A Structural Analysis of Risk Selection in Medicare Advantage

Keith Brand[†]
Bureau of Economics
Federal Trade Commission

December 2006

Abstract

I examine the joint determination of health insurance choice and subsequent health care utilization in the Medicare managed care and supplemental insurance markets in the United States. The objective is to evaluate the welfare impact of the Medicare Advantage, the Medicare managed care program, taking into account the effect of risk selection.

I model health care demand as simultaneously determined in five dimensions: inpatient care, outpatient care, doctor visits, prescription drugs, and dental care. The model incorporates uncertainty about efficacy when treatment decisions are made, limits on the efficacy of treatment, and diminishing marginal product both intensively and extensively.

I employ a mixed multinomial logit approach to health plan choice. Risk averse consumers choose health plans by taking expectations of indirect utilities over a known distribution of health states having private information about the mean. The model uses a unified framework in that the health insurance and utilization decisions are based on the same underlying preferences.

I use individual level data on health insurance choice and subsequent utilization, and data on aggregate enrollment in Medicare Advantage. The aggregate data is observed at the county/age/gender group level. The decomposition to the age/gender level allows me to solve for a vector of unobservables that is specific to every managed care plan/county combination which are used to construct moment conditions that supplement the likelihood function in estimation.

The results suggest that while each age/gender group benefits from Medicare Advantage on average, younger age groups benefit more. The results also indicate that a small percentage of each age/gender group is adversely affected by Medicare Advantage and this share is larger in older age groups.

JEL Classification: C34, H51, I10, I18
Keywords: Medicare, Health Insurance, Health Care Demand, Demand Systems

---

[†] The views expressed herein are not purported to be those of the Federal Trade Commission or any of its Commissioners. I thank my advisers Steven Stern, Leora Friedberg, and John Pepper for their suggestions and support. Financial support from the Bankard Fund for Political Economy and the University of Virginia Graduate School of Arts and Sciences is gratefully acknowledged. I also thank Amalia Miller for helpful comments and the CMS, Resource Data Assistance Center, Robert Town, and Su Liu for their assistance with data. Correspondence: 601 New Jersey Ave Washington, DC 20580. Email: kbrand@ftc.gov.

1 Introduction

The related trends of an aging population and health care sector spending growth that continues to outpace the economy as a whole pose a significant fiscal problem for the Medicare program in the United States. In their 2006 annual report, the Medicare Board of Trustees projected that Medicare's hospitalization insurance trust fund will remain solvent until 2018, two years earlier than projected in 2005. In addition, costs associated with physician and outpatient services as well as the recently enacted prescription drug benefit, which are projected to grow at more than double the pace of gross domestic product over the period 2006-2015, will place an increasing burden on the Federal budget under current policy.[1] As policymakers seek to balance the provision of meaningful public health insurance with other policy goals, understanding health care utilization in the Medicare population is of significant policy and research interest.

Much of the debate on Medicare reform has centered on the role of private health plans. Although Medicare was designed as a single-payer health plan at its inception in 1965, it has contracted with private health plans to provide health care services and insurance since 1982. Currently, Medicare offers a private managed care option to its beneficiaries through the Medicare Advantage program. Contracting out Medicare's insurance role to private managed care organizations (MCOs) was thought to have the potential to reduce costs for two reasons. First, private provision of health insurance may be more technically efficient than public provision. Second, the vertical integration of risk-bearing and health care provision that characterizes managed care may be an effective tool in combating ex post moral hazard.

The effectiveness of managed care in reducing costs is difficult to verify because of risk selection. The conventional wisdom about managed care generally is that it is more attractive to lower risk types, so lower costs may not indicate that the program is achieving its policy objectives. However, many Medicare Advantage plans offer benefits not included in traditional Medicare and so may experience adverse selection as well. Consequently, while

[1] 2006 Annual Report of the Boards of Trustees of the Federal Hospital Insurance and Federal Supplemental Medical Insurance Trust Funds, available at http://www.cms.hhs.gov/ ReportsTrustFunds/downloads/.

most studies (for example, Call et al. (1999) and Feldman et al. (2003)) have found evidence that Medicare MCOs generally experience favorable selection, some (for example, Atherly, Dowd, and Feldman (2003)) find evidence that Medicare MCOs experience adverse selection in particular types of care.

In this paper, I assess the welfare impact of the Medicare Advantage taking into account risk selection. I develop a structural model of Medicare beneficiary health insurance and health care utilization decisions in the Medicare managed care market and the nongroup market for supplemental health insurance known as Medigap. The model examines the joint determination of five types of health care utilization and characterizes risk type in five dimensions to capture the complicated relationship between risk type and health insurance choice in this market. Using the estimated model, I estimate distributions of the change in consumer surplus within age/gender groups brought about by Medicare Advantage.

Assuming favorable selection, Medicare Advantage can affect beneficiary welfare in several ways. First, those who choose traditional Medicare with a Medigap supplement may pay a higher premium than they would have in the absence of Medicare Advantage. Second, some of those enrolled in Medicare Advantage may have preferred traditional Medicare with a Medigap supplement if the supplement had been priced as it would have been in the absence of Medicare Advantage. On the other hand, Medicare Advantage may improve the welfare of lower risk types in that it affords them the opportunity to separate themselves from the higher risk types. Finally, setting aside the issue of risk selection, all beneficiaries, irrespective of risk type, would benefit from Medicare Advantage in that a wider variety of health insurance products are available.

To my knowledge, there has been no research that attempts to quantify the welfare implications of this issue. Town and Liu (2003) estimate aggregate consumer and producer surplus due to Medicare Advantage and conclude that the total surplus generated net of the associated tax burden is significant.[2] However, their study treats the Medicare population as homogenous with respect to risk type and, therefore, does not consider risk selection and assumes that the change in consumer surplus is uniform across different risk groups.

[2] Under their conservative estimate, net welfare over the period 1993-2000 was $24.8 billion.

Preview of the Model The underlying process that generates data on health insurance choice and health care utilization is complicated and presents significant challenges to applied economic analysis. Uncertainty at several levels, budget constraints that are functions of the observed health insurance choice, consumer preferences and health endowments, and technology are all important factors to be considered. Technology itself is a complicated process involving different types of health care that may be substitutable to varying degrees and marginal products that may depend upon, among other things, the intensity of the treatment and the pretreatment health state of the patient.

Since Grossman (1972), the theoretical literature has treated health care utilization as a derived demand, i.e., there is some technological relationship between the commodity demanded, health care, and the utility generating commodity, health. One of my principal contributions is to directly apply the notion of health care as a derived demand in an empirical application. I model several aspects of health production commonly not treated in the applied literature. These include limits on the efficacy of treatment, uncertainty about the efficacy of treatment, and diminishing marginal product on both intensive and extensive margins.[3]

To facilitate the incorporation of these technological properties into the consumer choice problem, I depart in a fundamental way from the previous literature on health care demand. Most of the previous literature has modeled health care demand by expressing some outcome variable, such as expenditures, doctor visits, or inpatient days, as an explicit function of data, parameters, and unobservables. Following the existing literature on discrete/continuous choice, one could then apply Roy's identity to the specified demand function to find the corresponding indirect utility.[4] However, specifying a demand function that captures the intuitive aspects of health production and yields an analytical solution for the indirect utility may be very difficult. In this research, I take the opposite approach. I first specify the consumer choice problem by defining preferences and constraints in a way that captures basic intuition, and then define demand implicitly by the first order conditions of the consumer.

[3] Intensive diminishing marginal product is the usual notion that marginal product declines with inputs. Extensive diminishing marginal product states that the marginal product at any given quantity demanded is lower if the consumer is in a healthier pretreatment state.

[4] See Dubin and McFadden (1984) and Hanemann (1984) for examples.

I also depart from the previous literature by modeling jointly five types of health care utilization: inpatient hospital care, outpatient hospital care, doctor visits, prescription drugs, and dental care. While this creates significant computational burden, there are three important benefits. First, it captures the intuition that consumers often purchase health care in bundles for a given pretreatment state. Second, since the model captures a large share of expenditures, it gives a relatively complete picture of consumer demand and, as will be discussed below, consumer valuation of health plans. Third, it allows the model to capture risk, or propensity of illness, in multiple dimensions. This is important because while Medicare Advantage plans may generally experience favorable selection, plans that offer extra benefits such as a prescription drug or dental plan may experience adverse selection among beneficiaries who are high-risk for those types of treatment.

The model of health care demand based on utility maximization provides an intuitive basis for health plan choice. In the model, the value of a health plan for a risk averse consumer is an expected indirect utility where expectations are taken over a known distribution of health states. The parameters of this distribution depend on the risk type of the consumer, which is a function of observed and unobserved characteristics. As Cameron et al (1988) point out, this approach, while intuitive, is computationally burdensome in the absence of simplifying assumptions on preferences or health production that permit analytical solutions to the consumer choice problem on utilization.[5] These assumptions may be used in conjunction with distributional assumptions that allow closed form integration.[6] In order to model the intuitive properties of health care demand and health plan choice described above, I avoid such assumptions in this research. Instead, I rely on numerical and simulation methods to estimate the model.

Although the model to be presented is a microeconomic model, I use both individual and aggregate level data in estimation. Individual level data on health plan choice and subsequent health care utilization is contained in the 2000 Medicare Current Beneficiary Survey Cost

[5] A common assumption in both the theoretical and applied literatures is a deterministic, linear relationship between health care and health. See Dardanonni and Wagstaff (1991), Zabinski (1994), and Blomqvist (1997) for examples.

[6] For example, Zabinski (1994).

and Use file. I also use aggregate data on Medicare beneficiary managed care enrollment and eligibility from the year 2000 at the county/age/gender group level. The aggregate level performs three important functions in this research. First, the decomposition by demographic group is important in that it reveals risk selection on observable characteristics at the market level in addition to what is observed in the individual level data. Second, the aggregate level data is used to construct market shares equations that allow me to solve for health plan specific unobservables that will play a role in the individual health plan choice problem. Third, the aggregate data allows me to construct additional moments that permit more precise estimation of the model parameters.

The results are consistent with the expected effects of risk selection on welfare, although the effect is small. To the extent that older beneficiaries are higher risk, the benefits of Medicare Advantage should accrue disproportionately to younger beneficiaries. The average annual Medigap premium is predicted to decrease by less than \$5.00 if the Medicare Advantage program were discontinued, a reduction of less than one percent. While Medicare Advantage is estimated to increase consumer surplus in all age/gender groups on average, the change in consumer surplus is not uniform across age/gender groups and that higher risk beneficiaries are more likely to be adversely affected by Medicare Advantage. The average within age/gender group change in consumer surplus ranges from \$134.42 for women age 85 and over to \$277.39 for women age 65-69.

The remainder of the paper is organized as follows. Section 2 provides background on Medicare, Medicare Advantage, and Medigap. Section 3 describes the data. Sections 4 and 5 present the structural model and empirical strategy, respectively. Section 6 gives the results, specification testing, and a discussion. Section 7 concludes.

2 Background

2.1 Medicare

Medicare provides public health insurance to over 42 million Americans who are either over 65 years of age, disabled, or have end stage renal disease (ESRD). Medicare is the single largest payer of health care services in the United States, covering more than 16.8% of national health expenditures in 2003.[7] In fiscal 2005, Medicare outlays were $333.1 billion, making it the third largest federal program, behind only Social Security and Defense. Spending on the Medicare program comprised 13.5% of all federal outlays and 2.7% of GDP in 2005.[8]

Medicare was set up in two main parts, and its benefits package was based on that of traditional indemnity insurance plans that dominated the private sector in the mid 1960s. Part A, which covers hospitalization and skilled nursing care, is compulsory and is financed by a 2.9% payroll tax. Part B, which covers physician services and most outpatient care, is voluntary and is financed through general revenues and a monthly premium ($45.50 in 2000, $88.50 in 2006). The Part B premium is approximately 25% of the average cost of services under Part B and 10% of the average cost of all Medicare services.

Both parts of Medicare leave the beneficiary exposed to significant risk. Part A coverage has a $840 per event deductible ($776 in 2000) and runs out after 150 hospital days. Part B has a $100 annual deductible, a 20% coinsurance rate for covered medical care services and no maximum annual out-of-pocket expenditure limit. In addition, it excludes dental care, eye care, many types of preventive care, long-term care, and, until 2005, outpatient prescription drugs.

Because of this exposure to risk, over 90% of beneficiaries have some type of supplemental coverage. Common sources are Medicaid, employer provided retiree coverage, and nongroup Medigap policies. Another option for Medicare beneficiaries to obtain extra coverage is to enroll in a Medicare MCO or private FFS plan through the Medicare Advantage program.[9]

2.2 Medicare Advantage

Since the passage of the Tax Equity and Fiscal Responsibility Act (TEFRA) in 1982, the Centers for Medicare and Medicaid Services (CMS) has contracted with federally qualified

[7] Smith, et al. (2005)

[8] Congressional Budget Office: www.cbo.gov.

[9] The distribution of sources of supplementary coverage in 2000 was: Employer provided coverage (36%), Medigap (27%), Medicare Advantage (17%), Medicaid (11%), No supplementary coverage (9%). (Newhouse, 2001)

MCOs to bear risk and provide health care services. Under this arrangement, the MCO agrees to provide all necessary care to any beneficiary who wishes to enroll in a specified service area (usually a county) in exchange for a capitated, risk-adjusted, monthly payment per enrollee. The risk adjustment is based on age, gender, Medicaid eligibility, and residence in a long-term care facility. Until 1998, the base reimbursement rate, known as the Adjusted Area Per Capita Cost (AAPCC), was based on a five year moving average of 95% of the average expenditure in that county by beneficiaries enrolled in FFS Medicare.

One of the attractions of the Medicare Advantage program is that premiums are typically much lower than Medigap premiums while still providing comparable supplemental coverage. Another attraction is that Medicare Advantage commonly offer some coverage for services not included in traditional Medicare such as dental care and outpatient prescription drugs. However, beneficiaries sacrifice open provider choice and subject themselves to more extensive supply side controls on utilization.

2.3 Medigap

Medigap policies generally cover the consumer cost sharing associated with Medicare Parts A and B and are sold in the nongroup market. Since 1992, Medigap policies have been regulated in that benefits are standardized by federal law. Except in Massachusetts, Minnesota, and Wisconsin, insurers may offer Medigap supplements in only ten standardized plans, referred to as Medigap Plans A through J. All of these plans cover Medicare Parts A and B coinsurance and most cover the Part A deductible as well. The two most commonly purchased, Plans C and F, cover the deductible and coinsurance in both Parts A and B, reducing the first-dollar and marginal prices for any Medicare covered service to zero.

3 Data

The data for this research comes from multiple sources with three primary data sets provided by the CMS. These data sets include information on health insurance choice and health care utilization of individual Medicare beneficiaries, characteristics and counties of operation of all Medicare managed care organizations (MCOs), aggregate enrollment of these MCOs at the county/age/gender level, and county level information on Medigap premiums and per beneficiary Parts A and B spending in the fee-for-service (FFS) Medicare program.

3.1 Description

3.1.1 Individual Beneficiary Data

The individual level data for this research comes from the Medicare Current Beneficiary Survey (MCBS) 2000. The MCBS is a nationally representative survey of 12,305 Medicare beneficiaries containing data on health care utilization, health insurance choice, and demographic data such as age, gender, income, and zip code. Health insurance data provides information on all sources of coverage in which the beneficiary was enrolled and the identity of the firm providing coverage if the beneficiary was enrolled in a Medicare MCO.

Survey respondents are asked if their former (or current) employer has provided health insurance or has paid the cost of a Medigap supplement or Medicare MCO. The MCBS also provides data on Medicaid eligibility and on the source of Medicare eligibility (age, disability, or ESRD). Beneficiaries were removed from the sample if they reported having an employer sponsored health plan[10] (n=3838) or were eligible for Medicaid[11] (n=2407). Beneficiaries were also dropped if the source of Medicare eligibility is disability or ESRD (n=1944). With 1,135 beneficiaries in more than one of these categories, this leaves a sample

[10] I drop those with employer sponored coverage because no data on these health plans is observed so the choice set cannot be well-specified. The assumption is that anyone who is offered such coverage takes it up. In addition, dropping beneficiaries with employer sponsored coverage avoids complications associated with Medicare as secondary payer legislation. See Glied and Stabile (2001) for a discussion.

[11] I drop all individuals who are dual eligible or who are a Qualified Medicare Beneficiary (QMB). QMBs are low income beneficiaries who do not qualify for full Medicaid benefits but are eligible for Medicaid coverage of the Medicare Part B premium and all cost sharing of Medicare Parts A and B. Since this is roughly equivalent to a Medigap supplement at a zero premium, I assume that all QMB eligibles take up this coverage. I retain all Specified Low-Income Beneficaries (SLMB). These are low income beneficiaries who are eligible for a Medicaid subsidy of the Medicare Part B premium but still face the cost sharing associated with Medicare Parts A and B.

of 5,251 beneficiaries.

The MCBS contains data on health care utilization in nine categories five of which are used in this research: inpatient hospital care, outpatient hospital care, doctor visits, prescription drugs, and dental care. Together these five types of care comprise 67.4% of all health care spending documented in the MCBS data. Most of the remaining expenses (80.9%), take place in long-term care facilities. So the types of care examined in this study comprise 91.6% of all nonfacility based health care expenditures.

In specifying the structural model, an important issue is selecting an appropriate choice variable for the consumer. Many studies on health care utilization, for example Dowd, et al. (1991), Cameron, et al. (1988), and Deb and Trivedi (1997 and 2002), use event counts such as the number of doctor visits or the number of inpatient days as the dependent variable. Hunt-McCool, et al. (1994) estimate models explaining the number of doctor visits or hospital admissions as well as budget shares of expenditures on doctor visits or inpatient care. In presenting the results of the RAND Health Insurance Experiment, Manning, et al. (1987) use expenses as the dependent variable.

In the MCBS Cost and Use file, utilization is documented at the event level with a dollar amount associated with each event. I use total expenditures as opposed to event counts because there is considerable variation in spending within a given event count for each category of utilization. This variation seems to be important in estimating the distribution of unobserved components of consumer health characteristics. Since the quantities and intensity of specific treatments at each event are not observed, I use total expenditures for the year as a reasonable proxy for both the quantity and intensity of treatment. I account for price variation across regions by deflating total expenditures by a county level price index. The price index I use is the hospital wage index provided by the CMS.

3.1.2 Aggregate Enrollment Data

The Medicare Managed Care Quarterly/State/County/Plan database provides enrollment and eligibility counts from January 2000 at the county level for each Medicare MCO. This data has been made available by the CMS disaggregated into twelve gender/age groups. The age groups are: less than 65, 65-69, 70-74, 75-79, 80-84, greater than 84. I use all age groups except the less than 65 group, since, by definition, the source of Medicare eligibility is either disability or ESRD.

3.1.3 Medicare Managed Care Data

The Medicare Health Plan Compare database provides detailed information on product characteristics of all Medicare Advantage plans in every county such as the premium, cost sharing arrangements for each category of health care, and the scope of services provided. This data allows me to construct the complete choice set of all Medicare Advantage health plan options, with the relevant product characteristics, for all Medicare beneficiaries.

3.2 Data Limitations

There are three important data limitations with respect to choice of supplemental insurance coverage. First, counties typically have many firms offering multiple Medigap plans, but premiums and market shares for these plans are not available. Second, the share of Medicare beneficiaries enrolled in Medicaid is observed at the state level but not the county level. Third, market shares of non-Medicare Advantage private plans, such as those provided as part of a retiree benefits package, and their product characteristics, are not observed.

I take the following steps to account for these limitations. I follow Town and Liu (2003) and Atherly et al. (2004) by reducing the choice set among Medigap options to a single plan. Following Town and Liu (2003), I have obtained premium data from the AARP on its Plan F Medigap policy. Plan F is the most popular Medigap option nationally and the AARP, with over 2 million enrollees, is one of the largest providers of Medigap insurance (Town and Liu, 2003). So for each beneficiary in the model, the choice set includes FFS Medicare with the proxy Medigap supplement, all of the Medicare Advantage plans offered in that county, and FFS Medicare with no supplemental coverage.

The aggregate enrollment data will be used to construct market shares within age/gender groups for each Medicare MCO in each county of operation. Consequently, the size of the market in each county is adjusted to reflect Medicaid eligibility and employer sponsored supplemental health insurance. To account for Medicaid eligibility, I again follow Town and Liu (2003) by using information on state level Medicaid enrollments to deflate the size of the Medicare market of each county within that state. I assume that the state level enrollment rates are invariant across the age/gender groups. To account for employer sponsorship, I use the MCBS data to adjust the market size by using the within age/gender group percentage of those who reported having an employer sponsored health plan. Since the MCBS data set is small, I use the same rate to adjust the market size of each age/gender group across all counties in the country.

3.3 Descriptive Statistics

Figure 1 plots the distribution of log-expenditures of the sum of the five types of health care examined in this study for the Medicare Advantage and FFS Medicare groups in the full MCBS. The left intercept is the share of beneficiaries in each group that consumed no health care. This figure gives a clear picture of first order stochastic domination by the FFS Medicare group, implying that beneficiaries enrolled in Medicare Advantage utilize less health care, on average. Over 8% of the Medicare Advantage group consumed no health care compared to only 4% of the FFS group.

Of the 5,251 beneficiaries in the sample, 3,103 (59.1%) chose FFS Medicare with a Medigap supplement, 709 (13.5%) chose FFS Medicare only, and 1,439 (27.4%) chose a Medicare Advantage plan. Figure 2 gives the log-expenditure distributions for the Medigap, Medicare Advantage, and FFS-only subsamples used in this research. First order stochastic domination of the Medigap group is evident, but now the Medicare Advantage group dominates the FFS-only group through the 70th percentile. At that point, the FFS-only group dominates the Medicare Advantage group. These patterns are not surprising since Medigap enrollees generally face no cost sharing for inpatient care, outpatient care, or doctor visits, and while Medicare Advantage enrollees often face very little cost sharing, they are also subject to supply side controls that may hold down expenditures.

Table 1 gives mean expenditures and shares of total health care expenditures for the five types of health care used in this study by insurance group. On average, health care expenditures by beneficiaries in the Medigap subsample was \$7777, follow by the FFS-only group with \$5485, and the Medicare Advantage group with \$4860. Note the shift toward prescription drugs and dental care in the Medicare Advantage group. 17.8% and 6.0% of expenditures went towards prescription drugs and dental care, respectively, while the corresponding percentages for the Medigap and FFS-only groups are 12.2% and 3.3%, and 9.6% and 1.8%.

Table 2 gives the demographic characteristics of the Medigap, Medicare Advantage, and FFS-only groups. For comparison, I've included the same data for the two groups excluded from the sample: those with employer provided health insurance and those eligible for full Medicaid benefits or QMB status. The table excludes those whose Medicare eligibility is based on either disability or ESRD. Of the three included groups, Medigap enrollees, on average, are older, have the highest income, and are most commonly female. Table 3 gives average total health care spending by age/gender group. In four of the five age groups, men spend more than women, on average, and for both men and women, average expenditure peaks in the 75-79 age group.

Table 4 gives the basic cost sharing terms of Medigap Plan F and the descriptive statistics of the Medicare Health Plan Compare database. The national Medicare population weighted average of the monthly Medigap premium is \$131.24. The same average in just the counties with at least one Medicare Advantage plan is \$136.27. Medicare Advantage premiums are much lower, on average. The weighted average is \$37.17 with nearly a third of the plans being offered at a zero premium. 73.8% of all plans offer some prescription drug coverage and 20.9% offer some dental care coverage. Copayments for the services that are covered by FFS Medicare are generally very low. More than 86% of plans have no copayment for inpatient or outpatient care services and the average copayment for physician visits is \$7.89.

Turning to the aggregate enrollment data, there were 3782 county/MCO/plan combinations in 910 different counties in the Medicare Advantage program in 2000. The total Medicare population in these counties is 24,075,451, or roughly 60% of the Medicare population nationwide. I drop all counties in which the total Medicare Advantage market share is less than one percent, leaving a sample of 839 counties with 1765 county/MCO combinations and 3266 county/MCO/plan combinations. The total Medicare population in 839 counties is 23,624,176, or 98.1% of all beneficiaries with access to at least one Medicare Advantage plan.

Figure 3 gives the share of men and women enrolled in Medicare Advantage by age group. Aside from the less than 65 age group, which is excluded in this study, the pattern shows enrollment shares falling in age for both men and women. Overall, 16.7% of women were enrolled in a Medicare Advantage health plan as were 16.5% of men.

Figure 4 plots the proxy Medigap premium on Medicare Advantage market share for the 839 counties included in this research. The dashed line is the average Medigap premium for those counties with no Medicare MCOs. The figure indicates that MCOs enter the Medicare Advantage program in counties with generally higher Medigap premiums. Since MCOs generally cluster in urban areas, where establishing provider networks is less costly, this is not surprising. The solid curved line is a sixth order polynomial trend. The upward trend is consistent with the hypothesized effect of risk selection in this market: Medigap premiums would be higher in markets with large manage care penetration if lower risk beneficiaries systematically enroll into managed care. Of course, the figure cannot be taken as evidence of this causal relationship because Medigap premiums and Medicare Advantage market shares are jointly determined.

Both the individual and aggregate data are broadly consistent with previous findings that Medicare Advantage plans experience favorable selection. Generally, Medicare Advantage plans are more likely to attract younger Medicare beneficiaries and, on average, health care expenditures are lower for Medicare Advantage enrollees. There is also evidence that Medigap premiums are higher in counties with greater Medicare Advantage market share, a result that is consistent with the effect of risk selection. As noted earlier, lower expenses in the Medicare Advantage group could result from lower risk beneficiaries choosing Medicare Advantage, differences in consumer cost sharing arrangements, or unobserved MCO characteristics such as stringent utilization review practices.

4 Economic Model

The consumer's problem is modeled in two stages involving a discrete and then a continuous choice. In the first stage, the consumer chooses from available health plans given an income and health endowment, knowledge of the distribution of health states with private information about the mean, and complete information on the product characteristics of each health plan available in the market. In the second stage, the consumer receives a draw from the distribution of health states and then chooses a vector of health care inputs subject to the budget and productivity constraints implied by the choice of health plan. The consumer has imperfect information about the efficacy of treatment and so chooses health care inputs to maximize expected utility, where expectations are taken over a known distribution of outcomes for a given level of treatment.

4.1 Utilization Choice

I develop the consumer choice model of health care utilization based on some intuitive properties about health production and basic consumer theory. The model is developed to capture the following properties about health production:

i) Limits on the efficacy of treatment.

ii) Uncertainty about the efficacy of treatment for given quantities of health care.

iii) Diminishing marginal product in the production of health both intensively and extensively.

No objective measures of posttreatment health are used in this analysis. So the structure of the model, in particular the health production function, is guided by intuition only and is not testable in the usual sense of how well it explains the observed relationship between health care inputs and health output. As will be discussed later, the estimation algorithm is based on fitting the model to the utilization data by solving for the set of the unobserved components of pretreatment health that is consistent with defining the observed data as the solution to the constrained optimization problem faced by the consumer.

4.1.1 Basic Model

To simplify the exposition, I first consider the simple consumer choice problem with one type of health care

$$\underset{m}{MAX}U(C,H) \;\; s.t.\, C = y - pm \;\; \text{and} \;\; H = H(m,\theta).$$

Here, θ denotes pretreatment health state with larger values indicating worse states of health, m denotes a scalar quantity of health care sold at unit price p, y denotes income, $H(\cdot)$ is the health production function, H is posttreatment health, and C is a numeraire.

For simplicity, I assume that $H(\cdot)$ is a second order polynomial in m. Higher order polynomials introduce the possibility of nonconvex upper contour sets, which would greatly complicate estimation. I also assume the initial condition

$$H(0,\theta) = -\theta.$$

I assume that consumers are risk averse in both C and H. Hence, $U_C > 0$, $U_H > 0$, $U_{CC} < 0$, $U_{HH} < 0$ for all values of C and H. Finally, for simplicity, I assume $U_{CH} = 0$.[12]

The first property I consider is limits on the efficacy of treatment. To ensure that quantity demanded is finite when the marginal price is zero, a circumstance that occurs frequently when individuals have health insurance, I make the stronger assumption that there is some finite value of health care, denoted $\bar{m}$ such that $H_m > 0$ for $m < \bar{m}$, $H_m < 0$ for $m > \bar{m}$, and $H_m = 0$ for $m = \bar{m}$. If we consider consumer preferences on C and m (as opposed to C and H), the value $\bar{m}$ can thought of as a satiation point in m because the indifference curves slope up for $m > \bar{m}$, i.e., health care becomes a "bad".

The consumer choice problem is illustrated in Figure 5. The slope of the budget frontier is $-p$ and the marginal rate of substitution is $-\frac{U_H H_m}{U_C}$. If the consumer is facing a positive price, the consumer will choose $m^* < \bar{m}$, irrespective of income. This is illustrated by the solid, downward sloping budget line. If the consumer has full insurance and, hence, paid nothing for health care, the budget frontier would be the horizontal, dashed line. The budget

[12] The assumption of a zero cross-partial derivative is consistent with Zabinski (1994) and Hubbard-Rennhoff (2005), but not Cardon and Hendel (2001), who assume the utility is a second order polynomial in health care and the numeraire.

set would contain arbitrarily large quantities of health care. If preferences were monotone in health care, the model would predict these arbitrarily large quantities for those with full insurance. Satiation in preferences rules this out. In this case, the consumer would choose $m^* = \bar{m}$, again, irrespective of income.

The assumption that $H(\cdot)$ is a second order polynomial in m implies

$$H_m(m, \theta) = \tau(\bar{m} - m), \quad \text{for some } \tau > 0.$$

The parameter τ scales the height of the production function for different levels of m for a given satiation point $\bar{m}$. Integrating and applying the initial condition gives

$$H(m, \theta) = \tau \left(\bar{m}m - \frac{1}{2}m^2 \right) - \theta.$$

Diminishing marginal product on the extensive margin can be incorporated by assuming that $\bar{m}$ is an increasing function of θ, i.e., the marginal product at any given level of treatment is higher if the consumer is in a worse pretreatment state. Uncertainty about the efficacy of treatment can be added by assuming the satiation point, given pretreatment health, is stochastic. Hence,

$$H(m, \theta) = \tau \left(\bar{m}(\theta, \nu)m - \frac{1}{2}m^2 \right) - \theta, \quad \text{with } \frac{\partial \bar{m}(\theta, \nu)}{\partial \theta} > 0, \frac{\partial \bar{m}(\theta, \nu)}{\partial \nu} > 0, \tag{1}$$

where ν is unknown when the treatment decision is made and is drawn from a known distribution F_ν. Note that H is deterministic when $m = 0$. The function (1) is the basis for the detailed model of health production described below.

4.1.2 Full Model

In this section, I describe the full model of consumer choice with multiple types of health care. I begin with a description of consumer preferences and endowments, and then extend the simple model of health production presented above to the higher dimensional setting. As noted earlier, five categories of health care utilization are used in this model: inpatient hospital care, outpatient hospital care, doctor visits, prescription drugs, and dental care. I index these categories with the set $T = \{IP, OP, DV, PD, DC\}$. In addition to the expansion of the consumer choice problem, the full model also presents several concepts absent in the simple model. These include imperfect and asymmetric information about the pretreatment health state, nonpecuniary costs of health care, and heterogeneity in the production of health that is health insurer specific.

Preferences As in the basic model, consumer i has preferences over posttreatment health H_i and the consumption of a numeraire C_i. Health care affects utility through C_i via the budget constraint and through H_i via the health production function. In the full model, however, health care also has a direct effect on utility that captures the nonpecuniary costs, or disutility, associated with the consumption of health care the intuition for which is explained below. Let $m_{ijk} = \{m_{ijkt}\}_{t \in T}$ denote a 5×1 vector of health care inputs purchased from health plan k offered by firm j (hereafter referred to as plan jk). I assume that preferences can be represented by the utility function

$$U(C_i, H_i, m_{ijk}) = \frac{C_i^{1-\gamma_1} - 1}{1 - \gamma_1} + \delta \frac{H_i^{1-\gamma_2} - 1}{1 - \gamma_2} - \kappa_1' m_{ijk} + \kappa_2' m_{ijk}^2 - \kappa_{IP}^F 1\left[m_{ijkIP} > 0\right]$$

where δ, γ_1, γ_2, κ_1, κ_2, κ_{IP}^F are parameters to be estimated. The parameters γ_1 and γ_2 measure the degree of risk aversion in the consumption of the numeraire C and the level of health H, respectively. The parameter δ measures the rate at which the consumer is willing to give up utility derived from C for utility derived from H.

The parameters κ_1 and κ_2 are 5×1 vectors that capture continuous nonpecuniary costs of health care. The parameter κ_{IP}^F denotes a fixed nonpecuniary cost that is incurred only if some inpatient hospital care is consumed. In addition to being intuitive, the disutility parameters are important if the model is to fit the data well. It is not uncommon in the data for a consumer to choose a zero amount of some type of health care despite a zero first dollar price. As long as the first unit of health care is productive, some nonpecuniary cost is required if the model is to explain the data. The coefficients on m_{ijk} and m_{ijk}^2 capture this nonpecuniary cost. Intuition suggests that each element of κ_1 is positive but each element of κ_2 could be positive or negative. The fixed cost associated with inpatient care κ_{IP}^F is included because the lowest nonzero inpatient care expenditure exceeds \$300 while each of the other four types of care have positive observations below \$5. The fixed cost is included to explain this discontinuity.

Endowments Let pretreatment health state of consumer i be denoted by the 5×1 vector θ_i with elements $\{\theta_{it}\}_{t \in T}$. I decompose the vector θ_i into two 5×1 component vectors, one of which is known to the consumer when the health insurance decision is made, ξ_i, and the other represents a health shock that is revealed after the health insurance choice is made, ϵ_i. Therefore,

$$\theta_i = \xi_i + \epsilon_i.$$

Each component of ξ_i is assumed be the sum of a linear-in-parameters index of observable characteristics that are related with health X_i and heterogeneity that is unobserved to health plans and the researcher but known to the consumer. Define the vector of these unobservables as η_i and each component of ξ_i as

$$\xi_{it} = X_i \beta_t + \eta_{it}, \ \forall t \in T,$$

where β_t is a parameter vector to be estimated. The vector X_i is composed of age and its square and gender. The vector ξ_i can be thought of as the risk type of the consumer and the vector η_i captures the consumer's private information about risk type.

Finally, consumer i is exogenously endowed with an income y_i. For simplicity, I assume that y_i does not depend on ξ_i.

The Production of Health In this section, I extend the simple model of health production in (1) to include health plan specific heterogeneity and multiple types of health care. In the model, the consumer has a pretreatment health state vector θ_i and can increase health by purchasing a health care vector m_{ijk}, which will map into a posttreatment health state vector, denoted h_i. This vector will then be mapped to the scalar H_i, which is an argument of the utility function.

To introduce plan specific heterogeneity in the production of health, I assume the satiation point $\bar{m}$ and the parameter τ, first presented in section 4.1.1, vary over insurers and over types of treatment within insurers. The intuition behind this heterogeneity is discussed in the next section. For simplicity, I assume $\tau_{jt} > 0$ and $\bar{m}_{jt} > 0$. This assumption implies that the first arbitrarily small amount of health care has a positive marginal benefit, irrespective of pretreatment state, i.e., the satiation point and the total product at the satiation point are always positive. Given this assumption, let

$$\begin{aligned}
\tau_{jt} &= \exp\{\chi_{1jt}\} \\
\bar{m}_{jt} &= \exp\{\chi_{2jt} + \nu_{it} + \zeta_t g(\theta_{it})\},
\end{aligned}$$

where χ_{1jt} and χ_{2jt} denote firm specific unobserved characteristics associated with firm j and treatment t. Note that χ_{1jt} and χ_{2jt} are not indexed by k, so the unobserved product characteristics are assumed to be constant across plans within a firm. The firm specific unobservable χ_{2jt} shifts the expected satiation point and the firm specific unobservable χ_{1jt} determines the total product for a given satiation point and m_{ijkt}. Diminishing marginal product on the extensive margin is captured in $g(\theta_{it})$, an increasing function which will be specified below, and ζ_t, a parameter to be estimated. As in section 4.1.1, the health production function exhibits diminishing marginal product on the extensive margin if the marginal product is increasing in θ_{it}. Since the exponential is an increasing function, this will be true if $g'(\theta_{it}) > 0$ and $\zeta_t > 0$. The stochastic element ν_{it} represents uncertainty about the efficacy of treatment for health care category t that is unknown when the utilization decision is made.

Given this assumption, I define h_{it}, the posttreatment health for consumer i in dimension t from consuming health care purchased from plan jk, as

$$h_{it} = \exp\{\chi_{1jt}\}\left(\exp\{\chi_{2jt} + \nu_{it} + \zeta_t g(\theta_{it})\}m_{ijkt} - \frac{1}{2}m_{ijkt}^2\right) - \theta_{it}. \tag{2}$$

To map the five-element set $\{h_{it}\}_{t\in T}$ to the scalar argument H_i, I use the constant elasticity of substitution (CES) function

$$H_i = \left(\sum_{t\in T}\alpha_t h_{it}^{\rho}\right)^{\frac{1}{\rho}}, \tag{3}$$

where $1/(1-\rho)$ is the elasticity of substitution and the parameters $\{\alpha_t\}_{t\in T}$ denote weights associated with the elements of $\{h_{it}\}_{t\in T}$. The elements of $\{\alpha_t\}_{t\in T}$ are assumed to follow the restrictions $\alpha_t \geq 0,\ \forall t \in T$, and $\sum_{t \in T} \alpha_t = 1$. I have no particular reason in selecting the CES function that is relevant to health or health care. The rationale behind choosing it is only to permit a flexible relationship between the set $\{h_{it}\}_{t \in T}$ and the scalar H_i. For example, the CES function includes perfect substitute, perfect complement, and Cobb-Douglas functions as special cases:

$$H_i = \sum_{t \in T} \alpha_t h_{it},\ \text{for } \rho = 1,\ \lim_{\rho \to -\infty} H_i = \min_t \{\alpha_t h_{it}\},\ \text{and } \lim_{\rho \to 0} H_i = \prod_{t \in T} h_{it}^{\alpha_t}.$$

Unobserved Product Characteristics As described above, each health plan has two unobserved (to the researcher) product characteristics, χ_{1jt} and χ_{2jt}, that affect the productivity of health care in each dimension t. The intuition behind these product characteristics is based on work by Jin (2005) and Cutler et al (2005). In both studies, MCOs are differentiated in terms of the quality of health care services provided. Since many MCOs have a restricted provider network or may limit access to specific technologies, it seems natural to differentiate health plans in dimensions other than the premium, benefits provided, and cost sharing arrangements and to suppose that this differentiation will affect consumer choice. The role of the vectors $\{\chi_{1tj}, \chi_{2tj}\}_{t \in T}$ is to capture such differentiation.

This model of product differentiation is inadequate in the following sense. Some health plans may be characterized as providing a low quality product and not very restrictive in the provision of services. Others may exercise considerable control on access to services, which acts as a constraint on the consumer choice problem, but these services may be of a high quality. Unfortunately, there is nothing in the data that will allow me to separate these plan characteristics. In the model here, consumers are assumed to solve an optimization problem free of constraints imposed by the health plan that could lead to inequalities in the first order conditions at interior solutions. Hence, this model of product differentiation confounds a unrestrictive low quality product and a restrictive high quality product, both of which would lead to reduced utilization.

Finally, I assume that the vectors $\{\chi_{1tj}, \chi_{2tj}\}_{t \in T}$ is the same for the Medigap option and the FFS-only option in each county. This is intuitive because both Medigap enrollees and FFS-only beneficiaries have open provider choice, so these choices should be differentiated by cost sharing arrangements only. As will be discussed in the section on estimation, the aggregate data on health plan choice in the Medicare Advantage program will allow me to solve for the vectors $\{\chi_{1tj}, \chi_{2tj}\}_{t \in T}$ at the MCO/county level. However, this data will not allow me solve for the corresponding vector for FFS Medicare in each county. Instead, I use additional data on county level Medicare spending on Parts A and B services two solve for two county-specific fixed effects that will capture variation in productivity across counties in the FFS Medicare sector.

Specification of $g(\theta_{it})$ In capturing the property of diminishing marginal product on the extensive margin, the specification of $g(\theta_{it})$ requires some care if the model is to retain some intuitive comparative statics. Specifically, intuition suggests that consumers in more adverse pretreatment state will consume more health care, ceteris paribus, i.e., $\frac{\partial m^*_{ijkt}}{\partial \theta_{it}} > 0$. However, if health production grows too quickly in θ_{it}, then a marginal increase in θ_{it} could lead to lower optimal quantity of health care. Intuitively, this occurs because if the marginal product is higher at some level of m_{ijkt}, then the marginal product is higher at the inframarginal units as well. In the extreme, it could be that health care becomes so productive that the model may imply that a consumer is better off in more adverse pretreatment states. Therefore, if the model is to maintain the comparative statics that a more adverse state of pretreatment health leads to greater optimal quantities of health care and lower indirect utility, then the expression

$$\frac{\partial \exp\{\chi_{2jt} + \nu_{it} + \zeta_t g(\theta_{it})\}}{\partial \theta_{it}} = \exp\{\chi_{2jt} + \nu_{it} + \zeta_t g(\theta_{it})\}\zeta_t g'(\theta_{it}) \tag{4}$$

will need to be carefully specified.

Early experiments indicated that the identity function $g(\theta_{it}) = \theta_{it}$ would be a poor choice for the reasons described above. Simulations of optimal health care choices under counterfactual health states often resulted in a positive relationship between indirect utility and θ_i for large values of θ_i. The problem is that, in this case, $\bar{m}_{itj}$ explodes in θ_{it} leading to large values of H_i even for very small values of ζ_t.

After a number of other experiments, I concluded that a reasonable choice is

$$g(\theta_{it}) = \left(1 + \exp\left(-\frac{\theta_{it} - a_t}{b_t}\right)\right)^{-1}$$

where a_t and $b_t > 0$ are parameters to be estimated. The purpose of a_t and b_t is to scale θ_{it} so that $g'(\theta_{it})$ is not close to zero over a broad range of the values of θ_{it}.[13] Under this specification, (4) is

$$\frac{\partial \exp\{\chi_{2jt} + \nu_{it} + \zeta_t g(\theta_{it})\}}{\partial \theta_{it}} = \exp\{\chi_{2jt} + \nu_{it} + \zeta_t g(\theta_{it})\}\zeta_t g(\theta_{it})(1 - g(\theta_{it}))/b_t \tag{5}$$

which is strictly positive, bounded above for given values of χ_{2jt} and ν_{it} (since $g(\theta_{it}) \in (0,1)$), and approaches zero as θ_{it} becomes arbitrarily large or small. Early tests using this specification were successful in that the optimal quantities are increasing and indirect utilities are decreasing in θ_{it}. The property of diminishing marginal product on the extensive margin is maintained because (5) is strictly positive. A more detailed discussion is presented in Appendix 8.1.

The Budget Constraint As noted in section 3, I use deflated expenditures, which proxies for quantity and intensity of treatment, as the choice variable for the consumer. For many choices of health plans, this will present complication in defining out-of-pocket expenses and marginal prices. If the consumer is facing a simple coinsurance rate such as 20% of expenses after some annual deductible, then defining out-of-pocket expenses and marginal prices is straightforward. If, however, the consumer pays a flat fee per event, then defining total out-of-pockets and marginal prices is not obvious. Some relationship must be established between deflated expenditures and number of events if the model is to conform to the data.

Cardon and Hendel (2001) address this by specifying a constant marginal price of health care that is a parameter to be estimated and is constant across all managed care plans in the data. This approach ignores all of the observed data on plan specific cost sharing terms and so the model would lose some important heterogeneity. Instead, I use the MCBS data to pre-estimate a reduced form relationship between event counts and a polynomial in expenditures,

[13] For example, if $a_t = 0$ and $b_t = 1$ and $g(\theta_{it})$ is evaluated using double precision, then $g(\theta_{it}) = 1$ if $\theta_{it} > 308$ and $g(\theta_{it}) = 0$ if $\theta_{it} < -50$.

types of insurance, and exogenous county characteristics. I then use the parameter estimates to evaluate the conditional expectation of event counts given the observed utilization data. Total out-of-pockets for each type of care is the copayment times the predicted number of events and the marginal price is the copayment times the derivative of the conditional expectation of events with respect to utilization. I present the details in Appendix 8.2.

For simplicity of notation, let $p_{jkt}(m_{ijkt})$ denote the appropriately defined marginal price for health care type t in health plan jk face by the consumer. Let $oop_{it}(m_{ijkt})$ denote total out-of-pockets incurred by consumer i from health care utilization m_{ijkt} and

$$oop_i(m_{ijk}) = \sum_{t \in T} oop_{it}(m_{ijkt})$$

denote total out-of-pockets. Finally, let π_{jk} denote the premium for plan jk and π_B denote the Medicare Part B premium. The budget constraint is then

$$C_i = y_i - \pi_{jk} - \pi_B - oop_i(m_{ijk}).$$

4.1.3 Optimization

Given the realization of ϵ_i, the beneficiary chooses m_{ijk} so as to maximize U given the constraints imposed by the choice of health plan jk and income y_i. The vector m^*_{ijk} therefore satisfies the vector of first order conditions

$$\begin{aligned} U_{m_{ijk}t}(m_{ijk}, \epsilon_i, jk)\big|_{m^*_{ijk}} &= -\frac{p_{jkt}(m_{ijkt})}{[y_i - \pi_{jk} - \pi_B - oop_i(m_{ijk})]^{\gamma_1}} - \kappa_{it} + 2\kappa_{2t} m_{itjk} \\ &\quad + \delta \int \cdots \int \frac{\alpha_t h_{it}^{\rho-1} \frac{\partial h_{it}}{\partial m_{ijkt}}}{H_i^{\gamma_2+\rho-1}} dF_\nu(\nu_i) \\ &\leq 0, \ \ \forall t \in T, \ \text{ where} \\ \frac{\partial h_{ir}}{\partial m_{ijkt}} &= \exp\left\{\chi_{1jt}\right\} \left(\exp\{\chi_{2jt} + \nu_{it} + \zeta_t g(\theta_{it})\} - m_{ijkt}\right). \end{aligned} \tag{6}$$

The terms in the first line of (6) denote the marginal costs, pecuniary and nonpecuniary, of health care. The term in the second line of (6) denotes the expected marginal benefit where expectations are taken over F_ν. The components of the system $U^*_{m_{ijk}} \leq 0$ corresponding to types of care for which m^*_{ijkt} is positive implicitly define demand functions in a neighborhood of m^*_{ijk}.

4.1.4 Corner Solutions

As in most data on health care utilization, corner solutions are common in the MCBS utilization data. Only 288 of 5,251 beneficiaries in the MCBS included in my sample consume a positive amount of all five types of medical care modeled in this study. Many studies have used a two-part approach to censoring, i.e., first estimating the probability of any care and then estimating a linear model with the appropriate error correction, (for example, Dowd, et al. (1991)). The relationship between the unobservables and observed quantities implied by the structure of this model suggests that a different method will be required.

The approach is analogous to a simple tobit model which integrates the density of the unobserved component of a linear model over the region of the support that is consistent with some observed, censored outcome. The task here is to integrate the joint density of ϵ_i over the region of $\mathbb{R}^5$ that is consistent with an observed vector of quantity choices that may be a mix of interior and corner solutions, given a vector ξ_i. If the beneficiary consumes a positive amount of all five types of care, there is one vector $\epsilon_i \in \mathbb{R}^5$ that is consistent with the observed utilization data. However, if at least one component of m^*_{ijk} is zero, there will be a region, rather than a point, in $\mathbb{R}^5$ that will be consistent with the observed outcome. To establish some notation, let Θ denote the set of model parameters, ϵ_{i_C} the elements of ϵ_i for which a corner solution is observed, F_{ϵ_C} the marginal distribution of ϵ_{i_C}, and $F_{\epsilon_I}(\cdot|\epsilon_{i_C})$ the conditional distribution of the elements of ϵ_i for which an interior solution is observed.

This structure is similar to the demand system analyzed in Wales and Woodland (1983). However, in their model the unobservables enter the demand equations linearly and the t^{th} unobservable enters the t^{th} demand equation only. These restrictions do not hold in this demand model, causing some complication in integrating over the region of interest. First, the limits of integration are defined by the first order conditions and so are functions of the model parameters and the data. Second, the optimality conditions $U^*_{m_{ijk}} \leq 0$ are nonlinear in the elements of ϵ_i, implying that the region of integration is not Euclidean, but rather a nonlinear manifold. This manifold is defined by the system $U^*_{m_{itjk}} = 0$ for all $t \in T$ such that $m^*_{ijkt} > 0$ and its dimension equals the number of corner solutions.

For example, suppose $m^*_{ijkIP} = m^*_{ijkOP} = m^*_{ijkDC} = 0$, while $m^*_{ijkDV} > 0$, $m^*_{ijkPD} > 0$, and consider the vector $\tilde{\epsilon}_i$ that solves the full system $U_{m_{ijk}} = 0$. Then the components of $\tilde{\epsilon}_i$ corresponding to IP, OP, and DC can be thought of as reservation states, realizations of ϵ_i at which the consumer is just indifferent between purchasing the first increment of health care and not in the dimensions corresponding to the corner solutions, given the observed positive quantities purchased. So the values of $\tilde{\epsilon}_i$ corresponding to IP, OP, and DC define the apex of a non-Euclidean subspace of $\mathbb{R}^3$, denoted $A_{jk}(\Theta)$, that is comprised of all points ϵ_{i_C} that are consistent with the observed data.

The region of integration $A_{jk}(\Theta)$ is indexed by jk since each point on the frontier is a function of plan characteristics via the system $U^*_{m_{ijk}} = 0$. Different product characteristics result in different censoring points for the same observed m^*_{ijk}. In this way, the MCO characteristics influence the probability (loosely speaking) of any care as well as the quantity given some care.

For any realization $\epsilon_{i_C} \in A_{jk}(\Theta)$, there exists some pair $(\epsilon_{iDV}, \epsilon_{iPD})$ that solves the system $U^*_{m_{ijk}}|_{\forall t \in T: m^*_{ijkt} > 0} = 0$. The set of these pairs comprise the region of interest and are functions of the model parameters and ϵ_{i_C}. So the expression for the joint density of ϵ_i that is consistent with the data is

$$\iiint_{\epsilon_{i_C} \in A_{jk}(\Theta)} f_{\epsilon_I}(\epsilon_{iDV}(\Theta, \epsilon_{i_C}), \epsilon_{iPD}(\Theta, \epsilon_{i_C}) | \epsilon_{i_C}) \left\| \mathbb{J}_{ijk}(\Theta, \epsilon_{i_C}) \right\| dF_{\epsilon_C}(\epsilon_{i_C}) \tag{7}$$

where $(\epsilon_{iDV}(\Theta, \epsilon_{i_C}), \epsilon_{iPD}(\Theta, \epsilon_{i_C}))$ denote implicit functions relating the elements of ϵ_i along the intersection of the manifolds defined by the first order conditions corresponding to the types of health care for which there is positive consumption, DV and PD. The Jacobian matrix is denoted by $\mathbb{J}_{ijk}(\Theta, \epsilon_{i_C})$ (in this case a 2×2 matrix), and $\|\mathbb{J}_{ijk}(\Theta, \epsilon_{i_C})\|$ is the Jacobian of the transformation from ϵ_i to m_{ijk} when $m^*_{ijkIP} = m^*_{ijkOP} = m^*_{ijkDC} = 0$.[14]

[14] If five interior solutions are observed, the Jacobian matrix is $\mathbb{J}_{ijk} = -\left[\frac{\partial U^*_{m_{ijk}}}{\partial \epsilon_i}\right]^{-1} \left[\frac{\partial U^*_{m_{ijk}}}{\partial m_{ijk}}\right]$. For observations with a mix of corner and interior solutions, the Jacobian matrix is the analog to the above expression using only the rows and columns of the two matrices corresponding to interior solutions.

4.1.5 Value of a Health Plan

Finally, I define the value function of plan jk to beneficiary i, given the endowment (ξ_i, y_i), as the expected indirect utility where expectations are taken over the known distribution of health states:

$$V_{ijk}(\eta_i) = \int \cdots \int U\left(m^*_{ijk}, \theta_i, y_i\right) dF_\epsilon(\epsilon_i). \tag{8}$$

Recall that $\theta_i = \xi_i + \epsilon_i$ and $\xi_i = X_i\beta + \eta_i$. In this notation, I emphasize the dependence on the vector η_i because it is the unobserved component of the consumer's health endowment and is integrated out of the likelihood contribution. For each value of ϵ_i, the consumer chooses the health care vector m^*_{ijk} conditional on having selected plan jk and having the endowment (ξ_i, y_i). The indirect utility for that draw of ϵ_i is $U\left(m^*_{ijk}, \theta_i, y_i\right)$. The value of health plan jk is a weighted average of these indirect utilities where the weights are given by $dF_\epsilon(\epsilon_i)$.

4.2 Health Insurance Choice

I use a standard discrete choice model in which the choice probabilities are based on the expected indirect utility (8) for each health plan in the choice set. I apply in the same model to both the individual and aggregate health insurance choice data.

A brief note on notation before proceeding: there is a distinction in applying the forthcoming discrete choice model to the individual and aggregate level data. The individual characteristics age and income are observed in the MCBS but not in the aggregate Medicare managed care enrollment data. The aggregate data is defined at the age/gender group level with age groups: 65-69, 70-74, 75-79, 80-84, 85 and over. So income and age within a group must be integrated out of the choice probabilities. However, the vector η_i is, of course, not observed in either data set and so must be integrated out in both individual and aggregate choice probabilities. To make notation clear, I adopt the following convention. In as (8), the value of plan jk to an individual beneficiary i, for whom X_i and y_i are observed in the data, is denoted $V_{ijk}(\eta_i)$. In contrast, the value of plan jk to a simulated beneficiary with draws X_i, y_i, and η_i is denoted $V_{jk}(\xi_i, y_i)$. In each case, the dependence emphasizes the variables that the value function is integrated over.

4.2.1 Individual Health Insurance Choice

In addition to the systematic health plan valuation described in the previous section, the beneficiary's valuation of a health plan is modeled as varying for idiosyncratic reasons. Let e_{ijk} denote this idiosyncratic variation in the value of beneficiary i for plan jk. Define e_i as a $2+\sum_{j=1}^{J} K_j$ vector of idiosyncratic terms where J is the number of Medicare MCOs in the county and K_j is the number of plan offered by MCO j. I assume the value of plan jk to beneficiary i conditional on the unobserved component of the consumer's health endowment η_i is

$$\check{V}_{ijk}(\eta_i) = V_{ijk}(\eta_i) + e_{ijk}.$$

Since a component of this idiosyncratic variation may reflect the beneficiary's general attitude towards managed care, it is not appropriate to assume that the draws e_{ijk} are mutually independent across MCOs or across plans within MCOs. Therefore, I assume a nested multinomial logit choice probability conditional on the vector η_i. In this nesting structure, the beneficiary is modeled as first choosing between the Medigap plan, no supplemental coverage, and a Medicare Advantage option. If Medicare Advantage is chosen, the beneficiary then chooses among MCOs and then among plans if the selected MCO offers multiple plans.

With these assumptions, and under the assumption that the beneficiary chooses the health insurance plan with the largest $\check{V}_{ijk}(\eta_i)$, the probability that beneficiary i chooses plan jk conditional on η_i is

$$\begin{aligned} P_{ijk}(\eta_i) &= \frac{\exp\left\{\frac{V_{ijk}(\eta_i)}{(1-\sigma_a)(1-\sigma_w)}\right\}}{\mathcal{I}_j^{\sigma_w}\mathcal{J}^{\sigma_a}\left[\exp\left\{V_{i,medigap}(\eta_i)\right\}+\exp\left\{V_{i,FFSonly}(\eta_i)\right\}+\mathcal{J}^{1-\sigma_a}\right]} \qquad (9) \\ \text{where} \quad \mathcal{I}_j &= \sum_{k=1}^{Kj}\exp\left\{\frac{V_{ijk}(\eta_i)}{(1-\sigma_w)(1-\sigma_a)}\right\} \text{and } \mathcal{J} = \sum_{j=1}^{J}\mathcal{I}_j^{1-\sigma_w}. \end{aligned}$$

4.2.2 Aggregate Health Insurance Choice

The definition of the choice probability at the individual level is also used to model the aggregate health plan choice data. I take the mixed logit approach similar to Berry, Levinsohn, and Pakes (BLP, 1995), Brownstone and Train (1999), and related literature. In this model, mixing distributions are based on both observable and unobservable characteristics

Health plan choice at the aggregate level is observed for ten age/gender groups in each county. I index these age/gender groups by g. As noted in the introduction, the aggregate health plan choice data is observed at the level of the MCO, not health plan. So the model prediction of the observed market share of MCO j in demographic group g is

$$P_{gj} = \int \cdots \int \left[\sum_{k=1}^{Kj} P_{jk}(\xi_i, y_i) \right] dF_{\eta}(\eta_i) dF_{y,X|g}(y_i, X_i) \tag{10}$$

where

$$P_{jk}(\xi_i, y_i) = \frac{\exp\left\{ \frac{V_{jk}(\xi_i, y_i)}{(1-\sigma_a)(1-\sigma_w)} \right\}}{\mathcal{I}_j^{\sigma_w} \mathcal{J}^{\sigma_a} \left[\exp\left\{ V_{medigap}(\xi_i, y_i) \right\} + \exp\left\{ V_{FFSonly}(\xi_i, y_i) \right\} + \mathcal{J}^{1-\sigma_a} \right]}$$

and $F_{y,X|g}$ denotes the distribution of consumer characteristics income and age given demographic group g.[15] This distribution is constructed empirically and will be detailed in the section on estimation. Note that the limitation of observing the aggregate data at the level of the MCO as opposed to plan is accounted for by aggregating the predicted market share over the set of plans offered by the MCO in that county. This is in contrast to Town and Liu (2003) who assume that all beneficiaries choose the plan with the lowest premium.

5 Estimation

5.1 Distributional Assumptions

There are four sets of stochastic elements in this model for which a parametric distribution will be assumed: the vector of unobserved components of the beneficiary's health endowment η_i, the vector of health shocks ϵ_i, the vector of health productivity shocks ν_i, and the vector defining the idiosyncratic component of beneficiary health plan valuation, e_i. As noted in section 4, I assume the vector e_i, is distributed GEV. I assume that the vectors ϵ_i, η_i, and ν_i are each distributed multivariate normal but the vectors are mutually independent. The assumption that ν_i is independent of ϵ_i and η_i is made for simplicity. The assumption that

[15] The dependence on y_i and X_i is added here because these characteristics are observed in the individual level data, but not in the aggregate level data and so must be integrated out conditional on being in demographic group g.

η_i and ϵ_i are independent can be made without loss of generality because they are additive in the definition of the pretreatment health state vector θ_i and the sum of two normal variables can always be defined as the sum of two orthogonal components. The assumption that ϵ_i and η_i are additive in the definition of θ_i also implies that the means of ϵ_i and η_i are not separately identified. As will be discussed below, the variances of ϵ_i and η_i are identified because both health insurance and health care utilization decisions are observed. In addition, the variance of ν_i is identified but the mean is not separately identified from the mean of the unobserved health productivity vector χ_2. Therefore, I assume

$$\begin{aligned}
\epsilon_i &\sim N(\mu, \Omega_\epsilon) \\
\eta_i &\sim N(0, \Omega_\eta) \\
\nu_i &\sim N(0, \Omega_\nu).
\end{aligned}$$

For simplicity, I assume that Ω_ν is a diagonal matrix.

5.2 Aggregate Data

I use the aggregate health plan choice data to uncover the Medicare MCO specific vector of unobserved product characteristics $\chi_j = (\chi_{1j}, \chi_{2j})$ and form moment conditions in a manner analogous to Berry, Levinsohn, and Pakes (BLP,1995). The algorithm involves first using data on aggregate FFS Medicare expenditures in each county to solve for the vector $(\chi_{1FFS}, \chi_{2FFS})$. Second, given the values $(\chi_{1FFS}, \chi_{2FFS})$, I solve for the vector χ_j that makes the observed market shares fit the theoretical market shares for each MCO j in the county. I then form moment conditions by assuming the vector χ_j is mean independent of a set of instruments. Finally, the set of unobserved product characteristics for all health insurance plans available in each county is brought to the individual data.

5.2.1 FFS Medicare

As noted in the section 4, the absence of county level Medigap enrollment data prevents me from solving for the vector (χ_1, χ_2) for the Medigap and FFS-only health insurance plans in each county. However, I use additional county level data from Medicare to solve for this vector under some additional restrictions. The CMS makes available the per beneficiary Medicare spending in Parts A and B at the county level. This average is taken over only those in FFS Medicare, with or without a Medigap supplement. These two values allow me solve for two parameters that are specific to a county and its FFS Medicare population. My approach is to set parameter vectors, denoted $(\bar{\chi}_{1FFS}, \bar{\chi}_{2FFS})$, that are common across all counties and then scale these vectors at the county level using the two county-specific parameters that will be found using the FFS Medicare spending data.

Let $\lambda = (\lambda_1, \lambda_2)$ and define for each county

$$\begin{aligned} \chi_{1FFS} &= \lambda_1 \bar{\chi}_{1FFS} \\ \chi_{2FFS} &= \lambda_2 \bar{\chi}_{2FFS}. \end{aligned}$$

Denote the observed per beneficiary Medicare Parts A and B payments as M_A^o and M_B^o, respectively. In the model, the theoretical analogs for these values are

$$\begin{aligned} M_A(\lambda) &= \sum_g w_g^{FFS} \int \cdots \int \left(P\left(\xi_i, y_i; \lambda\right) E_\epsilon\left[M_A\left(\theta_i, y_i, \lambda\right) \mid medigap\right] + \right. \\ &\quad \left. (1 - P\left(\xi_i, y_i; \lambda\right)) E_\epsilon\left[M_A\left(\theta_i, y_i; \lambda\right) \mid FFS\ only\right]\right) dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i) \\ M_B(\lambda) &= \sum_g w_g^{FFS} \int \cdots \int \left(P\left(\xi_i, y_i; \lambda\right) E_\epsilon\left[M_B\left(\theta_i, y_i, \lambda\right) \mid medigap\right] + \right. \\ &\quad \left. (1 - P\left(\xi_i, y_i; \lambda\right)) E_\epsilon\left[M_B\left(\theta_i, y_i; \lambda\right) \mid FFS\ only\right]\right) dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i) \end{aligned} \tag{11}$$

where $E_\epsilon\left[M_A\left(\theta_i, y_i; \lambda\right) \mid medigap\right]$ denotes the expected Part A payment for a beneficiary with the Medigap supplement and characteristics (θ_i, y_i) for a given value of λ and expectations are taken over the distribution of the health shock ϵ. The corresponding expression for the expected Part B payment is $E_\epsilon\left[M_B\left(\theta_i, y_i; \lambda\right) \mid medigap\right]$. Similarly, $E_\epsilon\left[M_A\left(\theta_i, y_i; \lambda\right) \mid FFS\ only\right]$ and $E_\epsilon\left[M_B\left(\theta_i, y_i; \lambda\right) \mid FFS\ only\right]$ denote the expected Parts A and B payments for a beneficiary without the Medigap supplement and characteristics (θ_i, y_i) for a given value of λ and expectations are taken over the distribution of the health shock ϵ. Since I do not observe market shares for Medigap at the county level, I use the theoretical probability that a beneficiary with endowment (ξ_i, y_i) chooses the Medigap supplement for a given value of λ,

$$P\left(\xi_i, y_i; \lambda\right) = \frac{\exp\left\{V_{medigap}\left(\xi_i, y_i; \lambda\right)\right\}}{\exp\left\{V_{medigap}\left(\xi_i, y_i; \lambda\right)\right\} + \exp\left\{V_{FFSonly}\left(\xi_i, y_i; \lambda\right)\right\}},$$

to weight the Medigap and FFS-only contributions to the two expected Medicare payments. These values are then integrated over (ξ_i, y_i) conditional on being in demographic group g. Finally, these demographic group contributions are weighted by w_g^{FFS}, the share of Medicare population in that county that is not enrolled in Medicare Advantage and in demographic group $\dot{g}$. These weights $\left\{w_g^{FFS}\right\}_{g=1,10}$ are observed in the data because the total Medicare Advantage enrollment and the total number of eligibles in each demographic group g is observed in each county. Hence,

$$w_g^{FFS} = \frac{\text{Eligible}_g - \text{Medicare Advantage Enrolled}_g}{\sum_{k=1}^{10} \text{Eligible}_k - \text{Medicare Advantage Enrolled}_k}.$$

The first step in the estimation algorithm is to find the vector λ^* that equates the observed and predicted Medicare payments,

$$\begin{aligned} M_A^o &= M_A(\lambda^*) \\ M_B^o &= M_B(\lambda^*). \end{aligned}$$

This search is carried out for each of the 839 counties with Medicare Advantage activity as well as 159 other counties in which there are no Medicare Advantage participants but are represented by individuals in the MCBS. While I suppress the notation, it should be noted that the vector λ^* is an implicit function of the model parameters. Finally, the search for the vector λ^* over the 998 counties is subject to the normalization

$$\frac{1}{998}\sum_{c=1}^{998} \lambda_{1c}^* = \frac{1}{998}\sum_{c=1}^{998} \lambda_{2c}^* = 1$$

where c indexes counties. This normalization is necessary because the means of λ_1^* and λ_2^* are not separately identified from the scale of the vectors $\bar{\chi}_{1FFS}$ and $\bar{\chi}_{2FFS}$.

5.2.2 Medicare Advantage

Given the vector λ^*, the next step in the estimation algorithm is to solve for the vector (χ_1, χ_2) for each Medicare Advantage MCO in each county using the data on demographic group enrollments. Let J denotes the number of MCOs in the county and index these by j. Let χ_j denote $\left(\chi_{1j}, \chi_{2j}\right)$ and the10J vector χ denote $(\chi_1, ..., \chi_J)$. For each MCO in a county, there are ten elements in χ_j and ten observed age/gender group market shares. Define the theoretical market share for MCO j in demographic group g as

$$P_{gj}(\chi, \lambda^*) = \int \cdots \int \left[\sum_{k=1}^{Kj} P_{gjk}(\xi_i, y_i; \chi, \lambda^*) \right] dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i) \tag{12}$$

where

$$P_{gjk}(\xi_i, y_i; \chi, \lambda^*) = \frac{\exp\left\{ \frac{V_{jk}(\xi_i, y_i; \chi_j)}{(1-\sigma_a)(1-\sigma_w)} \right\}}{\mathcal{I}_j(\chi_j)^{\sigma_w} \mathcal{J}(\chi)^{\sigma_a} \left[\exp\{V_{medigap}(\xi_i, y_i; \lambda^*)\} + \exp\{V_{FFSonly}(\xi_i, y_i; \lambda^*)\} + \mathcal{J}(\chi)^{1-\sigma_a} \right]}. \tag{13}$$

This definition is the same as (10) except the dependence on χ and λ^* has been added. Let $P(\chi, \lambda^*)$ denote the $10J$ vector of theoretical market shares with typical element $P_{gj}(\chi, \lambda^*)$. Similarly, let P^o denote the $10J$ vector of observed market shares. The second step in the estimation algorithm is to search for the $10J$ vector χ^* such that

$$P^o = P(\chi^*, \lambda^*). \tag{14}$$

This search is carried out for each of the 839 counties with some Medicare Advantage activity. This results in a 1765×10 array of vectors (χ_1, χ_2), each row corresponding to a unique MCO/county combination. As with λ^*, the vector χ^* is an implicit function of the model parameters.

Moment Conditions Following BLP, I construct moments conditions using the vector of MCO specific unobservables. I assume the vector χ_j^* is mean independent of a set of instruments denoted $\dot{Z}$

$$E\left[\chi_j^* | Z\right] = 0. \tag{15a}$$

As in BLP, the assumption here is that the vector χ_j^* is known to the MCO when it determines the premium. Hence, the premium is endogenous and instrumental variables methods are required. Plausible instruments must be exogenous and excluded from the market shares equations. I use data from the Area Resource File (ARF) 2001 as instruments. The ARF provides county level data on health care provider counts, health care facilities, the number of health maintenance organizations (HMOs) in the commercial sector, and the market share of the commercial HMOs. I use data on physicians per 1000 Medicare beneficiaries, dentists per 1000, hospital beds per 1000, registered nurses per 1000, licensed practical nurses per 1000, ambulatory surgical centers per 1000, and HMO market penetration in the commercial sector. The identifying assumption is that each Medicare MCO cannot affect these broader market outcomes through its premium decision. Hence, the ARF data is plausibly exogenous. It also seems reasonable that the ARF data should not directly affect Medicare beneficiary health insurance choice, but only indirectly through premiums and prices for health care services.

Table 5 provides descriptive statistics of the ARF data. Table 6 provides results of a least squares regression of Medicare MCO premiums on the instruments. While the R^2 is relatively low at 0.038, the instruments have a strong statistical association with MCO premiums.

With 10 elements in the vector χ_j^* and 7 instruments, there are 70 moments conditions that are used the estimate the model parameters. Let $\overrightarrow{\chi^*}$ denote the 1765×10 array with typical row χ_j^* and $\vec{Z}$ denote the 1765×7 array of instruments. Finally, define the 70×1 vector of moments as

$$\Lambda(\Theta) = \frac{vec\left(\left[\overrightarrow{\chi^*}\right]'\left[\overrightarrow{Z}\right]\right)}{1765},$$

where, as in section 4, Θ denotes the vector of model parameters.

5.2.3 Simulation Methods

Numerically solving for the vector (λ^*, χ^*) for each county requires simulating the integrals in (11) and (12). Simulating these integrals also involves numerically solving for optimal health care utilization for simulated pretreatment health states θ_i and incomes y_i over all health insurance plans.

To replicate the age component of X_i given age/gender group g, I use the observed age distribution conditional on age/gender group in the MCBS. To replicate y_i, I use data from the 2000 Census Summary File 3 on county level median household income by age group.[16]

[16] The age groups are not as fine in the Census data as in the Medicare Managed Care enrollment data.

I assume that $\ln y_i$ is normally distributed with the mean given by the median household income from the Census data and a variance that it estimated using the MCBS. I assume this variance to be constant across all age/gender groups and counties. I apply antithetic acceleration to all replications. Since each function simulated is monotone in the simulated terms, antithetic acceleration will reduce the variance associated with simulation error.[17] For the research, I use $R_\eta = 8$ and $R_\epsilon = 4$.

5.3 Individual Data

The third step in the estimation algorithm brings the values $(\lambda^*,\ \chi^*)$ for each county to the MCBS data. Since I observe the county of residence for each individual in the MCBS, I can connect the information in the county level FFS Medicare expenditures and aggregate managed care enrollment to the model of individual health insurance choice and health care utilization..

The likelihood contribution of an individual is the probability of choosing the observed health insurance plan multiplied by the joint density of ϵ_i that is consistent with the observed utilization data given the health plan choice, integrated over the joint density of η_i. The likelihood contribution of an individual with the configuration of interior and corner utilization solutions from section 4.1.4 and having selected health plan jk is

$$\mathcal{L}_i\left(\Theta\right) = \int\cdots\int\left[P_{ijk}(\eta_i)\iiint_{\epsilon_{i_C}\in A_{jk}} f_{\epsilon_I}(\epsilon_{iDV}(\epsilon_{i_C}),\epsilon_{iPD}(\epsilon_{i_C})|\epsilon_{i_C})\,\|\mathbb{J}_{ijk}(\epsilon_{i_C})\|\,dF_{\epsilon_C}(\epsilon_{i_C})\right]dF_\eta(\eta_i). \tag{16}$$

A component of this likelihood contribution is the vector $(\lambda^*,\ \chi^*)$ that corresponds to the beneficiary's county of residence. This includes all the vectors χ_j^* over all J MCOs participating in Medicare Advantage in the county. Recall that the pair $(\epsilon_{iDV}(\epsilon_{i_C}),\epsilon_{iPD}(\epsilon_{i_C}))$ is a function of the model parameters through the first order conditions. This includes the vector (χ_1,χ_2) that appears in the health production function. If the beneficiary chose Medigap or

The age groups are 65-74 and 75 and over.

[17] See Stern (1997) for a discussion of antithetic acceleration and other simulation methods.

FFS-only, then the vector (χ_1, χ_2) in the joint density of ϵ_i is

$$\begin{aligned}\chi_1 &= \lambda_1^* \bar{\chi}_{1FFS} \\ \chi_2 &= \lambda_2^* \bar{\chi}_{2FFS}.\end{aligned}$$

If the beneficiary chose a Medicare Advantage, then the vector (χ_1, χ_2) in the joint density of ϵ_i is

$$\begin{aligned}\chi_1 &= \chi_{1j}^* \\ \chi_2 &= \chi_{2j}^*,\end{aligned}$$

where j corresponds to the selected MCO. All elements of the vector (λ^*, χ^*) appear in the health plan choice probability $P_{ijk}(\eta_i)$ with λ^* used to evaluate the Medigap and FFS-only value functions and each vector χ_j^* used to evaluate the value function for each plan offered by MCO j.

The stochastic dependence between the health insurance choice and health care utilization decisions is captured in the vector η_i. This is because the beneficiary unobservable in the health insurance choice probability is η_i (ϵ_i and e_i are integrated out), while in the health care utilization model, the beneficiary unobservable is $\eta_i + \epsilon_i$ (e_i does not play a role in the utilization decision). Assuming that η_i and ϵ_i are drawn from independent distributions, we have

$$COV(\eta_i, \eta_i + \epsilon_i) = VAR(\eta_i).$$

As in the previous studies that model health plan choice and health care utilization jointly, asymmetric information in this market and the endogeneity of the health plan characteristics in the utilization model are important to the extent that the distribution of η_i is nondegenerate.

5.4 Objective Function

To develop an objective function for the model, I follow Imbens and Lancaster (1994) supplement the likelihood function with the vector of moment conditions derived from the aggregate enrollment data as follows,

$$\Gamma(\Theta) = \left[\begin{array}{cc} N^{-1} \sum_i \frac{\partial \ln \mathcal{L}_i(\Theta)}{\partial \Theta'} & \Lambda(\Theta)' \end{array} \right]' W \left[\begin{array}{cc} N^{-1} \sum_i \frac{\partial \ln \mathcal{L}_i(\Theta)}{\partial \Theta'} & \Lambda(\Theta)' \end{array} \right], \tag{17}$$

where W is a positive definite weighting matrix. Following Hansen (1982), the optimal weighting matrix is

$$W = E\left[\left[\begin{array}{cc} N^{-1} \sum_i \frac{\partial \ln \mathcal{L}_i(\Theta)}{\partial \Theta'} & \Lambda(\Theta)' \end{array} \right] \left[\begin{array}{cc} N^{-1} \sum_i \frac{\partial \ln \mathcal{L}_i(\Theta)}{\partial \Theta'} & \Lambda(\Theta)' \end{array} \right]' \right]^{-1}.$$

In the model, there are 101 parameters and 171 moments. I solve the model in two rounds using a block diagonal weighting matrix where the upper left block is an identity matrix and the lower right block is the inverse of the expected inner product of the instruments in the first round to obtain a consistent estimate $\hat{\Theta}$ and the sample analog of the optimal weighting matrix evaluated at $\hat{\Theta}$ in the second.

The estimation algorithm follows this broad outline:

1) Condition on an initial guess of Θ.

2) Solve for the vector λ^* in each of the 998 counties as discussed in section 5.2.1.

3) Given the values of λ^*, solve for the vector χ^* in each of the 839 counties with at least one Medicare Advantage plan as discussed in section 5.2.2.

4) Given the values of λ^* and χ^*, evaluate the likelihood contribution for each of the 5251 individuals in the MCBS.

5) Evaluate $\Gamma(\Theta)$ and $\frac{\partial \Gamma(\Theta)}{\partial \Theta}$ and update the guess of the model parameters.

6) Repeat 2) through 5) until $\left\| \frac{\partial \Gamma(\Theta)}{\partial \Theta} \right\|$ is within a tolerance of zero.

5.5 Identification

As noted earlier, health insurance characteristics in a model of health care demand are endogenous if the unobservables in the health plan choice and health care demand equations are correlated. A model of joint determination captures the stochastic dependence but still leaves a question of identifying the effect of a covariate on the two components of the model. For example, the effect of an exogenous variable such as age or income on health plan choice may differ from its effect on health care demand. An instrumental variables approach could use exclusion restrictions to derive instruments, but often plausible exclusion restrictions may not exist. In this case, it would be difficult to assume, for example, that age or gender affects health care demand but does not affect health plan choice. Another method requiring stronger distributional assumptions is the error correction method as described in Heckman (1979). Cameron et al (1988) employ a method less dependent on distributional assumptions by estimating a plan choice probability and then using this probability as an instrument in the health care demand equation.

The unified framework in this model in which both the health insurance and health care utilization decisions are based on the same preferences implies parameter restrictions that permit the identification these two marginal effects in the absence of exclusion restrictions. This is because a covariate, say age, affects both the utilization and health insurance choice decisions only through the pretreatment health state vector θ_i and the parameter vector giving the marginal effect of age on θ_i, β_{age}, is the same in both decisions. For example, the marginal effect of the covariate age on health care demand given health plan choice is given by the vector

$$\begin{aligned}
\frac{\partial m^*_{ijk}}{\partial X_{i,age}} &= -\int \cdots \int \left[U^*_{m_{ijk}m_{ijk}}\right]^{-1} \left[U^*_{m_{ijk}\theta_i}\frac{\partial \theta_i}{\partial X_{i,age}}\right] dF_\eta(\eta_i) \\
&= -\int \cdots \int \left[U^*_{m_{ijk}m_{ijk}}\right]^{-1} \left[U^*_{m_{ijk}\theta_i}\beta_{age}\right] dF_\eta(\eta_i) \\
&= \int \cdots \int \mathbb{J}_{ijk} dF_\eta(\eta_i)\beta_{age}
\end{aligned}$$

where $\mathbb{J}_{ijk}$ is the Jacobian of the transformation of m^*_{ijk} to ϵ_i, and the marginal effect of age on the probability of choosing health plan jk is

$$\begin{aligned}
\frac{\partial \int \cdots \int P_{ijk}(\eta_i) dF_\eta(\eta_i)}{\partial X_{i,age}} &= \int \cdots \int \frac{\partial P_{ijk}(\eta_i)}{\partial V_{ijk}(\eta_i)} \sum_{t\in T} \frac{\partial V_{ijk}(\eta_i)}{\partial \theta_{it}} \frac{\partial \theta_{it}}{\partial X_{i,age}} dF_\eta(\eta_i) \\
&= \int \cdots \int \frac{\partial P_{ijk}(\eta_i)}{\partial V_{ijk}(\eta_i)} \sum_{t\in T} \frac{\partial V_{ijk}(\eta_i)}{\partial \theta_{it}} dF_\eta(\eta_i)\beta_{age,t}.
\end{aligned}$$

The parameter vector of interest β_{age} is same in both of these marginal effects, so parameter restrictions, as opposed to exclusion restrictions, are used to identify the parameters.

5.6 Computation

Estimating the model parameters is computationally intensive. Each iteration of the model parameters requires 3-6 hours of CPU time, depending on the norm of the step in the parameter space. The computer code used to estimate the model was written in Fortran 90 and estimation was carried out on the Cedar and Dogwood Linux Clusters at the University of Virginia. Because of the time required for each evaluation, parallel processing was used to reduce the time required to estimate the model. Because the shape of the objective function in the parameters space is difficult to determine, a grid search was carried out before Newton based updating was used. The value of $\left\| \frac{\partial \Gamma(\Theta)}{\partial \Theta} \right\|$ at convergence is 0.000619.

6 Results

6.1 Estimates of Structural Parameters

The estimates of the structural parameters are presented in tables 7-9. Generally, the parameters are precisely estimated. Of the 101 free parameters in the model, 80 of the estimates are significant at the 1% level, 2 are significant at the 5% level, 19 are not significant at conventional levels. Throughout this discussion, I refer to the standard errors of the parameter estimates, which are reported in parentheses.

6.1.1 Preferences and Endowments: δ, γ_1, γ_2, κ_1, κ_2, κ_{IP}^F, β

The estimates of the parameters describing preferences and the health endowment are given in Table 7. The estimates of the two risk aversion parameters γ_1, on consumption C_i, and γ_2, on health H_i, are similar, 0.6333 (0.0012) and 0.6460 (0.0053), respectively. The estimates of the continuous nonpecuniary cost parameters κ_1 and κ_2 indicate that the marginal nonpecuniary cost is increasing in all five types of health care. Doctor visits and outpatient care have the lowest nonpecuniary cost followed by inpatient care and dental care. The estimate of the fixed nonpecuniary cost associated with inpatient care is 4.957 (0.3635). For comparison, the typical expected indirect utility lies in the $450 - 500$ range, so the fixed nonpecuniary cost is roughly 1% of a typical expected indirect utility.

The estimates of the parameters β that capture the marginal effect of the observed characteristics X_i on the pretreatment health state θ_i, suggest a slightly concave relationship between the pretreatment health state and age. (Recall that larger values of θ_i indicate worse pretreatment health states.) Both the first and second order effects are significant in all dimensions of health except inpatient care. The coefficient on female is significant in dental care and doctor visits but not in inpatient care, outpatient care, or prescription drugs.

6.1.2 Production and CES Functions: ζ, a, b, Ω_ν, $\bar{\chi}_{1FFS}$, $\bar{\chi}_{2FFS}$, α, ρ

The estimates of the parameters of the production function in (2) and CES function in (3) are given in Table 8. The weights in the CES function α indicate that the least weight is placed on the outpatient care dimension, 0.1513 (0.0021). It is interesting that the greatest weights are placed on the doctor visit, 0.2542 (0.0013), and prescription drug, 0.2535 (0.0014), dimensions and not the inpatient care dimension, 0.1793 (0.0014). The estimate in the dental care dimension is 0.1617. The estimate of ρ is 0.4111 (0.0013).

The parameters ζ, which capture diminishing marginal product on the extensive margin, are all significant at 1% and are much larger for inpatient, outpatient, doctor visits, and prescription dugs than for dental care. The parameters that capture uncertainty about the efficacy of treatment Ω_ν are also all significant at 1% for those elements not on the boundary of the parameter space. The variance is relatively small for doctor visits 0.0524 (0.0078) and is insignificantly different from zero for dental care. As may be expected, the dimension with the largest variance is inpatient care, 0.3692 (0.0032). The estimates of the parameters that denote the mean of the FFS Medicare productivity vector $\bar{\chi}_{2FFS}$ are highly significant in all types of care , but the estimates of $\bar{\chi}_{1FFS}$ are not are not significant in any type of care.

6.1.3 Parametric Distributions: μ, Ω_ϵ, Ω_η, σ_a, σ_w

The estimates of the parameters of the parametric distributions are given in Table 9. As noted in section 5, the importance of asymmetric information in the health insurance choice is captured in the variance of the private information of the beneficiary about risk type

Ω_η. Many studies that have examined the joint determination of health insurance choice and health care utilization, including Dowd et al. (1991), Cardon and Hendel (2001), and Hubbard-Rennhoff (2005), have not found statistically significant estimates of the variance of the private information. The results here are somewhat mixed. Of the fifteen elements in the matrix Ω_η, six are significant at the 1%, including all five of the diagonal elements, and the other nine are insignificant at conventional levels.

In contrast, all of the elements of the covariance matrix of the health shocks Ω_ϵ are significant at the 1% level. In comparing the diagonal elements, the largest variance is in the inpatient care dimension, followed by doctor visits and prescription drugs. Dental care has by far the smallest variance.

The estimate of σ_a, the correlation of the GEV errors across MCOs within a county is 0.1743 (0.0832), while the estimate of σ_w, the correlation of the GEV errors across plans within an MCO is much higher: 0.8208 (0.1563). These results are intuitive in that many of the components of the GEV errors discussed in section 4.2.1, such as a convenient location of a clinic or the presence of a familiar provider, should be very similar across different plans offered by the same MCO.

6.1.4 Managed Care Unobserved Product Characteristics: χ_1^*, χ_2^*

Recall that the elements of χ_2^* shift the satiation point in the health production function and χ_1^* determines the height of the production function at the satiation point for each type of health care. Higher values of each χ_1^* and χ_2^* indicate more productive health care. The role of these unobservables in the model is to differentiate Medicare MCOs from FFS Medicare in terms of the quality of services provided or access to care. The expectation is that, on average, Medicare MCOs will have lower values of (χ_1^*, χ_2^*) relative to $(\bar{\chi}_{1FFS}, \bar{\chi}_{2FFS})$. This is because while the Medicare Advantage plans offer extra benefits and are, on average, much less expensive than Medigap supplements, the overall Medicare Advantage market share is quite low. The differences between the managed care and FFS unobserved product characteristics explain these differences in market shares.

Table 10 gives the descriptive statistics of (χ_1^*, χ_2^*) within types of health care. For convenience, I have added the estimates of $(\bar{\chi}_{1FFSt}, \bar{\chi}_{2FFSt})$ below in table 11. In almost all cases, the means of the managed care production unobservables are less than the estimates of the corresponding elements of $(\bar{\chi}_{1FFS}, \bar{\chi}_{2FFS})$. The two exceptions are that the means of χ^*_{1DC} and χ^*_{1OP} are greater than the estimates of $\bar{\chi}_{1FFS,DC}$ and $\bar{\chi}_{1FFS,OP}$. To give a better sense of the distribution of (χ^*_1, χ^*_2), I plot the joint distribution of $(\chi^*_{1t}, \chi^*_{2t})$ for each type of health care in figures 6 through 10. I use crude histograms that divide the pairs $(\chi^*_{1t}, \chi^*_{2t})$ into two-way bins. The bin marked with an "X" contains the pair $(\bar{\chi}_{1FFSt}, \bar{\chi}_{2FFSt})$.

Two patterns emerge from the figures. First, the relationship between $(\chi^*_{1t}, \chi^*_{2t})$ in each type of health care is remarkably linear, with correlations close to -1. Second, the managed care pairs $(\chi^*_{1t}, \chi^*_{2t})$ generally fall on a line lying southwest of the bin containing the pair $(\bar{\chi}_{1FFSt}, \bar{\chi}_{2FFSt})$. This fits the intuition described above. The Medicare MCOs must have some attributes unattractive to beneficiaries in order to explain their low market shares given their premiums and benefits packages.

6.2 Specification Testing

To test the model, I apply standard nonparametric goodness-of-fit tests to the data on health insurance choice and health care utilization given the observed health insurance choice. The tests involve comparing the distributions given by the data to the distributions predicted by the model.

6.2.1 Health Care Utilization

Throughout estimation, I approximate health insurance choice probabilities using simulation methods. The value functions in these choice probabilities are approximated by replicating many counterfactual health states, solving for optimal health care utilization for each of these replications, and then taking the mean of the indirect utilities. I use the health care utilization outcomes pertaining only to the observed health insurance choice to construct a distribution of predicted health care utilization conditional on the observed health insurance choice.

I compare this predicted distribution to the empirical distribution in the MCBS using

Pearson's Chi-Squared goodness-of-fit test. The test consists of breaking the distributions into K bins and comparing the number of individuals observed to be in each bin to the number predicted to be in each bin given the sample size. Let N denote the total number of individuals in the MCBS sample, N_k the total number of individuals in the sample whose utilization is observed to be in the k^{th} bin, and P_k the share of the predicted utilization data in the k^{th} bin. The test statistic is

$$\sum_{k=1}^{K} \frac{(N_k - NP_k)^2}{NP_k}$$

which is distributed Chi-Squared with $K - 1$ degrees of freedom. For each type of health care, the first bin consists of corner solutions only, and the remaining bins are defined by dividing the subsample of interior solutions so that there are roughly an equal number of individuals in each bin.

The results are given in figures 11 through 15. Each figure gives the CDF of observed and predicted health care utilization for one type of care. The inset box gives the mean, standard deviation, maximum, the percentage of corner solutions, and the result of the Pearson's test. For each type of health care utilization, the null that the observed data is drawn from the predicted distribution is rejected at less than 1%. Nonetheless, the model is reasonably effective in predicting some of the basic descriptive statistics of the data. The difference between the predicted and observed means is around 1% of the observed mean for inpatient care and doctor visits, 2% for prescription drugs, and 9% for outpatient care. Only in dental care is there a very large difference (61% of the observed mean). The model predictions for the percentage of corner solutions is good for inpatient care (a difference of 4.1% between the predicted and observed rates), outpatient care (4.1%), and dental care (0.2%). The model does less well in predicting corner solutions in prescription drugs (6.5%) and doctor visits (11.0%).

6.2.2 Health Insurance Choice

I conduct a similar test on the health insurance choice model. I construct the test based on the probability that an individual chooses FFS Medicare with the Medigap supplement.

The test is based on dividing the sample into K bins and comparing the number predicted to choose Medigap to the number observed to choose Medigap in each bin.

Let N_k denote the total number of individual in the k^{th} bin, M_k denote the number of individuals in the k^{th} bin who chose Medigap, and P_{ki} the predicted probability that individual i in the k^{th} chooses Medigap. The test statistic is

$$\sum_{k=1}^{K} \frac{\left(M_k - \sum_{i=1}^{N_k} P_{ki}\right)^2}{\sum_{i=1}^{N_k} P_{ki}}$$

which is also distributed Chi-Squared with $K-1$ degrees of freedom. Here, I use the predicted Medigap probabilities to define the bins in constructing the test statistics. For each beneficiary in the MCBS sample, I evaluate the corresponding Medigap probability. I then order these probabilities and divide them into 10 bins with approximately 525 beneficiaries per bin.

The results are given in figure 16. The graph gives the distribution of the predicted Medigap choice probabilities and the inset box provides the means and standard deviations of the predicted Medigap choice probabilities and the observed, binary outcome. The distribution of predicted probabilities is bimodal with peaks in the 0.00-0.05 and 0.80-0.85 bins. The model somewhat under-predicts Medigap enrollment; the mean predicted probability is 0.513 and the observed Medigap share is 0.593. The model does produce the intuitive result that the mean predicted Medigap probability conditional on actually having chosen Medigap is much higher than that conditional on not choosing Medigap: 0.709 versus 0.213. Again, the null hypothesis that the observed Medigap choices are being drawn from the distribution of predicted Medigap choice probabilities is rejected at less than 1% with a Pearson's Chi-Squared test statistic of 1885.70.

6.3 Policy Experiment

The goal in this research is to estimate the change in consumer surplus brought about by Medicare Advantage within the 10 age/gender groups for which Medicare Advantage enrollment is observed. Theory suggests that older age groups, who on average consume more health care and hence would benefit the most from pooling with younger beneficiaries, would have the least favorable distribution. However, as discussed in section 3, Medicare Advantage plans are generally less expensive than Medigap supplements and often include extra benefits. Since older age groups have lower income and are higher risk on average, they may benefit more than younger groups.

The key element in answering this question is finding the Medigap premium that would obtain under the counterfactual of the elimination of the Medicare Advantage program. If Medicare Advantage does systematically enroll lower risk types, then the Medigap premium should be lower in the absence of Medicare Advantage.

In the absence of modeling the objectives and behavior of Medigap insurers, an assumption must be made in order to assess the counterfactual Medigap premium. The assumption I make in this experiment is that the Medigap loss ratio, the ratio of expected pay-outs to the premium, would the same under the counterfactual as under the status quo. First, I predict expected Medigap payments in the model for each county under the status quo. I then search for the counterfactual premium that produces the same loss ratio and is consistent with how Medicare beneficiaries would sort between FFS Medicare only and with the Medigap supplement in a world with no Medicare Advantage options.

The assumption of a fixed loss ratio warrants some discussion. There are three main reasons why expected payouts would be less than the premium: administrative costs, market power, and a fee for risk bearing. If all of the difference was due to administrative costs, and administrative costs were roughly proportional, then the assumption of a fixed loss ratio would be reasonable. However, if there is a significant fixed administrative cost, then the loss ratio could be higher under the counterfactual because Medigap would have a larger insurance pool. In the absence of insurer specific cost data, it is difficult to assess how the assumption of a fixed loss ratio will affect the experiment.

If the Medigap market is close to perfectly competitive, then market power would not be an issue and the fixed loss ratio assumption would be reasonable. However, if Medigap insurers have some market power that is diminished by Medicare Advantage plans, then the counterfactual loss ratio would be lower than under the status quo since the premium could rise because of reduced competition. This in turn could result in a lower expected payout due to an income effect, further reducing the loss ratio. Assessing the level of competition within the Medigap market is difficult for the reasons discussed in the section 3. Generally, Medigap premiums and market shares are unobserved.

Finally, signing the impact of the counterfactual on the fee for risk bearing is difficult because it depends on the higher moments of the Medigap payout distribution. Generally, the fee for risk bearing should fall under the counterfactual because the variance of the average payout declines with enrollment. However, in a model with heterogenous risk types, the effect on the variance is difficult to predict. For example, if Medigap enrollment expanded and the new enrollees were on average lower risk than under the status quo, then, while the average payout would fall, the variance could rise.

These issues suggest that a more compelling assessment of this policy question requires a model of insurers as well as consumers. As already noted, data limitations preclude a complete model that captures insurers costs and strategic behavior. I thus proceed with the experiment using the fixed loss ratio assumption as an approximation to the potentially offsetting effects just described.

Let π_m^o denote the observed annual Medigap premium for a given county and $M_{medigap}(\pi_m^o)$ denote the expected Medigap payout given that premium. The model prediction of the expected payout is

$$M_{medigap}(\pi_m^o) = \sum_g w_g^{FFS} \frac{\int \cdots \int P\left(\xi_i, y_i; \pi_m^o, \lambda^*\right) E_\epsilon \left[M_{medigap}\left(\theta_i, y_i; \pi_m^o, \lambda^*\right)\right] dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i)}{\int \cdots \int P\left(\xi_i, y_i; \pi_m^o, \lambda^*\right) dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i)} \tag{18}$$

where $M_{medigap}\left(\theta_i, y_i; \pi_m^o, \lambda^*\right)$ denotes the Medigap payout for a beneficiary with pretreatment health state and income (θ_i, y_i) given the premium π_m^o and the value of λ^* that was found for that county in the estimation algorithm. As in section 6, $P\left(\xi_i, y_i; \pi_m^o, \lambda^*\right)$ denotes the probability that a beneficiary with endowment (ξ_i, y_i) chooses Medigap over FFS-only and w_g^{FFS} denotes the share of Medicare population that is not enrolled in Medicare Advantage in age/gender group g

$$w_g^{FFS} = \frac{\text{Eligible}_g - \text{Medicare Advantage Enrolled}_g}{\sum_{k=1}^{10} \text{Eligible}_k - \text{Medicare Advantage Enrolled}_k}.$$

Let $\tilde{\pi}_m$ denote the Medigap premium under the counterfactual of the elimination of the Medicare Advantage program. Under the counterfactual, the expected Medigap payout is

$$M_{medigap}(\tilde{\pi}_m) = \sum_g w_g^{pop} \frac{\int \cdots \int P\left(\xi_i, y_i; \tilde{\pi}_m, \lambda^*\right) E_\epsilon \left[M_{medigap}\left(\theta_i, y_i; \tilde{\pi}_m, \lambda^*\right)\right] dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i)}{\int \cdots \int P\left(\xi_i, y_i; \tilde{\pi}_m, \lambda^*\right) dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i)}. \tag{19}$$

The critical distinction between (18) and (19) is that the age/gender groups are weighted by the total number of eligibles, not the number not enrolled in Medicare Advantage. That is

$$w_g^{pop} = \frac{\text{Eligible}_g}{\sum_{k=1}^{10} \text{Eligible}_k}.$$

Under the fixed loss ratio assumption, the policy experiment involves solving for the value of $\tilde{\pi}_m$ such that

$$\frac{M_{medigap}(\tilde{\pi}_m)}{\tilde{\pi}_m} = \frac{M_{medigap}(\pi_m^o)}{\pi_m^o}.$$

The search involves simulating the integrals in (18) and (19) as discussed in section 6. For each age/gender group g, I replicate R_η draws of (ξ_i, y_i) from F_η and $F_{y,X|g}$. For each replication of (ξ_i, y_i), I replicate R_ϵ draws of ϵ_i from F_ϵ. For each replication of ϵ_i, I construct the vector (θ_i, y_i) and numerically solve for the optimal amount of health care utilization for the Medigap and FFS-only options as well as the Medigap payout given the optimal amount health care utilization under the Medigap option. Taking an average of the Medigap payouts gives an approximation of $E_\epsilon \left[M_{medigap}\left(\theta_i, y_i; \pi_m^o, \lambda^*\right)\right]$. Taking averages of the indirect utilities gives approximations of $V_{medigap}\left(\xi_i, y_i; \pi_m^o, \lambda^*\right)$ and $V_{FFSonly}\left(\xi_i, y_i; \lambda^*\right)$ for each replication of (ξ_i, y_i). Using the approximations of the value functions yields the value of the Medigap choice probability

$$P\left(\xi_i, y_i; \pi_m^o, \lambda\right) = \frac{\exp\left\{V_{medigap}\left(\xi_i, y_i; \pi_m^o, \lambda^*\right)\right\}}{\exp\left\{V_{medigap}\left(\xi_i, y_i; \pi_m^o, \lambda^*\right)\right\} + \exp\left\{V_{FFSonly}\left(\xi_i, y_i; \lambda^*\right)\right\}}.$$

Taking averages over the replications of (ξ_i, y_i) and taking the weighted average over $\left\{w_g^{FFS}\right\}_{g=1,10}$ provides the prediction of the status quo loss ratio

$$\frac{M_{medigap}(\pi_m^o)}{\pi_m^o}.$$

I solve for $\tilde{\pi}_m$ by simulating the integrals in (19) applying the same replications used in the approximation of (18), repeating the process and updating guesses of $\tilde{\pi}_m$ until $\left|\frac{M_{medigap}(\tilde{\pi}_m)}{\tilde{\pi}_m} - \frac{M_{medigap}(\pi^o_m)}{\pi^o_m}\right|$ is within a tolerance of zero.

Results Figure 17 gives the distribution of the change in the annual Medigap premium under the counterfactual of the elimination of the Medicare Advantage program $\tilde{\pi}_m - \pi^o_m$. The histogram is weighted to account for the Medicare population in each county. The inset box gives descriptive statistics. The results are consistent with the hypothesis that Medigap premiums would fall under the counterfactual, although the magnitude of the effect is small. On average, the annual Medigap would fall by \$4.38 if the Medicare Advantage program was eliminated. This is less than 1% of the average annual Medigap premium. In many counties, the Medigap premium is predicted to increase under the counterfactual, reflecting the fact that while Medicare Advantage may enroll lower risk types on average, this is not the case in all counties. The largest decrease is \$87.99 and the largest increase is \$59.28.

With the value of $\tilde{\pi}_m$ for each county, I evaluate the change in consumer surplus resulting from Medicare Advantage for each age/gender group in each county. Following Small and Rosen (1981), the status quo consumer surplus in age/gender group g is

$$\begin{aligned} CS_g(\chi^*, \lambda^*, \pi^o_m) &= \int \cdots \int \frac{1}{U_{y_i}} \ln \left[\left(\sum_{j=1}^{J} \left[\sum_{k=1}^{K_j} \exp\left\{ \frac{V_{jk}\left(\xi_i, y_i; \chi^*_j\right)}{(1-\sigma_a)(1-\sigma_w)} \right\} \right]^{1-\sigma_w} \right)^{1-\sigma_a} + \right. \\ & \left. \exp\left\{V_{medigap}\left(\xi_i, y_i; \pi^o_m, \lambda^*\right)\right\} + \exp\left\{V_{FFSonly}\left(\xi_i, y_i; \lambda^*\right)\right\} \right] dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i). \end{aligned}$$

Note that I define consumer surplus at the ex ante stage, i.e., when the health insurance choice is made but before the value of the health shock ϵ_i is realized. Given the value of $\tilde{\pi}_m$, the counterfactual consumer surplus is

$$\begin{aligned} CS_g(\lambda^*, \tilde{\pi}_m) &= \int \cdots \int \frac{1}{U_{y_i}} \ln \left[\exp\left\{V_{medigap}\left(\xi_i, y_i; \tilde{\pi}_m, \lambda^*\right)\right\} + \right. \\ & \left. \exp\left\{V_{FFSonly}\left(\xi_i, y_i; \lambda^*\right)\right\} \right] dF_\eta(\eta_i) dF_{X,y|g}(X_i, y_i). \end{aligned}$$

In assessing the welfare effect of Medicare Advantage taking into account risk selection, I construct the distribution of

$$CS_g(\lambda^*, \tilde{\pi}_m) - CS_g(\chi^*, \lambda^*, \pi^o_m) \tag{20}$$

across counties for each age/gender group. I apply weights to (20) where the weights are the number of Medicare eligibles in age/gender group g in that county. The descriptive statistics are also summarized in table 12. The results are consistent with the theoretical predictions. While the average change in consumer surplus due to Medicare Advantage is positive for all age/gender groups, it generally decreases in age group for both men and women. The exception is that for both men and women, the average change in consumer surplus is lower in the 65-69 group than in the 70-74 group. Figures 18 and 19 depict the CDFs of the changes in consumer surplus for men and women, respectively. While first order stochastic domination is not exhibited between every age group pair, the figures make it clear that younger age groups are more positively affected by Medicare Advantage than older age groups.

In each age/gender group, there is some share of the population that is negatively affected by Medicare Advantage. These shares are summarized in Table 13. Consistent with intuition, a larger share of older age groups are negatively affected for both men and women. However, these shares are very small with a maximum across age/gender groups of 0.391%. In addition, the values of the losses are relatively small. The maximum loss across all counties and age/gender groups is \$21.45, which is experienced by men age 85 and over in Los Alomos county, New Mexico. The maximum gain across all counties and age/gender groups is \$1226.27, which is experienced by women age 70-74 in Sonoma county, California.

In another measure of the welfare effect of Medicare Advantage, Table 14 provides the predicted share of the Medicare Advantage enrollment that would have selected Medigap if it were priced as it would have been in the absence of Medicare Advantage. The results show that 0.0021%-0.0049% would have done so, and although the relationship is not monotone in age, it is generally higher in older age groups.

7 Conclusion

This research studies the welfare trade-off between providing more health plan options through the Medicare Advantage program and the possibility that risk selection into Medicare Advantage raises premiums for Medigap supplemental coverage. Theory predicts that the welfare effect will vary across risk types with higher risk types being more likely to be adversely affected.

To address this issue, this research has presented a structural model of health plan and health care utilization that contributes to the existing literature by incorporating the notion of health care as a derived demand into the consumer choice problem. The model captures the technological relationship between health care and health by including many of the aspects thought to be important in the production of health but commonly not treated in the applied literature.

The results suggest that the diversity of health plan choices brought about by Medicare Advantage has resulted in significant increases in consumer surplus across all age/gender groups. However, the evidence also suggests that the change in consumer surplus is not uniform within or across age/gender groups and that higher risk beneficiaries are more likely to be adversely affected by Medicare Advantage, although the adverse effects are very small. Finally, the evidence also suggests that less than 0.005% of all Medicare Advantage enrollees would have preferred to be enrolled in traditional Medicare with a Medigap supplement if the Medigap supplement were priced as it would be in the absence of Medicare Advantage.

8 Appendices

8.1 Specification of $g(\theta)$

In this appendix, I describe the issues considered and some results of early experiments that led to my choice of $g(\theta)$. For ease of exposition, I examine a simple one dimensional, deterministic case. The issue is a potential conflict between incorporating the notion of diminishing marginal product on the extensive margin and maintaining the intuitive comparative static that indirect utilities should fall with the pretreatment state θ. Consider posttreatment health in the simple model described in section 4.1.1,

$$H = \tau \left(\bar{m}(\theta)m - \frac{1}{2}m^2 \right) - \theta$$

where

$$\bar{m}(\theta) = \exp\left\{\chi_2 + \zeta g(\theta)\right\}.$$

At an optimum, we have

$$\frac{\partial H^*}{\partial \theta} = \tau \exp\{\chi_2 + \zeta g(\theta)\}\zeta g'(\theta)m^* - 1$$

and

$$\frac{\partial U^*}{\partial \theta} = \frac{\partial U}{\partial H}\frac{\partial H}{\partial \theta}\bigg|_{m^*} = \frac{\partial U^*}{\partial H}[\tau \exp\{\chi_2 + \zeta g(\theta)\}\zeta g'(\theta)m^* - 1].$$

So $\frac{\partial U^*}{\partial \theta} < 0$ if and only if

$$\tau \exp\{\chi_2 + \zeta g(\theta)\}\zeta g'(\theta)m^* < 1.$$

The intuition is that the consumer should not be healthier after treatment resulting from an increase in θ than if the consumer hadn't experienced the increase in θ to begin with. This can be guaranteed by removing the property of diminishing marginal product on the extensive margin, i.e., $\zeta g'(\theta) = 0$. If this property is maintained, it is possible that $\frac{\partial U^*}{\partial \theta} > 0$ would be true at some data points for some values of the model parameters. My object is to choose $g'(\theta)$ so as to minimize this possibility.

My first choice was $g(\theta) = \theta$ with a small starting guess of $\zeta = 0.001$. I tested this by plotting the path of θ that makes an observed value of m utility maximizing and the corresponding indirect utility. I use the values $\gamma_1 = 0.4, \gamma_2 = .6$, $\delta = 2$, $\tau = 1$, $\chi_2 = 4$, and $y = 24$. I input medical expenses m from 1 to 100 assuming the consumer pays a constant coinsurance rate of 0.2. For each value of m, I solve for the θ that satisfies the consumer's first order condition, and the resulting indirect utility, and the value of $\frac{\partial H^*}{\partial \theta}$. Figure 21 plots the path of m^* and U^* on θ. The path of m^* has the intuitive upward slope. The path of U^* is initially downward sloping but later rises. Figure 22 illustrates the problem. The marginal effect of pretreatment health on posttreatment health is not only positive, it eventually explodes. So even though the consumer spend more of his income on health care, the care is becoming so productive that the consumer actually is better off being very sick and requiring a large amount of health care.

Figures 22 and 23 give the same results for $g(\theta) = \left(1 + \exp\left(\frac{\theta - a}{b}\right)\right)^{-1}$ and $a = 0,\ b = 10,000$, and $\zeta = 1$. As before, the path of m^* has the intuitive upward slope. In this case, the path of U^* also is intuitive in that it slopes down and is monotone. This is because the path of $\frac{\partial H^*}{\partial \theta}$ is always negative and goes to -1 as θ becomes very large or small.

8.2 Out-of-Pocket Expenses

In this appendix, I discuss the reduced form models used to estimate the relationship between health care expenditures and out-of-pocket expenses and marginal prices for Medicare Advantage enrollees. This is required because I use deflated health care expenditures as the choice variable for the consumer and health insurance cost sharing is usually defined as a fixed copayment per event. Therefore, some relationship must established between health care expenditures and the variable relevant for consumer cost sharing in order to define an appropriate total out-of-pocket and marginal price for the consumer choice problem. This is unnecessary for Medigap and FFS-only enrollees because, for all types of care other than inpatient hospital care, either consumer cost sharing is defined in terms of coinsurance rates or the consumer faces the full marginal price.

8.2.1 Doctor Visits and Outpatient Care

For doctor visits and outpatient events, I use ordinary least squares as opposed to a count data model since there are a large number of trips and the distribution is relatively continuous. In addition, Poisson and negative binomial regressions resulted in implausibly large predicted values for the number of events in some cases. For example, the maximum number of doctor visit events in the MCBS data is 444. The maximum predicted value is 314 using least squares but 56,820 using a negative binomial model and 750 using a Poisson regression. Since there is high degree of correlation between doctor visit events and outpatient expenditures and vice versa, I use both types of expenditures to predict both types of event counts.. I use the all of the MCBS data and county level data from the Area Resource File. The results are given in Table 15.

8.2.2 Prescription Drugs

For prescription drug events, there is further complication in that copays are different for branded and generic drugs and much of the total expenditure is on off-formulary drugs. Classifying each event as either generic, branded, or off-formulary is impractical because the data required is not available. I make the simplifying assumption that the consumer's share of the total expenditure is a function of consumer and area characteristics and the particular cost sharing terms of the health plan. These cost sharing terms include an indicator of branded drug coverage, and indicator for the type of cost sharing for generic and branded drugs (coinsurance or copay) and the amount of the cost sharing given the type. I also include an indicator of the presence of an annual maximum benefit and the level of the maximum if one exists. Defining y as the consumer's share of total expenditures, I regress $\ln y - \ln(1 - y)$ on the above covariates. The results are given in Table 16. Note that this assumption implies that, in the model, the beneficiary faces a constant marginal price for prescription drugs given health plan choice.

8.2.3 Dental Care

Dental care events are complicated by the fact that, in all cases, Medicare MCOs only cover preventive dental events and charge a copay per event. Fortunately, the data allows me to separate preventive from nonpreventive events. The MCBS records if specific services were performed at each event. I classify an event as nonpreventive if any of the following services were performed: extractions, fillings, root canals, crowns, bridge work, orthodontics. I then use total expenditures and total preventive expenditures to predict the number of preventive events using a Poisson regression. I use Poisson rather than negative binomial because a test that the dispersion parameter is zero is not rejected at the 0.50 level. The results are given in Table 17. I assume the beneficiary pays the full cost of the nonpreventive events plus the predicted number of preventive events times the preventive event copay.

8.2.4 Inpatient Care

Out-of-pocket expenses for inpatient events are significantly massed at zero. 90% of all Medicare Advantage enrollees who consumed some inpatient care paid nothing. This percentage is not significantly different if the MCO charges an admission fee (89.5%) or not (90.3%). 89.9% of Medicare Advantage enrollees were in a plan that did not charge an inpatient admission fee. 94.6% of all Medigap enrollees and 66.0% of all FFS-only beneficiaries who consumed some inpatient care paid nothing. Of the FFS-only beneficiaries who consumed some inpatient care, 19.4% paid the inpatient event deductible of $776. Count model regressions of inpatient expenditures on inpatient events yielded nothing significant. Because if this, I make the simplifying assumption that the inpatient admission fee, FFS-only or Medicare Advantage, is an annual fixed cost. If a beneficiary consumed any inpatient care, I assume he/she paid the admission fee, but the marginal price is zero.

References

[1] Atherly, Adam (2002), "The Effect of Medicare Supplemental Insurance on Medicare Expenditures", International Journal Of Health Care Finance and Economics, 2, 137-162.

[2] Atherly, A., B. Dowd, and R. Feldman (2004), "The Effect of Benefits, Premiums, and Health Risk on Health Plan Choice on the Medicare Program", Health Services Research, 39(4), 847-864.

[3] Beaulieu, Nancy Dean (2002), "Quality Information and Consumer Health Plan Choices", Journal of Health Economics, 21, 43-63.

[4] Berry, Steven T. (1994),"Estimating Discrete-Choice Models of Product Differentiation", RAND Journal of Economics, 25(2), 242-262.

[5] Berry, Steven, James Levinsohn, Arial Pakes (1995), "Automobile Prices in Market Equilibrium", Econometrica, 63(4), 841-890.

[6] Berry, Steven, James Levinsohn, Arial Pakes (1998), "Differentiated Product Demand Systems from a Combination of Micro and Macro Data: The New Car Market", NBER Working Paper 6481.

[7] Blomqvist, Ake (1997), "Optimal Non-linear Health Insurance", Journal of Health Economics, 16, 303-321.

[8] Börsch-Suppon, Axel and Vassilis Halivassiliou (1993), "Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models", Journal of Econometrics, 58, 347-368.

[9] Brownstone, D., and K. Train (1999), "Forecasting New Product Penetration with Flexible Substitution Patterns", Journal of Econometrics, 89, 109-129.

[10] Buchmueller, Thomas (2006), "Price and Health Plan Choices of Retirees", Journal of Health Economics, 25(4), 81-101.

[11] Call, Kathleen Thiede, Bryan Dowd, Roger Feldman, Matthew Maciejewski (1999), "Selection Experiences in Medicare HMOs: Pre-Enrollment Expenditures", Health Care Financing Review, 20(4), 197-209.

[12] Cameron A., Pravin Trivedi, Frank Milne, and J. Piggott (1988), "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia", Review of Economic Studies, 55, 85-106.

[13] Cardon, James H. and Igal Hendel (2001), "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey", RAND Journal of Economics, 32(3), 408-427.

[14] Cutler, David (2000), "The Anatomy of Health Insurance," in: A.J. Culyer and J.P. Newhouse, eds., The Handbook of Health Economics (Elsevier, Amsterdam), 563-644.

[15] Cutler, David, Sarah J. Reber (1998), "Paying for Health Insurance: the Trade-off Between Competition and Adverse Selection", The Quarterly Journal of Economics, 113(2), 433-465.

[16] Cutler, D., N. Beaulieu, and K. Ho (2005), "The Business case for Diabetes Disease Management at Two Managed Care Organization", Working Paper, Harvard University.

[17] Dardanoni, Valentino and Adam Wagstaff (1990), "Uncertainty and the Demand for Medical Care", Journal of Health Economics, 9, 23-38.

[18] Deb, Partha and Pravin Trivedi (1997), "Demand for Medical Care by the Elderly: A Finite Mixture Approach", Journal of Applied Econometrics, 12, 313-336.

[19] Deb, Partha and Pravin Trivedi (2002), "The Structure of Demand for Health Care: Latent Class versus Two Part Models", Journal of Health Economics, 21(4), 601-628.

[20] Dowd, Bryan, Roger Feldman, Steven Cassou, and Michael Finch (1991), "Health Plan Choice and the Utilization of Health Care Services", Review of Economics and Statistics, 73(1), 85-93.

[21] Dubin, Jeffrey A. and Daniel L. McFadden (1984), "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption", Econometrica, 52(2), 345-362.

[22] Feldman, Roger, Bryan E. Dowd, and Marian Wrobel (2003), "Risk Selection and Benefits in the Medicare+Choice Program", Health Care Financing Review, 25(1), 23-36.

[23] Feldman, Roger, Michael Finch, Bryan Dowd, and Steven Cassou (1989), "The Demand for Employment-Based Health Insurance Plans", Journal of Human Resources, 24(1), 115-142.

[24] Finkelstein, Amy and Kathleen McGarry (2003), "Private Information and its Effect on Market Equilibrium: New Evidence from Long-Term Care Insurance", NBER Working Paper 9957.

[25] Grossman, Michael (1972), "On the Concept of Human Capital and the Demand for Health", Journal of Political Economy, 80, 223-255.

[26] Hanemann, W. Michael (1984), "Discrete/Continuous Models of Consumer Demand", Econometrica, 52(3), 541-562.

[27] Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moment Estimators", Econometrica, 50, 1029-1054.

[28] Harris, Kathleen M. and Michael P. Keane (1999), "A Model of Health Plan Choice: Inferring Preferences and Perceptions from a Combination of Revealed Preference and Attitudinal Data", Journal of Econometrics, 89, 131-157.

[29] Heckman, James (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", Econometrica, 46(6), 931-959.

[30] Heckman, James (1979), "Sample Selection Bias as a Specification Error", Econometrica, 47(1), 153-161.

[31] Hellinger, Fred and Herbert Wong (2000), "Selection Bias in HMOs: A Review of the Evidence", Medical Care Research and Review, 57(4), 405-439.

[32] Hubbard-Rennhoff, Christina (2005), "Employee Choice of Health Plan Choice and Consumption of Health Care under Employer-Provided Plans", Ph. D. Dissertation, University of Virginia.

[33] Hunt-McCool, Janet, B. Kiker, and Ying Chu Ng (1994), "Estimates of the Demand for Medical Care under Different Functional Forms", Journal of Applied Econometrics, 9, 201-218.

[34] Imbens, Guido and Tony Lancaster (1994), "Combining Micro and Macro Data in Microeconometric Models", Review of Economic Studies, 61, 655-680.

[35] Lee, Lung-Fei and Mark M. Pitt (1986), "Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach", Econometrica, 54(5), 1237-1242.

[36] Manning, W., J.P.Newhouse, N. Duan, E. Keeler and A. Leibowitz (1987), "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment", American Economic Review, 77(3), 251-277.

[37] McCellan, Mark (2000), "Medicare Reform: Fundamental Problems, Incremental Steps", Journal of Economic Perspectives, 14(2), 21-44.

[38] Newhouse, Joseph P. (2001), "Medicare Policy in the 1990s", NBER Working Paper 8531.

[39] Pohlmeier, Winfried and Volker Ulrich (1995), "An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care", Journal of Human Resources, 30(0), 339-361.

[40] Rothschild, Michael and Joseph Stiglitz (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", Quarterly Journal of Economics, 90(4), 629-649.

[41] Small, Kenneth and Harvey Rosen (1981), "Applied Welfare Economics with Discrete Choice Models", Econometrica, 49(1), 105-1300.

[42] Smith, Cynthia, Cathy Cowen, Art Sensenig, and Aaron Catlig (2005), "Health Spending Growth Slows in 2003", Health Affairs, 24(1), 185-194.

[43] Stern, Steven (1997), "Simulation-Based Estimation", Journal of Economic Literature, 35, 2006-2039.

[44] Strombom, B., T. Buchmueller, and P. Feldstein (2002), "Switching Costs, Price Sensitivity, and Health Plan Choice, Journal of Health Economics, 21(1), 89-116.

[45] Town, Robert and Su Liu, (2003), "The Welfare Impact of Medicare HMOs", RAND Journal of Economics, 34(4), 719-736.

[46] Wales, T.J. and A.D. Woodland (1983), "Estimation of Consumer Demand Systems with Binding Non-negativity Constraints", Journal of Econometrics, 21(3), 263-285.

[47] Zabinski, Daniel (1994), "Reform of the Tax Preference for Employer Provided Health Insurance", Ph.D. Dissertation, University of Virginia.

Table 1 Health Care Expenses and Budget Share
MCBS Sample (N=5251)

	Medigap		Medicare Advantage		FFS Only	
	Mean (Stan Dev)	Share	Mean (Stan Dev)	Share	Mean (Stan Dev)	Share
Inpatient Care	3086.21 (9372.51)	39.7%	2164.41 (7901.79)	44.5%	2778.17 (10260.50)	50.1%
Outpatient Care	790.29 (2023.57)	10.2%	581.90 (4552.42)	12.0%	529.04 (4552.42)	9.6%
Doctor Visits	2696.72 (3844.04)	34.7%	952.42 (1778.12)	19.6%	1551.57 (2729.00)	28.3%
Prescription Drugs	944.96 (1090.95)	12.2%	867.50 (1490.41)	17.8%	527.53 (864.59)	9.6%
Dental Care	258.82 (1123.94)	3.3%	293.95 (814.16)	6.0%	99.02 (737.68)	1.8%
Total	7777.00 (12648.88)	100%	4860.176 (10473.85)	100%	5485.319 (12587.74)	100%

Table 2 Age, Gender, and Income by Insurance Type
MCBS, Age Entitlement Only

	N	Share of Sample	Age	Female	Income
			Mean (Standard Deviation)		
Medigap	3103	30.4%	78.72 (7.58)	.61 (.49)	28,087.65 (49,066.55)
Medicare Advantage	1439	14.1%	76.49 (7.07)	.59 (.49)	24,650.54 (24,009.07)
FFS Only	709	7.0%	77.91 (8.40)	.50 (.50)	17,852.31 (15,543.33)
Employer Provided	3519	34.5%	76.37 (6.79)	.54 (.50)	34,999.07 (40,964.77)
Medicaid or QMB	1430	14.0%	79.80 (8.78)	.73 (.44)	11,811.97 (41,201.77)
Full MCBS	10,361	100%	77.70 (7.61)	.59 (.49)	26,964.91 (41,125.48)

Table 3 Mean Total Health Care Expenditures by Age/Gender
MCBS Sample (N=5251)

Mean/(Stan Dev)/ N

	65-69	70-74	75-79	80-84	85 & over
Female	5634.49 (10,547.75) 416	6248.629 (12,241.93) 649	6855.212 (11,381.55) 589	6386.13 (8913.13) 656	6337.73 (10,096.61) 800
Male	6210.49 (16,665.41) 385	5939.152 (11,345.19) 567	8465.763 (15,738.76) 463	7475.57 (11430.83) 387	8103.799 (14,884.14) 339

Table 4 Descriptive Statistics of Medigap and
Medicare Advantage Plans

Medigap Basic Benefit:

i) Part A Coinsurance. Coverage is extended for an additional 365 days.
ii) Part B Coinsurance.
iii) First three pints of blood.

Medigap Additional Benefits:

BENEFIT	PLAN A	PLAN B	PLAN C	PLAN F
Skilled Nursing Coinsurance			X	X
Part A Deductible		X	X	X
Part B Deductible			X	X
Part B Excess				X
Foreign Travel Emergency			X	X

Plan F Premiums (weight=Medicare eligibles):
National Weighted Average: $131.24
Weighted Average in Medicare Advantage Counties: $136.27

Medicare Advantage:

Premium, National Weighted Average	$31.70
Percentage of plans offered at a zero premium	32.5%
Percentage of plans that offer a drug benefit	73.8%
Average generic copay	$7.45
Branded drug benefit given a drug benefit	82.9%
Average branded copay	$17.70
Percentage of plans that offer a dental benefit	20.9%
Average preventive dental copay	$7.33
Average inpatient admission fee	$28.83
Percentage of plans with no inpatient admission fee	86.6%
Average outpatient admission fee	$26.86
Percentage of plans with no outpatient admission fee	87.4%
Average physician copay	$7.89
Percentage of plans with no physician copay	7.5%

Table 5 Descriptive Statistics of the Instruments

	Mean	Stan Dev	Min	Max
Physicians	19.85	18.96	0	175.74
Dentists	3.90	2.28	0	15.27
Hospital Beds	0.04	0.21	0	4.32
RNs	22.29	20.39	0	334.44
LPNs	3.44	4.47	0	115.59
Ambulatory Surgical Centers	0.07	0.11	0	1.71
Commercial HMO Share	0.28	0.15	0	0.96

All variables, except commercial sector HMO market share, are defined as per 1000 Medicare beneficiaries.

Table 6 Regression of MCO Premium on Instruments

Variable	Coefficient	Stan Error	p value
Constant	35.27	2.17	0.000
Physicians	-0.34	0.09	0.000
Dentists	1.84	0.61	0.003
Hospital Beds	17.03	4.33	0.000
RNs	0.12	0.07	0.077
LPNs	-0.82	0.24	0.001
Ambulatory Surgical Centers	-21.61	7.79	0.006
Commercial HMO Share	-16.03	5.82	0.006

R^2 =0.036, Root MSE = 35.04, N = 1765, F(7,1757) = 9.37

Estimates of Structural Parameters
(Standard Errors in Parentheses)

Table 7 Preferences and Endowments

γ_1	0.6333** (0.0120)	δ	2.9804** (0.1209)
γ_2	0.6460** (0.0053)	$\kappa_{IP}{}^{F}$	4.9572** (0.3635)

	Dental Care	Prescription Drugs	Doctor Visits	Outpatient Care	Inpatient Care
κ_1	0.1098** (0.0049)	0.1488** (0.0013)	0.0471** (0.0016)	0.0341** (0.0008)	0.1061** (0.0227)
κ_2	-0.0916** (0.0055)	-0.1936** (0.0049)	-0.0942** (0.0018)	-0.0679** (0.0021)	-0.0637** (0.0253)
β_{female}	606.533** (79.976)	423.072* (268.82)	-84.010 (316.85)	-180.897 (231.78)	-743.238 (879.03)
β_{age}	2002.053** (16.533)	506.624** (78.706)	2021.927** (98.409)	364.043** (39.819)	290.100** (85.814)
$\beta_{age}{}^{2}$	-13.7981** (0.1174)	-3.4699** (0.5346)	-11.9319** (0.6532)	-1.8373** (0.3022)	-1.4620 (0.8786)

Table 8 Production and CES Functions

ρ	0.4111** (0.0013)

	Dental Care	Prescription Drugs	Doctor Visits	Outpatient Care	Inpatient Care
α	0.1617 ----	0.2535** (0.0014)	0.2542** (0.0013)	0.1513** (0.0021)	0.1793** (0.0014)
ζ	0.6480** (0.0225)	2.4445** (0.130)	2.5206** (0.0296)	2.7272** (0.0245)	2.9810** (0.0070)
Ω_v	0.00002 (0.4413)	0.1849** (0.0086)	0.0524** (0.0078)	0.1764** (0.0193)	0.3692** (0.0032)
a	-5563.21** (401.54)	-5901.25** (116.88)	-17,895.02** (497.04)	-35,214.55** (417.51)	-51,256.14** (468.6238)
b	2856.70** (246.33)	1367.52** (72.730)	10,205.32** (280.90)	4256.07** (130.86)	6056.48** (286.94)
$\bar{\chi}_{1FFS}$	-0.0030 (0.1109)	-0.0017 (0.4274)	-0.0091 (0.1157)	-0.0438 (0.9425)	-0.00030 (0.3181)
$\bar{\chi}_{2FFS}$	4.6456** (0.1026)	4.3571** (0.4173)	3.8480** (0.1065)	4.3066** (0.9392)	5.4793** (0.3126)

** Statistically Significant at 1% * Statistically Significant at 5%

Table 9 Parametric Distributions

	Dental Care	Prescription Drugs	Doctor Visits	Outpatient Care	Inpatient Care
μ	-77,888.36** (604.77)	-27,414.58** (3009.25)	-101,009.4** (3847.43)	-26,019.17** (1353.64)	-67,194.14** (1928.20)

Estimates of Ω_ε

	Dental Care	Prescription Drugs	Doctor Visits	Outpatient Care	Inpatient Care
Dental Care	637.93E+04** (2.320E+04)				
Prescription Drugs	228.49E+04** (10.61E+04)	73.304E+06** (16.67E+04)			
Doctor Visits	305.01E+04** (7.142E+04)	15.849E+06** (99.48E+04)	10.735E+07** (44.09E+04)		
Outpatient Care	388.38E+04** (6.559E+04)	13.899E+06** (56.98E+04)	52.745E+06** (107.6E+04)	10.871+07** (63.16E+04)	
Inpatient Care	134.26E+04** (10.45E+04)	19.211E+06** (114.4E+04)	12.160E+07** (110.7E+04)	10.151E+07** (73.26E+04)	59.961E+07** (127.8E+04)

Estimates of Ω_η

	Dental Care	Prescription Drugs	Doctor Visits	Outpatient Care	Inpatient Care
Dental Care	115.28E+04** (7.692E+04)				
Prescription Drugs	56.638E+04 (54.91E+04)	1857.4E+04** (122.0E+04)			
Doctor Visits	80.154E+04** (11.86E+04)	21.698E+04 (124.3E+04)	1397.0E+04** (118.6E+04)		
Outpatient Care	21.352E+04 (34.76E+04)	72.421E+04 (68.69E+04)	21.539E+04 (65.32E+04)	1524.4E+04** (81.36E+04)	
Inpatient Care	18.231E+04 (204.0E+04)	240.58E+04 (419.8E+04)	224.30E+04 (370.7E+04)	158.65E+04 (399.5E+04)	5563.5E+04** (481.6E+04)

σ_a	0.1743* (0.0832)	σ_w	0.8208** (0.1563)

** Statistically Significant at 1%   * Statistically Significant at 5%

Table 10 Managed Care χ_1 and χ_2

		Dental Care	Prescription Drugs	Doctor Visits	Outpatient Care	Inpatient Care
χ_1	Mean (Stan Dev)	0.2056 (0.5321)	-0.1500 (0.8477)	-0.1349 (0.6324)	-0.0010 (0.4966)	-0.4739 (0.5643)
	Min Max	-1.782 2.0532	-2.015 2.0021	-1.621 1.874	-1.2180 2.3650	-1.8540 0.9531
χ_2	Mean (Stan Dev)	3.3540 (0.6879)	3.4129 (0.7460)	3.0452 (0.6054)	3.8795 (0.4566)	5.1002 (0.4095)
	Min Max	2.0519 5.9777	1.4251 5.9563	1.4120 4.9251	1.8420 5.5011	3.2003 5.4421

Table 11 Estimates of $\overline{\chi}_{1FFS}$ and $\overline{\chi}_{2FFS}$
(Standard Errors in Parentheses)

	Dental Care	Prescription Drugs	Doctor Visits	Outpatient Care	Inpatient Care
$\overline{\chi}_{1FFS}$	-0.0030 (0.1109)	-0.0017 (0.4274)	-0.0091 (0.1157)	-0.0438 (0.9425)	-0.00030 (0.3181)
$\overline{\chi}_{2FFS}$	4.6456** (0.1026)	4.3571** (0.4173)	3.8480** (0.1065)	4.3066** (0.9392)	5.4793** (0.3126)

** Statistically Significant at 1%   * Statistically Significant at 5%

Table 12 Estimated change in Consumer Surplus
due to Medicare Advantage
Mean (Standard Deviation) Median

AGE GROUP	MEN			WOMEN		
65-69	255.32	(184.10)	235.32	276.32	(191.54)	237.41
70-74	302.30	(214.56)	245.54	289.21	(209.57)	239.58
75-79	214.63	(158.01)	174.20	206.97	(153.89)	162.44
80-84	200.48	(159.63)	158.98	180.14	(140.21)	146.28
85 and over	162.47	(135.12)	121.26	132.38	(129.25)	94.63

Table 13 Percentage of Medicare Population for whom the estimated change in Consumer Surplus in negative

AGE GROUP	MEN	WOMEN
65-69	0.009%	0.014%
70-74	0.004%	0.021%
75-79	0.025%	0.113%
80-84	0.198%	0.226%
85 and over	0.510%	0.390%

Table 14 Estimated percentage of Medicare Advantage enrollees that would switch to Medigap if Medigap were priced as it would be in the absence of Medicare Advantage

AGE GROUP	MEN	WOMEN
65-69	0.0023%	0.0024%
70-74	0.0019%	0.0026%
75-79	0.0035%	0.0038%
80-84	0.0040%	0.0041%
85 and over	0.0052%	0.0041%

Figure 1 Distributions of ln(Health Care Expenditures), Medicare Advantage and FFS Medicare (Full MCBS)

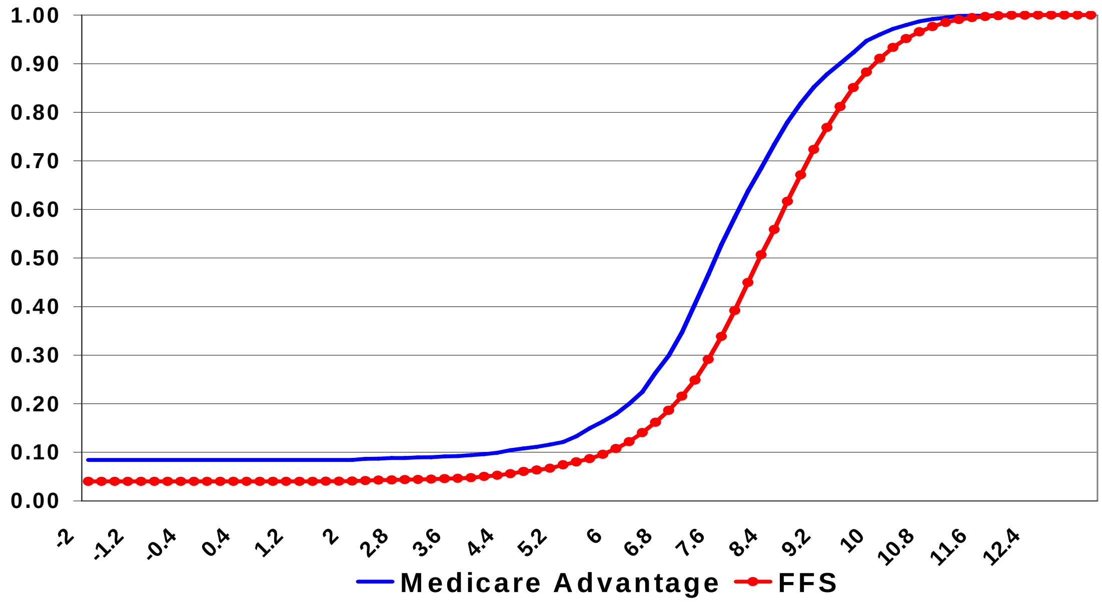

Figure 2 Distributions of ln(Health Care Expenditures), Medicare Advantage, Medigap, and FFS Medicare (MCBS Sample)

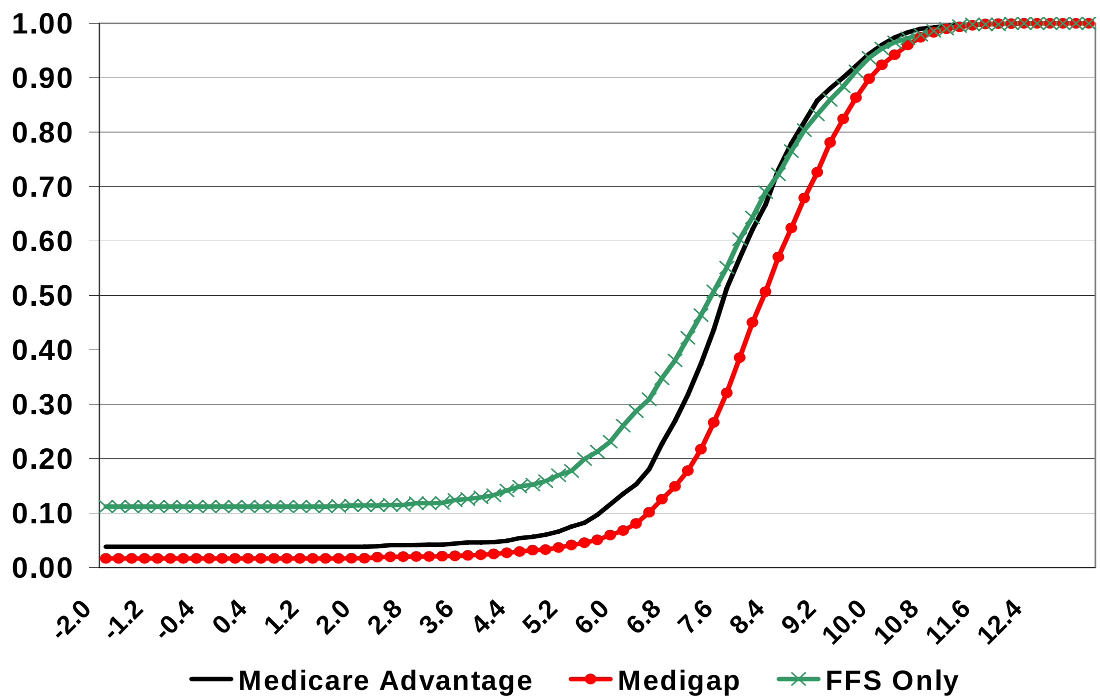

Figure 3 Medicare Advantage Enrollment
Shares by Age/Gender
Medicare Managed Care Quarterly/State/County/Plan Database

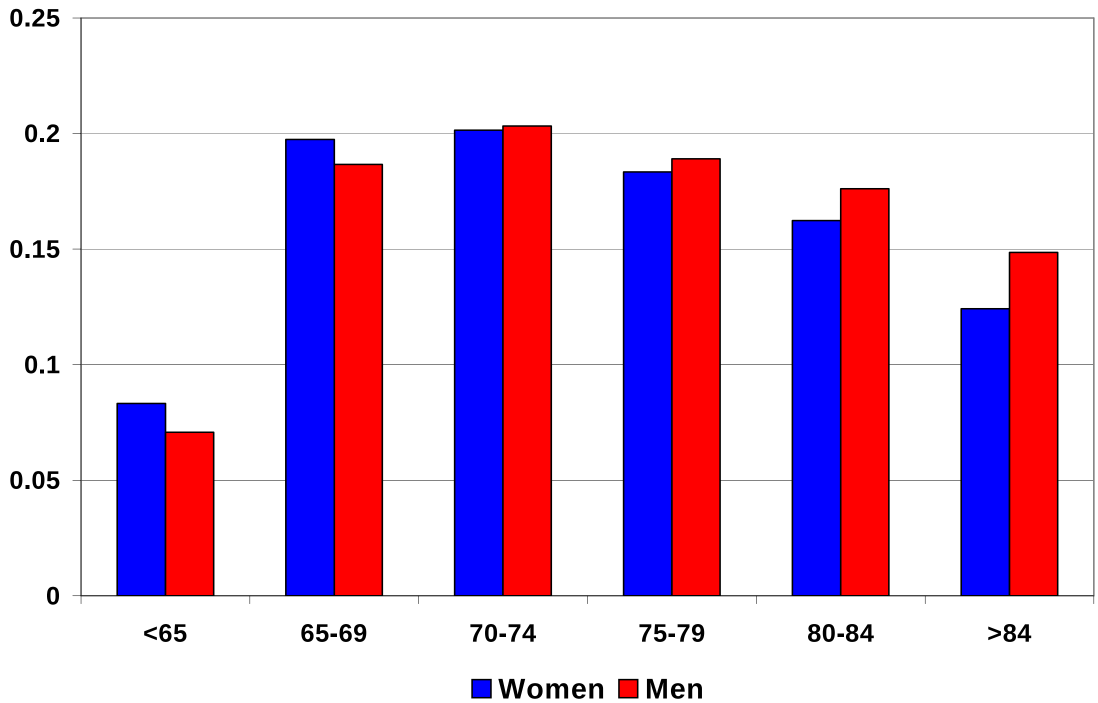

Figure 4 Proxy Medigap Premium on
Medicare Advantage Market Share

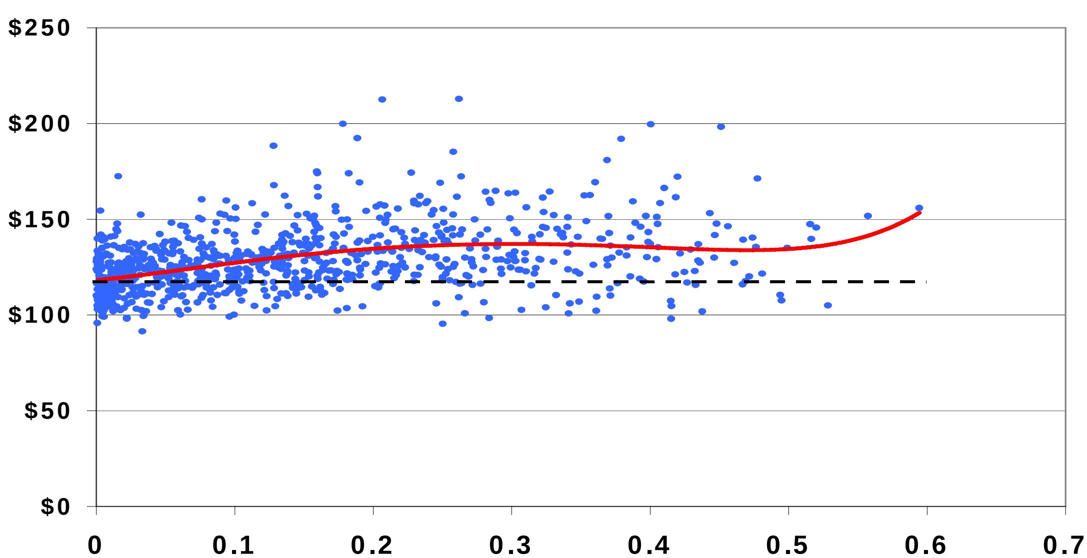

Figure 5 Satiation in Preferences because of Limits on the Efficacy of Treatment

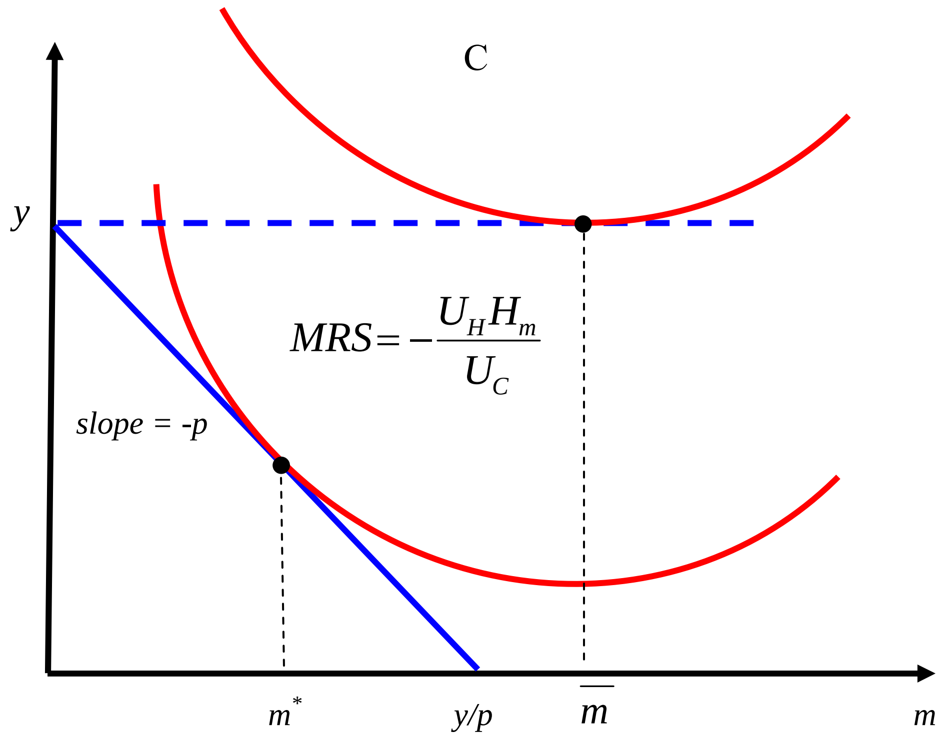

Figure 6 Joint Distribution of χ_1 and χ_2: Inpatient Care

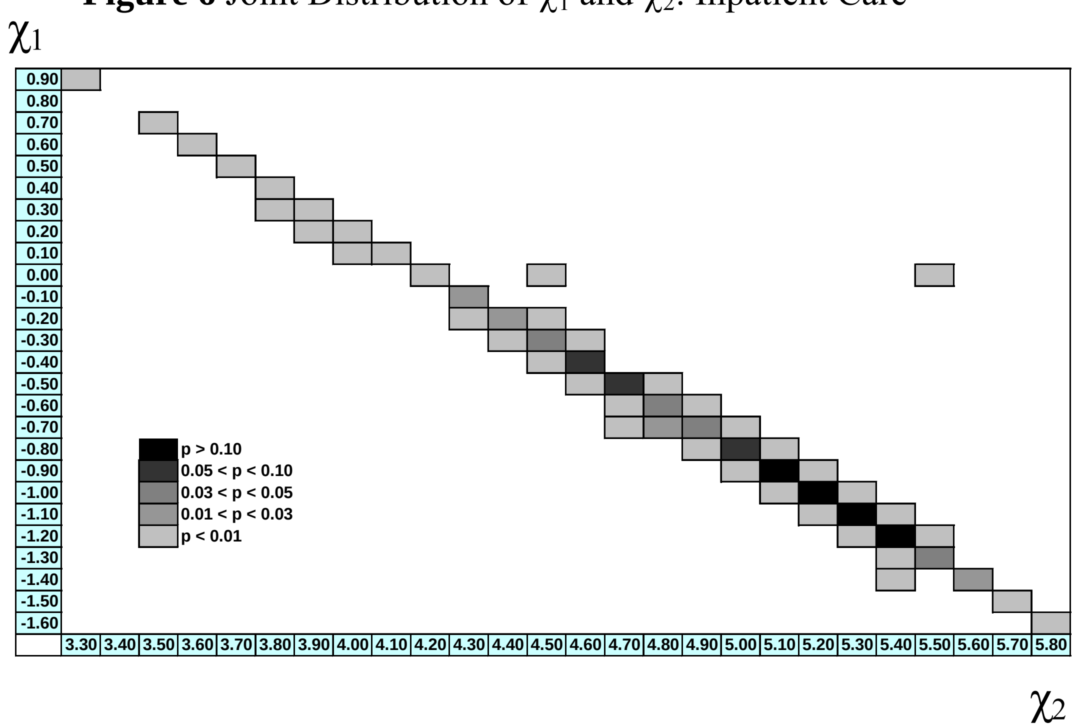

Figure 7 Joint Distribution of χ_1 and χ_2: Outpatient Care

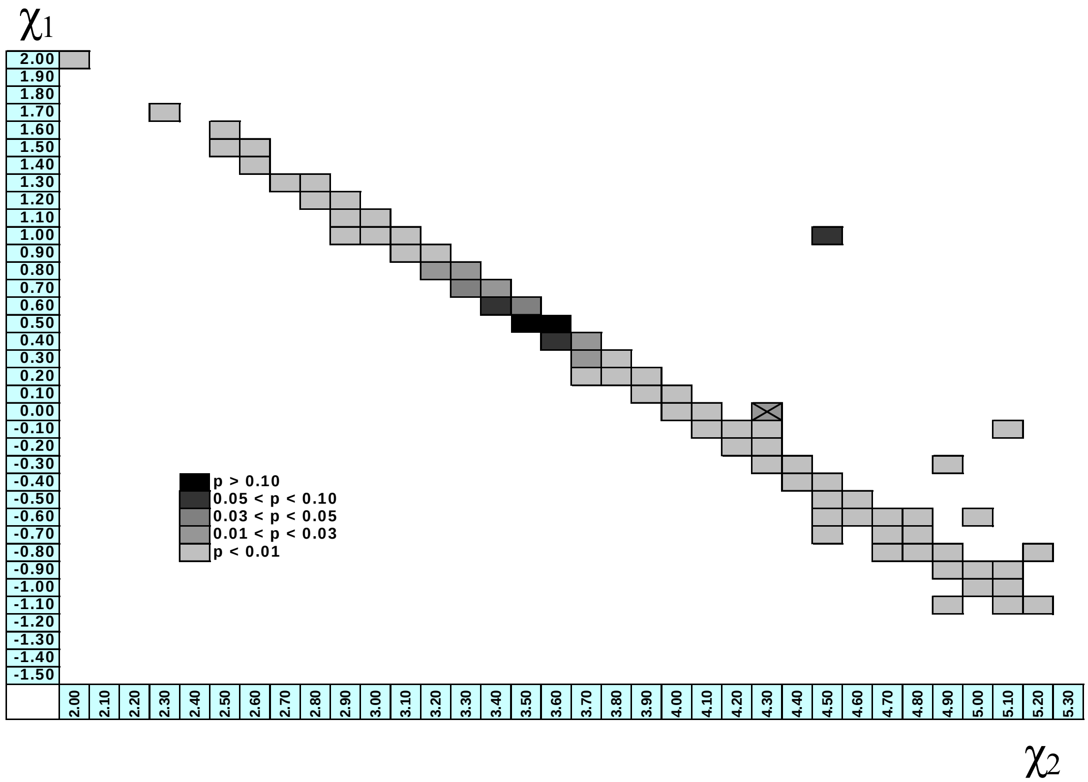

Figure 7 Joint Distribution of χ_1 and χ_2: Doctor Visits

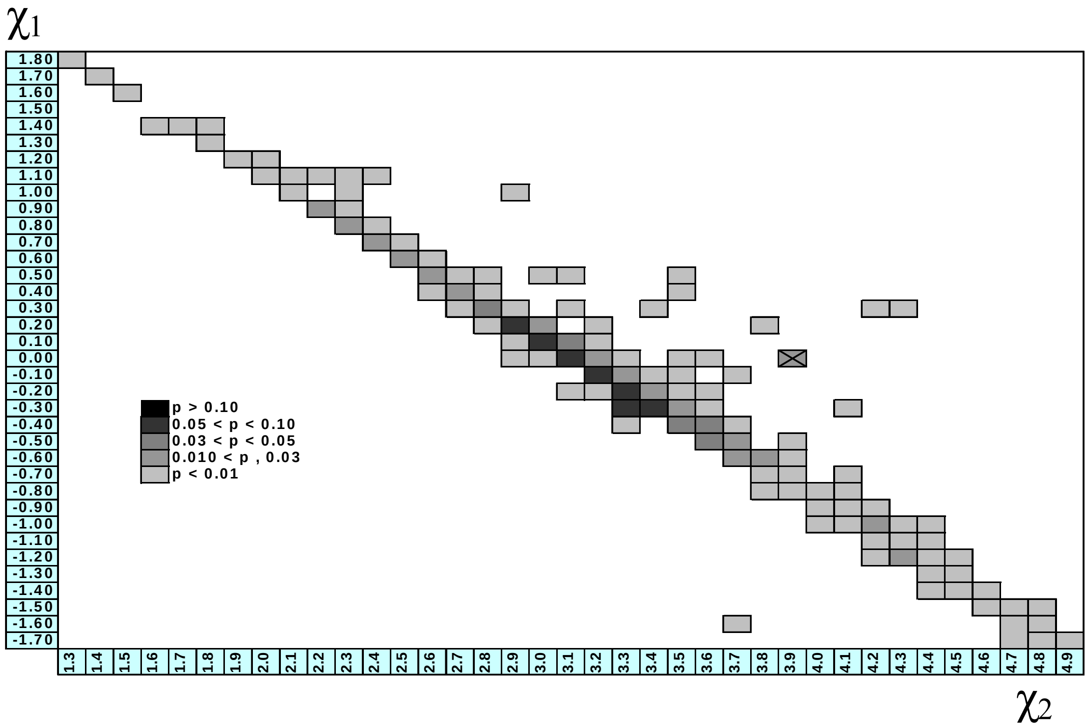

Figure 9 Joint Distribution of χ_1 and χ_2: Prescription Drugs

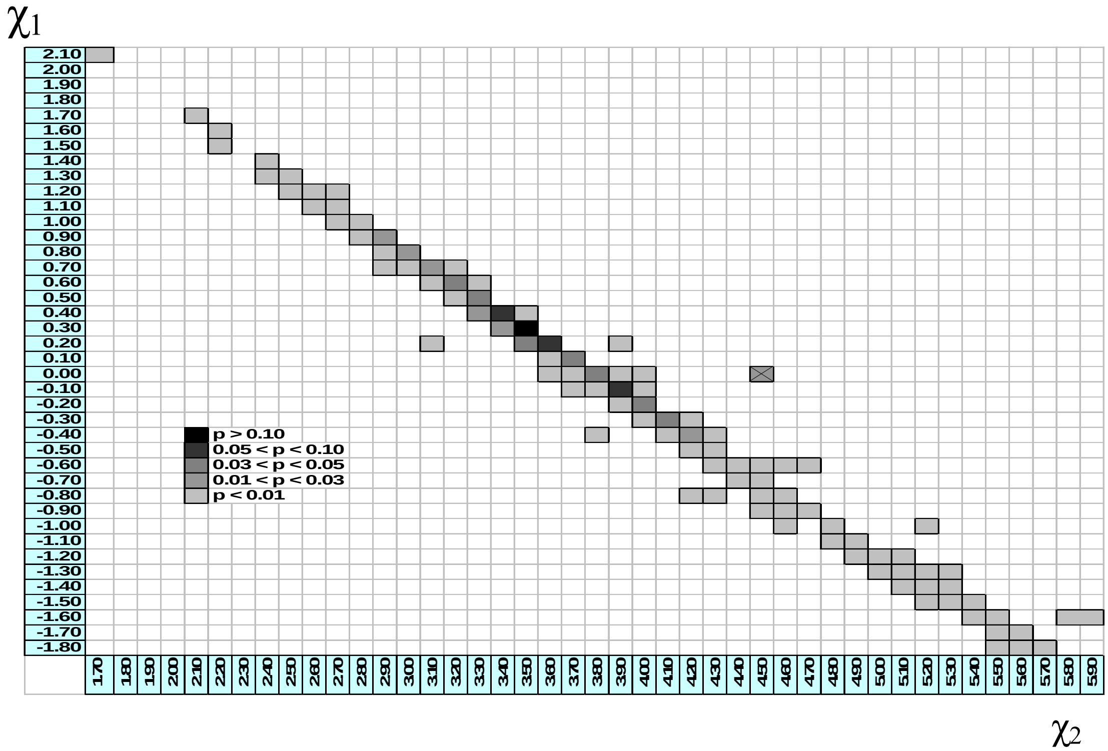

Figure 10 Joint Distribution of χ_1 and χ_2: Dental Care

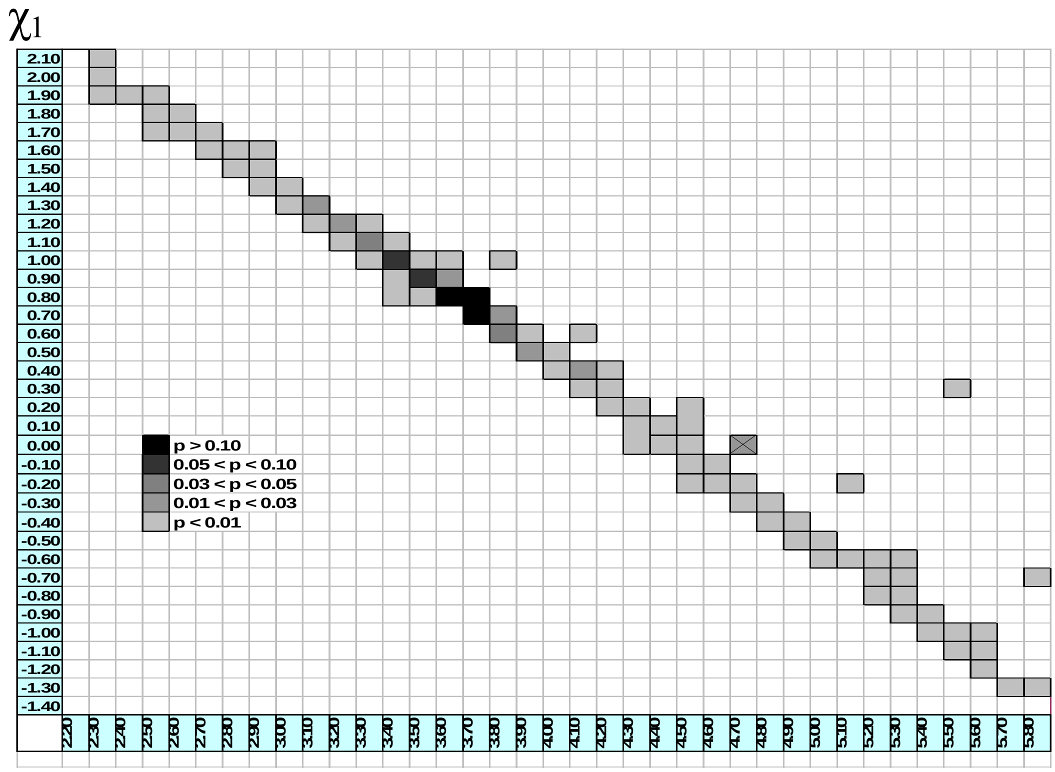

Figure 11 CDF of Inpatient Care Expenditures

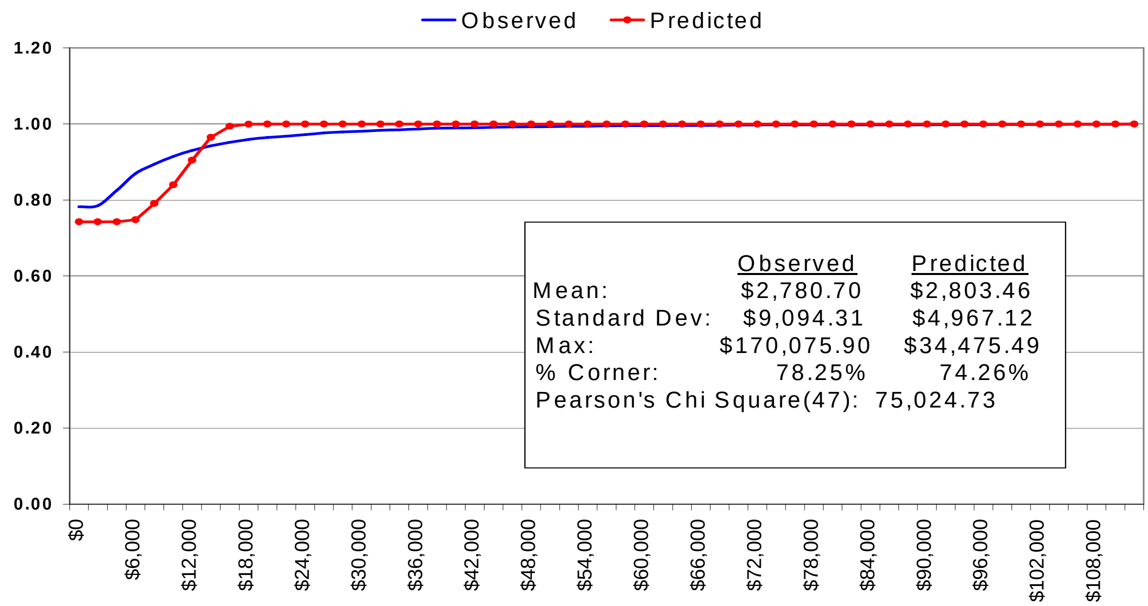

Figure 12 CDF of Outpatient Care Expenditures

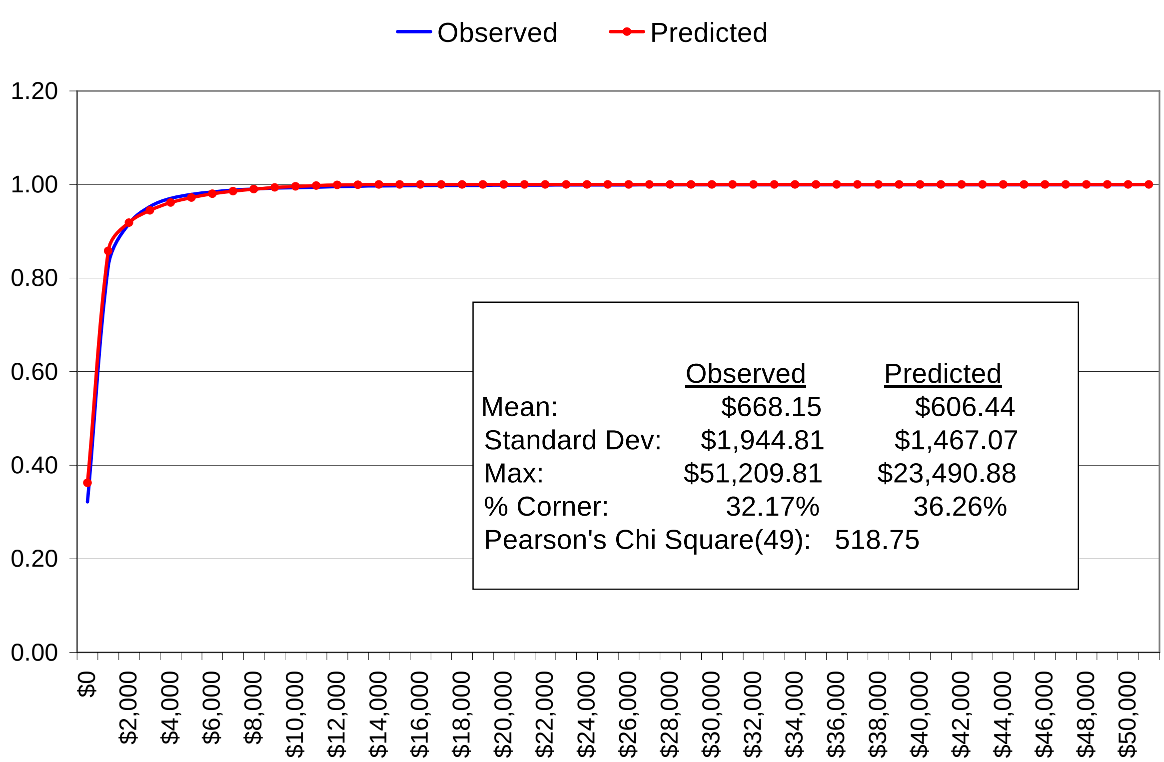

Figure 13 CDF of Doctor Visit Expenditures

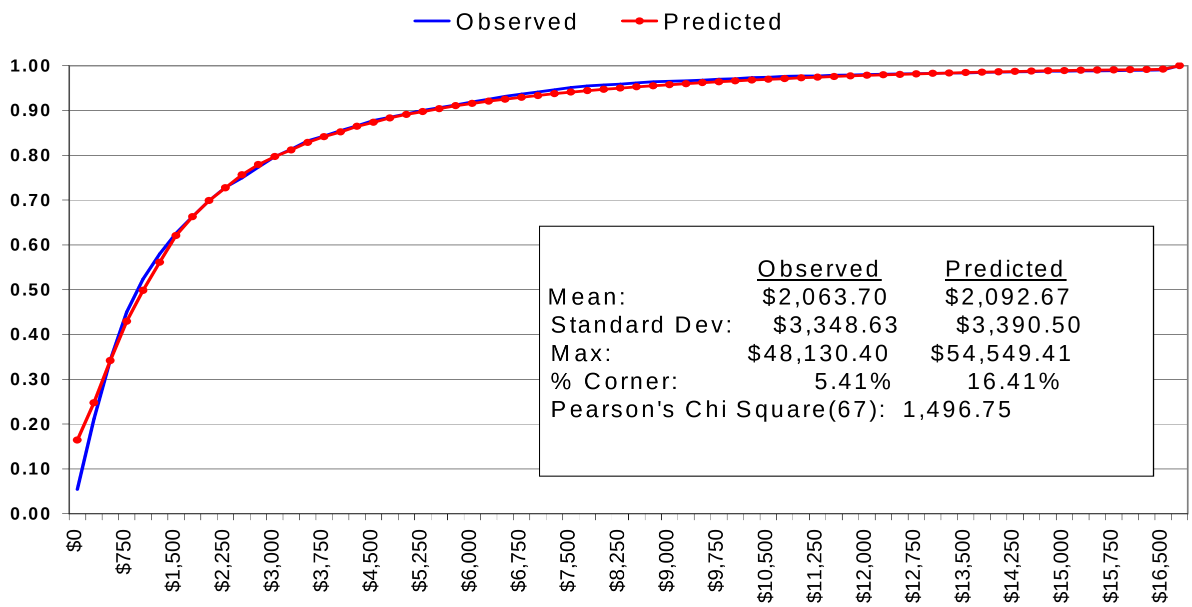

Figure 14 CDF of Prescription Drug Expenditures

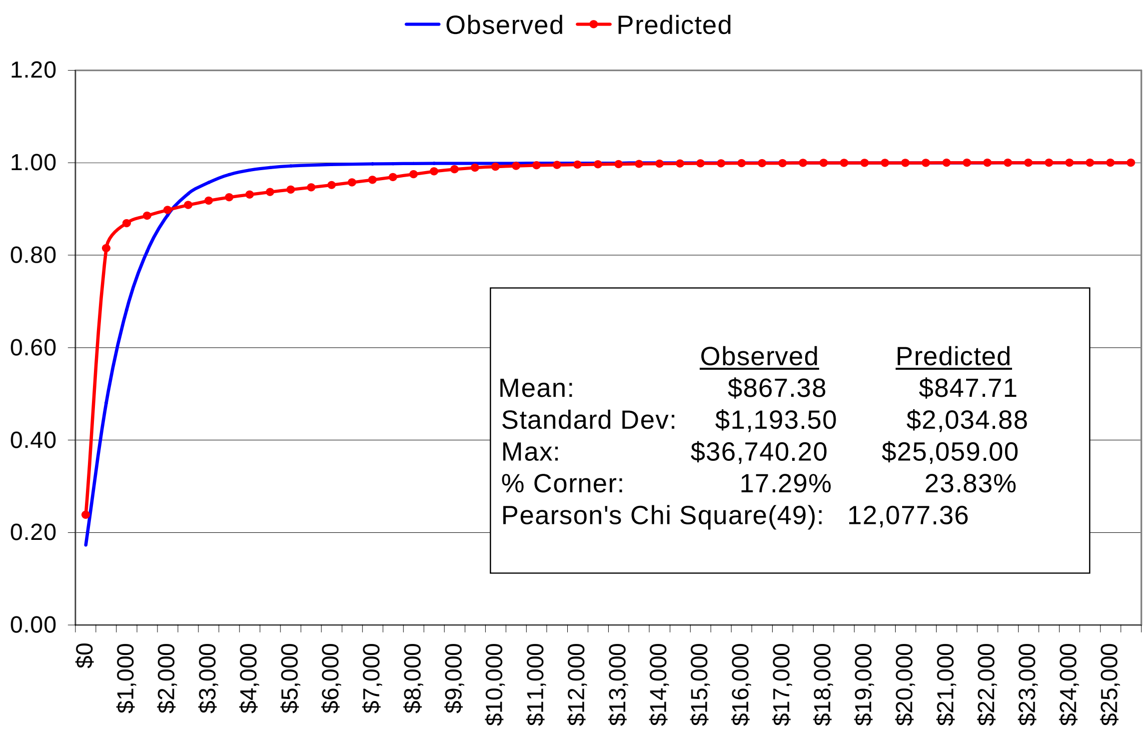

Figure 15 CDF of Dental Care Expenditures

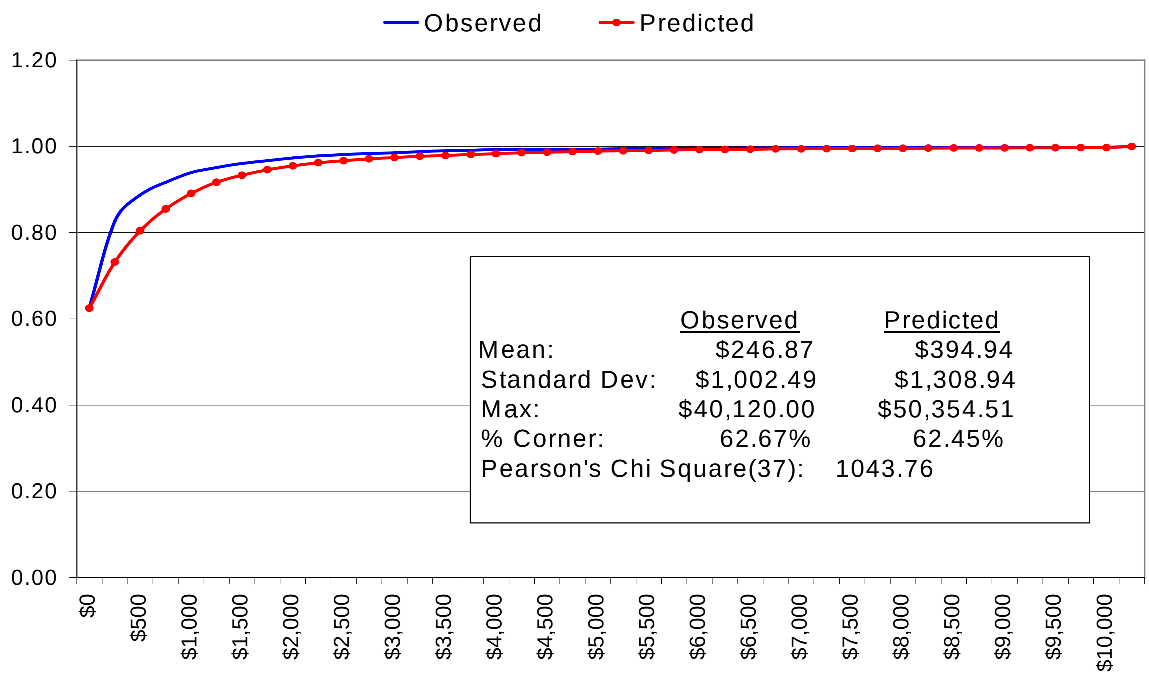

Figure 16 Predicted Probability of Choosing Medigap

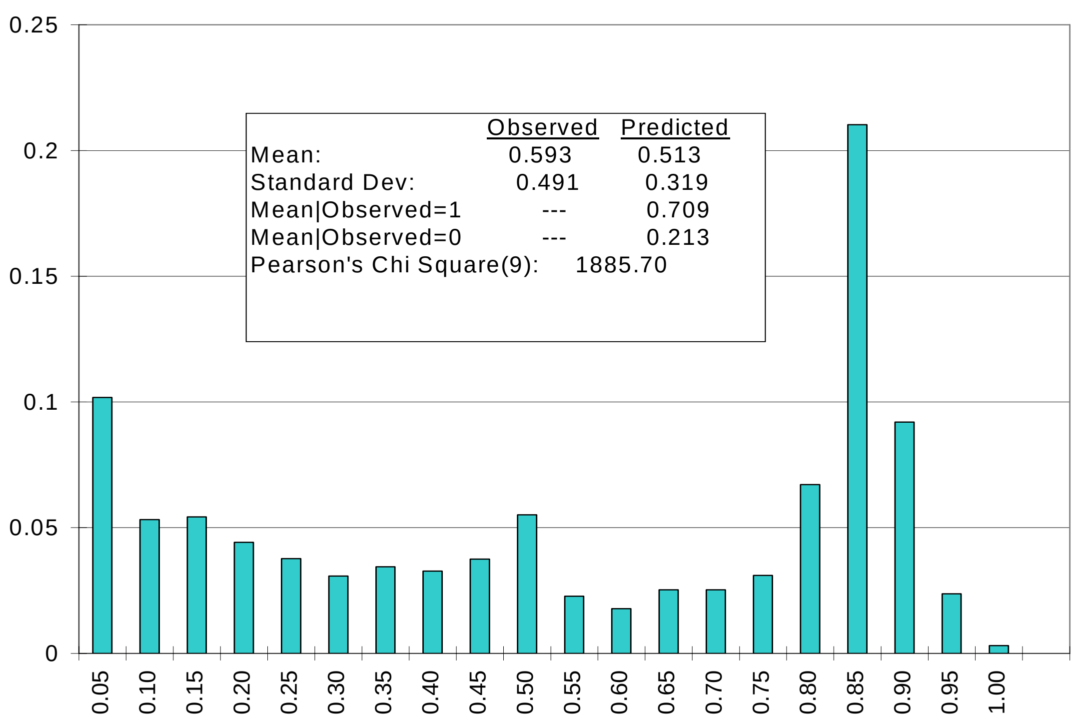

Figure 17 Distribution of Change in Medigap Premium

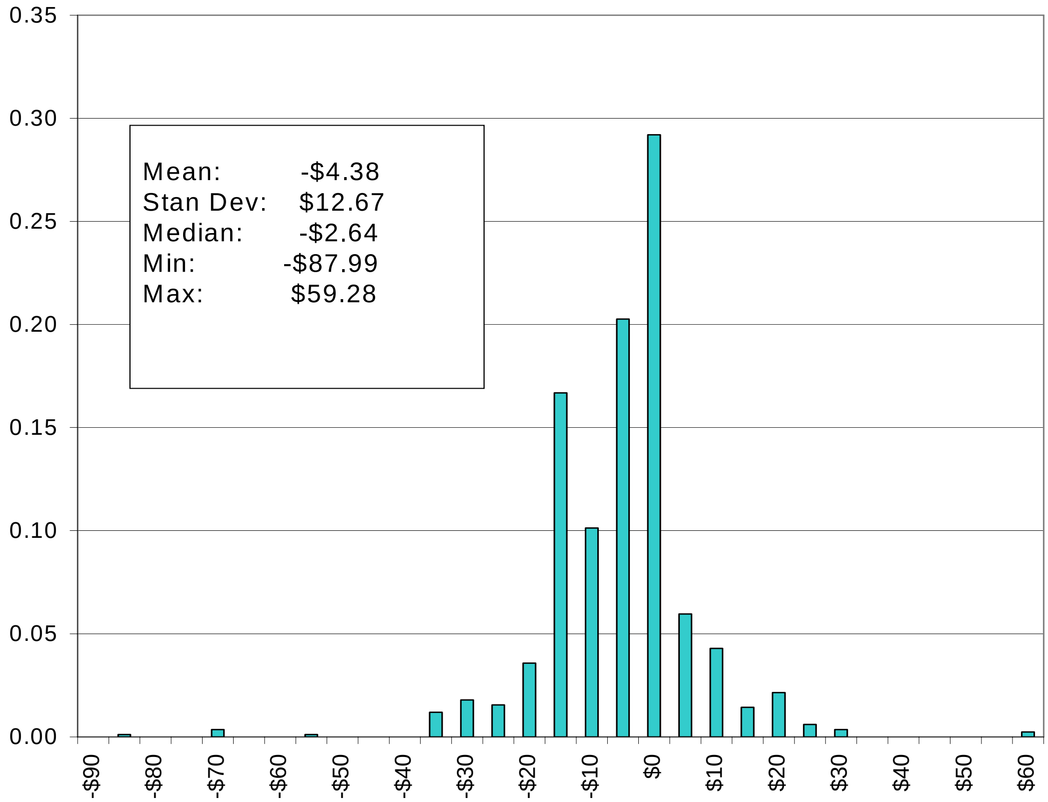

Figure 18 CDFs of Change in Consumer Surplus, Men

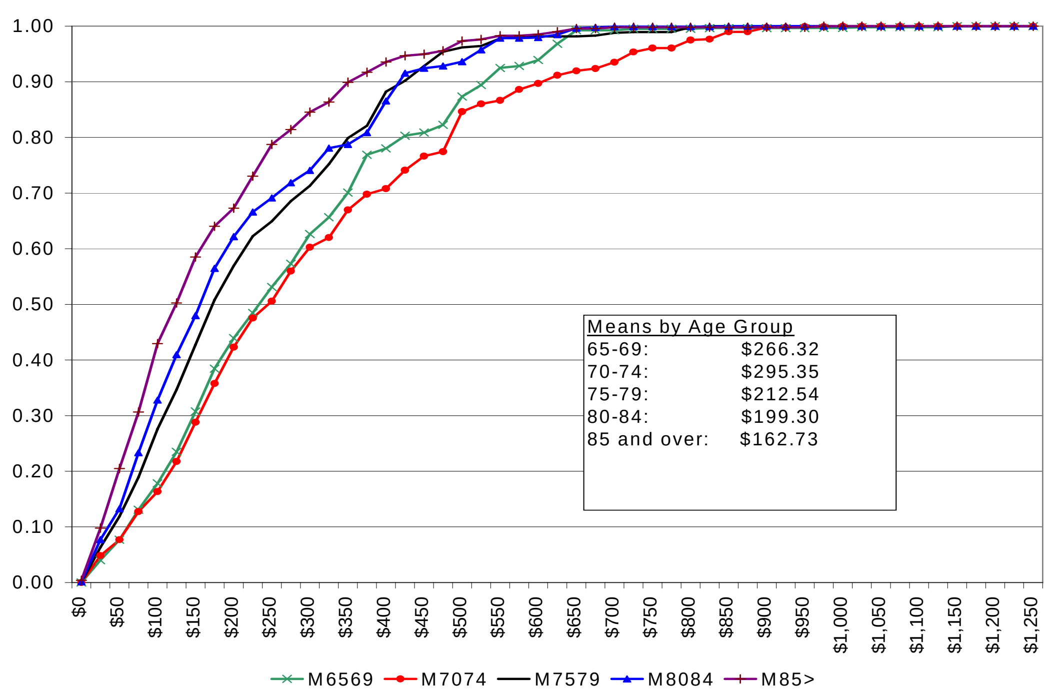

Figure 19 CDFs of Change in Consumer Surplus, Women

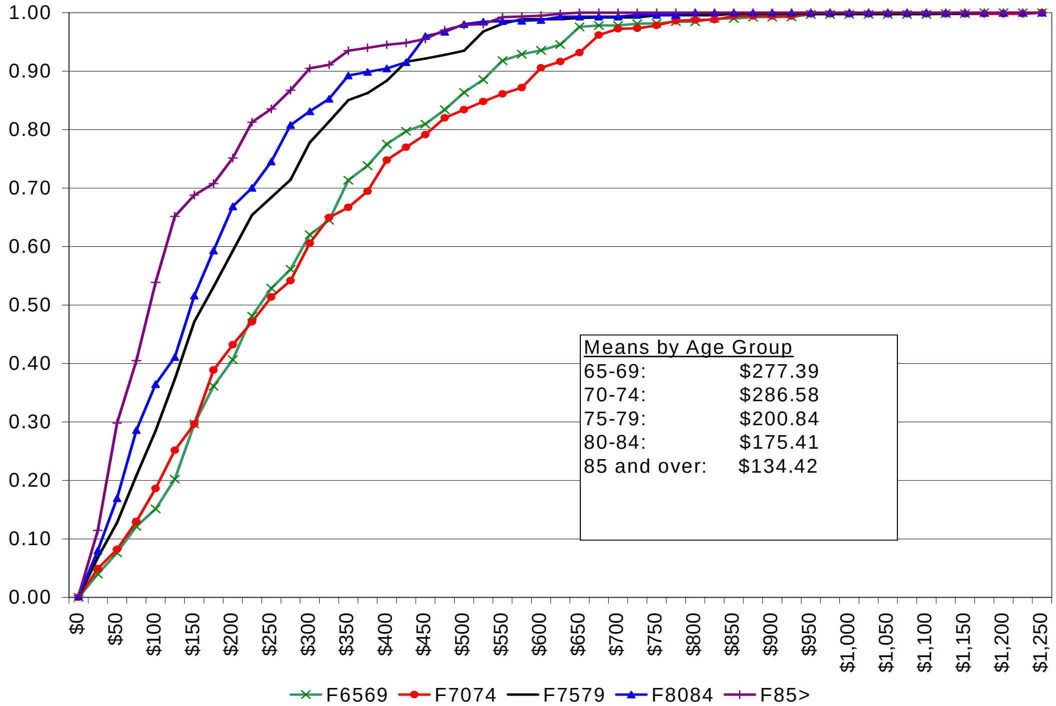

Figure 20 Indirect Utility and Optimal Quantity on Pretreatment State for $g(\theta)=\theta$

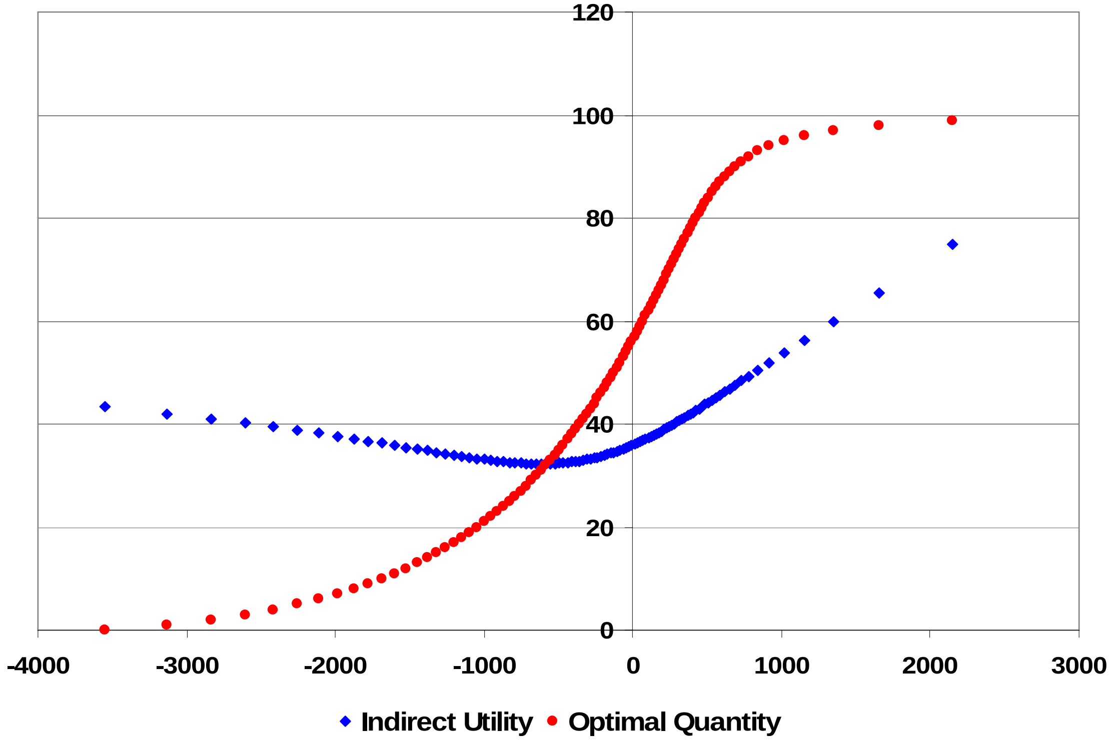

Figure 21 $\dfrac{\partial H^*}{\partial \theta}$ $for\ \ g(\theta)=\theta$

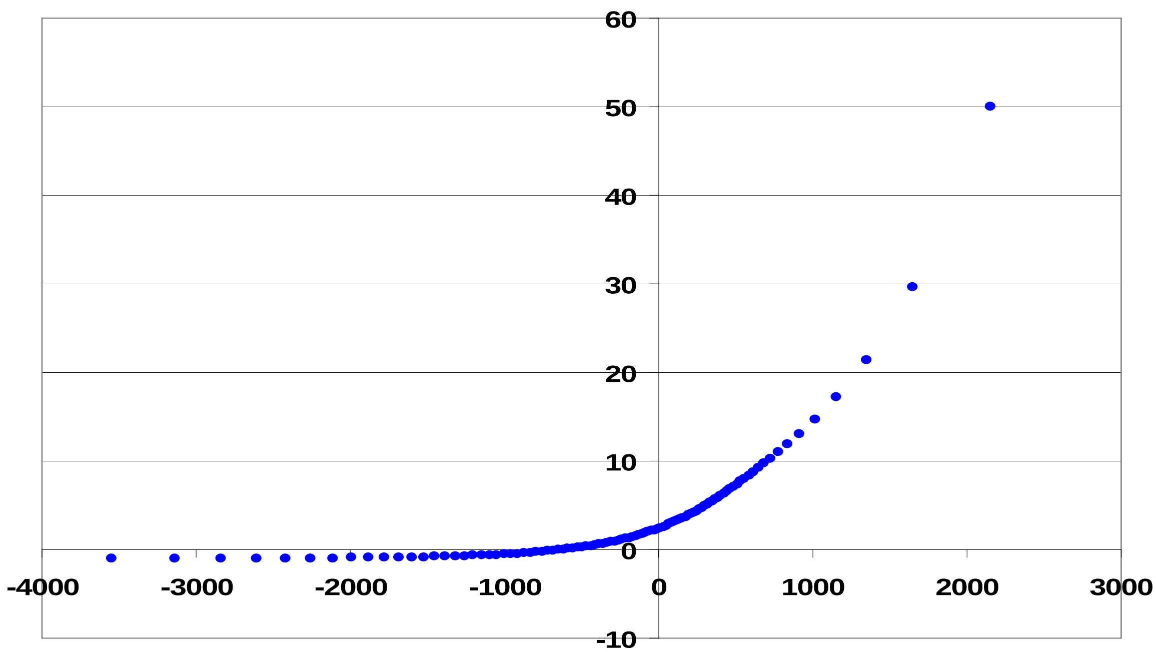

Figure 22 Indirect Utility and Optimal Quantity on Pretreatment State for

$$g(\theta) = \left(1 + \exp\left(-\frac{\theta - a}{b}\right)\right)^{-1}$$

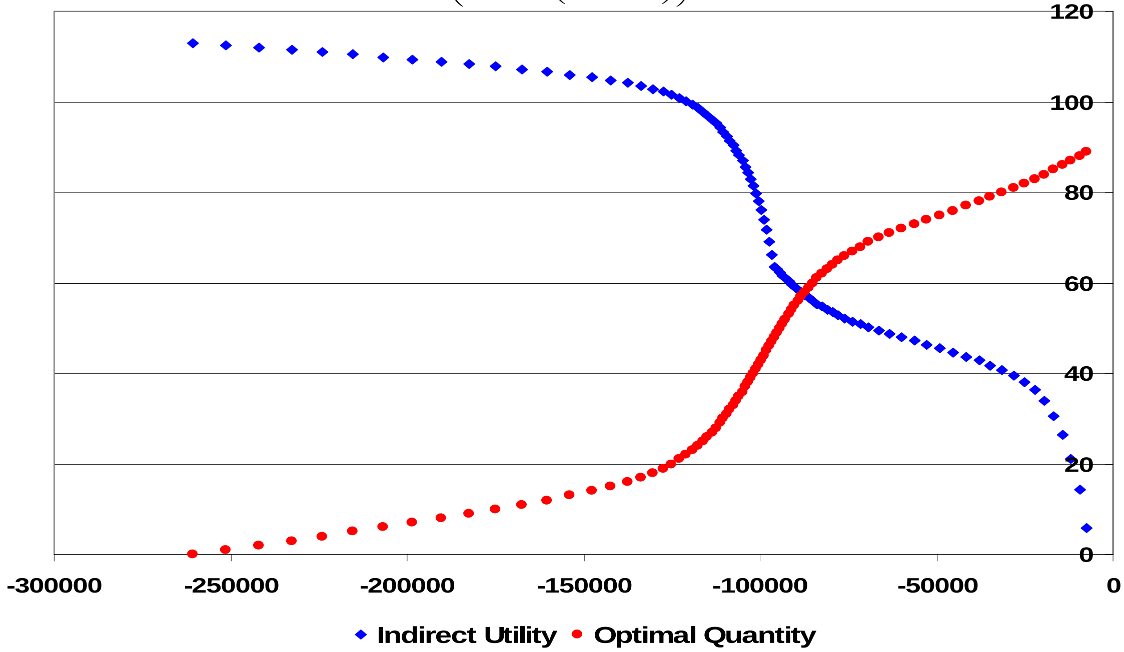

Figure 23 $\dfrac{\partial H^*}{\partial \theta}$ *for* $g(\theta) = \left(1 + \exp\left(-\frac{\theta - a}{b}\right)\right)^{-1}$