models. The goal was to use adjusted Wald test to assess the *p* value of the null hypothesis, "$\hat{\delta}_2 = \hat{\delta}_3 = \hat{\delta}_4 = 0$". If the p value was greater than 0.05, the Ramsey's RESET test suggested that there was no problem in model fit. Although STATA provided the syntax, "*ovtest*" to conduct this test, the adjustment for survey design did not support this syntax directly. See Deb, Manning, et al. (2011) for detail.

The *p* values of these tests suggested whether the predicted values of each model yielded similar variation as the observed ones. If the variation of the predicted values in each model differed from the actual variation, the p values less than 0.05 indicated that there were problems in model fit and rejected the null hypothesis that the observed values did not differ significantly from the predicted values. In Table B.5, adjusted Wald tests could not be performed in most OLS models and did not provide the *f* statistics for Pregibon's Link test. Among all models, one- and two-part Poisson GLM seemed to perform the best and had p value greater than 0.05 in two tests, modified Hosmer-Lemeshow test and Link test. However, correlation tests suggested systemic bias problems in all models (Table B.5).

Copas test for model over fitting

Over fitting occurred when a model was well tailored for a specific dataset but the predictive power for other similar data decreased (Deb, Manning et al. 2011). Copas test was recommended for this issue (Copas 1983; Blough et al. 1999; Deb, Manning et al. 2011)[61]. The results of Copas test were not conclusive because the majority of the models did not achieve convergence or become problem-free in Copas tests.[62] Even if convergence was achieved, there was no insignificant test results in the models that

[61] Copas test worked by first predicting the values of half of the observations based on a chosen model, then generating the estimates for these observations, and regressing these predicted values on the values of the other half observations (Deb, Manning, et al. 2011). If the adjusted Wald test for the null hypothesis that regression coefficient equaled to one showed significant result ($p<0.05$), this showed that might be a problem in overfitting. After repeating this procedure for 1000 times, a summary measure was to show how many times the adjusted Wald test was significant per 1000 loops. These procedures were modified from those used in Copas (1983) and Blough et al. (1999).

[62] A model became problematic when there was no single over-fitting test with p value large than 0.05 after executing this test for 1000 times. The problematic models included one- and two-part OLS model, one- and two-part transformed OLS models, one-part GLM (Poisson, Gamma, and Gaussian) and two-part Gaussian GLM. The convergence was not achieved in the other models, two-part transformed OLS model, and two-part GLM (Poisson and Gamma family).

achieved convergence, including one- and two-part OLS models, one-part transformed OLS model and two-part transformed OLS model with one smearing factors. The use of Copas test for this dissertation was not conclusive.

Conclusion for model comparison

Based on the descriptive analysis in this chapter, the one-part Gaussian GLM was first excluded for providing imprecise and biased estimation (Table B.2 and B.3) and the one- and two-part transformed OLS models were also excluded for consistently underestimating the levels of health spending (Table B.4). The use of smearing factor for transformed two-part model actually improved the estimation and enabled the model to capture some high-spending individuals (Figure B.2). However, predicting health expenditure from 1996 to 2008 in a single estimator was no easy task and the mean absolute prediction error (MAPE) and mean square error (MSE) were large, compared to those in the health expenditure modeling literature (Buntin and Zaslavsky 2004).

To choose the best performing models for the next section, model fit tests in Table B.5 mattered the most because these tests were widely used and examined in the literature. One- and two-part Poisson GLM performed the best with the least number of problems, only in link test and Ramsey's RESET test. However, if this study aimed to focus on some specific groups, such as high-cost individuals or Medicare enrollees with disabilities, certain types of tests, such as the illustration of the AAPCC ratios that indicated which models could identify high-cost individuals, might be better methods to assess the predictive power of different models.

Results of logit model and best fitting models

Logit model for the probability of incurring health spending

Result interpretation

In Table B.6, the coefficients of the first column showed the effects of individual characteristics, relative to the reference group. In the first row, the coefficient of each year relative to 1996 was listed. The cells in the middle intersected by different years and individual characteristics were the interaction terms that showed how the effects of these characteristics changed in specific years relative to these characteristics' coefficients in 1996.

Coefficients of individual characteristics in 1996 and the year main effects

In Table B.6, being black (-0.86, $p < 0.05$), very good mental health status (-0.52, $p<0.05$) and being divorced (-0.96, $p<0.05$) were negatively associated with the probability of incurring health spending in 1996. On the contrary, the years of education, very good to fair health status, activity limitation, any limitation and private coverage were significantly associated with higher probability of incurring health spending in 1996.

For the coefficients of years from 1997 to 2008, the earlier years, especially before 2001, and 2008 were significantly associated with a lower probability of incurring health spending. In most of the years, these years had negative regression coefficients and this meant that the probability of incurring any health spending was lower in most years after 1996.

Individual characteristics' interactions with years

From 1997 to 2008, most of the individual characteristics did not have significant coefficients in multiple years. The distribution of significant coefficients also varied. For example, age and being female (both positively) were more likely to have statistically significant coefficients in recent years, as very good mental health status (positively) and widowhood (negatively) were more likely to have significant coefficients in earlier years (Table B.6).

Health expenditure regression results

The regression coefficients of one-part Poisson GLM were listed in Table B.7 and those of two-part Poisson GLM were neglected for their similarity.

Coefficients of individual characteristics in one- and two-part expenditure models

In Table B.7 and the results from two-part Poisson GLM, there were some common traits for one- and two-part model. Both indicated that the year main effects were more likely to be higher and statistically significant in recent years (especially after 2006). Individual characteristics, including being female and residence in the West, were significantly associated with lower total health expenditure in 1996, while other factors, including worse health status, years of education, being divorced, having activity limitation or any limitation, and being covered by Medicaid, were significantly associated with higher total health spending.

Coefficients of individual characteristics in different years

The coefficients in both models showed how the effects of individual characteristics changed after 1996. As not all of these coefficients were statistically significant, many of them fluctuated from 1997 to 2008. However, some individual characteristics were significant in multiple years, especially fair or poor health status, being divorced, and needing assistance in ADL. Their effects might be associated with more (for example, poor health status in 2003) or less (for example, one more year of education in 2002) total health spending.

Predicted coefficients in one- and two-part models

In one-part Poisson GLM (log link), the predicted coefficients of variables were smaller than those in two-part Poisson GLM because the only difference between these two models was the exclusion of individuals without any health spending in two-part Poisson GLM. The health spending for the reference groups was also large in two-part model ($1380.4, $e^{7.23}$) than that in one-part Poisson GLM ($1199.9, $e^{7.09}$)

The direction and magnitude of the coefficients of individual characteristics in 1996 provided an interesting comparison with the coefficients after 1996. For example, divorced enrollees had higher effect in 1996 compared to those married, $4072.2 ($p<0.05$), and this showed that they had a higher average health spending than the married Medicare enrollees, after controlling for other socioeconomic and health characteristics. However, the coefficient of being divorced in 2008, $5953.9 less ($p<0.05$), was much less than that in 1996, $4072.2. This interaction indicated that those divorced had less health spending than the married in 2008. This piece of information could be interpreted as that the divorced Medicare enrollees age 65 years and over had a tendency to decrease their spending from 1996 to 2008 that resulted in a lower health spending level than the married in 2008, although the divorced had a higher level of spending in 1996.

Reduced model for health expenditure

After choosing the model that fit the best, another concern was the model consistency across a variety of specifications. To respond to this issue, a reduced model with less number of covariates and an extended model with extra control of chronic health conditions in a shorter time frame were introduced. The regression coefficients of the reduced model were listed in Table B.8 and those of the extended model in Table B.9. It was obvious to notice the significant coefficients in recent years, including education, health status and the constant in full and reduced models. However, there were more covariates that differed in statistical significance in these two models. For example, the coefficient of the West in 1996 and the coefficients of poor health status in recent years were not significant in the reduced model.

Extended model of health expenditure

The extended model in Table B.9 found different sets of significant variables for health expenditure. Age, races, residence in the Midwest, widowhood, other marital status, and cognitive limitation were significant factors to determine not only the spending level relative to the reference groups in 2000, but also the coefficients after 2000. This model also showed significance in several chronic conditions, including diabetes, asthma, stroke, arthritis, angina, emphysema, and heart attack. These conditions were related to higher spending in 2000 and a slower rate of increase after 2000. Moreover, smoking was not a significant factor from 2000 to 2008 after controlling for other characteristics.

Discussion

Model selection

Model prediction error summary

In this study, individual characteristics that might influence the historical growth of health expenditures for Medicare were tested with multiple regression models. Except for one-part Gaussian GLM, one- and two-part transformed OLS models, the regression models performed well in terms of mean prediction, ratio of prediction to AAPCC amount, mean absolute prediction error (MAPE) and mean square error (MSE). One-part Gaussian GLM was first excluded for its imprecision of estimation. One- and two-part transformed OLS models were then excluded for consistently underestimate the absolute amount of total health spending.

Model summary with R^2

However, these tests mentioned above were not widely used indicators to assess how close these predictions were to the actual amount. One solution was to regress the actual spending with the predicted expenditure (Zheng and Agresti 2000). The estimated amount from each model could be assigned a summary score, R^2, which suggested how much of the variation in the actual spending could be explained by the predicted values. Although two-part Gaussian GLM had the highest R^2, 0.17, this model consistently underestimated health spending and provided a regression coefficient lower than one, 0.93. The models that well predicted the spending level among Medicare enrollees age 65 years and over[63] included one- and two-part OLS models, and one- and two-part Poisson GLM (p>0.05 for the adjusted Wald test with the hypothesis, β_1=1).

Model fit tests

The other approaches could not replace the formal model fit tests frequently used in the literature. In Table B.5, all models failed to pass correlation test and this fact suggested the model prediction might be subject to systemic bias. The results in Table B.5 suggested that one- and two-part Poisson GLM (log link) might be the least problematic, only in link tests and RESET tests. They did not violate the assumed model fit in these two types of tests. The one- and two-part OLS models did not have problem in the RESET tests, as one-part Gamma GLM did not have problem in the modified Hosmer-Lemeshow test.

Estimation of individual spending

Two of the better performing models were selected, one- and two-part Poisson GLM. They had fewer problems in model fit tests than the others and fit the spending distribution well according to the comparison of the R^2 produced by all models. Hence, they were used to estimate the regression coefficients of individual characteristics. In the individual level, these two models had many common statistically significant variables. The magnitude of coefficients of two-part Poisson GLM (among those actually incurring health spending only) tended to be larger than that in one-part Poisson GLM.

The relative weight of these characteristics' coefficients across years differed in the individual and population level. The characteristics had large coefficients in the

---

[63] Good prediction meant the regression coefficient, β_1, equaled one in the regression used to predict the actual observations, $y = \beta_0 + \beta_1\hat{y} + \varepsilon$.

individual estimates, such as worse health status and mental health status, did not translate into larger spending for all Medicare enrollees. Although these variables meant higher amount of health spending, the aggregate sum of these high-spending characteristics in Medicare population did not put much weight on the Medicare system.

Complex survey design in MEPS datasets

One of the challenges to choose the best performing model was the complex survey design that did not permit the production of some vital statistics, such as *f* statistic in GLM, which helped researchers to know how well these models worked. This survey design also made some changes to the execution of model fit tests used in Table B.5. For example, the Pearson's rho (ρ) in the correlation test was not obtained directly through correlation analysis because a correlation test was not officially supported in the STATA survey analysis toolkit. Instead, the *p* values of the correlation tests in this table were calculated by using regression models, suggested by Sribney (2005).

Appendix C: propensity score matching methods and the results in Chapter 3

Method

Data linkage

Matching algorithm settings

The propensity score estimated by these two logit models were used to compare the health spending level between the treatment and control groups. In Table C.1, the settings of different matching algorithms were listed. The matching algorithms determined which observations from the control groups (neighbors for the observations in the treatment group) were compared with. The matching algorithms used in this chapter were tested for how the results resembled the theoretical predictions in Table 3.1.

Algorithm selection to draw conclusions

Another issue for matching was which algorithms to be used to draw a conclusion. Among these matching algorithms, the efficacy of each matching algorithm could be evaluated by the bias reduction and the sensitivity analysis with Rosenbaum bounds (Rosenbaum 2002). The biases in the matched and unmatched differences were defined as the mean standardized differences between the treated and control groups (Rosenbaum and Rubin 1983). By definition, the size of bias could be assessed by the ratio of standard deviation to the mean differences (Leuven and Sianesi 2003) and the bias reduction could be calculated based on the changes in these ratios (DiPrete and Gangl 2004). The matching algorithms not only changed the balance of propensity score and the percentage of bias, but also the percentage of bias reduction.

Moreover, the sensitivity analysis with Rosenbaum bounds illustrated different levels of unobserved heterogeneity (hidden bias)[64] and their upper and lower (95%) bounds of statistical significance (Gangl 2004). Different levels of heterogeneity, gamma (Γ), were applied to find how much heterogeneity alone could cause the same matched difference with these matching algorithms. The critical value of gamma (Γ) meant that, if the scale of hidden bias reached to this level, the matched difference could be contributed by the hidden bias. The higher the critical value of gamma, the less likely the matched difference was caused by hidden bias alone.

Variable balance after propensity score matching

The other way to assess the matching quality included checking the balance in all variables with another user-defined program for STATA, *pstest* (Leuven and Sianesi 2003). The program examined the differences in mean values of all independent variables that were used to predict the propensity score. If the independent variables were balanced after matching, the mean values in the treated and control groups were not significantly different.

Ranges of common support

Finally, the ranges of common support in the matching algorithms were illustrated by the other user-defined program for STATA, *psgraph* (Leuven and Sianesi, 2003). The

[64] See Rosenbaum (2002) and DiPrete and Gangl (2004) for detail in quantifying the hidden biases and gamma (Γ) values in different matching algorithms. The equations to calculate the gamma (Γ) values along with its t and p values in different matching algorithms were listed in detail in DiPrete and Gangl (2004).

treated or control observation were within the range of common support only if they could be matched with one or more neighbors in the other group. For those cases that could not be matched because their propensity scores were not similar to any neighbors, they were not used for this matching method and were drawn outside the range of common support.

This program drew two histograms with a common X axis, the propensity score from zero to one. The histogram that showed the number of observation in the treatment group was plotted above the X axis. The distribution of observations in the control group was plotted below the X axis. The graph aimed to compare the distribution of observation in the treatment and control groups according to the propensity score. This graph was design to find the skewed distribution of any group that might lead to poorly matched pairs.

Quantifying the average treatment effect on the treated (ATT)

After choosing one matching method that performed the best, the difference in the values between the matched treated and control pairs in the chosen matching algorithm could be taken as the average treatment effect on the treated (ATT) because the unobserved endogeneity was in theory removed.

Results

Matching algorithms and common support

Because of different matching algorithms, the "neighbors" for comparisons in one matching method were very likely to be different from those in others. In Table C.2 (total spending model) and C.3 (out-of-pocket spending model), not all individuals eligible for generating propensity score were included for matching algorithms in Table C.2 This sample attrition was because some individuals had propensity score outside "common support".

Illustration of common support

In Figure C.1, the propensity score from zero to one was drawn on the x-axis and the number of treated individuals (enrolling in HMOs, Medicare Advantage/Part C) was illustrated above the horizon, as the controls (not enrolling in HMO, traditional Medicare) were below the horizon. The range of common support was the overlap area of the treated and controls. Three (all matching algorithms for total health spending except for nearest neighbor [1]) and zero (all matching algorithms for out-of-pocket health spending) observations were excluded for the lack of matching from the control group and placed outside the range of common support.

Average treatment effect on the treated (ATT) based on matching algorithms

Matching results

In Table C.2, the matched differences in total health spending were not uniformly greater or less than the unmatched spending difference ($2,411.0, less spending for HMO coverage) over the first three to four years of Medicare coverage. However, all of the matched out-of-pocket spending differences (in absolute values) were less than the unmatched difference ($1,943.0, less for HMO coverage) over the first three to four years in HRS (Table C.3). The concern that these different matching results brought was which matching algorithms were the most robust and the least vulnerable from hidden bias.

Precision of different matching algorithms

Imprecision of nearest neighbor matching with one neighbor (without replacement)

Following the theoretic framework in Chapter 3, the nearest neighbor matching with one neighbor (without replacement) that did not reuse the matched neighbors in the control group could always found one closest neighbor in the control group for each treated enrollees so that there was no sample attrition in the total and out-of-pocket models (Table C.2 and C.3). However, the concern was to match poorly by inadequately assigning one closest control that was actually too distant to the treated observation. This inadequacy could be found in total spending model, in which nearest neighbor (one neighbor) matching had a matched difference not similar to the results in other (five an ten neighbors) nearest neighbor matching algorithms. In contrast, this imprecision in nearest neighbor model (one neighbor) was not prominent in out-of-pocket model (Table C.3).

Precision and the number of neighbors in the nearest neighbor matching

Despite of the danger of bias due to more neighbors in Table 3.1, the nearest neighbor matching with more neighbors (from one to five and ten with replacement) yielded smaller standard errors in total and OOP spending models. But there were observations dropped and the *p* values were not always smaller with more neighbors. In fact, the *p* values of nearest neighbor matching with one neighbor were lower in the total spending model and higher in the OOP spending model, compared to matching methods with more neighbors.

Moreover, the bias reduction and critical value of gamma (Γ) in the sensitivity test (Table C.2 and C.3) showed that bias reduction did not improve a lot with more neighbors but the critical values of gamma were higher in both models. The sensitivity analysis indicated that the nearest matching algorithms with more neighbors (from one to ten) provided a more robust estimation of matched difference.

Radius matching

Radius matching that matched all neighbors within a selected range (default value as 0.06 propensity score) did not always generate efficient results (smaller variance). The standard error in out-of-pocket spending models was larger than those in the nearest neighbor (five and ten neighbors) matching, as the theoretical framework in Table 3.1 predicted a more precise estimation in radius matching. The variance in radius matching was between those generated by nearest matching with one and five neighbors in OOP spending models. Radius matching did not outperform nearest neighbor matching in bias reduction, but the sensitivity analysis showed radius-matching estimates had a higher critical value than those in nearest neighbor matching (one, five or ten neighbors).

Kernel matching and bandwidth

For kernel matching that weighted the distance of propensity score to the neighbors, a wider bandwidth, from the default value (0.06) to 0.1, increased the number of neighbors for comparison and helped to generate a smaller variance, as predicted in Table 3.1. However, the variances predicted based on kernel matching was not necessarily smaller than those generated based on nearest neighbor matching in the total and OOP spending models.

The critical values of gamma in the sensitivity analysis indicated kernel matching were higher than all nearest neighbor matching (from one to five and ten) in total and OOP spending models. More specifically, the kernel matching had higher gamma values than those obtained from the nearest neighbor matching methods in the total spending model and the highest value (kernel matching [0.1]) among all matching algorithms in OOP spending model.

Local linear matching

Because local linear matching that is a way of weighting neighbors based on non-parametric methods, its standard errors and t statistics could not be produced. However, local linear matching did not produce the highest bootstrapped z statistics (absolute value) in total spending model or the highest bias reduction in both models.

Matching results based on the sensitivity analysis

In Table C.2 and C.3, the results showed that local linear matching provided the highest percent of bias[65] reduction for total and out-of-pocket health expenditure, although the bias reductions in all matching methods were larger than 79% (nearest neighbor matching with one neighbor in total spending model).

More importantly, the critical values of gamma (Γ), obtained from Rosenbaum bonds (the sensitivity analysis for propensity score matching) provided an easy-to-use measurement for conclusion. For the total spending model, the critical value in kernel matching with bandwidth as 0.6 was the highest, between 2.62 and 2.63 (Table C.2). For

[65] Bias was defined as "the difference of the sample means in the treated and non-treated (full or matched) sub-samples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups" according to Leuven's webpage (http://fmwww.bc.edu/repec/bocode/p/pstest.html).

OOP spending model, the highest critical value was between 3.20 and 3.21 in kernel matching with bandwidth 0.1. The HMO's effects (average treatment effects on the treated, ATT) on total health spending in the first three to four year of Medicare coverage was $2,651.0 (nominal dollar) less than traditional Medicare coverage ($p = 0.49$ or 0.46 in Kernel matching with caliper as 0.1 and Kernel matching with bandwidth as 0.6 respectively) and $1,943.0 (nominal dollar) less for out-of-pocket health spending ($p < 0.01$).

Variable balance between the treated and control groups

After observations being matched, it would be important to look for variables with inadequate balancing (uneven distribution in the treated and control groups). In total spending model, there were no imbalanced[66] variables in kernel matching with bandwidth as 0.6 (with or without Caliper). The mean values or percentages of the independent variables in the matched neighbors in the control group (enrolled in traditional Medicare) were not different from those in the treatment group (enrolled in Medicare Advantage/Part C).

With kernel matching with bandwidth (0.1), there was no imbalanced variable in out-of-pocket spending model. The p values of the statistical tests in the mean values or percentages of the variables used to predict propensity scores all became larger than 0.05.

[66] If variables were imbalanced, the difference of their mean values between the treated and control groups (or the matched neighbor from the control groups after matching) were statistically significant ($p < 0.05$).

Effects of HMO estimated by regression models

Choosing GLM family for the expenditure variance structure

Modified Park tests were performed to know the variance was constant to (λ=0), or proportional to (λ=1), or proportional to the square of (λ=2), or proportional to the cube (λ=3) of the spending mean (total or OOP spending models). For total and out-of-pocket spending, the variance function seemed to be proportional to the square (λ=2) or cube (λ=3) of the mean (λ=2.22 and 2.16 respectively). Moreover, both total and out-of-pocket spending had lower Chi-square values (0.73 and 0.41 respectively) and large *p* values (0.39 and 0.52 respectively) when λ equaled two. This suggested that Gamma GLM would be optimal for both spending distribution. The regression coefficients of gamma GLM (log link) were listed in Table C.4. The derived marginal effect for HMO coverage was $1,772.5 ($p < 0.01$) less out-of-pocket spending than traditional Medicare in the first three to four years of Medicare coverage, as the effect on total spending was $1,515.5 ($p = 0.58$) less.

Effects of HMO coverage estimated by regression models after controlling for chronic conditions and death events

Moreover, the model in Chapter 4 added chronic conditions to predict their effects on health returns. This chapter also attempted to quantify the financial impact from death events that were statistically significant in the probability of mortality within three to four years of Medicare coverage. By adding chronic conditions and death events as independent variables in the model in Table C.4, the results of this new regression model was listed in Table C.5. The estimated marginal effects of HMO coverage, hypertension, arthritis and death events were $1,896.2 less, $1,632.2 more, $785.4 more and $1,974.0

more ($p < 0.01$ for all) on out-of-pocket spending respectively, as their effects on total spending were \$1,012.3 less ($p = 0.72$), \$3,186.3 more ($p = 0.23$), \$997.1 (p 0.51), and \$17,669.5 more ($p < 0.01$) respectively.

Tables

Table 2.1. The national health spending (billions) by Medicare enrollees aged 65 years and over from 1996 to 2008.

Categories	Subgroups	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Annual Growth Rate (%)
Total health expenditure		183	194	204	208	203	245	274	291	322	333	333	364	366	5.8
Sex	Male	79.3	91.5	72	91.1	91.1	108	122	127	140	141	142	155	154	5.5
	Female	104	102	132	117	112	137	152	164	181	193	191	208	212	5.9
Age	65-74	73.5	69.9	75.4	86.8	94.4	101	115	139	142	140	139	156	163	6.6
	75-84	64.2	79.7	69.2	70.4	65.3	86.2	103	107	114	126	121	123	123	5.4
	85 and over	45.2	44.1	59.3	50.5	43.2	57.9	56.3	45.4	65.7	67.6	71.7	85	79.9	4.7
Race	White	164	175	181	189	181	219	243	256	287	290	287	317	315	5.4
	Black	15.2	16	18.8	15.8	16.5	20.1	22.3	26.6	25.8	32.8	30.6	31	35.7	7.1
	Other	3.5	2.4	4.1	3.3	5.1	6	8.6	8.4	8.7	11.1	14.9	15.5	14.9	12.1
Ethnicity	Hispanic	8.5	9.8	9.3	11.2	9.5	11.2	15.9	14.5	15	17.3	21.7	26.5	24.9	9.0
Years of education															
	0-8	29.9	44.2	51.7	30.7	44.7	48.6	54.4	46	43.7	44.3	48.9	50.5	44.3	3.3
	9-12	84.7	94.8	93.8	113	92	117	134	142	150	164	155	170	171	5.9
	>13	68.3	54.6	58.5	64.3	66.2	79.4	86.1	103	128	125	129	144	150	6.6
Income relative to poverty line															
	Poor/ Negative	30.1	25.1	22.2	26.4	23.5	33.4	34.7	33.3	31.9	39.7	38.4	50.3	37.6	1.9
	Near poor	17	14.5	21.7	11.9	13.4	21.8	19.2	23.5	20.9	29.6	36.9	32	34.6	5.9
	Low income	40.5	38.7	46.6	42.9	43.6	51.6	56.5	61.7	68.3	67.4	60.7	76.3	67.7	4.3
	Middle income	52.4	72.2	66.4	71.8	67.9	75.6	79.6	87	99.6	90.2	85.2	93.2	107	5.9
	High income	43	43.1	47	54.6	54.5	62.5	84.1	86	101	106	111	112	119	8.5

Continued in the next page

Categories	Subgroups	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Annual Growth Rate (%)
Regions															
	Northeast	36.1	30.2	32.8	39.7	36.6	47.5	56.4	58.3	62.8	56.5	58	66	74.5	6.0
	Midwest	46.7	35	38.8	36.6	44	51.5	59.2	69.2	69.6	78.4	71.6	77	75.4	4.0
	South	49.5	67.4	62.7	65	69.4	80.4	84	98.3	106	107	118	126	117	7.2
	West	26.3	36.8	35.7	36.2	31.3	37.9	47.4	47.7	59	59	54.8	63.1	72.8	8.5
Marital status															
	Married	74.7	98.9	88.8	103	111	125	138	147	164	172	171	183	187	7.6
	Widowed	57.6	73.9	91.8	80.9	68.5	87.1	95.7	109	119	117	116	117	128	6.7
	Divorced	19	13.5	16.2	17.2	16.1	20.5	27.6	23	25.9	31.4	31.4	41.4	32	4.3
	Other	7.3	7.4	7.1	6.9	7.8	11.8	12.7	12.7	12.3	12.9	13.5	21.9	18.8	7.9
Self-rated health status															
	Excellent	18.5	11.1	10.5	11.2	11.7	14.1	17.3	17.3	21.4	22	26	27.1	24.7	2.4
	Very good	28.6	29	25.6	29.8	31.9	40.2	42.4	49.2	56.6	56.5	58.5	61.6	68.4	7.3
	Good	48.7	47.5	49.4	52.7	50.6	65.5	79.9	84.6	85.6	89.4	91.1	100	103	6.2
	Fair	48.8	35.7	40	43.1	49.6	50	64.7	63.2	73.6	74.5	70.7	82.5	79.7	4.1
	Poor	30.9	37.3	32.5	29.9	31.4	33.2	34	44.6	46.1	48.1	46.6	48.6	45.8	3.3
Self-rated mental health status															
	Excellent	48	31.9	29.7	32.3	36.4	37.7	47.7	56	60.9	50.4	60.6	68.4	75.1	3.7
	Very good	39.6	46.1	36.5	42.3	40.7	52.9	64.5	73.1	73.5	73.2	77.4	85.3	83	6.2
	Good	55.1	46.9	59.1	63.9	58.9	69.9	84.1	80.3	107	107	108	119	106	5.5
	Fair/poor	32.5	35.5	32.7	28.7	39.4	42.5	42	49.4	41.9	59.4	47.1	46.7	57.8	4.8
ADL help		38.4	41.1	29.7	27.3	30.4	41.3	42.2	41.9	46.8	54.5	56.5	53.2	53.6	2.8
IADL help		56.3	52.3	42.7	43.6	42.6	62.9	63.9	61.9	71.8	82.3	79.6	77.8	77.2	2.6
Activity limitation		84.5	70.5	68	65	69.8	86.9	96.1	102	112	129	127	122	115	2.6
Cognitive limitation		46.8	33.6	34	28.6	34.4	46	46.8	50.1	56.9	70.9	67.8	64.9	65.9	2.9

Continued in the next page

Categories	Subgroups	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Annual Growth Rate (%)
Any limitation		136	149	158	146	151	188	206	223	246	257	260	273	261	5.4
Medicare only		27.8	42.1	59.9	59.1	58.8	73.8	79.1	83.9	89.7	101.8	103.1	128.7	145.4	13.8
Medicaid		30.2	28.9	32.1	32.9	28.2	31.2	36.9	38.1	46.3	49.2	44.9	56.3	48.6	4.0
Private insurance		125	123	112	116	116	140	158	169	186	182	185	179	172	2.7
Chronic conditions															
	Diabetes					46.1	49.7	61.2	65.2	77.3	90.7	84.9	96.1	108	10.6
	Asthma					23.8	22.2	34.9	34.9	29.8	44.5	42.7	49.1	37.5	5.7
	Angina					31.2	31.1	32.9	39.4	46	42.3	35	48.8	52.5	6.5
	Stroke					35.9	42.2	39.8	46.7	44.2	39.9	44.6	68.1	68.8	8.1
	Emphysema					15.1	19.6	24.4	25.5	19	19.6	28.4	29.8	31.5	9.2
	Hypertension					111	135	157	180	199	205	216	250	273	11.2
	Coronary heart disease					42	45.2	56.2	56.5	71.5	64.2	63.4	85.3	121	13.2
	Heart attack					44.8	39.3	52.5	51.5	53.7	57.2	54.3	64	72.1	5.9
	Other heart disease					46.3	54.3	62.1	62.4	72.9	82.4	72.7	107	131	13.0
	Joint pain					114	131	156	168	177	188	188	212	210	7.6
Health behavior															
	Smoking					18.3	18.9	23.8	23.2	26.3	24.8	25.8	22.3	26.7	4.7

Note: statistics estimated based on annual MEPS-HC (household component) datasets from 1996 to 2008.

Table 2.2. Characteristics of Medicare enrollees age 65 and over from 1996 to 2008.

Categories		**1996**	**2000**	**2008**
Population (million)		34	35	39
Health spending	None	4.16%	4.34%	2.98%
	Any	95.85%	95.65%	96.95%
Sex	Male	42.43%	42.90%	43.26%
	Female	57.57%	57.39%	56.74%
Age	65-74	53.41%	51.88%	50.38%
	75-84	32.94%	33.33%	33.33%
	85 and over	13.55%	14.73%	16.19%
Race	White	89.02%	88.70%	86.01%
	Black	8.09%	8.43%	8.89%
	Other	2.86%	2.99%	5.09%
Ethnicity	Hispanic	4.76%	5.35%	6.99%
Years of education	0-8	19.05%	18.07%	11.19%
	9-12	48.96%	49.86%	47.84%
	>13	32.05%	32.17%	40.97%
Income relative to poverty line	Poor/ Negative	12.13%	10.78%	10.44%
	Near poor	8.15%	7.01%	7.78%
	Low income	21.93%	19.87%	18.35%
	Middle income	33.53%	35.07%	29.01%
	High income	24.40%	27.37%	34.35%
Regions	Northeast	19.90%	19.59%	18.41%
	Midwest	23.37%	21.35%	21.13%
	South	33.23%	35.07%	36.13%
	West	18.41%	19.51%	20.22%
Marital status	Married	52.82%	54.20%	53.44%
	Widowed	31.16%	33.62%	31.04%
	Divorced	6.87%	7.72%	10.64%
	Other	3.91%	4.40%	4.98%
Self-rated health status	Excellent	18.52%	14.22%	14.12%
	Very good	25.30%	24.97%	28.50%
	Good	28.02%	31.59%	30.79%
	Fair	18.14%	16.78%	15.01%
	Poor	8.63%	6.72%	6.34%
Self-rated mental health status	Excellent	32.05%	25.56%	26.97%
	Very good	27.76%	27.74%	28.75%
	Good	27.03%	29.86%	28.50%
	Fair/poor	11.84%	11.30%	10.34%

Continued in the next page

Year		1996	2000	2008
ADL help		8.34%	6.00%	6.52%
IADL help		14.44%	10.27%	11.33%
Activity limitation		23.24%	19.61%	20.02%
Cognitive limitation		12.19%	9.80%	11.51%
Any limitation		55.19%	55.07%	57.76%
Medicaid		11.71%	9.99%	9.58%
Private insurance		67.06%	55.94%	47.33%
Chronic conditions	Diabetes		14.99%	22.25%
	Asthma		7.98%	7.74%
	Angina		8.45%	10.08%
	Stroke		9.11%	13.74%
	Emphysema		4.75%	6.71%
	Hypertension		49.86%	67.18%
	Coronary heart disease		11.30%	22.35%
	Heart attack		11.22%	13.08%
	Other heart disease		14.83%	27.99%
	Joint pain		51.88%	53.44%
Health behavior	Smoking		11.40%	8.58%

Note: statistics estimated based on the annual MEPS-HC (household component) datasets from 1996 to 2008.

Table 2.3. The estimated annual growth of the aggregate Medicare health spending (billions) grouped by individual characteristics from 1996 to 2008.

Categories	Subgroups	1996	2008	Change in spending	Annual growth rate (%)	Percentage of change (%)
Population (million)		34	39	5	1.1	114.7
Actual spending		183.0	366.0	183	5.8	200.0
Adjusted spending*		231.0	327.0	96	2.9	141.6
Sex	Male	104.0	138.0	34	2.4	132.7
	Female	117.0	170.0	53	3.1	145.3
Age**	65-74	105.0	156.0	51	3.3	148.6
	75-84	80.7	115.0	34.3	3.0	142.5
	85 and over	43.4	53.4	10	1.7	123.0
Race	White	205.0	280.0	75	2.6	136.6
	Black	20.6	29.8	9.2	3.1	144.7
	Other	5.7	15.7	9.96	8.4	273.5
Ethnicity	Hispanic	10.4	24.8	14.4	7.2	238.5
Years of education	0-8	49.4	39.4	-10	-1.9	79.8
	9-12	107.0	156.0	49	3.1	145.8
	>13	73.3	130.0	56.7	4.8	177.4
Income relative to poverty line	Poor/ Negative	31.6	35.9	4.3	1.1	113.6
	Near poor	21.5	26.8	5.3	1.8	124.7
	Low income	55.1	60.7	5.6	0.8	110.2
	Middle income	71.9	92.8	20.9	2.1	129.1
	High income	48.7	108.0	59.3	6.6	221.8
Regions	Northeast	48.1	66.0	17.9	2.6	137.2
	Midwest	63.2	67.6	4.4	0.6	107.0
	South	63.6	104.0	40.4	4.1	163.5
	West	31.4	67.0	35.6	6.3	213.4
Marital status	Married	103.0	171.0	68	4.2	166.0
	Widowed	79.6	107.0	27.4	2.5	134.4
	Divorced	26.8	25.7	-1.1	-0.3	95.9
	Other	10.2	15.3	5.1	3.4	150.0
Self-rated health status	Excellent	23.6	24.3	0.7	0.2	103.0
	Very good	41.6	95.3	53.7	6.9	229.1
	Good	78.8	153.0	74.2	5.5	194.2
	Fair	69.7	111.0	41.3	3.9	159.3
	Poor	41.5	55.7	14.2	2.5	134.2

Continued in the next page

Categories	Subgroups	1996	2008	Change in spending	Annual growth rate (%)	Percentage of change
Self-rated mental health status	Excellent	64.9	74.6	9.7	1.2	114.9
	Very good	48.5	79.8	31.3	4.1	164.5
	Good	68.2	102.0	33.8	3.4	149.6
	Fair/poor	38.8	46.4	7.6	1.5	119.6
ADL help		41.0	43.9	2.9	0.6	107.1
IADL help		67.7	65.0	-2.7	-0.3	96.0
Activity limitation		108.0	104.0	-4.0	-0.3	96.3
Cognitive limitation		56.7	53.9	-2.8	-0.4	95.1
Any limitation		195.0	248.0	53.0	2.0	127.2
Medicaid		40.7	39.7	-1.0	-0.2	97.5
Private insurance		172.0	163.0	-9.0	-0.4	94.8

Note: The spending prediction included the amount incurred by the average values of other associated characteristics in subgroups. For example, when summing the adjusted spending incurred by the females, the coefficients of being female was first added to the spending levels obtained from the reference groups (male) to estimate female spending. Then, the spending levels related to the mean income, years of education and other associated variables were assigned to observations in the same subgroups to estimate the level of Medicare spending. This helped to capture the associated changes in factors other than the grouping variables.

* Adjusted spending was obtained by multiplying the population size and the spending levels that were predicted by the mean values of all independent variables.

** Age was right-censored up to 90 years until 2000 and up to 85 years after 2000.

Table 2.4. The estimated annual growth in aggregate Medicare health spending (billions) grouped by individual characteristics and chronic health conditions from 2000 to 2008 (the extended model)

Categories	Subgroups	2000	2008	Change in spending	Annual growth rate (%)	Percentage of change
Population (million)		34	39	5	1.4	114.7
Actual spending		203	366	163	7.4	180.3
Sex	Male	127	152	25	2.2	119.7
	Female	174	208	34	2.2	119.5
Age*	65-74	148	171	23	1.8	115.5
	75-84	111	128	17	1.8	115.3
	85 and over	42.1	60.8	18.7	4.6	144.4
Race	White	276	311	35	1.5	112.7
	Black	23.5	32.6	9.1	4.1	138.7
	Other	1.7	15.6	13.94	28.0	939.8
Ethnicity	Hispanic	15.6	27.1	11.5	6.9	173.7
Years of education	0-8	62.3	44.2	-18.1	-4.3	70.9
	9-12	148	173	25	2.0	116.9
	>13	90.2	143	52.8	5.8	158.5
Income relative to poverty line	Poor/ Negative	33	39.9	6.9	2.4	120.9
	Near poor	22.4	29.9	7.5	3.6	133.5
	Low income	63.5	67.1	3.6	0.7	105.7
	Middle income	103	104	1	0.1	101
	High income	79.3	119	39.7	5.1	150.1
Regions	Northeast	56.8	74	17.2	3.3	130.3
	Midwest	74.1	76.7	2.6	0.4	103.5
	South	112	124	12	1.3	110.7
	West	52.8	75.1	22.3	4.4	142.2
Marital status	Married	160	188	28	2.0	117.5
	Widowed	107	120	13	1.4	112.1
	Divorced	23.7	33.5	9.8	4.3	141.4
	Other	10.2	17.7	7.5	6.9	173.5

Continued in the next page

Categories	Subgroups	2000	2008	Change in spending	Annual growth rate (%)	Percentage of change
Self-rated health status	Excellent	20.5	27.3	6.8	3.6	133.2
	Very good	62.4	82.6	20.2	3.5	132.4
	Good	94.2	119	24.8	2.9	126.3
	Fair	78.6	78.8	0.2	0.0	100.3
	Poor	37.9	39.3	1.4	0.5	103.7
Self-rated mental health status	Excellent	64.3	82.3	18	3.1	128
	Very good	76.2	93	16.8	2.5	122
	Good	99.5	117	17.5	2.0	117.6
	Fair/poor	54	54.1	0.1	0.0	100.2
ADL help		35.9	43.8	7.9	2.5	122
IADL help		55.8	67.7	11.9	2.4	121.3
Activity limitation		98	108	10	1.2	110.2
Cognitive limitation		46.3	61.5	15.2	3.5	132.8
Any limitation		218	247	29	1.6	113.3
Medicaid		33.7	41	7.3	2.5	121.7
Private insurance		177	172	-5	-0.4	97.2
Chronic conditions						
	Diabetes	67.6	98.7	31.1	4.7	146
	Asthma	42.9	47.3	4.4	1.2	110.3
	Angina	45.4	63.4	18	4.2	139.6
	Stroke	22.2	31.2	9	4.3	140.5
	Emphysema	35.8	34.1	-1.7	-0.6	95.3
	Hypertension	179	269	90	5.1	150.3
	Coronary heart disease	57.4	112	54.6	8.4	195.1
	Heart attack	61	63.9	2.9	0.6	104.8
	Other heart disease	69.1	126	56.9	7.5	182.3
	Joint pain	194	218	24	1.5	112.4
Health Behaviors	Smoking	31.6	28.2	-3.4	-1.4	89.2

Note: The spending prediction included the amount incurred by the average values of other associated characteristics in subgroups. For example, when summing the adjusted spending incurred by the females, the coefficients of being female was first added to the spending levels obtained from the reference groups (male) to estimate female spending. Then, the spending levels related to the mean income, years of education and other associated variables were assigned to observations in the same subgroups to estimate the level of Medicare spending. This helped to capture the associated changes in factors other than the grouping variables.

* Age was right-censored up to 90 years until 2000 and up to 85 years after 2000.

Table 3.1. The comparison of the estimation bias and precision between matching algorithms.

Decision	Bias	Variance
Nearest neighbour matching:		
multiple neighbours / single neighbour	(+)/(-)	(-)/(+)
with caliper / without caliper	(-)/(+)	(+)/(-)
Use of control individuals:		
with replacement / without replacement	(-)/(+)	(+)/(-)
Choosing method:		
NN-matching / Radius-matching	(-)/(+)	(+)/(-)
KM or LLM / NN-methods	(+)/(-)	(-)/(+)
Bandwidth choice with KM:		
small / large	(-)/(+)	(+)/(-)

KM: Kernel Matching, LLM: Local Linear Matching
NN: Nearest Neighbour
Increase: (+), Decrease: (-)

Source: Caliendo and Kopeinig, 2008

Table 3.2. Individual characteristics of eligible Medicare enrollees in HRS data set.

	Total health expenditure (N = 1841)		Out-of-pocket health expenditure (N=4126)	
	No. of obs.	Mean (S. D.)	No. of obs.	Mean (S. D.)
Under Medicare				
Total Health Expenditure (thousands)*	1841	26,350.71	1820	25,933.13
		64,906.73		64,425.12
Out-of-pocket Health Expenditure (thousands)*	1820	5,762.32	4126	6,514.04
		12,452.13		19,347.96
HMO coverage (%)	1841	29.66%	4126	26.56%
Pre-Medicare characteristics				
Age (years)	1841	64.57	4126	64.34
		0.53		0.68
Female (%)	1841	54.16%	4126	56.23%
Race (%)				
Black	1841	13.80%	4126	14.01%
Other	1841	2.88%	4126	3.47%
Hispanic (%)	1841	7.66%	4126	8.75%
Regions (%)				
Midwest	1841	24.01%	4126	23.63%
South	1841	39.76%	4126	40.55%
West	1841	19.23%	4126	19.85%
Years of education	1841	12.02	4126	12.20
		3.12		3.09
Income (thousands)*	1841	7973.76	4126	10,275.19
		19835.74		28,137.00
Self-rated health status (%)				
Very good	1841	28.57%	4126	28.99%
Good	1841	33.19%	4126	32.84%
Fair	1841	17.33%	4126	18.37%
Poor	1841	5.59%	4126	5.50%
CESD score (0 to 8)	1841	1.30	4126	1.31
		1.79		1.85

Continued in the next page

	Total health expenditure (N = 1438)		Out-of-pocket health expenditure (N=3580)	
	No. of obs.	Mean (S. D.)	No. of obs.	Mean (S. D.)
Difficulty in ADL (0 to 5) (%)				
1	1841	6.41%	4126	5.84%
2	1841	2.12%	4126	2.01%
3	1841	1.20%	4126	1.09%
4-5'	1841	0.81%	4126	0.80%
Difficulty in IADL (0 to 5) (%)				
1	1841	4.13%	4126	3.34%
2-3'	1841	0.60%	4126	0.51%
Difficulty in mobility (0 to 5) (%)				
1	1841	18.14%	4126	19.75%
2	1841	9.61%	4126	10.49%
3	1841	4.89%	4126	5.26%
4	1841	3.59%	4126	4.02%
5	1841	1.96%	4126	1.87%
Marital status (%)				
Separated/Divorced	1841	8.91%	4126	11.17%
Widowed	1841	13.63%	4126	12.92%
Never married	1841	2.82%	4126	2.86%
Insurance (%)				
Medicaid	1841	3.15%	4126	4.00%
Champus/VA	1841	4.83%	4126	6.11%
Private insurance (from self)	1841	28.03%	4126	30.20%
Private insurance from spouses)	1841	13.80%	4126	15.15%
Pre-Medicare interview year (1992 as reference) (%)	1841	1,997.37	4126	2,000.33
		1.84		3.13
Medicare interview year (1994 as reference) (%)	1841	1,999.38	4126	2,002.34
		1.83		3.13
Birth year (1926 as reference) (%)	1841	1,932.27	4126	1,935.47
		1.72		3.35

Note: * nominal dollars.

Table 3.3. The Logit models predicting the propensity score of selecting Medicare Advantage/Part C among individuals with information on total or out-of-pocket health spending.

Model summary	**Total health spending**	**OOP health spending**
Eligible observations	1841	4126
Likelihood ratio	261.63	530.65
p	< 0.01	< 0.01
Pseudo R^2	0.12	0.11
	Coefficients (S. E.)	**Coefficients (S. E.)**
Female	0.04	-0.03
	(0.12)	(0.08)
Race		
Black	0.38*	0.29*
	(0.17)	(0.12)
Other	0.37	0.30
	(0.33)	(0.20)
Hispanic	0.50*	0.32*
	(0.23)	(0.15)
Regions		
Midwest	-1.08**	-1.31**
	(0.17)	(0.12)
South	-0.95**	-0.95**
	(0.15)	(0.11)
West	0.66**	0.45**
	(0.17)	(0.11)
Other	(omitted)	(omitted)
Years of education	0.03	0.02
	(0.02)	(0.01)
Income (log scale)***	0.0091	-0.0017
	(0.0127)	(0.0087)
Self-rated health status		
Very good	0.14	-0.07
	(0.17)	(0.12)
Good	0.13	-0.02
	(0.17)	(0.12)
Fair	0.24	-0.14
	(0.22)	(0.15)
Poor	-0.06	-0.31
	(0.34)	(0.24)

Continued in the next page

	Total health spending	OOP health spending
	Coefficients (S. E.)	Coefficients (S. E.)
CESD score (0 to 8)	-0.07	-0.02
	(0.04)	(0.02)
Difficulty in ADL (0 to 5)		
1	0.31	0.25
	(0.25)	(0.18)
2	0.00	0.15
	(0.46)	(0.30)
3	0.81	0.41
	(0.56)	(0.41)
4/5	-0.32	0.39
	(0.91)	(0.52)
Difficulty in IADL (0 to 5)		
1	0.03	0.24
	(0.30)	(0.21)
2	-0.15	-0.36
	(0.87)	(0.63)
Difficulty in mobility (0 to 5)		
1	0.05	-0.10
	(0.15)	(0.10)
2	0.24	0.07
	(0.21)	(0.14)
3	0.04	0.03
	(0.29)	(0.20)
4	-0.63	-0.49*
	(0.38)	(0.24)
5	-1.01	-0.31
	(0.61)	(0.39)
Marital status (married as reference)		
Separated/Divorced	0.16	0.13
	(0.20)	(0.12)
Widowed	-0.36	-0.32*
	(0.18)	(0.13)
Never married	0.03	0.42
	(0.34)	(0.22)
Pre-Medicare health coverage		
Medicaid	0.16	-0.08
	(0.34)	(0.22)

Continued in the next page

	Total health spending	OOP health spending
	Coefficients (S. E.)	**Coefficients (S. E.)**
Champus/VA	-0.79**	-1.31**
	(0.30)	(0.23)
Private insurance (from self)	0.45**	0.45**
	(0.13)	(0.09)
Private insurance (from spouse)	0.39*	0.45**
	(0.16)	(0.11)
Pre-Medicare interview year	0.17	0.14
	(0.32)	(0.25)
Medicare interview year	-0.29	-0.29
	(0.31)	(0.25)
Birth year	0.07	0.08
	(0.09)	(0.05)
Constant	110.27	130.14**
	(64.74)	(25.41)

Note: * $p<0.05$; ** $p<0.01$; *** denoted that the log measures were transformed from the amount in nominal dollars.

Table 3.4. The cost saving effect of HMO coverage under Medicare predicted by GLM (gamma) regressions, compared with the expenditure difference estimated by propensity score matching.

	Total health spending difference			OOP health spending difference		
Part 1: Gamma GLM estimates	**Marginal effect**	**S.E.**	**p**	**Marginal effect**	**S.E.**	**p**
	-1,515.47	2,743.89	0.58	-1,772.52	445.67	< 0.01
Part 2:	**Differences**	**S. E.**	**p**	**Differences**	**S. E.**	**p**
Matched for propensity score	-2,650.95	3,761.16	0.46*	-1,411.45	620.83	<0.01**
Unmatched	-2,411.02	3,312.39	0.78	-1,943.00	681.39	<0.01

Note: * Kernel matching with caliper (0.1) or Kernel matching with bandwidth (0.06). ** Kernel matching with bandwidth (0.1).

Table 4.1. The observed length of time (months) for Medicare enrollees in HRS, categorized by death and HMO coverage under Medicare.

Groups	**Survival subgroup**	**No. of obs.**	**Length of observation (months)**				**Comparison of survival length**
			Mean	Std. Dev.	Min	Max	p
Alive		3732	47.90	3.02	37	58	< 0.01
Deceased		300	47.46	2.93	40	56	
All		4032	47.86	3.01	37	58	
Traditional Medicare	Alive	2732	47.96	3.07	38	58	0.04
	Deceased	235	47.56	2.91	40	55	
	Subtotal	2967	47.93	3.06	38	58	
Medicare Advantage/ Part C	Alive	1000	47.73	2.85	37	58	0.02
	Deceased	65	47.11	3.01	41	56	
	Subtotal	1065	47.69	2.87	37	58	

Note: the p values were derived from the t tests that compared the lengths of observation between those survived and deceased.

Table 4.2. Characteristics of those surviving and the deceased covered in the first three to four years of Medicare coverage.

	Surviving individuals		Deceased individuals		P*
	No. of obs.	Mean (S. D.)	No. of obs.	Mean (S. D.)	
Chronic conditions					
Hypertension (%)	3629	0.45	466	0.51	0.02
		(0.5)		(0.5)	
Arthritis (%)	3541	0.6	435	0.55	0.03
		(0.71)		(0.68)	
HMO coverage (%)	2591	0.4	352	0.34	0.05
		(0.49)		(0.48)	
Female (%)	3632	0.54	467	0.35	<0.001
		(0.5)		(0.48)	
Health expenditure					
Total-spending (thousands)**	1174	21.31	264	39.55	<0.001
		(57.74)		(69.32)	
Out-of-pocket spending (thousands)**	3279	5.48	301	6.74	0.12
		(13.61)		(14.15)	
Race (%)					
Black	3631	0.14	466	0.19	<0.01
		(0.35)		(0.39)	
Other	3631	0.04	466	0.02	
		(0.19)		(0.15)	
Hispanic (%)	3630	0.09	466	0.08	0.47
		(0.29)		(0.27)	
Regions (%)					
Midwest	3630	0.25	467	0.24	0.04
		(0.43)		(0.43)	
South	3630	0.4	467	0.47	
		(0.49)		(0.5)	
West	3630	0.17	467	0.15	
		(0.38)		(0.36)	
Other	3630	0	467	0	
		(0.03)		(0.05)	
Years of education	3625	12.47	465	11.5	<0.001
		(3.12)		(3.36)	
Income (thousands)	3632	15.04	467	10.16	<0.01
		(33.74)		(24.7)	

Continued in the next page

	Surviving individuals		Deceased individuals		p
	No. of obs.	Mean (S. D.)	No. of obs.	Mean (S. D.)	
Self-rated health status (%)					
Very good	3629	0.33	467	0.19	<0.001
		(0.47)		(0.39)	
Good	3629	0.33	467	0.35	
		(0.47)		(0.48)	
Fair	3629	0.15	467	0.25	
		(0.35)		(0.44)	
Poor	3629	0.04	467	0.13	
		(0.2)		(0.34)	
CESD score (0 to 8)	3293	1.2	380	1.71	<0.001
		(1.8)		(2.01)	
Difficulty in ADL (0 to 5) (%)					
1	3535	0.05	412	0.08	<0.001
		(0.21)		(0.27)	
2	3535	0.01	412	0.05	
		(0.11)		(0.21)	
3	3535	0.01	412	0.02	
		(0.09)		(0.13)	
4-5'	3535	0.01	412	0.03	
		(0.08)		(0.17)	
Difficulty in IADL (0 to 5) (%)					
1	3533	0.03	412	0.05	<0.001
		(0.16)		(0.23)	
2-3'	3533	0	412	0.03	
		(0.06)		(0.18)	
Difficulty in mobility (0 to 5) (%)					
1	3506	0.19	401	0.19	<0.001
		(0.39)		(0.39)	
2	3506	0.09	401	0.14	
		(0.29)		(0.35)	
3	3506	0.04	401	0.06	
		(0.19)		(0.24)	
4	3506	0.03	401	0.06	
		(0.17)		(0.23)	
5	3506	0.01	401	0.05	
		(0.11)		(0.21)	

Continued in the next page

	Surviving individuals		Deceased individuals		P
	No. of obs.	Mean (S. D.)	No. of obs.	Mean (S. D.)	
Marital status (married as reference) (%)					
Separated/Divorced	3629	0.11	466	0.13	0.57
		(0.31)		(0.34)	
Widowed	3629	0.1	466	0.1	
		(0.3)		(0.3)	
Never married	3629	0.03	466	0.03	
		(0.17)		(0.16)	
Pre-Medicare health coverage (%)					
Medicaid	3618	0.04	462	0.09	<0.001
		(0.19)		(0.28)	
Champus/VA	3620	0.06	462	0.06	0.79
		(0.23)		(0.23)	
Private insurance (from self)	3532	0.48	443	0.44	0.04
		(0.5)		(0.5)	
Private insurance (from spouse)	3547	0.23	446	0.18	0.03
		(0.42)		(0.39)	
Pre-Medicare interview year	3632	2,000.24	467	1,997.18	<0.001
		(3.44)		(3.43)	
Medicare interview year	3338	2,002.38	439	1,999.23	<0.001
		(3.41)		(3.45)	
Birth year	3632	1,935.87	467	1,932.68	<0.001
		(3.63)		(3.54)	

Note: S.D: standard deviation. * The p values for continuous outcomes were obtained through t tests and those for categorical outcomes were through Chi-square tests. ** The spending was the amount of money spent on health care after being covered or Medicare for three to four years.

Table 4.3. The results of logit model predicting the probability of mortality after being enrolled in the first three to four years of Medicare coverage.

	Total health expenditure	**Out-of-pocket health expenditure**
Model summary		
No. of obs.	1752	4032
Likelihood ratio	216.98	467.22
p	< 0.01	< 0.01
Pseudo R2	0.14	0.22
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	-0.30	-0.47**
	0.18	(0.17)
Health expenditure (thousands)	0.0044**	0.0027
	(0.0010)	(0.0020)
Pre-Medicare characteristics		
Female	-0.93**	-0.72**
	(0.17)	(0.15)
Race		
Black	-0.03	0.08
	(0.21)	(0.19)
Other	-0.07	0.03
	(0.47)	(0.41)
Hispanic	-0.41	-0.44
	(0.33)	(0.30)
Regions		
Midwest	-0.16	-0.07
	(0.24)	(0.22)
South	0.21	0.11
	(0.21)	(0.20)
West	0.01	0.13
	(0.26)	(0.24)
Other	(omitted)	(omitted)
Years of education	0.02	0.01
	(0.03)	(0.02)
Income (thousands)	-0.0031	-0.0012
	(0.0051)	(0.0040)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Self-rated health status		
Very good	0.27	0.52
	(0.31)	(0.29)
Good	1.03**	1.19**
	(0.29)	(0.28)
Fair	1.22**	1.45**
	(0.32)	(0.31)
Poor	1.61**	1.73**
	(0.40)	(0.38)
CESD score (0 to 8)	-0.01	0.02
	(0.04)	(0.04)
Difficulty in ADL (0 to 5)		
1	-0.08	0.03
	(0.28)	(0.26)
2	0.20	0.45
	(0.44)	(0.38)
3	-0.28	0.04
	(0.63)	(0.57)
4/5	0.03	0.26
	(0.67)	(0.61)
Difficulty in IADL (0 to 5)		
1	-0.30	-0.54
	(0.35)	(0.34)
2	0.91	0.23
	(0.79)	(0.70)
Difficulty in mobility (0 to 5)		
1	0.42*	0.25
	(0.20)	(0.18)
2	0.69**	0.65**
	(0.24)	(0.21)
3	0.69*	0.36
	(0.32)	(0.30)
4	0.53	0.27
	(0.36)	(0.34)
5	1.42**	0.78
	(0.53)	(0.47)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	**Coefficients (S. E.)**
Marital status (married as reference)		
Separated/Divorced	0.23	0.32
	(0.25)	(0.21)
Widowed	0.55*	0.35
	(0.22)	(0.20)
Never married	0.22	0.09
	(0.42)	(0.40)
Pre-Medicare health coverage		
Medicaid	-0.41	-0.42
	(0.39)	(0.35)
Champus/VA	0.11	-0.04
	(0.32)	(0.29)
Private insurance (from self)	-0.21	-0.22
	(0.19)	(0.17)
Private insurance (from spouse)	0.05	-0.18
	(0.23)	(0.21)
Pre-Medicare interview year	-0.51	-0.47
	(0.38)	(0.36)
Medicare interview year	0.41	0.25
	(0.38)	(0.35)
Birth year	-0.12	-0.19*
	(0.11)	(0.10)
Hypertension	0.29	0.20
	(0.15)	(0.14)
Arthritis	0.02	-0.01
	(0.09)	(0.09)
Constant	422.23**	807.87**
	(88.75)	(53.32)

Note: Standard errors in parentheses. The coefficients are the log odds ratios of the independent variables. * p<0.05. ** p<0.01.

Table 4.4. The results of logit model predicting the incidence of hypertension in the first three to four years of Medicare coverage.

	Total health expenditure	**Out-of-pocket health expenditure**
Model summary		
No. of obs.	898	1905
Likelihood ratio	23.08	68.19
p	0.92	0.002
Pseudo R2	0.04	0.05
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	-0.09	-0.17
	(0.29)	(0.19)
Health expenditure (thousands)	-0.0018	0.0084
	(0.0025)	(0.0062)
Pre-Medicare characteristics		
Female	0.25	0.32
	(0.26)	(0.17)
Race		
Black	-0.29	0.45
	(0.48)	(0.25)
Other	-0.75	-0.02
	(0.81)	(0.44)
Hispanic	0.50	0.21
	(0.45)	(0.30)
Regions		
Midwest	0.44	0.32
	(0.40)	(0.26)
South	0.19	0.17
	(0.38)	(0.24)
West	0.26	-0.11
	(0.42)	(0.28)
Other	(omitted)	(omitted)
Years of education	0.04	0.01
	(0.05)	(0.03)
Income (thousands)	-0.0094	-0.0038
	(0.0094)	(0.0043)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	**Coefficients (S. E.)**
Self-rated health status		
Very good	0.11	0.52*
	(0.35)	(0.26)
Good	0.58	0.91**
	(0.34)	(0.26)
Fair	0.00	0.81*
	(0.53)	(0.32)
Poor	-0.46	-0.01
	(1.15)	(0.62)
CESD score (0 to 8)	0.04	0.03
	(0.08)	(0.05)
Difficulty in ADL (0 to 5)		
1	-0.26	-0.18
	(0.69)	(0.42)
2	1.90	0.98
	(1.02)	(0.65)
3	(omitted)	0.06
		(1.24)
4-5'	(omitted)	0.26
		(1.34)
Difficulty in IADL (0 to 5)		
1	0.04	0.14
	(0.67)	(0.45)
2	(omitted)	-0.10
		(1.11)
Difficulty in mobility (0 to 5)		
1	0.01	0.02
	(0.33)	(0.21)
2	-0.26	0.19
	(0.49)	(0.27)
3	-0.61	-0.30
	(0.83)	(0.48)
4	0.07	-0.11
	(0.86)	(0.52)
5	(omitted)	-1.37
		(1.30)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients(S. E.)
Marital status (married as reference)		
Separated/Divorced	0.00	-0.06
	(0.41)	(0.24)
Widowed	-0.07	-0.33
	(0.37)	(0.25)
Never married	-0.29	-0.27
	(0.80)	(0.44)
Pre-Medicare health coverage		
Medicaid	1.07	0.91*
	(0.64)	(0.36)
Champus/VA	0.79	0.57*
	(0.46)	(0.29)
Private insurance (from self)	-0.21	-0.14
	(0.31)	(0.19)
Private insurance (from spouse)	-0.62	-0.60*
	(0.41)	(0.26)
Pre-Medicare interview year	-0.31	0.17
	(0.64)	(0.50)
Medicare interview year	0.25	-0.02
	(0.63)	(0.50)
Birth year	0.04	-0.10
	(0.18)	(0.10)
Hypertension	(omitted)	(omitted)
Arthritis	-0.03	-0.07
	(0.13)	(0.10)
Constant	20.91	-107.95*
	(146.95)	(54.94)

Note: Standard errors in parentheses. The coefficients are the log odds ratios of the independent variables. * p<0.05. ** p<0.01.

Table 4.5. The characteristics associated with hypertension incidence in the first two years of Medicare coverage.

	Total health expenditure	Out-of-pocket health expenditure
Model summary		
No. of obs.	1455	2279
Likelihood ratio	54.48	65.58
p	0.03	< 0.01
Pseudo R2	0.06	0.04
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	-0.06	-0.13
	(0.22)	(0.16)
Health expenditure (thousands)	0.0009	0.0075
	(0.0015)	(0.0086)
Pre-Medicare characteristics		
Female	0.19	0.18
	(0.20)	(0.15)
Race		
Black	-0.42	-0.18
	(0.39)	(0.26)
Other	-0.46	0.08
	(0.63)	(0.37)
Hispanic	0.41	-0.12
	(0.37)	(0.28)
Regions		
Midwest	0.02	0.11
	(0.31)	(0.24)
South	0.25	0.36
	(0.29)	(0.22)
West	-0.22	0.28
	(0.33)	(0.24)
Other	(omitted)	(omitted)
Years of education	0.05	-0.02
	(0.04)	(0.03)
Income (thousands)	-0.0115	-0.0017
	(0.0067)	(0.0029)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Self-rated health status		
Very good	0.14	0.34
	(0.28)	(0.22)
Good	0.28	0.45*
	(0.29)	(0.23)
Fair	0.53	0.48
	(0.35)	(0.27)
Poor	0.24	0.60
	(0.60)	(0.41)
CESD score (0 to 8)	-0.08	-0.03
	(0.06)	(0.04)
Difficulty in ADL (0 to 5)		
1	0.37	0.27
	(0.41)	(0.32)
2	1.22	0.89
	(0.75)	(0.52)
3	(omitted)	0.54
		(0.76)
4-5'	1.18	0.90
	(1.15)	(0.85)
Difficulty in IADL (0 to 5)		
1	-0.13	-0.64
	(0.58)	(0.50)
2	-0.07	0.12
	(1.38)	(0.83)
Difficulty in mobility (0 to 5)		
1	0.63**	0.32
	(0.23)	(0.18)
2	0.29	0.10
	(0.35)	(0.26)
3	1.23**	0.49
	(0.44)	(0.36)
4	0.23	0.36
	(0.62)	(0.40)
5	0.11	-0.16
	(0.95)	(0.65)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients(S. E.)
Marital status (married as reference)		
Separated/Divorced	-0.77*	-0.61*
	(0.39)	(0.26)
Widowed	-0.29	0.03
	(0.30)	(0.21)
Never married	-1.54	-0.84
	(1.03)	(0.53
Pre-Medicare health coverage		
Medicaid	-1.14	0.47
	(1.05)	(0.35)
Champus/VA	-0.31	-0.22
	(0.43)	(0.31)
Private insurance (from self)	0.31	0.28
	(0.23)	(0.16)
Private insurance (from spouse)	0.15	0.08
	(0.26)	(0.20)
Pre-Medicare interview year	0.23	0.20
	(0.55)	(0.44)
Medicare interview year	-0.15	-0.02
	(0.55)	(0.44)
Birth year	0.02	-0.09
	(0.14)	(0.09)
Hypertension	(omitted)	(omitted)
Arthritis	-0.03	0.06
	(0.11)	(0.09)
Constant	-197.84*	-190.46**
	(89.29)	(49.74)

Note: Standard errors in parentheses. The coefficients are the log odds ratios of the independent variables. * p<0.05. ** p<0.01.

Table 4.6. The results of logit model predicting the arthritis incidence after three to four years of Medicare coverage.

	Total health expenditure	Out-of-pocket health expenditure
Model summary		
No. of obs.	610	1380
Likelihood ratio	24.27	23.97
p	0.83	0.87
Pseudo R2	0.06	0.03
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	0.08	-0.08
	(0.32)	(0.21)
Health expenditure (thousands)	0.0052*	0.0055
	(0.0018)	(0.0086)
Pre-Medicare characteristics		
Female	0.05	0.32
	(0.31)	(0.19)
Race		
Black	-0.04	-0.33
	(0.46)	(0.32)
Other	-0.57	-0.08
	(1.14)	(0.47)
Hispanic	-0.05	0.13
	(0.63)	(0.33)
Regions		
Midwest	0.21	-0.14
	(0.48)	(0.30)
South	0.22	0.09
	(0.45)	(0.27)
West	0.43	0.00
	(0.47)	(0.30)
Other	(omitted)	(omitted)
Years of education	0.00	0.02
	(0.05)	(0.03)
Income (thousands)	-0.0075	0.0014
	(0.0087)	(0.0022)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Self-rated health status		
Very good	0.16	0.35
	(0.41)	(0.26)
Good	0.39	0.51
	(0.41)	(0.27)
Fair	-0.21	0.46
	(0.63)	(0.36)
Poor	(omitted)	0.14
		(0.70)
CESD score (0 to 8)	-0.05	0.06
	(0.11)	(0.06)
Difficulty in ADL (0 to 5)		
1	-0.08	0.22
	(0.95)	(0.61)
2	(omitted)	(omitted)
3	(omitted)	(omitted)
4-5'	(omitted)	(omitted)
Difficulty in IADL (0 to 5)		
1	-0.44	-0.58
	(1.10)	(0.76)
2	(omitted)	(omitted)
Difficulty in mobility (0 to 5)		
1	0.11	-0.07
	(0.41)	(0.26)
2	0.73	0.52
	(0.66)	(0.37)
3	1.68	0.47
	(0.99)	(0.62)
4	-0.68	-0.06
	(1.50)	(0.79)
5	(omitted)	(omitted)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Marital status (married as reference)		
Separated/Divorced	0.17	-0.20
	(0.52)	(0.31)
Widowed	0.28	-0.05
	(0.44)	(0.29)
Never married	0.51	-0.61
	(0.82)	(0.76)
Pre-Medicare health coverage		
Medicaid	0.83	0.58
	(0.89)	(0.60)
Champus/VA	-0.27	-0.18
	(0.66)	(0.38)
Private insurance (from self)	0.33	-0.12
	(0.34)	(0.21)
Private insurance (from spouse)	0.69	0.15
	(0.39)	(0.25)
Pre-Medicare interview year	1.06	0.56
	(0.69)	(0.51)
Medicare interview year	-1.08	-0.47
	(0.68)	(0.51)
Birth year	-0.06	-0.08
	(0.22)	(0.12)
Hypertension	0.34	0.15
	(0.29)	(0.19)
Arthritis	(omitted)	(omitted)
Constant	154.20	-18.43
	(168.69)	(59.10)

Note: Standard errors in parentheses. The coefficients are the log odds ratios of the independent variables. * $p<0.01$.

Table 4.7. The results of logit model predicting the probability of being diagnosed with arthritis within first two years of Medicare coverage.

	Total health expenditure	**Out-of-pocket health expenditure**
Model summary		
No. of obs.	1067	1719
Likelihood ratio	37.42	33.72
p	0.45	0.62
Pseudo R2	0.046	0.026
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	0.19	0.09
	(0.22)	(0.18)
Health expenditure (thousands)	-0.0017	0.0215
	(0.0028)	(0.0136)
Pre-Medicare characteristics		
Female	0.31	0.23
	(0.21)	(0.16)
Race		
Black	0.52**	0.19
	(0.27)	(0.23)
Other	-0.50	-0.31
	(0.64)	(0.45)
Hispanic	0.27	0.13
	(0.40)	(0.29)
Regions		
Midwest	0.21	0.22
	(0.31)	(0.25)
South	0.17	0.20
	(0.29)	(0.23)
West	-0.17	0.12
	(0.33)	(0.26)
Other	(omitted)	(omitted)
Years of education	0.04	0.02
	(0.04)	(0.03)
Income (thousands)	-0.0018	-0.0023
	(0.0034)	(0.0030)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Self-rated health status		
Very good	0.48	0.45*
	(0.29)	(0.22)
Good	0.33	0.20
	(0.31)	(0.24)
Fair	0.48	0.10
	(0.39)	(0.31)
Poor	0.63	0.19
	(0.61)	(0.49)
CESD score (0 to 8)	0.08	0.08
	(0.06)	(0.05)
Difficulty in ADL (0 to 5)		
1	-0.02	0.47
	(0.69)	(0.50)
2	-0.91	-0.15
	(1.42)	(0.85)
3	-0.76	-0.13
	(1.54)	(1.26)
4-5'	(omitted)	(omitted)
Difficulty in IADL (0 to 5)		
1	-0.59	-0.44
	(0.66)	(0.56)
2	2.02	1.48
	(1.49)	(1.14)
Difficulty in mobility (0 to 5)		
1	0.27	0.29
	(0.25)	(0.20)
2	0.73	0.41
	(0.39)	(0.32)
3	-0.31	-0.11
	(0.81)	(0.54)
4	0.84	0.34
	(0.66)	(0.54)
5	1.09	0.16
	(1.72)	(1.38)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	**Coefficients (S. E.)**
Marital status (married as reference)		
Separated/Divorced	-0.75	-0.48
	(0.40)	(0.29)
Widowed	-0.20	-0.07
	(0.30)	(0.24)
Never married	-0.74	-0.32
	(0.77)	(0.55)
Pre-Medicare health coverage		
Medicaid	0.12	0.24
	(0.63)	(0.49)
Champus/VA	-0.63	-0.21
	(0.49)	(0.34)
Private insurance (from self)	0.10	-0.03
	(0.22)	(0.18)
Private insurance (from spouse)	-0.14	-0.13
	(0.29)	(0.23)
Pre-Medicare interview year	-0.27	-0.36
	(0.51)	(0.43)
Medicare interview year	0.29	0.38
	(0.51)	(0.43)
Birth year	0.03	-0.03
	(0.14)	(0.10)
Hypertension	0.11	0.08
	(0.20)	(0.16)
Arthritis	(omitted)	(omitted)
Constant	-104.66	6.50
	(84.66)	(49.40)

Note: Standard errors in parentheses. The coefficients are the log odds ratios of the independent variables. * $p<0.05$.

Table 4.8. The results of ordered logit model predicting the probability of having worse health status in the first three to four years of Medicare coverage.

	Total health expenditure	**Out-of-pocket health expenditure**
Model summary		
No. of obs.	1731	4029
Likelihood ratio	1136.2	2529.94
p	< 0.01	< 0.01
Pseudo R2	0.22	0.21
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	-0.05	-0.05
	(0.11)	(0.07)
Health expenditure (thousands)	0.0056**	0.0153**
	(0.0010)	(0.0030)
Pre-Medicare characteristics		
Female	-0.32**	-0.16*
	(0.10)	(0.07)
Race		
Black	0.08	0.12
	(0.14)	(0.09)
Other	0.41	0.26
	(0.30)	(0.17)
Hispanic	0.30	0.23
	(0.20)	(0.12)
Regions		
Midwest	-0.22	-0.10
	(0.15)	(0.10)
South	0.16	0.12
	(0.14)	(0.09)
West	0.00	0.04
	(0.16)	(0.10)
Other	(omitted)	(omitted)
Years of education	-0.03	-0.05**
	(0.02)	(0.01)
Income (thousands)	-0.0020	-0.0019
	(0.0029)	(0.0011)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Self-rated health status		
Very good	1.27**	1.40**
	(0.16)	(0.11)
Good	2.65**	2.70**
	(0.17)	(0.11)
Fair	3.80**	3.89**
	(0.21)	(0.14)
Poor	4.65**	5.09**
	(0.30)	(0.20)
CESD score (0 to 8)	0.10**	0.08**
	(0.03)	(0.02)
Difficulty in ADL (0 to 5)		
1	-0.14	0.09
	(0.21)	(0.14)
2	0.32	0.51*
	(0.35)	(0.24)
3	0.68	0.82*
	(0.47)	(0.34)
4-5'	-0.09	0.06
	(0.60)	(0.41)
Difficulty in IADL (0 to 5)		
1	-0.13	-0.11
	(0.25)	(0.18)
2	-0.11	0.29
	(0.67)	(0.46)
Difficulty in mobility (0 to 5)		
1	0.41**	0.30**
	(0.13)	(0.08)
2	0.71**	0.53**
	(0.17)	(0.11)
3	0.65**	0.36*
	(0.24)	(0.15)
4	1.05**	0.67**
	(0.29)	(0.18)
5	0.54	0.44
	(0.45)	(0.30)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Marital status (married as reference)		
Separated/Divorced	0.12	0.07
	(0.17)	(0.10)
Widowed	0.44**	0.08
	(0.14)	(0.09)
Never married	0.17	0.14
	(0.29)	(0.18)
Insurance		
Medicaid	0.17	0.54**
	(0.27)	(0.16)
Champus/VA	0.22	0.23
	(0.22)	(0.13)
Private insurance (from self)	-0.32**	-0.15*
	(0.11)	(0.07)
Private insurance (from spouse)	-0.29*	-0.09
	(0.14)	(0.09)
Pre-Medicare interview year	0.30	0.28
	(0.25)	(0.19)
Medicare interview year	-0.44	-0.28
	(0.24)	(0.19)
Birth year	0.05	0.00
	(0.07)	(0.04)
Hypertension	0.15	0.16**
	(0.10)	(0.06)
Arthritis	0.01	0.05
	(0.06)	(0.04)
Cut 1	-190.97	-1.04
	(57.24)	(20.89)
Cut 2	-188.50	1.41
	(57.23)	(20.89)
Cut 3	-186.35	3.54
	(57.23)	(20.89)
Cut 4	-184.15	5.75
	(57.23)	(20.89)

Note: Standard errors in parentheses. The coefficients are the log odds ratios of the independent variables. * p<0.05. ** p<0.01.

Table 4.9. The results of the ordered logit model predicting the probability of having one more score over the CESD scale (0 to 8) after the first three to four years of Medicare coverage.

	Total health expenditure	**Out-of-pocket health expenditure**
Model summary		
No. of obs.	1658	3922
Likelihood ratio	561.54	1353.13
p	< 0.01	< 0.01
Pseudo R2	0.11	0.11
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	-0.21	-0.12
	(0.11)	(0.08)
Health expenditure (thousands)	0.0018*	0.0029
	(0.0007)	(0.0015)
Pre-Medicare characteristics		
Female	0.20	0.17*
	(0.11)	(0.07)
Race		
Black	0.09	0.15
	(0.14)	(0.09)
Other	1.13**	0.52**
	(0.30)	(0.18)
Hispanic	-0.33	-0.18
	(0.21)	(0.13)
Regions		
Midwest	-0.23	-0.01
	(0.15)	(0.10)
South	-0.06	0.09
	(0.14)	(0.09)
West	-0.19	-0.02
	(0.16)	(0.11)
Other	(omitted)	(omitted)
Years of education	-0.02	-0.04**
	(0.02)	(0.01)
Income (thousands)	0.0032	-0.0015
	(0.0031)	(0.0015)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	**Coefficients (S. E.)**
Self-rated health status		
Very good	0.25	0.25*
	(0.16)	(0.11)
Good	0.54**	0.53**
	(0.16)	(0.11)
Fair	0.71**	0.80**
	(0.19)	(0.13)
Poor	0.93**	1.07**
	(0.28)	(0.18)
CESD score (0 to 8)	0.44**	0.43**
	(0.03)	(0.02)
Difficulty in ADL (0 to 5)		
1	0.22	0.20
	(0.20)	(0.14)
2	0.30	0.33
	(0.35)	(0.23)
3	0.59	0.29
	(0.42)	(0.30)
4-5'	0.23	0.18
	(0.51)	(0.36)
Difficulty in IADL (0 to 5)		
1	-0.18	-0.22
	(0.25)	(0.18)
2	-0.25	0.44
	(0.61)	(0.44)
Difficulty in mobility (0 to 5)		
1	0.24	0.27**
	(0.13)	(0.08)
2	0.61**	0.56**
	(0.17)	(0.11)
3	0.70**	0.47**
	(0.23)	(0.14)
4	0.80**	0.67**
	(0.27)	(0.17)
5	0.05	0.06
	(0.41)	(0.29)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Marital status (married as reference)		
Separated/Divorced	0.10	0.10
	(0.17)	(0.10)
Widowed	0.21	-0.03
	(0.14)	(0.10)
Never married	0.18	0.20
	(0.27)	(0.18)
Pre-Medicare health coverage		
Medicaid	0.15	0.29
	(0.28)	(0.16)
Champus/VA	0.03	0.05
	(0.23)	(0.13)
Private insurance (from self)	-0.02	-0.03
	(0.12)	(0.08)
Private insurance (from spouse)	0.14	0.18
	(0.15)	(0.09)
Pre-Medicare interview year	-0.17	-0.21
	(0.27)	(0.21)
Medicare interview year	0.17	0.18
	(0.27)	(0.20)
Birth year	-0.09	0.00
	(0.07)	(0.04)
Hypertension	-0.06	0.02
	(0.10)	(0.07)
Arthritis	0.01	0.01
	(0.06)	(0.04)
Cut 1	-173.17	-58.04
	(61.23)	(21.89)
Cut 2	-172.01	-56.88
	(61.23)	(21.89)
Cut 3	-171.17	-56.09
	(61.22)	(21.88)
Cut 4	-170.64	-55.52
	(61.22)	(21.88)
Cut 5	-169.93	-54.89
	(61.22)	(21.88)

Continued in the next page

	Total health expenditure	Out-of-pocket health expenditure
	Coefficients (S. E.)	**Coefficients (S. E.)**
Cut 6	-169.27	-54.30
	(61.22)	(21.88)
Cut 7	-168.39	-53.49
	(61.22)	(21.88)
Cut 8	-166.84	-52.16
	(61.22)	(21.89)

Note: Standard errors in parentheses. The coefficients are the log odds ratios of the independent variables. * p<0.05. ** p<0.01. There was no p value reported for cut one to eight.

Table A.1. The mean total health expenditure in different groups of Medicare enrollees age 65 and over in 1996, 2000 and 2008.

Categories	Subgroups	1996	2000	2008	Annual growth rate (%)
Total health expenditure		5,423	5,878	9,303	4.5
Sex	Male	5,546	6,173	9,068	4.1
	Female	5,332	5,659	9,482	4.8
Age	65-74	4,076	5,272	8,214	5.84
	75-84	5,774	5,662	9,375	4.04
	85 and over	9,895	8,508	12,549	1.98
Race	White	5,468	5,932	9,321	4.44
	Black	5,569	5,659	10,211	5.05
	Other	3,619	4,924	7,426	5.99
Ethnicity	Hispanic	5,322	5,149	9,075	4.45
Years of education	0-8	4,660	7,175	10,065	6.42
	9-12	5,121	5,341	9,087	4.78
	>13	6,343	5,984	9,348	3.23
Income relative to poverty line	Poor/Negative	7364	6329	9157	1.82
	Near poor	6,176	5,544	11,325	5.05
	Low income	5,480	6,359	9,384	4.48
	Middle income	4,643	5,620	9,394	5.87
	High income	5,227	5,768	8,771	4.31
Regions	North-east	5,382	5,420	10,297	5.41
	Midwest	5,930	5,975	9,081	3.55
	South	4,431	5,759	8,200	5.13
	West	4,232	4,648	9,154	6.43
Marital status	Married	4,195	5,907	8,928	6.29
	Widowed	5,478	5,894	10,465	5.39
	Divorced	8,195	6,039	7,645	-0.58
	Other	5,503	5,129	9,621	4.66
Self-rated health status	Excellent	2,969	2,392	4,454	3.38
	Very good	3,350	3,701	6,117	5.02
	Good	5,163	4,619	8,551	4.2
	Fair	7,979	8,563	13,508	4.39
	Poor	10,624	13,555	18,385	4.57
Self-rated mental health status	Excellent	4,461	4,125	7,091	3.86
	Very good	4,228	4,250	7,317	4.57
	Good	6,052	5,711	9,462	3.72
	Fair/poor	8,147	10,109	14,219	4.64
ADL helper		13,644	14,706	20,914	3.56
IADL helper		11,567	12,010	17,340	3.37
Activity limitation		10,795	10,322	14,610	2.52
Cognitive limitation		11,389	10,177	14,574	2.05
Any limitation		7,270	7,942	11,519	3.84
Medicaid		7,652	8,177	12,907	4.36
Private insurance		5,525	6,044	9,258	4.3

Continued in the next page

Categories	Subgroups	1996	2000	2008	Annual growth rate (%)*
Chronic conditions	Diabetes		8,913	12,317	4.04
	Asthma		8,654	12,316	4.41
	Angina		10,714	13,257	2.66
	Stroke		11,434	12,732	1.34
	Emphysema		9,185	11,952	3.29
	Hypertension		6,478	10,352	5.86
	Coronary heart disease		10,772	13,720	3.02
	Heart attack		11,582	14,024	2.39
	Other heart disease		9,052	11,917	3.44
	Joint pain		6,389	9,998	5.60
Health behavior	Smoking		4,660	7,931	6.65

Note: Statistics estimated based on the annual MEPS-HC (household component) datasets from 1996 to 2008. * Annual growth rate was calculated from 2000 to 2008 for chronic health conditions and the selected health behavior. The mean total Medicare spending grew for 58.3% from 2000 to 2008, 5.60% annually.

Table B.1. Results from the modified Park test for one- and two-part GLM estimations.

One-part expenditure model			
Regression coefficient	**SE**	***t***	***p***
1.77	0.06	31.57	<0.001
Adjusted Wald test		***f***	***p***
λ=0		996.77	<0.001
λ=1		187.68	<0.001
λ=2		17.41	<0.001

Two-part expenditure model			
Regression coefficient	**SE**	***t***	***p***
<0.001	<0.0001	26.86	<0.001
Adjusted Wald test		***f***	***p***
λ=0		721.52	<0.001
λ=1		>1000	<0.001
λ=2		>100000	<0.001

Table B.2. The amounts of health expenditure prediction and the mean expenditure among Medicare enrollees aged 65 years and over from 1996 to 2008, specified by different functional status.

	All Medicare enrollees age 65 and over	S.E.	Needing help in ADL	S.E.	Needing help in IADL	S.E.	Medicaid	S.E.	Limitation in activities	S.E.
Total health expenditure	7,045.0	87.0	17,581.4	562.0	14,303.9	345.8	9,378.9	304.4	12,404.3	242.7
AAPCC prediction	3,168.7	16.0	3,524.3	32.9	3,555.5	26.2	3,147.3	28.2	3,405.2	24.8
1-part OLS model	6,902.1	44.4	17,357.2	137.0	14,166.7	103.0	9,174.3	141.6	12,295.3	85.4
2-part OLS model	6,903.2	44.5	17,366.0	138.0	14,173.4	103.7	9,160.2	142.1	12,300.5	85.8
1-part transformed OLS model	3,687.2	26.5	9,894.6	121.0	7,812.9	86.8	4,572.6	88.4	6,795.6	66.3
2-part transformed OLS model	3,619.6	26.5	9,832.1	121.4	7,754.4	87.0	4,472.8	88.3	6,742.8	66.4
2-part transformed OLS model (1 smearing factor)	7,491.4	54.9	20,349.3	251.2	16,049.1	180.1	9,257.3	182.9	13,955.5	137.4
1-part Gaussian GLM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1-part Poisson GLM	6,894.0	44.8	17,339.5	190.6	14,157.3	129.6	9,166.4	153.6	12,293.0	99.7
1-part Gamma GLM	6,976.7	46.8	18,336.9	199.2	14,925.2	141.3	9,040.7	157.8	12,822.3	107.6
2-part Gaussian GLM	6,587.5	47.3	16,853.7	285.2	13,493.1	167.0	8,858.9	174.9	11,806.9	120.2
2-part Poisson GLM	6,897.9	44.8	17,345.6	190.3	14,162.9	129.1	9,157.0	154.1	12,296.6	99.6
2-part Gamma GLM	6,956.1	46.3	18,133.5	194.4	14,778.5	137.5	9,065.7	156.4	12,717.1	105.5

Table B.2. Continued.

	Limitation in cognition	S.E.	Any limitations	S.E.	Private coverage	S.E.	Poor health status	S.E.	Poor or fair mental health status	S.E.
Total health expenditure	12,758.1	316.7	9,136.0	128.2	6,948.8	118.3	16,148.0	573.1	11,197.7	343.2
AAPCC prediction	3,532.5	26.4	3,306.8	17.8	3,073.9	19.3	3,088.3	31.5	3,306.6	26.0
1-part OLS model	12,565.7	128.8	9,004.1	53.8	6,846.0	52.8	15,995.4	149.8	11,017.6	114.1
2-part OLS model	12,576.7	129.3	9,007.7	53.9	6,849.0	53.0	16,004.8	150.3	11,034.0	114.2
1-part transformed OLS model	6,992.6	100.4	4,874.8	36.3	3,735.3	34.3	8,998.5	133.5	5,887.4	82.1
2-part transformed OLS model	6,923.3	100.8	4,816.5	36.3	3,684.2	34.4	8,934.8	133.8	5,820.2	82.4
2-part transformed OLS model (1 smearing factor)	14,329.1	208.7	9,968.7	75.0	7,625.1	71.1	18,492.3	276.9	12,046.0	170.5
1-part Gaussian GLM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1-part Poisson GLM	12,559.6	148.7	9,004.2	57.5	6,844.7	56.3	16,009.8	194.6	11,019.5	125.4
1-part Gamma GLM	13,456.1	168.6	9,224.2	61.8	6,941.1	59.2	16,818.8	217.5	11,753.9	138.4
2-part Gaussian GLM	12,070.3	161.3	8,592.9	63.6	6,619.9	61.3	15,485.2	268.1	10,526.7	174.3
2-part Poisson GLM	12,572.5	147.8	9,006.8	57.4	6,846.9	55.9	16,016.6	192.6	11,033.6	125.1
2-part Gamma GLM	13,339.7	163.6	9,172.4	60.7	6,888.6	57.9	16,678.6	210.1	11,653.5	134.7

Table B.3. Comparison of the mean square errors (MSE) and mean absolute prediction errors (MAPE) in expenditure models that predicted the amount of health spending of Medicare enrollees age 65 and over.

	Mean absolute prediction error	Mean square error (×1,000,000)
1-part OLS model	5,927.66	131
2-part OLS model	5,923.29	131
1-part transformed OLS model	5,162.99	147
2-part transformed OLS model	5,167.25	147
2-part transformed OLS model (1 smearing factor)	6,137.48	136
1-part Gaussian GLM	5,846.21	130
1-part Poisson GLM	5,902.93	133
1-part Gamma GLM	5,793.98	127
2-part Gaussian GLM	5,846.47	130
2-part Poisson GLM	5,889.91	132
2-part Gamma GLM	7,045.00	203

Table B.4. Comparison of the R^2 produced after regressing the observed total health expenditure on the predicted values in all expenditure models.

Total health expenditure models				**Adjusted Wald test ($\boldsymbol{\beta_1} = 1$)**	
	R^2	Regression coefficient	S.E.	*f*	*p*
1-part OLS model	0.15	1	0.03	0	0.96
2-part OLS model	0.15	1	0.03	0	0.96
1-part transformed OLS model	0.14	1.54	0.04	174.95	<0.001
2-part transformed OLS model	0.13	1.55	0.04	173.73	<0.001
2-part transformed OLS model (1 smearing factor)	0.13	0.74	0.02	163.8	<0.001
1-part Gaussian GLM	0	-	-		-
1-part Poisson GLM	0.16	1	0.03	0	0.95
1-part Gamma GLM	0.15	0.89	0.02	20.6	<0.001
2-part Gaussian GLM	0.17	0.93	0.03	3.8	0.052
2-part Poisson GLM	0.16	1	0.03	0.01	0.92
2-part Gamma GLM	0.14	0.91	0.03	11.47	<0.001

Note: the functional form of the models, $(HealthExpenditure)=\beta_0+\beta_1\,(PredictedValues)+\varepsilon$.

Table B.5. Model fit tests for model for the expenditure models used in Chapter 2.

	Modified Homer-Lemeshow test		Correlation test		Link test		RESET test	
	f	***P***	**Pearson's ρ**	***P***	***f***	***P***	***f***	***p***
1-part OLS model	30.91	0	-0.11	1	-	0*	0.24	0.62
2-part OLS model	29.89	0	-0.1	0.97	-	0*	0.06	0.8
1-part transformed OLS model	230.64	0	0.17	0	130.36	0	292.07	0
2-part transformed OLS model	250.33	0	0.17	0	125.82	0	282.56	0
2-part transformed OLS model (1 smearing factor)	16.56	0	-0.13		-	0*	245.48	0
1-part Poisson GLM	1.55	0.12	-0.08	0.95	-	0*	197.97	0
1-part Gamma GLM	1.82	0.05	-0.03	0	-	0*	271.13	0
2-part Gaussian GLM	29.75	0	-0.26	0.05	-	0*	63	0
2-part Poisson GLM	1.61	0.1	-0.09	0.92	-	0*	197.87	0
2-part Gamma GLM	1.07	0.38	-0.05	0	-	0*	268.06	0

Table B.6. Results of the logit regression predicting the probability of incurring health spending among the Medicare enrollees age 65 and over.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Year main effects		-2.03*	-0.95	-1.92*	-1.75*	-0.39	0.54	0.01	0.24	-1.28	-0.99	-1.30	-1.96*
		Interaction terms											
Age	0.00	0.03	0.04	0.03	0.02	0.03	0.00	0.09*	0.04	0.05	0.09*	0.04	0.07*
Female	0.08	0.38	0.32	0.39	1.00*	0.34	0.49	0.68*	0.21	0.51	0.21	0.83*	0.71*
Race													
Black	-0.86*	0.25	0.54	1.10*	0.69	0.39	0.21	0.74	0.49	0.92*	0.27	0.65	0.25
Other	-0.11	-1.25	-0.44	-0.97	0.07	-0.86	-0.12	0.09	2.22	1.12	-0.07	-1.10	-0.44
Ethnicity													
Hispanic	-0.22	0.79	0.50	0.97	0.28	-0.73	-0.43	-0.34	-0.10	0.38	-0.32	0.03	0.29
Region													
Midwest	0.10	0.10	0.30	0.22	-0.71	-0.55	-0.59	-0.41	-0.19	0.63	-0.22	-0.20	-0.15
South	-0.15	0.55	1.24*	0.70	0.43	0.04	0.03	-0.29	-0.03	0.02	0.04	0.11	0.91
West	-0.38	0.61	0.91	0.81	0.08	0.51	0.48	0.49	0.00	0.69	0.63	0.51	0.24
Income (thousand)	0.00	0.02	0.01	0.02*	0.01	0.02	0.00	0.01	0.01	0.00	0.00	0.01	0.04*
Years of education	0.08*	0.07	0.04	0.07	0.11*	0.02	0.02	-0.01	0.04	0.08	0.05	0.08	0.05
Health status													
Very good	0.79*	-0.13	-1.03	-0.50	-0.45	0.12	-0.69	-0.30	-0.90*	0.11	0.49	-0.52	0.09
Good	1.23*	-0.16	-1.09	-0.55	-0.30	-0.10	-0.57	-0.24	-0.84	-0.01	-0.19	-0.42	-0.03
Fair	1.91*	-0.06	-1.21	-0.29	0.28	0.20	-0.01	0.66	-0.60	0.14	-0.04	-0.55	0.79
Poor	1.48	0.68	-0.57	-0.79	0.31	1.35	0.78	1.68	-1.02	1.22	0.25	1.61	(omitted)
Mental health status													
Very good	-0.52*	0.34	0.52	1.02*	0.82*	-0.12	0.88*	0.06	0.51	0.07	-0.18	0.18	0.37
Good	-0.48	0.65	0.37	0.61	0.66	-0.07	0.17	-0.52	0.15	-0.30	-0.02	0.06	0.26
Fair/Poor	-0.61	-0.32	0.37	0.85	0.49	-0.67	-0.27	-0.35	-0.31	0.18	0.57	-0.37	0.02
Marital status													
Widowed	0.22	-0.73	-0.40	-0.70	-0.74*	-0.18	-1.11*	-1.02*	0.02	-0.65	-0.42	-0.39	-0.43
Divorced	-0.96*	0.68	0.34	-0.01	0.66	1.32*	0.26	-0.29	-0.05	0.41	1.24*	0.87	0.12
Other	-0.75	-0.51	0.58	-0.11	-0.22	0.68	-0.02	-0.08	-0.18	-0.40	0.72	0.59	0.47

Continued in the next page

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
ADL helper	0.59	0.62	-1.58	-1.07	-0.88	1.25	0.77	-0.97	-0.67	-0.66	0.77	-0.27	1.43
IADL helper	-0.59	1.88*	1.54	1.28	2.15*	1.36*	0.77	0.42	0.67	1.19	-0.54	1.55	0.23
Activity limitation	1.28*	-0.97	-0.67	-1.18	-0.41	-0.22	-1.09	-1.35	0.54	-0.88	-0.35	-1.72*	-0.78
Cognitive limitation	-0.70	0.15	0.16	0.59	0.44	0.41	0.63	0.87	-0.31	0.41	0.60	1.85*	0.55
Any limitation	0.63*	0.19	-0.21	0.35	0.56	-0.21	0.24	0.66	-0.06	-0.01	0.44	0.35	-0.38
Medicaid	0.59	0.29	-0.33	-0.71	-0.62	-0.51	-0.73	-0.80	0.03	-0.39	-0.82	-0.02	-0.27
Private insurance	1.13*	-0.29	-0.65	-0.43	-1.01*	-0.64	-0.72	-0.32	-0.42	-0.55	-0.19	0.11	-0.44
Constant	1.04												

Note: * $p < 0.05$. The unweighted sample size is 41410 and the weighted sample size is 431,976,923 observations after adjusting for complex survey design. The *f* statistics and *p* value for the logit model are not reported due to the survey design adjustment.

Table B.7. The regression coefficients of one-part expenditure model (Poisson GLM with log link).

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Year main effects		0.21	-0.26	-0.09	0.41	0.32	0.83*	0.54	0.94*	0.53	0.75*	0.85*	0.84*
		Interaction terms											
Age	0.009	-0.002	-0.005	-0.011	-0.018	-0.013	-0.008	-0.020*	-0.015	-0.013	-0.013	-0.007	-0.01
Female	-0.21*	0	0.29*	0.03	0.13	0.1	0.1	0.13	0.16	0.23	0.2	0.19	0.22
Race													
Black	-0.07	-0.05	0.15	-0.18	-0.04	0.04	-0.11	-0.11	-0.06	-0.08	0.03	-0.01	0.07
Other	-0.2	-0.41	0.46	-0.38	-0.54	0.13	0.13	0.12	-0.4	-0.19	0.35	0.07	0
Hispanic	-0.01	-0.08	-0.3	0.12	-0.09	-0.03	-0.02	-0.08	-0.24	-0.1	0.02	-0.03	0.06
Region													
Midwest	0.09	-0.22	-0.06	-0.34*	-0.07	-0.12	-0.04	-0.03	-0.09	0.18	-0.08	-0.04	-0.18
South	-0.21	0.36*	0.15	-0.05	0.13	0.19	0.08	0.12	0.14	0.21	0.19	0.23	0.07
West	-0.26*	0.40*	0.26	0.17	0.04	0.13	0.21	-0.04	0.17	0.25	0.08	0.13	0.22
Income (thousands)	-0.003	0.007	0	0.005	0.004	0.002	0.004	0.001	0.005	0.002	0.005	0.003	0.005
Years of education	0.04*	-0.05	-0.01	0.01	-0.04	-0.02	-0.04*	0	-0.03	-0.01	-0.02	-0.02	-0.02
Health status													
Very good	0.14	0.17	0.25	0.21	0.26	0.15	0.12	0.22	0.15	0.1	0.01	-0.01	0.23
Good	0.42*	0.31	0.2	0.23	0.09	0.05	0.27	0.27	0.06	0.1	-0.02	0.07	0.23
Fair	0.58*	0.26	0.26	0.51	0.4	0.1	0.34	0.46*	0.39	0.2	0.19	0.3	0.43
Poor	0.66*	0.61*	0.64*	0.56	0.53*	0.17	0.39	0.79*	0.59*	0.36	0.39	0.58*	0.47*
Mental health status													
Very good	-0.15	0.17	0	0.01	0.02	0.13	0.08	0.12	0	0.09	0.06	0.11	0.03
Good	-0.11	-0.2	0.15	0.05	0.01	0.04	0	-0.05	0.04	0.1	0.06	0.05	0.03
Fair/Poor	-0.38	0.21	0.26	0.24	0.3	0.39	0.12	0.23	0.03	0.41	-0.09	-0.1	0.19

Continued in the next page

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
		Interaction terms											
Marital status													
Widowed	0.14	-0.19	-0.06	0.02	-0.17	-0.09	-0.2	-0.04	-0.11	-0.16	-0.15	-0.19	-0.15
Divorced	0.59*	-0.57	-0.44	-0.39	-0.52	-0.48	-0.41	-0.75*	-0.78*	-0.63*	-0.66*	-0.67*	-0.87*
Other	0.11	-0.35	-0.13	-0.06	-0.38	-0.09	-0.06	-0.26	-0.42	-0.23	-0.11	-0.03	-0.27
ADL helper	0.14	0.59*	0.24	-0.05	0.41	0.32	0.23	0.17	0.11	0.19	0.41*	0.33	0.27
IADL helper	0.26	0.06	-0.22	-0.03	-0.18	0.08	-0.13	-0.11	0.14	-0.06	-0.06	-0.19	-0.05
Activity limitation	0.45*	-0.23	-0.24	0.01	-0.26	-0.3	-0.13	-0.32	-0.48*	-0.18	-0.2	-0.34	-0.27
Cognitive limitation	0.35	-0.57	-0.26	-0.5	-0.43	-0.42	-0.38	-0.46	-0.35	-0.31	-0.34	-0.15	-0.41
Any limitation	0.40*	0.1	0.09	-0.08	0.16	0.2	-0.12	0.05	-0.01	-0.06	0.05	-0.06	-0.16
Medicaid	0.33*	-0.19	-0.16	0.08	-0.05	-0.23	-0.24	-0.34	-0.07	-0.12	-0.28	-0.2	-0.19
Private insurance	0.26*	-0.06	-0.08	-0.05	-0.03	-0.11	-0.14	-0.13	-0.09	-0.14	-0.08	-0.12	-0.17
Constant	7.09*												

Note: * denoted the statistical significance at the level of 0.05. The unweighted sample size is 41,634 and the weighted sample size is 434,421,492. There was no *f* statistics reported.

Table B.8. The reduced model (one-part Poisson GLM) of total health expenditure.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Year main effects		0.66	0.39	0.51	0.79*	0.76*	0.87*	0.80*	1.21*	0.73*	0.82*	0.93*	1.10*
		Interaction terms											
Age	0.030*	-0.002	-0.009	-0.015	-0.023*	-0.012	-0.014	-0.022*	-0.017	-0.018	-0.016	-0.015	-0.017*
Female	-0.10	-0.13	0.24	0.03	-0.01	0.07	0.01	0.08	0.06	0.16	0.14	0.09	0.12
Race													
Black	0.01	-0.11	0.09	-0.20	-0.13	0.02	-0.11	-0.04	-0.11	-0.09	-0.03	-0.06	0.09
Other	-0.27	-0.37	0.33	-0.35	-0.58	0.11	0.10	0.17	-0.29	-0.17	0.30	0.09	0.12
Hispanic	0.04	-0.14	-0.41	0.01	-0.17	-0.12	-0.10	-0.21	-0.40	-0.21	-0.09	-0.04	-0.06
Region													
Midwest	0.10	-0.14	-0.04	-0.28	-0.02	-0.11	-0.08	-0.03	-0.12	0.12	-0.01	-0.04	-0.22
South	-0.20	0.38*	0.21	0.07	0.20	0.17	0.06	0.11	0.11	0.23	0.24	0.22	-0.01
West	-0.25	0.52*	0.45*	0.20	0.07	0.11	0.19	-0.03	0.19	0.31	0.18	0.15	0.16
Income (thousand)	-0.003	0.004	-0.001	0.003	0.005	0.001	0.004	0.001	0.004	0.000	0.004	0.003	0.005
Years of education	0.04*	-0.05	-0.02	-0.01	-0.05*	-0.01	-0.03	0.00	-0.03*	-0.01	-0.02	-0.02	-0.02
Health status													
Very good	0.11	-0.31	-0.35	-0.34	-0.04	-0.32	-0.16	-0.19	-0.22	-0.14	-0.14	-0.21	-0.20
Good	0.51*	-0.34	-0.37*	-0.36	-0.21	-0.42*	-0.08	-0.23	-0.36*	-0.18	-0.21	-0.21	-0.26
Fair	0.95*	-0.42*	-0.41	-0.16	-0.04	-0.44*	-0.12	-0.24	-0.23	-0.10	-0.15	-0.22	-0.25
Poor	1.25*	0.03	-0.10	-0.14	0.12	-0.29	-0.11	0.07	-0.13	0.04	0.02	0.00	-0.21
Constant	7.45*												

Note: * indicated p <0.05. The unweighted sample size is 42,391 in this model and the weighted sample size is 441,883,397.6. There is no *f* statistics or *p* value reported.

Table B.9. The regression coefficients (Poisson GLM) in the extended model with chronic condition variables from 2000 to 2008.

		2000	2001	2002	2003	2004	2005	2006	2007	2008
Year main effects			0.01	0.50	0.32	0.87*	0.43	0.53	0.67*	0.63
			Interaction terms							
Age		0.000	-0.001	0.004	-0.013	-0.007	-0.002	-0.002	-0.002	-0.001
Female		0.00	0.00	-0.03	-0.02	0.03	0.08	0.09	0.03	0.03
Race	Black	-0.11	0.09	-0.06	0.00	-0.02	-0.02	0.10	0.08	0.08
	Other	-0.71*	0.65	0.59*	0.43*	0.14	0.34	0.80*	0.71*	0.53*
Hispanic		-0.01	0.05	0.16	-0.07	-0.23	-0.09	-0.05	-0.04	0.09
Region	Midwest	0.11	-0.19	-0.05	-0.05	-0.09	0.16	-0.06	-0.10	-0.24
	South	0.01	-0.07	-0.13	-0.12	-0.08	-0.02	0.01	-0.06	-0.16
	West	-0.06	-0.10	0.00	-0.25	-0.01	0.00	-0.02	-0.15	-0.01
Income		0.002	-0.002	0.001	-0.003	0.001	-0.003	0.000	0.000	0.001
Education		0.012	0.018	-0.001	0.031	-0.004	0.024	0.008	0.011	0.012
Health status	Very good	0.27	0.04	-0.04	0.03	-0.07	-0.09	-0.17	-0.18	0.02
	Good	0.31*	0.08	0.23	0.28	0.04	0.11	0.01	0.09	0.19
	Fair	0.63*	-0.05	0.10	0.21	0.16	0.01	-0.04	0.11	0.16
	Poor	0.76*	-0.17	-0.01	0.46	0.19	0.01	0.08	0.22	0.14
Mental health status	Very good	-0.09	0.09	0.05	0.11	-0.02	0.01	-0.01	0.08	-0.01
	Good	-0.07	0.02	-0.06	0.00	-0.01	0.08	0.03	0.02	0.04
	Fair/Poor	0.00	0.09	-0.16	-0.05	-0.27	0.03	-0.35	-0.37	-0.17
Marital status	Widowed	-0.04	0.04	-0.03	0.12	0.08	0.03	-0.05	-0.03	0.06
	Divorced	0.01	0.09	0.18	-0.14	-0.17	-0.07	-0.12	-0.08	-0.24
	Other	-0.06	0.15	0.09	-0.31	-0.27	-0.01	0.05	0.14	-0.11

Continued in the next page

		Main effects	2001	2002	2003	2004	2005	2006	2007	2008
			Interaction terms							
ADL helper		0.49*	-0.10	-0.10	-0.26	-0.27	-0.11	0.08	-0.06	-0.02
IADL helper		0.17	0.25	-0.06	-0.06	0.23	0.05	0.07	-0.10	0.02
Activity limitation		0.10	-0.05	0.16	-0.02	-0.16	0.08	0.08	-0.06	0.05
Cognitive limitation		-0.15	-0.02	0.05	0.12	0.09	0.17	0.03	0.33	0.11
Any limitation		0.35*	0.15	-0.18	0.03	-0.06	-0.10	0.04	-0.10	-0.18
Medicaid		0.09	0.16	0.06	-0.13	0.08	0.09	-0.11	0.00	-0.03
Private insurance		0.23*	-0.04	-0.12	-0.12	-0.11	-0.14	-0.02	-0.11	-0.15
Chronic conditions	Diabetes	0.27*	-0.09	-0.10	-0.17	-0.01	0.15	-0.01	-0.15	-0.20
	Asthma	0.30*	-0.36*	-0.05	-0.09	-0.24	-0.06	-0.17	-0.15	-0.22
	Angina	0.12	-0.08	-0.17	0.07	0.03	0.05	-0.11	-0.14	-0.27*
	Stroke	0.22	0.10	-0.09	0.01	-0.14	-0.39*	-0.09	0.00	-0.16
	Emphysema	0.20	0.10	-0.12	-0.17	-0.06	-0.19	0.05	-0.02	-0.06
	Hypertension	0.07	0.08	0.00	0.12	0.07	-0.06	0.06	0.16	0.11
	Coronary heart disease	0.13	0.17	0.21	-0.01	0.22	0.13	0.28	-0.07	0.18
	Heart attack	0.45*	-0.29	-0.20	-0.26	-0.44*	-0.31	-0.45*	-0.24	-0.40*
	Other heart disease	0.30*	0.05	-0.06	-0.18	-0.06	0.00	-0.03	-0.02	-0.20
	Arthritis	0.19*	-0.20	-0.11	-0.18	-0.21*	-0.22	-0.25*	-0.17	-0.09
Health behavior	Smoking	-0.12	-0.14	0.06	-0.16	-0.03	-0.08	-0.02	-0.17	0.03
Constant		7.16*								

Note: * indicated p <0.05. The unweighted sample size is 27,882 in this model and the weighted sample size is 282,932,617.7. There is no *f* statistics or *p* value reported.

Table C.1. The settings of all matching algorithms used in Chapter 3.

	Settings for matching algorithms
Nearest neighbor (1)	One neighbor for matching without replacement
Nearest neighbor (5)	Five neighbors with replacement
Nearest neighbor (5)/Caliper	Five neighbors with replacement within caliper as 0.1
Nearest neighbor (10)	Ten neighbors with replacement
Radius	Radius with caliper as 0.1
Kernel and Caliper	Kernel with caliper as 0.1
Kernel (0.06)	Kernel with band width as 0.06 (default)
Kernel (0.1)	Kernel with band width as 0.1
Local linear	Local linear

Table C.2. The comparison of matching results based on the matching algorithms used for total health spending.

Matching algorithms	Unmatched	Nearest neighbor (1)	Nearest neighbor (5)	Nearest neighbor (5)/ Caliper	Nearest neighbor (10)	Radius	Kernel and Caliper	Kernel (0.06)	Kernel (0.1)	Local linear
No. of obs.										
Control	1,295	1,295	1,295	1,295	1,295	1,295	1,295	1,295	1,295	1,295
Treated	546	546	543	543	543	543	543	543	543	543
Total	1,841	1,841	1,838	1,838	1,838	1,838	1,838	1,838	1,838	1,838
Mean expenditure										
Matched treated	24,654.74	24,654.74	24,600.62	24,600.62	24,600.62	24,600.62	24,600.62	24,600.62	24,600.62	24,600.62
Matched controls	27,065.76	27,691.32	26,260.13	26,260.13	25,811.61	26,525.82	27,251.57	27,251.57	26,741.18	27,019.68
Matched differences	-2,411.02	-3,036.57	-1,659.51	-1,659.51	-1,210.99	-1,925.20	-2,650.95	-2,650.95	-2,140.57	-2,419.07
Statistics										
Standard error	3,312.39	4,070.47	4,245.70	4,245.70	3,956.64	3,693.84	3,761.16	3,761.16	3,722.79	-
t	-0.73	-0.75	-0.39	-0.39	-0.31	-0.52	-0.70	-0.70	-0.57	-
Bootstrapped z	-	-1.03	-0.69	-0.38	-0.26	-0.56	-0.70	-0.74	-0.51	-0.63
p	0.78*	0.301	0.489	0.705	0.794	0.573	0.485	0.459	0.607	0.532
Propensity score										
Bias (%)		17.9	0.1	0.1	0.4	8.4	2	2	5.3	0
Bias reduction (%)		79	99.9	99.9	99.5	90.1	97.7	97.7	93.8	100
Sensitivity analysis										
Critical value of Γ		1.07-1.08	1.66-1.67	1.66-1.67	2.02-2.03	2.55-2.56	2.62-2.63	2.62-2.63	2.58-2.59	2.51-2.52

Note: * this p value was derived from the independent-sample t test, rather than the bootstrapped z statistics that were used in the other matching methods.

Table C.3. The comparison of matching results based on the matching algorithms used for out-of-pocket health spending.

Matching algorithms	**Unmatched**	**Nearest neighbor (1)**	**Nearest neighbor (5)**	**Nearest neighbor (5)/ Caliper**	**Nearest neighbor (10)**	**Radius**	**Kernel and Caliper**	**Kernel (0.06)**	**Kernel (0.1)**	**Local linear**
No. of obs.										
Controlled	3,030	3,030	3,030	3,030	3,030	3,030	3,030	3,030	3,030	3,030
Treated	1,096	1,096	1,096	1,096	1,096	1,096	1,096	1,096	1,096	1,096
Total	4,126	4,126	4,126	4,126	4,126	4,126	4,126	4,126	4,126	4,126
Mean expenditure										
Matched treated	5,087.16	5,087.16	5,087.16	5,087.16	5,087.16	5,087.16	5,087.16	5,087.16	5,087.16	5,087.16
Matched controls	7,030.16	6,511.90	6,572.35	6,572.35	6,497.54	6,465.93	6,469.47	6,469.47	6,498.61	6,446.31
Matched differences	-1,943.00	-1,424.74	-1,485.19	-1,485.19	-1,410.38	-1,378.77	-1,382.31	-1,382.31	-1,411.45	-1,359.15
Statistics										
S. E.	681.39	464.27	724.93	724.93	712.10	630.23	631.06	631.06	620.83	-
t	-2.85	-3.07	-2.05	-2.05	-1.98	-2.19	-2.19	-2.19	-2.27	-
Bootstrapped z	-	-2.42	-2.77	-3.35	-2.54	-3.41	-3.33	-3.41	-3.15	-2.92
p	0.05*	0.015	0.006	0.001	0.011	0.001	0.001	0.001	0.002	0.003
Propensity score										
Bias (%)	79.3	8.3	0.1	0.1	0.3	10.1	2.5	2.5	6.5	0
Bias reduction (%)		90	99.9	99.9	99.7	87.9	97	97	92.2	100
Sensitivity analysis										
Critical value of Γ	-	1.32-1.33	2.12-2.13	2.12-2.13	2.46-2.47	3.13-3.14	3.14-3.15	3.14 -3.15	3.20-3.21	3.04-3.05

Note: * this p value was derived from the independent-sample t test, rather than the bootstrapped z statistics that were used in the other matching methods.

Table C.4. The GLM (gamma) predicting total and out-of-pocket (OOP) health expenditure.

	Total Health Expenditure	Out-of-pocket Health Expenditure
No. of obs	1841	4126
Log likelihood	-20229.11	-39823.09
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	-0.06	-0.28**
	(0.10)	(0.07)
Pre-Medicare characteristics		
Female	-0.32**	0.05
	(0.10)	(0.06)
Race		
Black	-0.28*	-0.38**
	(0.14)	(0.08)
Other	0.09	-0.06
	(0.29)	(0.16)
Hispanic	-0.32	-0.29**
	(0.19)	(0.11)
Regions		
Midwest	0.10	0.10
	(0.15)	(0.09)
South	-0.19	0.17*
	(0.13)	(0.09)
West	0.05	0.19
	(0.16)	(0.10)
Other	(omitted)	(omitted)
Years of education	0.04*	0.03**
	(0.02)	(0.01)
Income (thousands)	-0.0031	-0.0008
	(0.0025)	(0.0010)

Continued in the next page

	Total health expenditure	OOP health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Self-rated health status		
Very good	0.11	0.11
	(0.14)	(0.09)
Good	0.42**	0.39**
	(0.14)	(0.09)
Fair	0.86**	0.65**
	(0.17)	(0.11)
Poor	1.17**	1.16**
	(0.25)	(0.17)
CESD score (0 to 8)	0.05	0.00
	(0.03)	(0.02)
Difficulty in ADL (0 to 5)		
1	0.25	-0.05
	(0.20)	(0.13)
2	0.23	0.28
	(0.37)	(0.22)
3	0.53	0.80*
	(0.45)	(0.34)
4-5	0.04	0.10
	(0.55)	(0.39)
Difficulty in IADL (0 to 5)		
1	0.07	0.02
	(0.24)	(0.16)
2	0.50	-0.22
	(0.62)	(0.41)
Difficulty in mobility (0 to 5)		
1	0.07	0.19*
	(0.12)	(0.08)
2	0.41*	0.20
	(0.17)	(0.10)
3	0.52*	0.22
	(0.24)	(0.15)
4	0.51	0.13
	(0.27)	(0.16)
5	0.53	0.35
	(0.40)	(0.30)

Continued in the next page

	Total health expenditure	OOP health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Marital status (married as reference)		
Separated/Divorced	0.31	-0.26**
	(0.16)	(0.09)
Widowed	0.07	-0.01
	(0.14)	(0.09)
Never married	0.02	-0.21
	(0.27)	(0.17)
Pre-Medicare health coverage		
Medicaid	-0.18	-0.59**
	(0.28)	(0.15)
Champus/VA	0.34	-0.64**
	(0.21)	(0.12)
Private insurance (from self)	-0.10	-0.02
	(0.11)	(0.07)
Private insurance (from spouse)	-0.07	-0.04
	(0.14)	(0.08)
Pre-Medicare interview year	0.03	0.43*
	(0.27)	(0.20)
Medicare interview year	-0.02	-0.39*
	(0.27)	(0.20)
Birth year	0.09	0.00
	(0.07)	(0.04)
Constant	-174.26**	-55.36**
	(53.83)	(19.85)

Note: * p<0.05; ** p<0.001.

Table C.5. The GLM (gamma) predicting total and out-of-pocket (OOP) health expenditure for Chapter 4.

	Total Health Expenditure	Out-of-pocket Health Expenditure
No. of obs	1752	4032
Log likelihood	-19229.14	-38864.20
	Coefficients (S. E.)	**Coefficients (S. E.)**
Under Medicare		
HMO coverage	-0.04	-0.29**
	(0.10)	(0.07)
Pre-Medicare characteristics		
Female	-0.23*	0.07
	(0.10)	(0.06)
Race		
Black	-0.33*	-0.47**
	(0.14)	(0.08)
Other	-0.14	-0.25
	(0.30)	(0.16)
Hispanic	-0.23	-0.24*
	(0.19)	(0.11)
Regions		
Midwest	0.04	0.08
	(0.14)	(0.09)
South	-0.23	0.15
	(0.13)	(0.08)
West	0.01	0.13
	(0.16)	(0.09)
Other	(omitted)	(omitted)
Years of education	0.04*	0.04**
	(0.02)	(0.01)
Income (thousands)	-0.0027	-0.0005
	(0.0029)	(0.0010)

Continued in the next page

	Total health expenditure	OOP health expenditure
	Coefficients (S. E.)	Coefficients (S. E.)
Self-rated health status		
Very good	0.10	0.05
	(0.14)	(0.09)
Good	0.33*	0.30**
	(0.14)	(0.09)
Fair	0.77**	0.57**
	(0.17)	(0.11)
Poor	0.95**	1.03**
	(0.26)	(0.17)
CESD score (0 to 8)	0.05	-0.01
	(0.03)	(0.02)
Difficulty in ADL (0 to 5)		
1	0.27	-0.03
	(0.21)	(0.13)
2	0.08	0.28
	(0.35)	(0.22)
3	0.31	0.75*
	(0.45)	(0.33)
4-5	-0.21	0.10
	(0.54)	(0.38)
Difficulty in IADL (0 to 5)		
1	0.07	0.05
	(0.24)	(0.16)
2	0.32	-0.15
	(0.63)	(0.41)
Difficulty in mobility (0 to 5)		
1	0.03	0.17**
	(0.12)	(0.07)
2	0.31	0.15
	(0.16)	(0.10)
3	0.39	0.14
	(0.23)	(0.14)
4	0.39	0.11
	(0.27)	(0.16)
5	0.52	0.33
	(0.41)	(0.29)

Continued in the next page

	Total health expenditure	OOP health expenditure
	Coefficients (S. E.)	**Coefficients (S. E.)**
Marital status (married as reference)		
Separated/Divorced	0.23	-0.26**
	(0.16)	(0.09)
Widowed	-0.03	0.01
	(0.14)	(0.09)
Never married	-0.15	-0.19
	(0.27)	(0.16)
Pre-Medicare health coverage		
Medicaid	-0.01	-0.57**
	(0.28)	(0.15)
Champus/VA	0.44**	-0.63**
	(0.22)	(0.12)
Private insurance (from self)	-0.11	0.01
	(0.11)	(0.07)
Private insurance (from spouse)	-0.11	-0.06
	(0.14)	(0.08)
Pre-Medicare interview year	0.01	0.46*
	(0.28)	(0.19)
Medicare interview year	-0.01	-0.43*
	(0.27)	(0.19)
Birth year	0.11	0.00
	(0.07)	(0.04)
Chronic conditions		
Hypertension	0.12	0.25**
	(0.10)	(0.06)
Arthritis	0.04	0.12**
	(0.06)	(0.04)
Death	0.65**	0.31**
	(0.13)	(0.11)
Constant	-211.60**	-59.83**
	(57.85)	(20.67)

Note: * p<0.05; ** p<0.001.

Figures

Figure 2.1. The illustration of the sources of the Medicare spending growth.

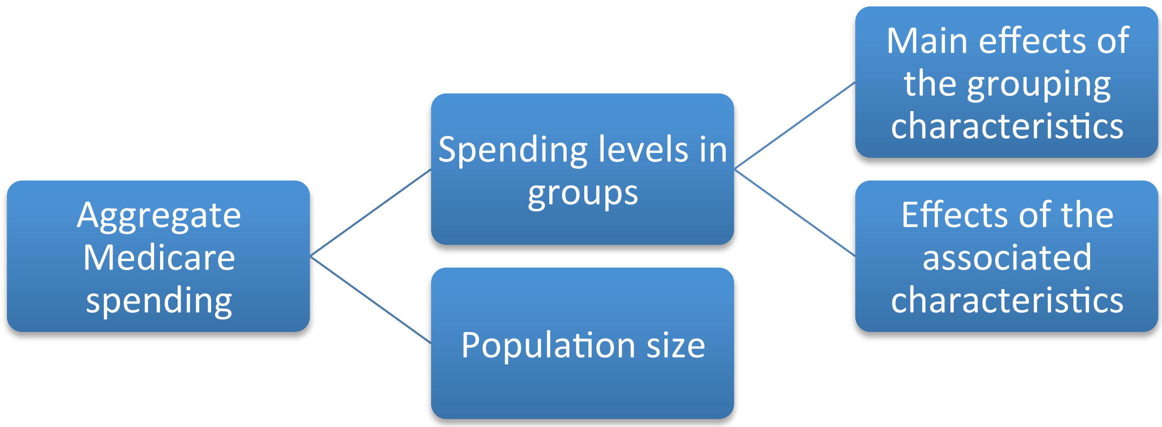


Figure 4.1. The probability of survival among those deceased in the first three to four years of Medicare coverage.

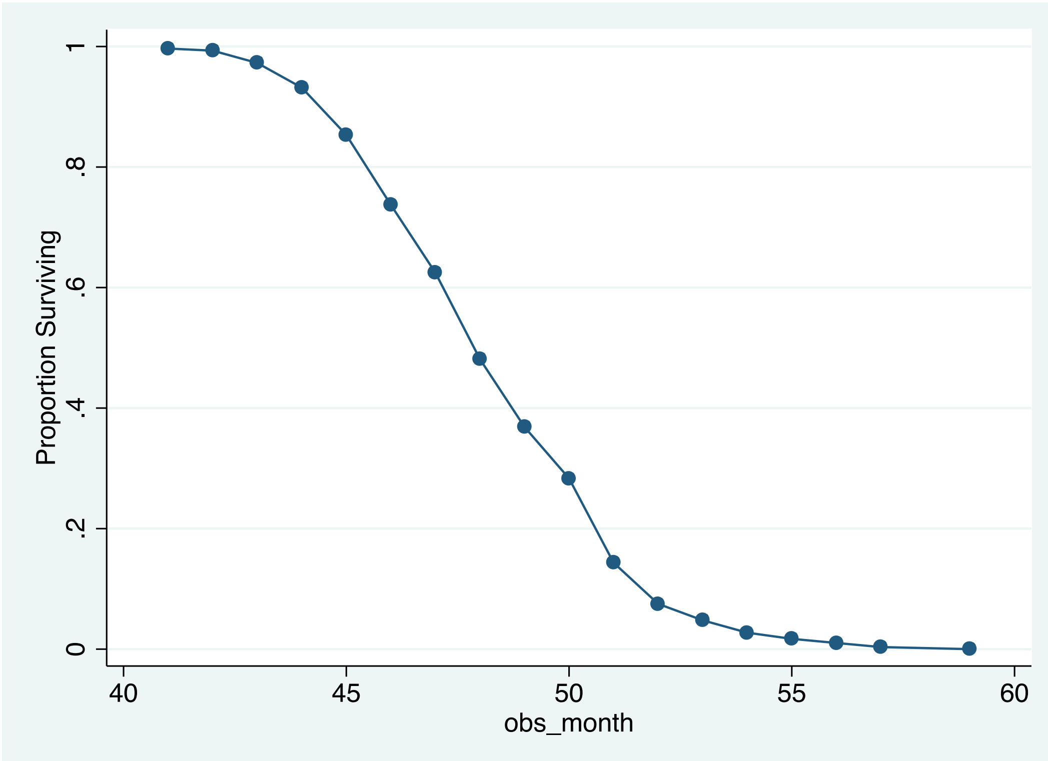
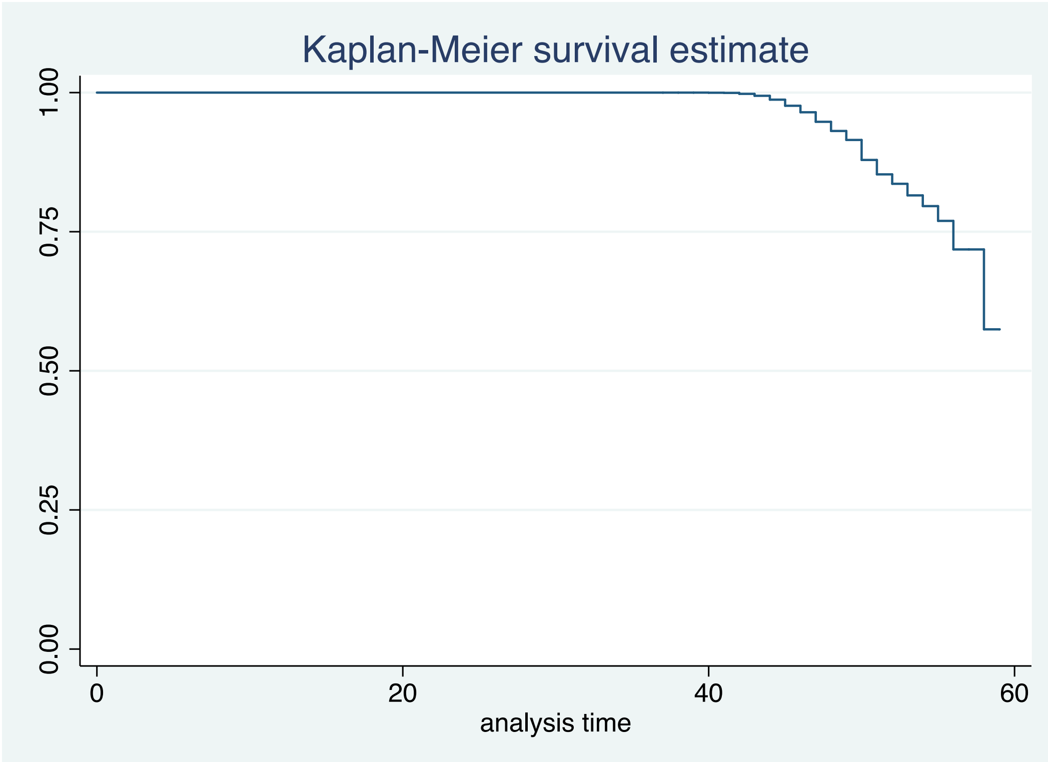


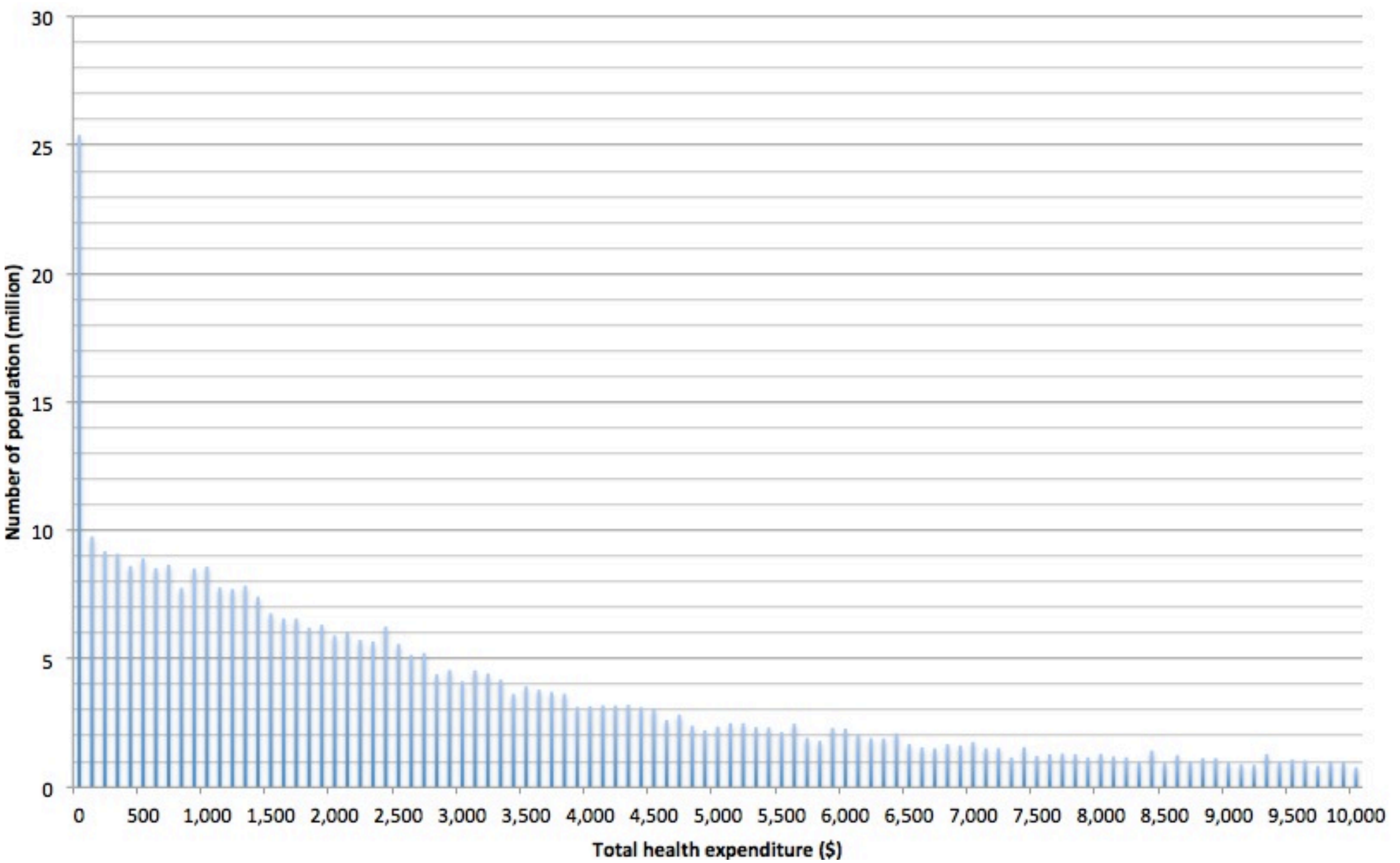
Figure 4.2. The Kaplan-Meier curve of the survival probability of Medicare enrollees in the first three to four years of Medicare coverage.



Note: analysis time was measured in months.

Figure A.1. Health expenditure distribution.

(a) Medicare enrollees age 65 years and over from 1996 to 2008, expenditure more than $10,000 omitted.



(b) Health expenditure distribution in 1996 and 2008, expenditures more than $7,000 omitted.

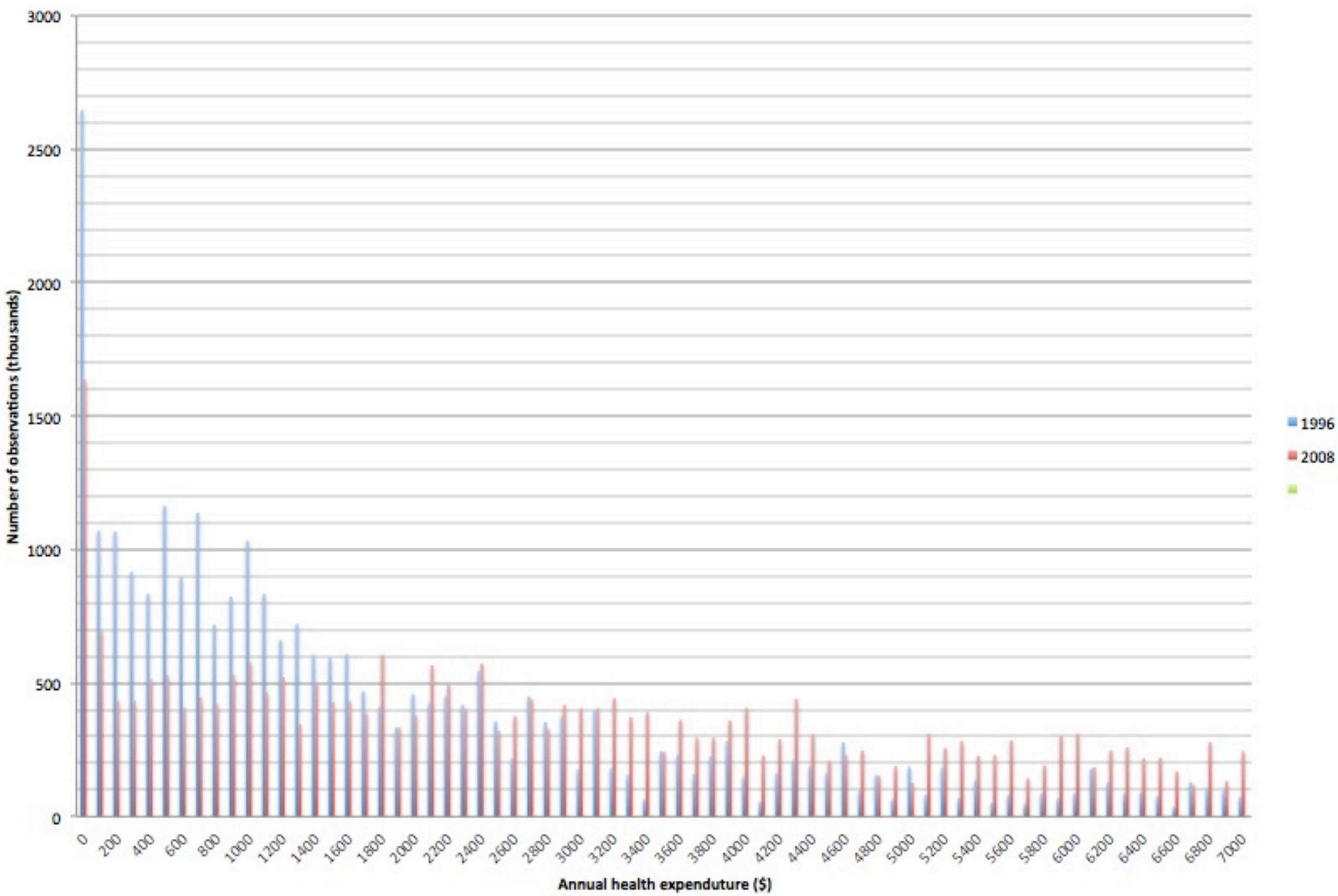


Figure A.2. The log-transformed amounts of total annual health spending among Medicare enrollees age 65 and over from 1996 to 2008 (zero spending excluded).

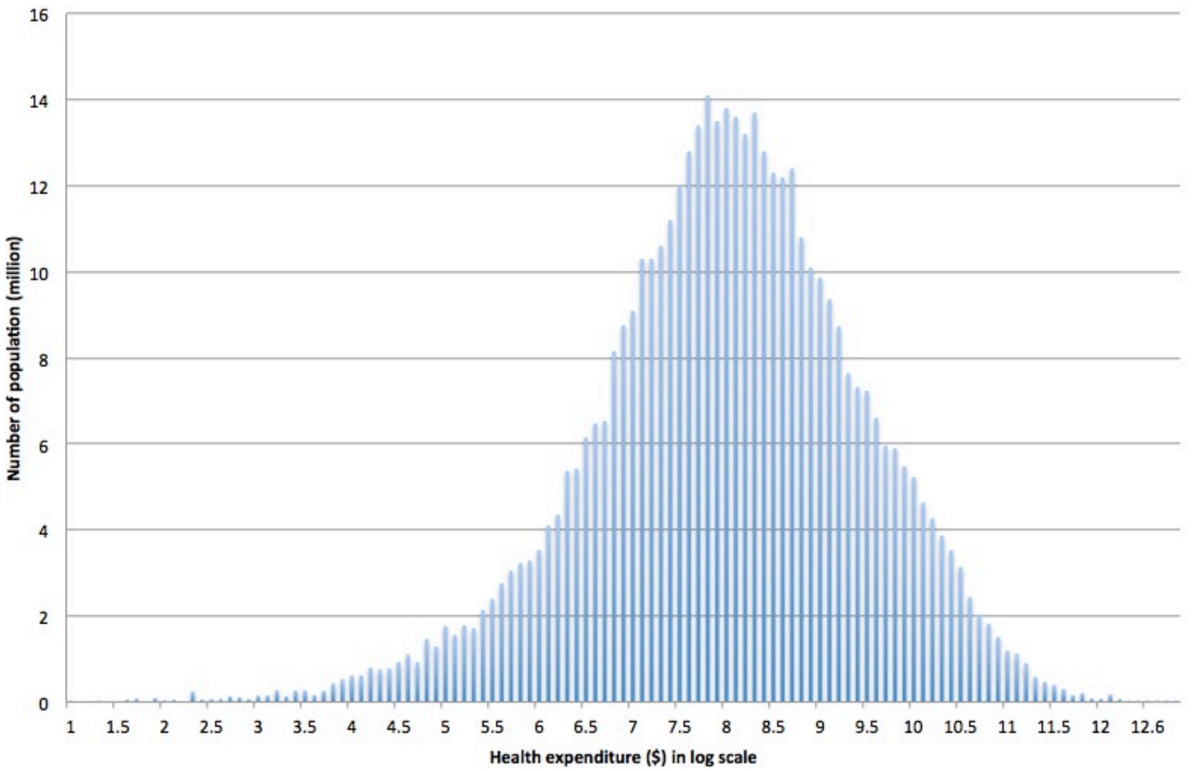
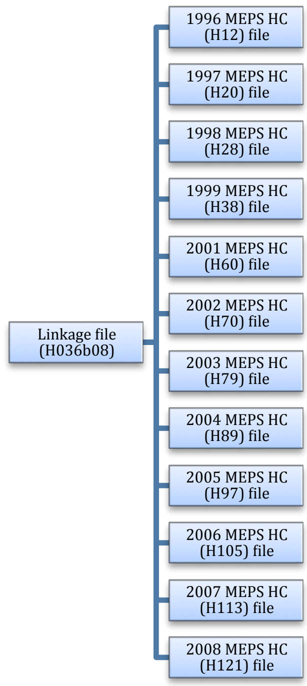


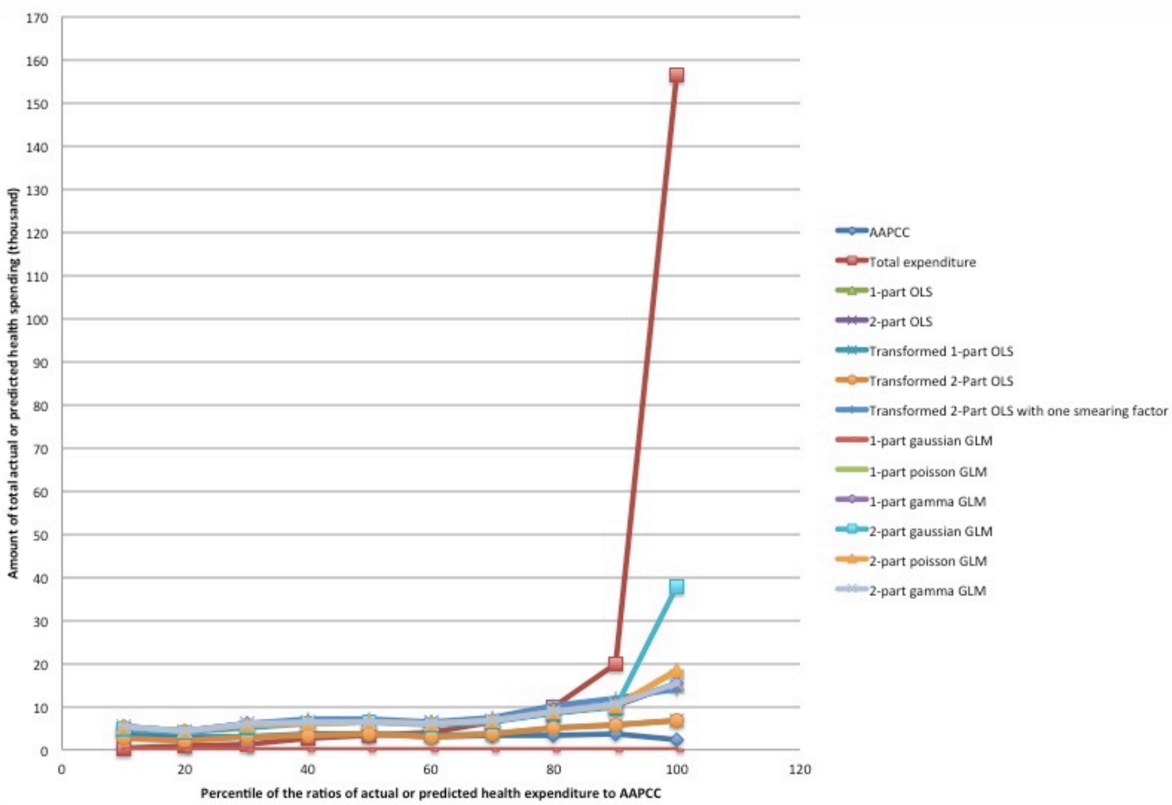
Figure B.1. The relationship of the MEPS linkage file and annual HC files (file names in parentheses).



Note: each observation in the HC files was assigned and id number (*dupersid*) and a panel number (*panel*) in the linkage file (*h036b08*) to identify their sampling units and strata.

Figure B.2. The amounts of observed and predicted annual health expenditure: the values of the deciles of the ratio of health spending to AAPCC were marked (2.9a and 2.9b).

(a) The values of all deciles.



(b) The values of deciles within $16,000.

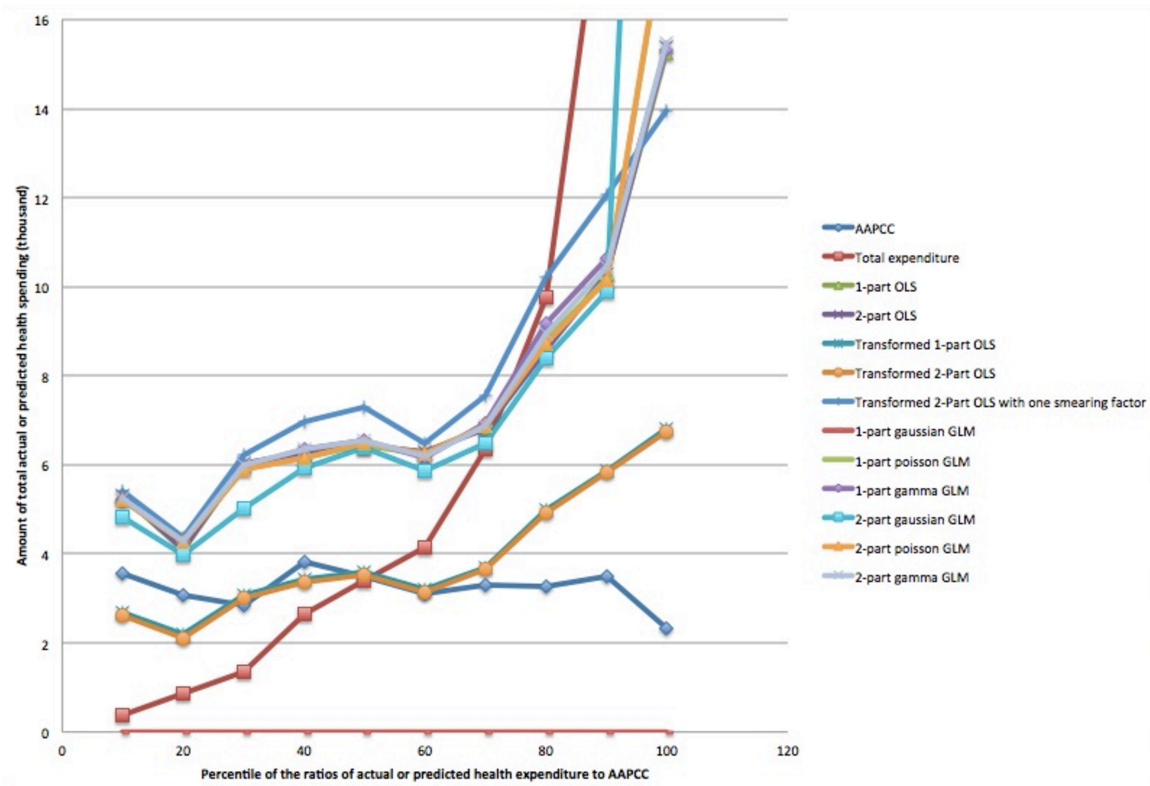
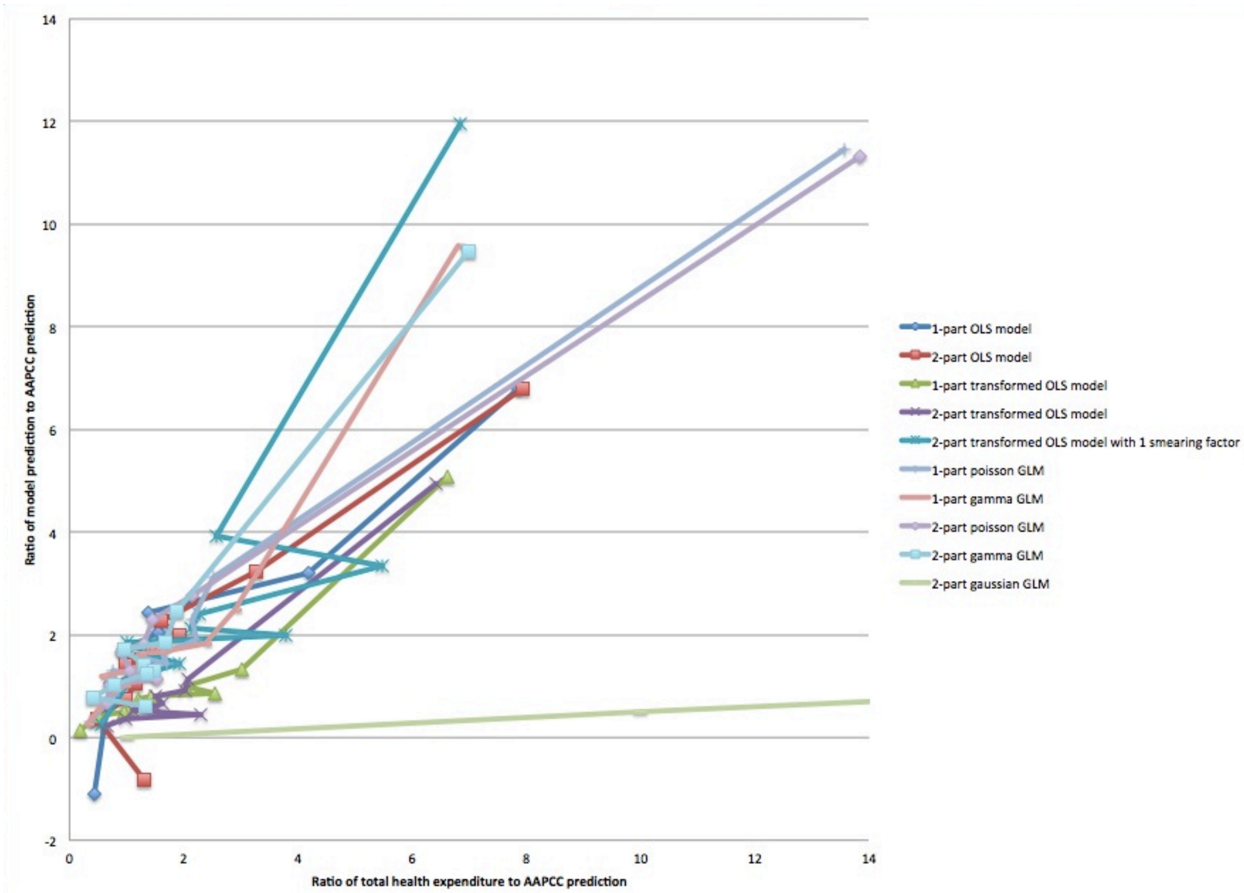


Figure B.3. The ratios of predicted health expenditure to AAPCC, plotted against the ratios of observed health spending to AAPCC.

(a) All ratios of predicted health expenditure to AAPCC.



(b) The ratios of predicted health expenditure to AAPCC less than four.

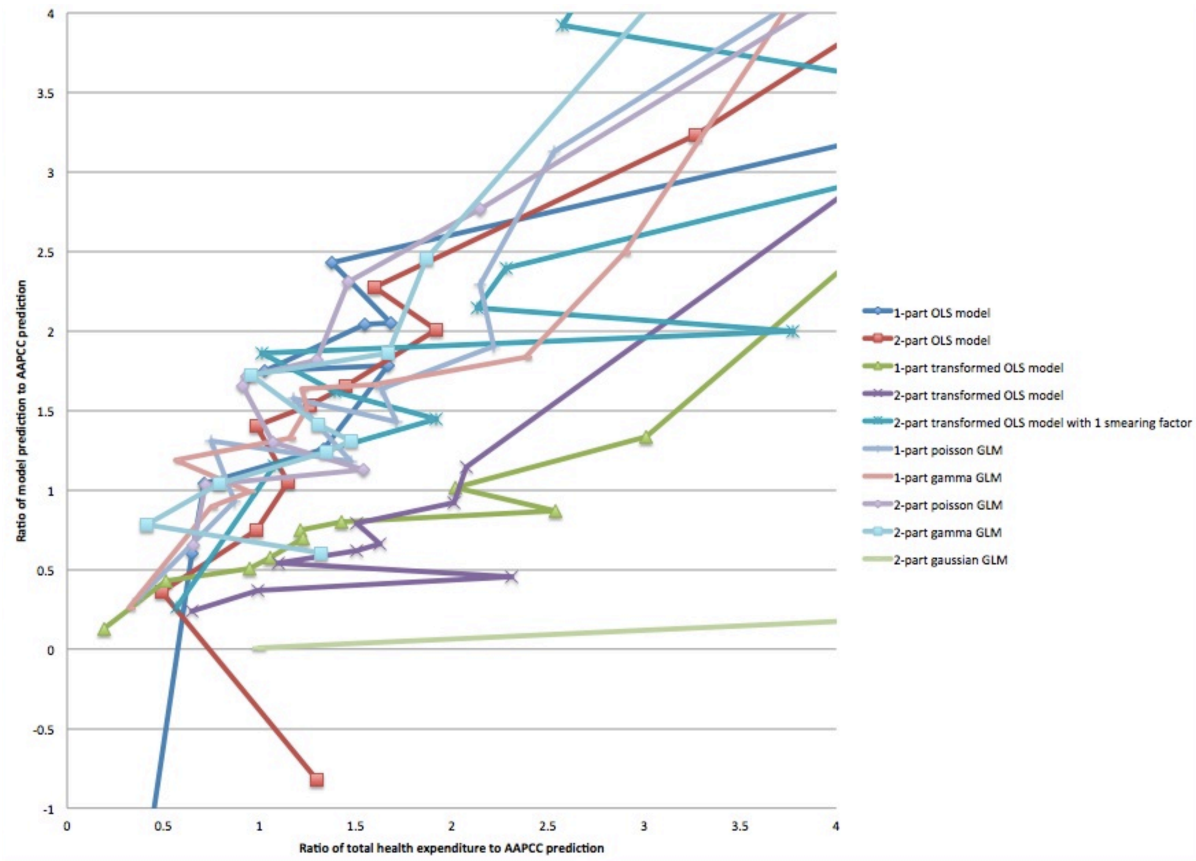
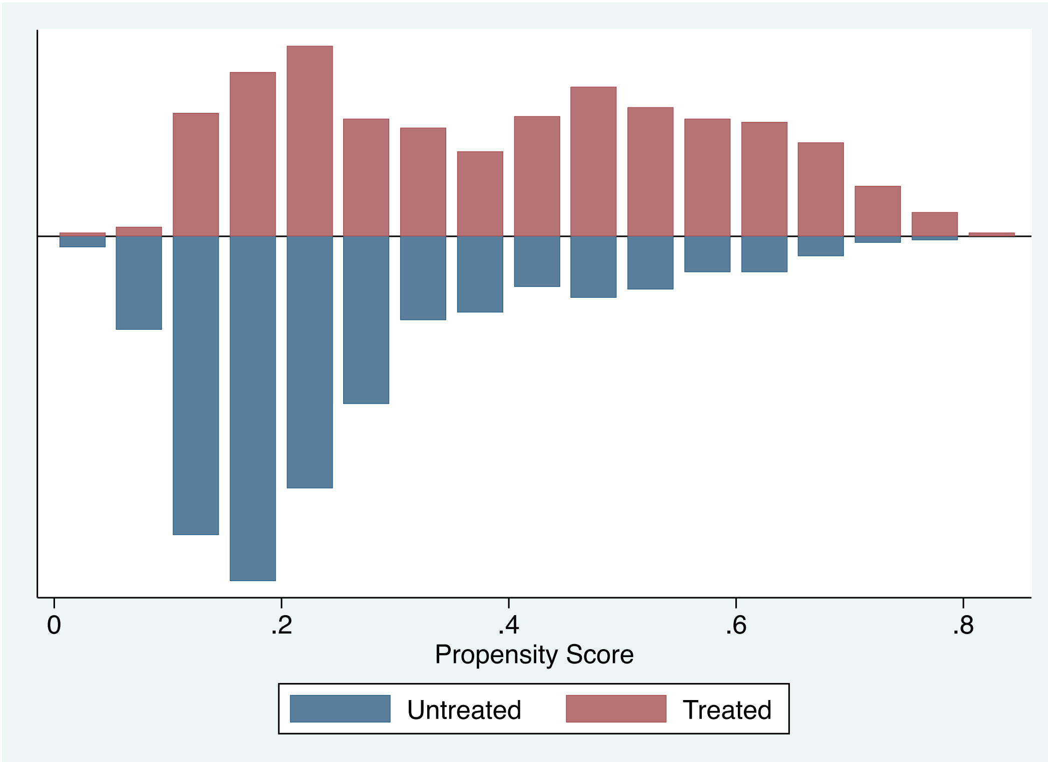


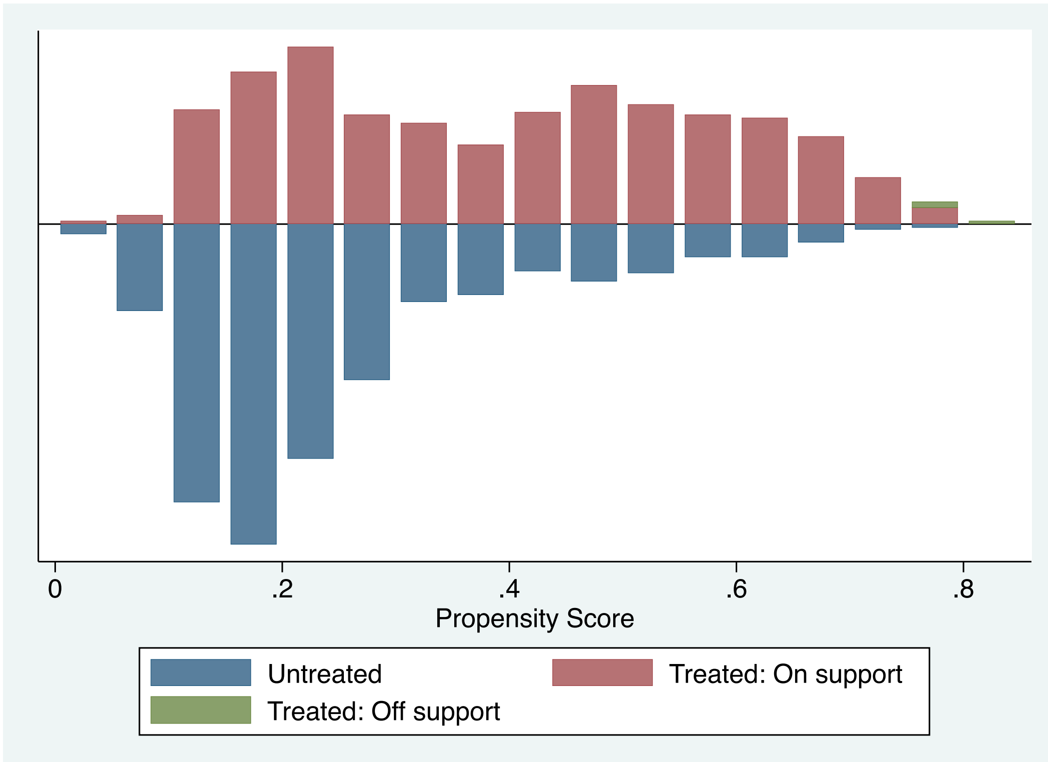
Figure C.1. The range of common support

(a) Nearest neighbor (1) matching without replacement for total health expenditure.



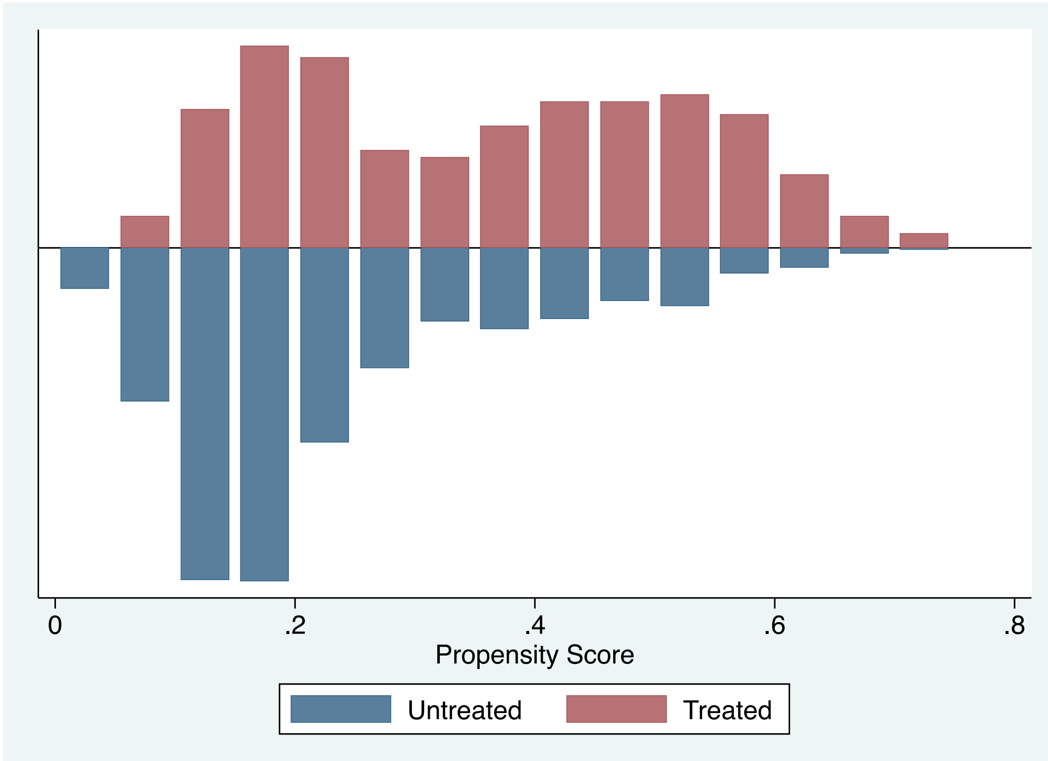
Note: all observations used for matching.

(b) Nearest neighbor (5 and 10), radius, kernel, and local linear matching for total health expenditure.



Note: three observations in Medicare Advantage/Part C were excluded for matching for being outside the range of common support.

(c) Nearest neighbor (1, 5, and 10), radius, kernel and local linear matching for out-of-pocket health expenditure.



Note: all observations used for matching.

References

Agency for Healthcare Research and Quality. (2003). *MEPS HC-050: 2000 Full Year Consolidated Data File*. Rockville, MD.

Agency for Healthcare Research and Quality. (2010). *MEPS HC-036: 1996-2008 Pooled Estimation File*. Rockville, MD.

Agency for Healthcare Research and Quality. (2011a). *Survey Background*. Rockville, MD. Retrieved March 10, 2011, from http://www.meps.ahrq.gov/mepsweb/about_meps/survey_back.jsp#household

Agency for Healthcare Research and Quality. (2011b). *Household Component*. Rockville, MD. Retrieved March 10, 2011, from http://www.meps.ahrq.gov/mepsweb/survey_comp/household.jsp

Agency for Healthcare Research and Quality. (2011c). *2010 National Healthcare Quality and Disparities Reports*. Rockville, MD.

Alemayehu, B., & Warner, K. E. (2004). The lifetime distribution of health care costs. *Health Serv Res,* 39(3), 627-642.

Basu, A., & Manning, W. G. (2009). Issues for the next generation of health care cost analyses. *Med Care,* 47(7 Suppl 1), S109-114.

Bhattacharya, J., & Lakdawalla, D. (2002). Does Medicare Benefit the Poor? New Answers to an Old Question. *NBER Working Papers*.

Biles, B., Dallek, G., & Nicholas, L. H. (2004). Medicare advantage: deja vu all over again? *Health Aff (Millwood), Suppl Web Exclusives*, W4-586-597.

Birnbaum, M., & Patchias, E. (2008, August 28) Medicare Coverage for Seniors: How Universal Is It and What Are the Implications? *Paper presented at the annual meeting of the APSA 2008 Annual Meeting, Hynes Convention Center, Boston, MA.* Retrieved March 11, 2011, from http://www.allacademic.com/meta/p281081_index.html

Blough, D. K., Madden C. W., et al. (1999). Modeling risk using generalized linear models. *J Health Econ,* 18(2), 153-171.

Buntin, M. B., & Zaslavsky, A. M. (2004). Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *J Health Econ, 23*(3), 525-542.

Cabral, M., & Mahoney, N. (2010). *Private Coverage and Public Costs: Identifying the Effect of Private Supplemental Insurance on Medicare Spending*. Unpublished manuscript, Palo Alto, CA.

Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys, 22*(1), 31-72.

Cameron, A. C., Trivedi, P. K., & Milne, F. (1988). A microeconometric model of the demand for health care and health insurance in Australia. *The Review of Economic Studies,* 55, 85-106.

Centers for Medicare & Medicaid Services. (2010a). Historical. *National Health Expenditure Data*. Baltimore, MD. Retrieved March 4, 2011, from http://www.cms.gov/NationalHealthExpendData/downloads/tables.pdf

Centers for Medicare & Medicaid Services. (2010b). Medicare Benefits. Baltimore, MD. Retrieved Dec 11, 2010, from http://www.medicare.gov/navigation/medicare-basics/medicare-benefits/medicare-benefits-overview.aspx?AspxAutoDetectCookieSupport=1

Chao, Y. S. (2010). Unfair Contribution and Consumption in Medicare: Results from the Medical Expenditure Panel Survey in 2006. *Online Journal of Health Ethics,* Retrieved June 1, 2011, from http://www.test2.ojhe.org/index.php/ojhe/article/viewArticle/142

Chen, L.-W., Zhang, W., et al. (2004). Pent-up demand: health care use of the uninsured elderly. *Economic Research Initiative on the Uninsured Working Paper Series*. Ann Arbor, MI.

Chernew, M., Cutler, D. M., et al. (2005). "Increasing health insurance costs and the decline in insurance coverage." *Health Serv Res* 40(4), 1021-1039.

Congressional Budget Office. (2008). *Technological Change and the Growth of Health Care Spending*. Washington, DC. Retrieved Maarch 20, 2011, from http://www.cbo.gov/ftpdocs/89xx/doc8947/01-31-TechHealth.pdf.

Cook, B. L., McGuire, T. G., Meara, E., & Zaslavsky, A. M. (2009). Adjusting for Health Status in Non-Linear Models of Health Care Disparities. *Health Serv Outcomes Res Methodol, 9*(1), 1-21.

Copas, J. B. (1983). Regression, Prediction and Shrinkage. *Journal of the Royal Statistical*, 45(3), 311-354.

Cox, D. F., & Hogan, C. (1997). Biased selection and Medicare HMOs: analysis of the 1989-1994 experience. *Med Care Res Rev*, 54(3), 259-274; discussion 275-285.

Cutler, D. M., & McClellan, M. (2001). Is technological change in medicine worth it? *Health Aff (Millwood),* 20(5), 11-29.

Cutler, D. M., Rosen, A. B., et al. (2006). The value of medical spending in the United States, 1960-2000. *N Engl J Med*, 355(9), 920-927.

Cutler, D. M., & Vigdor, E. R. (1999) Your Money and Your Life: The Value of Health and What Affects it. *NBER Working Paper Series*, w6895. Retrieved March 22, 2011, from http://ssrn.com/abstract=147383

D'Agostino, R. B., Jr. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med,* 17(19), 2265-2281.

Deb, P., W. Manning, et al. (2011). Preconference Course: Modeling Health Care Costs and Counts. iHEA - Toronto Conference, 2011, Toronto, ON.

Desmond, K. A., Rice, T., & Fox, P. D. (2006). Does greater Medicare HMO enrollment cause adverse selection into Medigap? *Health Econ Policy Law,* 1(1), 3-21.

DiPrete, T.A., Gangl, M. (2004) Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments. Discussion paper SP I 2004-101. Berlin: WZB.

Doyle, Y. G., Furey, A., et al. (2006). Sick individuals and sick populations: 20 years later. *J Epidemiol Community Health*, 60(5), 396-398.

Ellis, B. H., Bannister, W. M., Cox, J. K., Fowler, B. M., Shannon, E. D., Drachman, D., . . . Giordano, L. A. (2003). Utilization of the propensity score method: an exploratory comparison of proxy-completed to self-completed responses in the Medicare Health Outcomes Survey. *Health Qual Life Outcomes,* 1, 47.

Ettner, S. L. (1997). Adverse selection and the purchase of Medigap insurance by the elderly. *Journal of Health Economics,* 16(5), 543-562.

Farrell, D., Jensen, E., Kocher, B., Lovegrove, N., Melhem, F., Mendonca, L., et al. (2008). *Accounting for the cost of US health care: A new look at why Americans spend more*: McKinsey Global Institute.

Feldman, R., Dowd, B., et al. (2003). Risk selection and benefits in the Medicare+Choice program. *Health Care Financ Rev*, 25(1), 23-36.

Fisher, E. S. (2008). Building a Medical Neighborhood for the Medical Home. *New England Journal of Medicine*, 359(12), 1202-1205.

Folland, S., Goodman, A. C., & Stano, M. (1993) The Demand for Health Capital. in *The Economics of Health and Medical Care*. New York, NY: MacMillan, 136-141.

Gangl, M. (2004). RBOUNDS: Stata module to perform Rosenbaum sensitivity analysis for average treatment effects on the treated, Boston College Department of Economics, Boston, MA.

Gerdtham, U. G., Johannesson, M., et al. (1999). The demand for health: results from new measures of health capital. *European Journal of Political Economy*, 15(3), 501-521.

Gibson, R. (2011). Resource Use in the Last 6 Months of Life: What Does It Mean for Patients? *Archives of Internal Medicine,* 171(3), 194-195.

Ginsburg, P. B. (2008). *High and rising health care costs: Demystifying U.S. health care spending*. Princeton, NJ: The Robert Wood Johnson Foundation.

Glick, H. & Doshi, J. (2007). Analyzing Treatment Costs in Randomized Trials. Pittsburgh, PA: University of Pennsylvania.

Goldman, D. P., Zissimopoulos, J., et al. (2011). Medical Expenditure Measures in the Health and Retirement Study. *Forum for Health Economics & Policy,*14(3).

Greenwald, L. M., Levy, J. M., & Ingber, M. J. (2000). Favorable selection in the Medicare+Choice program: new evidence. *Health Care Financ Rev,* 21(3), 127-134.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy,* 80(2), 223-255.

Grossman, M. (1999). The Human Capital Model of the Demand for Health. National Bureau of Economic Research Working Paper Series No. 7078.

Hall, R. E., & Jones, C. I. (2007). The Value of Life and the Rise in Health Spending. *The Quarterly Journal of Economics,* 122(1), 39-72.

Heinrich, C., Maffioli, A., & Vázquez, G. (2010). A Primer for Applying Propensity-Score Matching. Washington, DC: Inter-American Development Bank, Office of Strategic Planning and Development Effectiveness (SPD).

Hill, S. C., & Miller, G. E. (2010). Health expenditure estimation and functional form: applications of the generalized gamma and extended estimating equations models. *Health Econ,* 19(5), 608-627.

Hsu, J., Price, M., Huang, J., Brand, R., Fung, V., Hui, R., et al. (2006). Unintended Consequences of Caps on Medicare Drug Benefits. *New England Journal of Medicine*, 354(22), 2349-2359.

Jones, A. M. (2010). Models For Health Care. HEDG Working Paper. York, UK: University of York.

Kaiser Family Foundation. (2010a). *Medicare A Primer* (No. 7615-03). Menlo Park, CA: The Henry J. Kaiser Family Foundation.

Kaiser Family Foundation. (2010b). Medicare: A Timeline of Key Developments. Menlo Park, CA. Retrieved Oct 31, 2010, from http://www.kff.org/medicare/timeline/pf_entire.htm

Kaul, P., McAlister, F. A., et al. (2011). Resource Use in the Last 6 Months of Life Among Patients With Heart Failure in Canada. *Archives of Internal Medicine*, 171(3), 211-217.

Kleinke, J. D. (2004). Access versus excess: value-based cost sharing for prescription drugs. *Health Aff (Millwood),* 23(1), 34-47.

Leuven, E. help for pstest. Retrieved Jan 9, 2012, from http://fmwww.bc.edu/repec/bocode/p/pstest.html.

Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, Boston College Department of Economics.

Levy, H., & Meltzer, D. (2004). What do we really know about whether health insurance affects health? Health policy and the uninsured. C. G. McLaughlin, Urban Institute Press.

Lo Sasso, A. T., & Buchmueller, T. C. (2004). The effect of the state children health insurance program on health insurance coverage. *Journal of Health Economics,* 23(5), 1059-1082.

Long, J. S. (1997). Regression Models for categorical and limited dependent variables. Thousand Oaks, CA: Sage Publications.

Lubitz, J. D., & Riley, G. F. (1993). Trends in Medicare Payments in the Last Year of Life*. New England Journal of Medicine,* 328(15), 1092-1096.

Luft, H.S. (1981). *Health Maintenance Organizations: Dimensions of Performance.* New York, NY: John Wiley and Son.

Manning, W. G. (1998). The logged dependent variable, heteroscedasticity, and the retransformation problem. *J Health Econ*, 17(3), 283-295.

Manning, W. G., Basu, A., & Mullahy, J. (2005). Generalized modeling approaches to risk adjustment of skewed outcomes data. *Journal of Health Economics*, 24(3), 465-488.

Manning, W. G., & Mullahy, J. (2001). Estimating log models: to transform or not to transform? *J Health Econ,* 20(4), 461-494.

Matsaganis, M., Mitrakos, T., et al. (2008). Modelling Household Expenditure On Hhealth Care In Greece. Working paper. Athens, Greece: BANK OF GREECE.

McBride, T. D., Penrod, J., et al. (1997). Volatility in Medicare AAPCC rates: 1990-1997. *Health Aff (Millwood)*, 16(5), 172-180.

McClellan, M., & Skinner, J. (1999). Medicare reform: who pays and who benefits? *Health Aff (Millwood),* 18(1), 48-62.

Mello, M. M., Stearns, S. C., et al. (2003). Understanding biased selection in Medicare HMOs. *Health Serv Res,* 38(3), 961-992.

Miller, R.H., & Luft, H. S. (1994). Managed Care Plan Performance Since 1980. *Journal of the American Medical Association,* 271(19), 1512-1519.

Monheit, A. C., Cantor, J. C., Koller, M., & Fox, K. S. (2004). Community rating and sustainable individual health insurance markets in New Jersey. *Health Aff (Millwood), 23*(4), 167-175.

Morrisey, M. A. (2007). *Health Insurance*. Chicago, IL: Health Administration Press.

Mullahy, J. (1998). Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *Journal of Health Economics,* 17, 247–281.

Newhouse, J. P. (2006). Reconsidering the moral hazard-risk avoidance tradeoff. *Journal of Health Economics*, 25(5), 1005-1014.

Newhouse, J. P., & the Insurance Experiment Group. (1993). *Free for All? Lessons from the RAND Health Experiment*. Cambridge, Mass: Harvard University Press.

Olin, G., & Lavis, A. (1998). Determinants of Enrollment and Disenrollment in Medicare HMOs. *Proceedings of the Survey Research Methods Section*, ASA(2): 168-171.

Palmer, J. L., & Saving, T. R. (2006). *A MESSAGE TO THE PUBLIC*. Washington, D.C.: The U.S. Social Security Administration.

Park, R. (1966). Estimation with heteroscedastic error terms. *Econometrica,* 34, 888.

RAND Center for the Study of Aging. (2010). The RAND HRS Data (Version J). Retrieved from http://hrsonline.isr.umich.edu/modules/meta/rand/randhrsj/rnd_Jdd.pdf

Peduzzi, P., Concato, J., et al. (1996). A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*, 49(12), 1373-1379.

Renton, A. (1994). Epidemiology and causation: a realist view. *J Epidemiol Community Health*, 48(1), 79-85.

Rose, G. (1985). Sick individuals and sick populations. *Int J Epidemiol,* 14(1), 32-38.

Rose, G. (2001). Sick individuals and sick populations. *Int J Epidemiol,* 30(3), 427-432; discussion 433-424.

Rosenbaum, P.R. (2002). Observational Studies. 2nd edition. New York, NY: Springer.

Rosenbaum, P.R. and Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrik,* 70(1), 41-55.

Rothman, K. J., & Greenland, S. (2005). Causation and causal inference in epidemiology. *Am J Public Health,* 95(Suppl 1), S144-150.

Rothschild, M., & Stiglitz, J. E. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics,* 90(4), 630-649.

Schimmel, J. (2006). Pent-Up Demand and the Discovery of New Health Conditions after Medicare Enrollment.

Schoder, J., & Zweifel, P. (2011). Flat-of-the-curve medicine: a new perspective on the production of health. *Health Economics Review,* 1(1), 2.

Selden, T. M. (1993). Uncertainty and health care spending by the poor: The health capital model revisited. *Journal of Health Economics,* 12(1), 109-115.

Sheldon, T. A., Long, A., et al. (1992). Health technology assessment. *Bmj,* 305(6850), 426.

Shen, Y., Hendricks, A., et al. (2005). VA-Medicare dual beneficiaries' enrollment in Medicare HMOs: access to VA, availability of HMOs, and favorable selection. *Med Care Res Rev*, 62(4), 479-495.

Smith, J. P. (2004). Unravelling the SES health connection, Institute for Fiscal Studies.

Smith, J. P. (2007). The Impact of Socioeconomic Status on Health over the Life-Course. *Journal of Human Resources,* 42(4).

Sribney, B. (2005). Estimating correlations with survey data. FAQs. Retrieved Jan 1, 2012, from http://www.stata.com/support/faqs/stat/survey.html.

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica,* 65(3), 557-586.

Stampfer, M. J., Hu, F. B., et al. (2000). Primary Prevention of Coronary Heart Disease in Women through Diet and Lifestyle. *New England Journal of Medicine,* 343(1), 16-22.

Stuart, B., Doshi, J. A., Briesacher, B., Wrobel, M. V., & Baysac, F. (2004). Impact of prescription coverage on hospital and physician costs: a case study of medicare beneficiaries with chronic obstructive pulmonary disease. *Clin Ther,* 26(10), 1688-1699.

Stukel, T. A., Fisher, E. S., Wennberg, D. E., Alter, D. A., Gottlieb, D. J., & Vermeulen, M. J. (2007). Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *Jama.* 297(3), 278-285.

Unroe, K. T., Greiner, M. A., et al. (2011). Resource Use in the Last 6 Months of Life Among Medicare Beneficiaries With Heart Failure, 2000-2007. *Archives of Internal Medicine*. 171(3): 196-203.

U. S. Agency for International Development. (2009). Palliative Care: A Continuum of Patient-Centered Care. Retrieved Jan 16, 2011, from http://www.usaid.gov/our_work/global_health/aids/TechAreas/caresupport/palcarefactsheet.html.

Van Houtven, C. H., Jeffreys, A. S., & Coffman, C. J. (2008). Home health care and patterns of subsequent VA and medicare health care utilization for veterans. *Gerontologist*. 48(5), 668-78.

White, C. (2008). Why did Medicare spending growth slow down? *Health Aff (Millwood)* 27(3), 793-802.

White, C., & Seagrave, S. (2005). What happens when hospital-based skilled nursing facilities close? A propensity score analysis. *Health Serv Res,* 40(6 Pt 1), 1883-1897.

Wilcox-Gok, V., & Rubin, J. (1990). *Private Health Insurance and the Utilization of Medical Care by the Elderly*. Retrieved April 3, 2011, from www.census.gov/sipp/workpapr/wp170.pdf.

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data, Cambridge, MA: Mit Press.

World Health Organization. (2012). The determinants of health. Health Impact Assessment (HIA). Retrieved Feb 2, 2012, from http://www.who.int/hia/evidence/doh/en/.

Yoo, M. (2011). Does Increased Education Lower Health Care Spending? Findings for Self-Managed Health Conditions. New Brunswick, NJ, Rutgers University.

Zhang, H., Kane, R. L., et al. (2008). Selection Bias and Utilization of the Dual Eligibles in Medicare and Medicaid HMOs. *Health Serv Res*.

Zheng, B., & Agresti, A. (2000). Summarizing the predictive power of a generalized linear model. *Statistics in Medicine,* 19(13), 1771-1781.

Zuckerman, S., & McFeeters, J. (2006). *Recent Growth in the Health Expenditures*: Commonwealth Fund.

Curriculum Vitae

Yi-Sheng Chao, MD MPH PhD

Education

Doctor of Philosophy, GPA: 3.305/ 4.0, Health Systems and Policy, UMDNJ/Rutgers University, 2012

Dissertation: Medicare Expenditures and its Health Returns Across Cohorts

Master of Public Health, GPA: 3.503/ 4.0, Public Health, Harvard University, 2009

Thesis topic: Reducing the regressiveness in access to health care - The review of Partners in Health

Doctor of Medicine, GPA: 3.23/ 4.0, Medicine, National Yang-Ming University, Taipei, Taiwan, 2006

Career History

2012-Present **Postdoctoral fellow**, University of Alberta, Edmonton, Alberta

2010-Present **Case Manager**, Gere Biotechnology, Taipei, Taiwan

2010-2011 **Intern**, World Health Organization, Geneva, Switzerland

2010 **Teaching Assistant**, UMDNJ, Piscataway, NJ

2009-2010 **Research Assistant**, UMDNJ, Piscataway, NJ

2009-2011 **Scientific Translator**, MedCom Asia, Xinbei, Taiwan

2008 **Medical Resident**, Jann-Ren Hospital, Kaohsiung, Taiwan

2007 **Medical Resident**, Mission Médicale de Taïwan au Burkina Faso, Koudougou, Burkina Faso

2005-2006 **Intern**, Taipei Veterans General Hospital, Taipei, Taiwan

Medical Relief Mission for on-site disaster healthcare

Physician, Taiwanroot Medical Relief Mission in Port-au-Prince, Haiti, 2010

Physician, Taiwanroot Medical Relief Mission in Pintung, Taiwan, 2008

Academic affiliation

Reviewer, Journal of Clinical Medicine and Research

Selected Publications and Presentations

Dolan GP, Harris RC, Clarkson M, Sokal R, Morgan G, Mukaigawara M, Horiuchi H, Hale R, Stormont L, Bechard-Evans L, **Chao YS**, Eremin S, Martins S, Tam J, Penalver J, Zanuzadana A, Nguyen-Van-Tam JS. The effectiveness of vaccination of healthcare workers for the protection of patients at higher risk of acute respiratory disease: a systematic review. *Emerging Infectious Diseases*. (accepted)

Chao, Y. S. (2010). "Unfair Contribution and Consumption in Medicare: Results from the Medical Expenditure Panel Survey in 2006". *Online Journal of Health Ethics*