

The Welfare Effects of Adverse Selection in Privatized Medicare*

Joshua Lustig[†]

Yale University

Department of Economics

Version: December 3, 2007

Job Market Paper

Abstract

This paper estimates the welfare losses from market failures caused by adverse selection in privatized Medicare. I model insurers' premium and coverage choices in an environment where consumers have heterogeneous preferences and may impose different costs on their insurers. The model generates predictions about insurers' costs and behavior under varying degrees of adverse selection. I use the model and exogenous variation in market structure to identify a causal link between consumers' types and insurers' costs. From the estimated parameters, I can infer whether consumers' preferences, which determine how much insurance they purchase, contain information about their expected health. The empirical results imply that adverse selection is indeed present in privatized Medicare. It is more costly to insure consumers with strong preferences for health insurance. With the estimated model, I simulate new equilibria after removing the distortionary effects of adverse selection from insurers' costs and incentives. The new equilibria exhibit more generous insurance coverage and lower premiums. These effects are particularly strong in markets with many insurers. The total surplus associated with privatized Medicare increases by 16.2%, suggesting that the welfare losses from adverse selection are substantial.

*I am indebted to my advisors, Steven Berry and Philip Haile for their continual guidance, support, and encouragement. I also benefited from constructive conversations with Fiona Scott-Morton, and my colleagues Jacob Gramlich, Katrina Jessoe, Ying Fan, Allesandro Bonatti, and Maher Said. Special thanks also go to Lyle Nelson at the Congressional Budget Office for greatly improving my understanding of Medicare.

[†]Department of Economics, Yale University, 28 Hillhouse Ave., New Haven, CT 06511; joshua.lustig@yale.edu; <http://pantheon.yale.edu/~jdl57/>

1 Introduction

Since Akerlof (1970) and Rothschild and Stiglitz (1976) first formalized the theory of adverse selection, theorists have emphasized the market failures that adverse selection can cause in insurance markets. Subsequent to these theoretical contributions, an empirical literature emerged that tests for adverse selection using consumers' observed insurance choices and their risk outcomes.¹ Recently, methods to detect adverse selection and distinguish it from other information asymmetries have grown increasingly sophisticated.² Despite these advances, the current literature remains unable to quantify the market failures emphasized in the theory.³ Market failures from adverse selection arise from distorted insurer behavior, and thus, measuring them necessitates a model of insurer behavior. This paper makes two contributions to the adverse selection literature: first, I provide evidence of adverse selection in privatized Medicare, and more importantly, I estimate the welfare losses from the resulting market failures.

I estimate a model in which insurers choose how much coverage to offer and what premiums to charge under varying degrees of adverse selection. Using insurers' optimal premium and coverage choices, I calculate insurers' costs of providing insurance coverage. To measure adverse selection, I relate these costs to enrollees' preferences for insurance, which are hidden from insurers but can be inferred using a model of consumer sorting. The exercise enables me to detect the presence of adverse selection. Specifically, do consumers' preferences for insurance, which determine how much health insurance they purchase, contain information about their expected health?

Estimates of insurers' costs confirm that consumers with strong preferences for generous insurance are more costly to insure. Market failures arise from this adverse selection because insurers are unable to charge different premiums to enrollees with different unobserved health. Inefficient pricing leads some consumers to purchase suboptimal levels of insurance to avoid subsidizing the unhealthy.⁴ Insurers have similar incentives to avoid the unhealthy, and may distort their plans away from generous coverage to deter their enrollment. To quantify these distortions, I remove the effects of adverse selection from insurers' costs and use the estimated model to simulate how insurers' plans and consumers' choices (and therefore, welfare) change in equilibrium. I find the effects of adverse selection to be substantial; after its removal total surplus increases by 16.2%.

I apply my model to an insurance market where policy makers are increasingly introducing competition: the market for the provision of managed care options to Medicare beneficiaries. Between 2000 and 2003,

¹ Standard models of insurance markets predict a positive correlation between insurance coverage and risk outcomes that can be tested with consumer data. See for example, Chiappori and Salanie (2000) and Cutler and Zeckhauser (2000). The latter documents a substantial literature confirming this positive correlation in health insurance markets.

²Finkelstein and Poterba (2006) test for adverse selection using consumer observables that *determine* insurance coverage but are not used by insurers when setting prices. Cardon and Hendel (2001) use a structural model of consumer behavior to separately identify moral hazard and adverse selection.

³One exception to this is Einav, Finkelstein, and Schrimpf (2007). This paper estimates the welfare costs of adverse selection in a UK annuity market.

⁴If adverse selection is severe, insurers may be unable to simultaneously offer generous insurance and attract healthy (i.e. profitable) enrollees. Markets for generous insurance may unravel. Cutler and Reber (1998) provide evidence of such an unravelling in a Boston insurance market.

the Medicare + Choice program (M + C) allowed Medicare enrollees to opt out of traditional, fee-for-service Medicare and enroll in plans offered by private insurers.⁵ Participating insurers, mostly Health Maintenance Organizations (HMOs), agreed to provide the same coverage as traditional Medicare in exchange for payments from the federal government. Insurers could also offer coverage not included in traditional Medicare, such as prescription drug coverage, and charge enrollees a monthly premium. Underlying the M + C program was the belief that competing HMOs could more efficiently provide insurance coverage to the nation's elderly.

The M + C market is one with endogenously differentiated products. HMOs choose how many plans to offer, how much insurance coverage to provide, and what premiums to charge. I assume HMOs' decisions to offer prescription drug insurance and other forms of coverage are made simultaneously in a static setting characterized by a Nash equilibrium. I model consumers' enrollment decisions with a discrete choice framework. Consumer sorting between HMO plans and traditional Medicare depends on a distribution of preferences for insurance as well as consumers' choice sets. The model's structure allows me to infer insurers' costs and to make predictions about consumer sorting across insurance plans. These form the basis for my measure of adverse selection.

To measure adverse selection, I estimate a causal relationship between consumers' preferences (implied by their enrollment decisions) and insurers' costs (implied by their coverage and premium choices). If attracting consumers with strong preferences for insurance coverage *causes* insurers to have high costs, then consumers with strong preferences must tend to have poor expected health. If consumers' preferences depend on characteristics such as risk aversion, in addition to unobserved health, my model could estimate a negative relationship between insurers' costs and consumers' preferences.⁶

My estimation strategy is complicated by unobserved factors that also affect insurers' costs, including moral hazard. Consider the set of HMOs that offer generous prescription drug coverage. These insurers will attract many consumers with strong preferences for insurance, who may or may not have poor expected health. If moral hazard is present, hidden action will make these consumers appear costly, independent of adverse selection. Similarly, insurers' implied costs become uninformative about adverse selection if unobserved costs influence insurers' coverage decisions. Insurers offering generous prescription drug coverage, for example, may have advantageous relationships with drug companies that I am unable to observe.

Thus, my strategy to identify and measure adverse selection exploits a relationship between adverse selection, insurers' costs, and variation in market structure. If and only if adverse selection exists, insurers that enroll many consumers with weak preferences for insurance (and therefore, good expected health) will have lower average costs. Markets with distinct structures provide insurers different opportunities to attract these consumers. In markets with few insurers, for example, HMOs offering generous insurance coverage attract more consumers with weak preferences for generous insurance. In markets with more insurers, however, the product space becomes saturated and these consumers are lost to plans that are less generous

⁵Medicare is a federal program that provides health insurance to the elderly and disabled.

⁶Fang, Keane, and Silverman (2004) and Finkelstein and McGarry (2006) find evidence of adverse selection in unobserved health, but advantageous selection in other characteristics, such as risk aversion, income, and cognitive ability. Advantageous selection exists when consumers enrolling in generous insurance have good unobserved health.

and less expensive. I use exogenous variation in market structure and insurers' characteristics to generate observable variation in sorting by consumers with weak and strong preferences for insurance. My measure of adverse selection is identified using only the portions of insurers' costs that are explained by this variation in preference-based sorting.⁷

Estimates of HMOs' cost functions provide evidence of adverse selection. An HMO's variable costs of coverage are 7.4% higher for a consumer whose preferences imply a willingness to pay for insurance that is one standard deviation above the median. Below, I use my model of insurer and consumer behavior to quantify the distortions that result from this difference in costs. I remove adverse selection from privatized Medicare by simulating policies that implement perfect risk adjustment. Government payments to each insurer are adjusted to reflect the characteristics of enrollees that they attract. Insurers that attract costly enrollees are compensated with higher payments, and insurers that attract inexpensive enrollees receive lower payments. This policy removes the distortionary effects of adverse selection on insurers' costs (and therefore, incentives), and is an extreme version of policies the government has experimented with in privatized Medicare in recent years.

Removing adverse selection requires government expenditures on privatized Medicare to increase by 1.7%. But I find that its removal induces insurers to expand their insurance coverage and reduce the premiums they charge, particularly for the most generous plans. Consumer surplus and insurer profits both increase substantially, offsetting the increase in government expenditures. Consumer surplus increases from expanded insurance coverage and lower premiums, and insurer profits, from expanded consumer participation in privatized Medicare and more generous government payment rates. The total surplus associated with privatized Medicare increases by 16.2% relative to an equilibrium with adverse selection. Equivalently, surplus increases by an amount equal to 1.1% of the government's total payments to insurers in the initial equilibrium. These surplus gains are concentrated in markets with many HMOs. Where one HMO operates, total surplus increases by \$6.46 per Medicare beneficiary per year. In markets with six or more HMOs, total surplus increases by \$29.19 per Medicare beneficiary per year. These results suggest that the government's recent attempts to implement risk adjustment in privatized Medicare are welfare enhancing.

The remainder of the paper proceeds as follows. Section 2 describes the M + C program's relevant institutional details and Section 3, the data sources that I exploit. Section 4 describes my model of consumer and insurer behavior in privatized Medicare and Section 5 discusses my estimation strategy. Section 6 provides a series of tables describing privatized Medicare between 2000 and 2003. Sections 7 and 8 present the parameter results and counterfactual simulations. Section 9 concludes.

⁷Moral hazard takes place after contracting is completed. Thus, an insurer's costs that are attributable to moral hazard do not depend on market structure. They depend only on the insurer's plans' characteristics. Similar intuition was used by Cardon and Hendel (2001) to distinguish between adverse selection and moral hazard.

2 Medicare + Choice Program

Introduced in 1965, Medicare is the primary form of health insurance for the elderly and disabled. It is presently one of the federal government's largest programs and constitutes a large portion of total health care spending.⁸

In 1982, Congress passed the Tax Equity and Fiscal Responsibility Act, mandating the provision of managed care options to Medicare beneficiaries.⁹ Since then, private insurers have played a continuous role in Medicare. This paper studies the years between 2000 and 2003 when the program governing privatized Medicare was called Medicare + Choice and HMOs were the dominant firm type.^{10,11} Under M + C, Medicare beneficiaries could opt out of traditional Medicare and receive health insurance from a qualified private insurer. Insurers wishing to enroll Medicare beneficiaries signed contracts with the Center for Medicare and Medicaid Services (CMS) describing what coverage they would provide, and at what costs. A minimum set of benefits were required, essentially equal to the coverage included in traditional fee-for-service Medicare.¹² In exchange, the CMS made per-capita payments to each insurer under contract. Insurers had the option of providing additional benefits, such as (but not limited to) prescription drug, dental and vision coverage, as well as preventative care, in exchange for monthly premiums paid by enrollees.

Under M + C, it was believed that competitive pressures would induce insurers to submit proposed contracts that would achieve cost savings for the Medicare program. In addition, Medicare beneficiaries would benefit from coverage more generous than traditional Medicare.¹³ This logic extended even to markets with few participating insurers, where insurers' plans still faced competition from traditional Medicare. Most participating insurers did offer plans with coverage not provided by traditional Medicare. In markets with one insurer, for example, 61% of insurers offered some supplemental prescription drug coverage between 2000 and 2003.

Until 2000, government payments to insurers were set equal to 95% of the expected cost of treating a beneficiary within traditional Medicare.¹⁴ Policy makers were aware of potential selection problems within privatized Medicare and between traditional and privatized Medicare. Between 2000 and 2003, the CMS began to experiment with expanded risk adjustment. The Balanced Budget Act of 1997 (BBA) dictated that payment rates to insurers reflect enrollees' lagged inpatient hospital experiences. During the sample

⁸Spending on Medicare in 2003 totaled \$315 billion and in 2002, Medicare accounted for 19% of total spending on personal health care and 2.6% of GDP (Medpac 2004).

⁹HMO participation in Medicare began in 1972. Participation, however, was minimal until TEFRA.

¹⁰Participation by Preferred Provider Organizations (PPOs) was beginning to develop and Private Fee for Service insurers (PFFS) were still in their infancy.

¹¹The 2003 Medicare Modernization Act redefined and expanded the role played by private insurers, relabeling the program "Medicare Advantage" and adding a prescription drug component to the program.

¹²Traditional Medicare is comprised of two parts, A and B. Part A covers hospital services and enrollment is automatic. Part B covers outpatient services. Enrollment is not automatic and requires a monthly payment by enrollees, but it is heavily subsidized by general tax revenue. Virtually all Medicare beneficiaries enroll in Part B.

¹³"Over time, participating plans will be under competitive pressure to improve their benefits reduce their premiums and cost sharing, and improve their networks and services, in order to gain or retain market share" (Medicare Managed Care Manual, CMS)

¹⁴Adjustments to payments were made according to enrollees' age, gender, and eligibility status. These adjustments accounted for little of the variation in realized costs. Most studies estimated that this crude system of risk adjustment accounted for only 1-2% of the variation in realized costs (Medpac 2000).

period, payments were weighted averages of the pre-BBA rate (90%) and the newer model (10%).¹⁵

To redistribute HMO participation geographically, the BBA reduced linkages between payment rates and the costs of care in traditional Medicare. Pre-BBA payment rates varied considerably across different portions of the country due to variation in demographics and medical provider characteristics. As a result, some markets were heavily underserved by Medicare HMOs. The BBA reduced this geographical variation in payment rates to increase HMO participation in rural areas.¹⁶

Contracts between the CMS and insurers were determined on a county and yearly basis after CMS payment rates were announced in each county. Once contracting between insurers and the CMS was completed, Medicare beneficiaries chose to remain in the traditional Medicare program or enroll in a private plan offered in their county. HMOs were not allowed to discriminate between beneficiaries within a market. Rather, insurers had to offer the same menu of plans to all individuals residing in the same county. Nor were they allowed to terminate a plan, increase premiums, or reduce coverage within a calendar year.¹⁷ Most beneficiaries that remained in traditional Medicare also enrolled in the voluntary part B component, and many enrolled in a supplementary Medigap plan.

3 Data Sources

This paper combines data from multiple sources. Insurers' plan characteristics (and beneficiaries' choice sets) are retrieved from the Medicare Compare databases for the years 2000-2003. Market share data and individuals' plan choices are taken from the CMS State-County-Plan (SCP) files and the Medicare Current Beneficiary Survey (MCBS). HMO operating and financial characteristics are retrieved from the Weiss Ratings Guide to HMOs and Health Insurers. I extract market characteristics from the US Census and the American Hospital Association Directory. Variable definitions and construction of the final data set are discussed in the appendix.

3.1 Medicare Compare Databases

The Medicare Compare Database is released each year to inform Medicare beneficiaries which private insurers are operating in their county, what plans they offer, and what benefits and costs are associated with each plan. For each plan, I collect information on dental coverage, vision coverage, brand and generic prescription drug coverage, and the copayments associated with prescription drugs, primary care doctor and specialist visits, and inpatient hospital admissions.

¹⁵The newer PIP-DCG risk adjustment model, which reflects inpatient hospital experiences, has been estimated to account for only 5-6% of variation in total health care costs (Medpac 2000).

¹⁶Payment rates were set to the maximum of a national floor rate, an updated pre-BBA rate, and a third blended rate. (Medpac, 2001)

¹⁷They could, however, increase coverage or reduce premiums.

3.2 CMS State-County-Plan Files

The SCP files provide market share data at the insurer-county-year level for all insurers in all counties. In each county-year, the SCP files also inform me how many residents are eligible for Medicare and the average CMS payment rate. Unfortunately, enrollment information is not provided at the plan level. For insurers offering multiple plans in a county, the market shares sum across the individual plan shares.

The SCP exaggerates entry because Medicare beneficiaries are not required to change insurers after moving across counties. In each county, the SCP lists all insurers with at least one enrollee, even those not under contract with the CMS. To alleviate this problem, I eliminate all insurers with less than a 1% county market share. In addition, to be included as a participant in a county, an insurer must be listed in the Medicare Compare Database.

3.3 MCBS Survey

The Medicare Current Beneficiary Survey tracks the behavior of a representative national sample of the Medicare population. Using the 2000-2003 versions of the MCBS, I match a sample of Medicare beneficiaries, many observed in multiple years, to specific insurance plans. The MCBS helps me overcome the lack of plan market share data in the SCP files.

MCBS administrative data informs me which M + C insurer each respondent is enrolled in. The Health Insurance section asks respondents specific questions about their Medicare managed care plan: whether dental, vision, or prescription drug coverage is provided, and what premiums are charged. Using answers to these questions, I identify which plan each respondent is enrolled in. The data set also informs me which county each respondent lives in. With the Medicare Compare Database, I construct choice sets for each MCBS respondent.

3.4 Weiss Ratings' Guides

Many types of insurers participate in M + C. Large, national companies such as Blue Cross and Blue Shield and Aetna offer plans in markets across the country. Smaller, regional HMOs offer plans in a few contiguous counties. To capture this heterogeneity, firm characteristics are extracted from the Weiss Ratings' Guides.

On a quarterly basis, the Weiss guides accumulate financial information from approximately 1500 United States HMOs and health insurers. This financial information includes total assets, capital, net income, etc. For all rated United States HMOs, Blue Cross and Blue Shield plans, and large health insurers, the Weiss guide gives information on administrative expenses, enrollment levels, average medical expenses per enrollee, and physician network sizes.

3.5 Additional Data Sources

Counties with distinct characteristics may impose different costs on insurers, and contain Medicare beneficiaries with distinct preferences. For each county, I retrieve population, income, and educational attainment data from the US Census. From the American Hospital Association Directory, I determine how many hospitals operate in each county.

4 Model

Insurers make entry decisions on a county and yearly basis and Medicare beneficiaries' choice sets include the plans available in the county they reside. On this basis, a market is defined as a county-year. The model below describes one market. Insurers decide whether to offer plans to Medicare beneficiaries, what coverage to provide, and what premiums to charge. Medicare beneficiaries choose between the plans offered and traditional Medicare.

4.1 Demand

Assume that consumer i , living in market m , obtains indirect utility from plan k offered by insurer j as follows:

$$u_{ijk} = (1 + v_{gi}) * g_{jk} + (1 + \theta_p D_m + v_{pi}) * \alpha p_{jk} + \xi_j^m + \varepsilon_{ijk}^m \quad (1)$$

$$u_{i0} = \varepsilon_{i0}$$

where

$$g_{jk} = x'_{jk} * \beta$$

$$x_{jk}, p_{jk} = jk's \text{ observable characteristics}$$

$$D_m = \text{market } m's \text{ demographic characteristics}$$

$$\xi_j^m = j's \text{ unobservable characteristics}$$

$$\varepsilon_{ijk} = \text{idiosyncratic } ijk \text{ horizontal preferences}$$

$$v_i \sim |Normal(0, \sigma^2)| = i's \text{ unobserved vertical preferences}$$

$\theta_p, \alpha, \xi, \beta$ and σ are the parameters to be estimated.

Each plan has observable characteristics x_{jk} and p_{jk} . p is the monthly premium. x describes $jk's$ coverage and cost sharing requirements. It includes doctor and hospital copayments and variables describing prescription drug, vision and dental coverage. The characteristics of each plan are collapsed into an index of overall generosity, $g_{jk} = x'_{jk} * \beta$.

Medicare beneficiaries have heterogeneous preferences for insurance generosity and premiums, determined by the absolute values of draws from a normal distribution and the demographic characteristics of market m .¹⁸ Individuals with large draws of v_g gain more utility from plans with generous coverage and are more likely to enroll in a high g plan than individuals with low v_g . Draws of v are positive to ensure that the marginal utility of insurance coverage and disutility of premiums are positive. The amount of heterogeneity in preferences is determined by the parameters σ_g^2 and σ_p^2 . Below, I discuss this heterogeneity further.

Consumers' horizontal preferences vary at the individual-insurer-plan and insurer-market levels. For the outside good and each plan jk , i receives idiosyncratic type-1 extreme value shocks to utility, ε_{ijk} . Each insurer j has a quality level ξ_j^m in market m that provides utility to all Medicare beneficiaries. It can contain information the econometrician does not observe (for example, physician network quality or insurers' experiences in particular markets) and information the econometrician does observe (insurer type). A regression of ξ_j^m on the latter class of data can suggest which variables are correlated with high utility provision.

4.2 Supply

Insurer j expects to incur a cost $c_{jk}(v_g, g, \Theta^{supply})$ from providing coverage with generosity g to a Medicare beneficiary with preferences v_g . There is a fixed cost per enrollee, $FC_{jk}(v_g, \Theta^{supply})$, that does not depend on g . The remaining variable costs of coverage, $MC_{jk}(v_g, \Theta^{supply}) * g$, are linear in g :

$$\begin{aligned} c_{jk}(v_g, g, \Theta^{supply}) &= FC_{jk}(v_g, \Theta^{supply}) + MC_{jk}(v_g, \Theta^{supply}) * g & (2) \\ &= (\pi_0 + \gamma_0 v_g + \delta'_0 X_m + \lambda'_0 Z_j + \psi_{jk}^0) + \\ &\quad (\pi_1 + \gamma_1 v_g + \delta'_1 X_m + \lambda'_1 Z_j + \psi_{jk}^1) * g \end{aligned}$$

where

$$\begin{aligned} g &= \text{amount of insurance coverage (generosity)} \\ X_m &= \text{market } m\text{'s characteristics} \\ Z_j &= \text{insurer } j\text{'s characteristics} \\ \psi_{jk}^0, \psi_{jk}^1 &= \text{cost shocks to plan } jk \\ v_g &= \text{consumers' preferences for } g \end{aligned}$$

π , γ , δ , and λ are parameters to be estimated.

This cost function nests adverse and advantageous selection. If consumers with strong preferences for

¹⁸ Assuming that i 's vertical preferences depend on observable market, rather than individual, characteristics is an obvious simplification. I make it because it is difficult to collect distributions of Medicare beneficiaries' observable characteristics at the county level. County distributions are necessary because insurers' choices are made at the county level. In principle, however, collecting this data is possible.

insurance coverage (i.e., those most likely to enroll in generous plans) are more costly to insure (adverse selection), γ_0 and γ_1 are positive. If there is a negative relationship between preferences and costs (advantageous selection), γ_0 and γ_1 are negative. The signs and magnitudes of γ_0 and γ_1 play an important role in determining the effects of selection in equilibrium.

Each insurer realizes two shocks to costs, ψ_{jk}^0 and ψ_{jk}^1 for each plan they offer. The shocks to FC and MC are perfectly observed by insurers, but not the econometrician, and are drawn from separate distributions that are *iid* and mean zero across insurers and markets.¹⁹

Finally, FC and MC depend on exogenous market and insurer characteristics, X_m and Z_j . X_m controls for cost differences across markets and includes information such as average income, population, etc. In wealthier markets, for example, insurers may need to pay higher wages to local employees. X_m also includes the number of hospitals in m , yearly dummy variables, current CMS payment rates, as well as the pre-BBA payment rate. Prior to 1997, this rate perfectly reflected the government's costs of providing insurance to beneficiaries in traditional Medicare.

Z_j includes insurers' financial and operating characteristics, such as lagged income and asset levels, administration costs, firm type, etc. Z controls for the possibility that heterogeneous insurers draw cost shocks from distributions with different means. For example, large insurers may be able to more effectively negotiate with pharmaceutical companies and therefore, have advantages offering generous drug coverage. Similarly, insurers with large physician networks may charge enrollees lower copayments for doctor office visits.

4.3 The Order of Decisions

I describe the game that insurers and Medicare beneficiaries play in four stages:

- Stage 1: Insurers decide whether to enter each market m , and if so, how many plans to offer. This decision conditions on observable market characteristics and the observable characteristics of all potential entrants. ξ is included in this set.
- Stage 2: After each insurer j commits to offer n_j^m plans in market m , cost shocks ψ_{jk}^0 and ψ_{jk}^1 are realized for each plan k .
- Stage 3: After observing competitors' entry decisions and ψ_{jk}^0 and ψ_{jk}^1 for each jk , each insurer simultaneously chooses g_{jk} and p_{jk} for each plan k . Insurer j chooses plan characteristics and premiums to maximize profits in market m .
- Stage 4: Consumers observe g and p for all plans. They choose to enroll in a private insurer's plan or remain in traditional Medicare.

¹⁹The mean zero assumption is without loss of generality.

Let $MS_{jk}(g, p, v \mid G, P, \Theta^{Demand})$ equal the percentage of Medicare beneficiaries with preferences, v_g , who enroll in plan jk , conditional on the full set of plans offered, (G, P) . S_m is the per-capita payment made by the government to insurers in market m . After each insurer chooses g and p and enrollment decisions are made, j earns profits:

$$\int_v \sum_{k=1}^{n_j} MS_{jk}(g, p, v \mid G, P, \Theta^{Demand}) * [S_m + p_{jk} - c_{jk}(v, g, \Theta^{Supply})] dF(v \mid \sigma)$$

I only make assumptions about the timing of insurers' decisions in stages one and two. I assume the following, however, about insurer behavior in stage three:

Assumption: Conditional on the outcomes of stages one and two, all participating insurers simultaneously choose g_{jk} and p_{jk} for $k = 1, \dots, n_j^m$. Choices satisfy a Nash equilibrium, which is assumed to exist.

Characterizing plan characteristics and premiums with a Nash equilibrium condition allows me to exploit insurers' first order conditions for g and p in the estimation routine.

Similarly, the game's four stage structure implies a set of moment conditions useful in estimation. In particular, Stage 2 implies that conditional on X and Z , the cost shocks, ψ^0 and ψ^1 , are independent of the number of plans and insurers in each market. It is important, therefore, that X and Z adequately control for differences in costs across insurers and markets.

4.4 Discussion

My model allows consumers only one dimension of preferences for insurance coverage. If there are two consumers, i and i' , with $v_{gi} > v_{gi'}$, then i has stronger preferences than i' for prescription drug and dental coverage. A more general environment might allow i to have stronger preferences for prescription drugs and i' to have stronger preferences for dental. I impose this restriction to simplify my analysis of insurer behavior. The restriction implies that the elements of x are perfect substitutes in each plan's market share function, since all consumers' willingness to trade between the elements of x depends on the same linear equation in β . In the appendix, I prove it is without additional loss of generality to assume insurers choose g , rather than x , for each plan.

A recent empirical literature emphasizes multiple dimensions of consumer heterogeneity, with adverse selection along some dimensions and advantageous selection along others. For example, risk aversion and cognitive ability, in addition to unobserved health, may determine preferences for insurance.²⁰ If consumers have T characteristics that determine their preferences (i.e. $v_g \equiv v_g(t_{i1}, t_{i2}, \dots, t_{iT})$), the elements of t may have different effects on insurers' costs. For example, t_{i1} might capture i 's unobserved health and t_{i2} , i 's risk aversion, which might be correlated with good health. In the appendix, I prove that under certain distributional assumptions on t and v , my model is equivalent to one with costs that depend on t (i.e.

²⁰See Fang, Keane, and Silverman (2006); Finkelstein and McGarry (2006)

$c(g, t) = (\gamma'_0 * t + \psi^0) + (\gamma'_1 * t + \psi^1) * g$. This result also follows from consumers having one dimension of preferences for insurance coverage.²¹

5 Estimation

Let Θ_0 denote the true set of parameters. They are estimated using a simulated method of moments framework. Estimation is done in one step, but below, I discuss supply and demand separately. Details are left for the appendix.

5.1 Demand

Let Ω_m denote the set of private insurers offering plans in market m , and Ω_{mj} the set of plans offered by insurer j . The probability that an individual enrolls in plan jk in market m is given by:

$$\begin{aligned} Prob(jk|m, \Theta) &= \int_v \frac{e^{(1+v_g)*g_{jk} - (1+\theta_p D_m + v_p)*\alpha p_{jk} + \xi_j^m}}{1 + \sum_{j' \in \Omega_m} \sum_{k' \in \Omega_{mj'}} e^{(1+v_g)*g_{j'k'} - (1+\theta_p D_m + v_p)*\alpha p_{j'k'} + \xi_{j'}^m}} dF(v|\sigma) & (3) \\ &= \int_v \frac{e^{u_{ijk}(v) + \xi_j^m}}{1 + \sum_{j' \in \Omega_m} \sum_{k' \in \Omega_{mj'}} e^{u_{ij'k'}(v) + \xi_{j'}^m}} dF(v|\sigma) \\ &= \int_v Prob(jk|v, m, \Theta) dF(v|\sigma) \end{aligned}$$

In many discrete choice models, equation (3) and market share data are used to identify ξ . Remaining parameters are then identified by assuming observed product characteristics are mean independent of ξ .²² This assumption is inappropriate here, because insurers observe ξ before choosing premiums and plan characteristics. Instead, my estimation strategy relies on two restrictions:

1. There is no unobserved quality at the plan level, i.e., $\xi_{jk}^m = \xi_j^m \forall jk$ and m .
2. Individuals' vertical preferences, v_{gi} and v_{pi} , do not change over time.

The first restriction allows me to identify θ , β , and α . The second restriction allows me to identify σ . Below I discuss identification of ξ . Then I describe how these two restrictions are used to identify θ , β , α , and σ .

²¹Since insurers can not discriminate against consumers with the same v , but distinct t , insurers' relevant cost functions reduce to $E[c_{jk}(t, g, \Theta^{Supply}) | v_g(t)]$. This expected cost is equal to $c_{jk}(v_g(t_{i1}, t_{i2}, \dots, t_{in}), g, \Theta^{Supply})$ if $\frac{dE[t_m|v]}{dv} = \Delta_m$ for all v and m .

²²See Berry, Levinsohn, and Pakes (2003). Other studies of the Medicare HMO market, such as Town & Liu (2003) and Maruyama (2006) make similar assumptions.

5.1.1 Identifying ξ

The SCP files provide market share data at the insurer-market level. Let s_j^m equal insurer j 's market share in m . Following Berry (1994), for any θ , β , α , and σ , there exists a unique ξ such that:²³

$$s_j^m = \sum_{k \in \Omega_{mj}} \text{Prob}(jk|m, \Theta) \quad \forall j \text{ and } m$$

5.1.2 Identifying θ , β , and α

If insurer j offers only one plan, then for any θ , β , and α , there is a ξ_j such that j 's predicted market share equals his observed market share. Thus, to identify θ , β , and α , I use the model to predict the choices made by MCBS respondents enrolled with insurers that offer more than one plan. The probability that an individual enrolls in plan jk , conditional on enrolling with insurer j in market m , is given by:

$$\text{Prob}(k|j, m, \Theta) = \int_v \frac{\text{Prob}(jk|v, m, \Theta)}{\sum_{k \in \Omega_{mj}} \text{Prob}(jk|v, m, \Theta)} dF(v|\sigma)$$

Because there is no unobserved quality at the plan level (the first restriction), it is easily shown that ξ does not enter $\text{Prob}(k|j, m, \Theta)$. The model is forced to rely on θ , β , and α to predict within-insurer choices.

For an individual i enrolled with insurer j in market m , let d_{ijk}^m equal one if i enrolled in plan k and zero otherwise. d_{ijk}^m is a random variable with expected value $\text{Prob}(k|j, m, \Theta_0)$. Across jk , the only source of variation in d_{ijk}^m are ε_{ijk} draws. Since ε draws are independent of plan characteristics and premiums, at the true parameters the residual $[d_{ijk}^m - \text{Prob}(k|j, m, \Theta)]$ has expected value zero and is uncorrelated with plan characteristics and premiums. The three moment conditions below exploit these properties.

$$E[x_{jk} * (d_{ijk}^m - \text{Prob}(k|j, m, \Theta_0)) | i, j, m] = 0 \quad (4)$$

$$E[x_{jk} x'_{jk} * (d_{ijk}^m - \text{Prob}(k|j, m, \Theta_0)) | i, j, m] = 0 \quad (5)$$

$$E[p_{jk} * (d_{ijk}^m - \text{Prob}(k|j, m, \Theta_0)) | i, j, m] = 0 \quad (6)$$

The intuition behind these moments is best understood after distributing the instruments inside of $d_{ijk}^m - \text{Prob}(k|j, m, \Theta)$. Equation (4) states that at the true parameters, the model accurately predicts the mean level of x consumed by j 's population of enrollees. If, for example, the model over predicts the share of consumers with dental coverage, β is adjusted to correct the discrepancy. Combined with (4), (5) states that at the true parameters, the model accurately predicts the covariance between characteristics in plans chosen by j 's population of enrollees.²⁴ The intuition behind (6) is similar to (4). The sample moments

²³ ξ is unique up to a normalization. I normalize the mean utility of traditional Medicare to be zero.

²⁴For example, do consumers always choose plans that offer dental *and* vision coverage?

sum over the set of MCBS respondents enrolled in each insurer, and then over insurers and markets.²⁵

5.1.3 Identifying σ

To identify σ , I use the model to predict the choices made by MCBS respondents observed in multiple markets (i.e. years). Suppose an individual is observed enrolling in plan jk in market m and $j'k'$ in market m' (one year later). The probability of this event is given by:

$$Prob(jk \& j'k'|m, m', \Theta) = \int_v Prob(jk|v, m, \Theta) * Prob(j'k'|v, m', \Theta) dF(v|\sigma)$$

Because individuals' preferences, v_g and v_p , do not change over time (the second restriction), $Prob(jk \& j'k'|m, m', \Theta) \neq Prob(jk|m, \Theta) * Prob(j'k'|m', \Theta)$. Let $d_{ijk, j'k'}^{m, m'}$ equal one if i is observed enrolling in plans jk and $j'k'$ in markets m and m' , and zero otherwise. As before, $[d_{ijk, j'k'}^{m, m'} - Prob(jk \& j'k'|m, m', \Theta)]$ is a random variable, and at the true parameters, has an expected value equal to zero. For each individual, the only source of variation in $[d_{ijk, j'k'}^{m, m'} - Prob(jk \& j'k'|m, m', \Theta)]$ across jk and $j'k'$ are draws of ε . This motivates the following moment condition:

$$E[x_{jk} * x_{j'k'} * (d_{ijk, j'k'}^{m, m'} - Prob(jk \& j'k'|m, m', \Theta_0)) | i, m, m'] = 0 \quad (7)$$

Condition (7) states that at the true parameter values, the model accurately predicts the covariance between characteristics of distinct plans chosen by the same individual in multiple years. These moment conditions are particularly useful for identifying heterogeneity in preferences. As σ_g^2 increases, the model predicts more individuals with strong preferences for insurance coverage. These individuals are unlikely to enroll in a plan with little coverage in one year and a plan with generous coverage in the next. If, for example, MCBS respondents are rarely observed switching from plans with unlimited prescription drug coverage to ones with no drug coverage, this behavior can be explained with a large σ_g^2 . A sample version of (7) is constructed by summing over combinations of plans available to each MCBS respondent, and then over MCBS respondents observed in multiple years.

5.1.4 Additional Moment Conditions

I construct a final set of demand-side moment conditions from the observation that at the true parameter values, functions of j 's competitors' (denoted by $-j$) plan characteristics and premiums are uncorrelated with the residual $[d_{ijk}^m - Prob(k|j, m, \Theta)]$. For example,

$$E[f(x_{-jk} | x_{jk}) * (d_{ijk}^m - Prob(k|j, m, \Theta_0)) | j] = 0 \quad (8)$$

²⁵The integrals used to construct $Prob(k|j, \Theta)$ do not have closed formed solutions and must be simulated using a sample of draws from $F(v|\sigma)$. The induced simulation error enters these moments linearly, and therefore does not generate any bias. The sample moments' variances are affected, however. Corrections to these variances are discussed in the appendix.

with

$$\begin{aligned}
 f_x(x_{-jk}|x_{jk}) &= \% \text{ of plans from } -j \text{ with } x_{-jk} < x_{jk} \\
 f_{p,x}(p_{-jk}, x_{-jk}|x_{jk}) &= \text{Mean premiums of } -jk \text{ with } x_{-jk} = x_{jk}
 \end{aligned}$$

At the true parameter values, equation (8) holds for the same reasons discussed above. At an incorrect set of parameter values, the residuals $\left[d_{ijk}^m - Prob(k|j, m, \Theta) \right]$ may be correlated with $-j$'s premiums and plan characteristics. This occurs because insurers condition on their own and competitors' ξ before designing their plans. There may be a relationship between x_{-jk} and ξ_j in equilibrium. For example, if ξ_j is large, j 's competitors might choose low premiums. At an incorrect Θ , the Berry contraction mapping generates an incorrect ξ , and the residuals $\left[d_{ijk}^m - Prob(k|j, m, \Theta) \right]$ become functions of the true ξ_j .

5.2 Supply

In each market the collection of premiums and plan characteristics, $\{g_{jk}, p_{jk}\}_{jk \in \Omega_m}$ are characterized by a Nash equilibrium. I assume insurers' first order conditions with respect to premiums and insurance coverage equal zero:

$$\frac{\partial Profit_j}{\partial g_{jk}} = \frac{\partial Profit_j}{\partial p_{jk}} = 0 \quad \forall \quad jk$$

A firm offering n_j plans has $2 * n_j$ first order conditions that reduce to $2 * n_j$ non-redundant linear equations in $2 * n_j$ cost shocks, ψ . For each choice of parameters, I invert the insurers' first order conditions to recover ψ_{jk}^0 and ψ_{jk}^1 for each jk . I construct moments from these recovered cost shocks to identify the parameters in costs, $(\pi_0, \gamma_0, \lambda_0, \delta_0, \pi_1, \gamma_1, \lambda_1, \delta_1)$.

5.2.1 Identifying γ

If adverse (or advantageous) selection exists, insurers' marginal costs are directly affected by the preferences of consumers they enroll. The parameters of interest, γ_0 and γ_1 , measure this causal effect. I use first order conditions for g and p to recover insurers' marginal costs, which are affected by adverse selection and insurers' cost shocks, ψ . Since insurers who realize different ψ may attract different types of consumers, I use exogenous variation in market structure to identify γ_0 and γ_1 . Below, I explain my identification strategy and the moment conditions it implies. First, consider a one-plan insurer's first order condition for generosity, g_{jk} :

$$\begin{aligned}
P_{jk} \frac{dMS_{jk}(g, p, \Theta)}{dg_{jk}} &= (\gamma_0 + \gamma_1 g_{jk}) \int_v v_g * \frac{dMS_{jk}(v|g, p, \Theta)}{dg_{jk}} dF(v|\sigma) \\
&+ \gamma_1 \int_v v_g * MS_{jk}(v|g, p, \Theta) dF(v|\sigma) \\
&+ (\tilde{\psi}_{0jk} + \tilde{\psi}_{1jk} * g_{jk}) \frac{dMS_{jk}(g, p, \Theta)}{dg_{jk}} + \tilde{\psi}_{1jk} MS_{jk}(g, p, \Theta)
\end{aligned} \tag{9}$$

The left hand side of equation (9) is the marginal revenue to j from an incremental change in g_{jk} . The right hand side captures the marginal cost from an incremental change in g_{jk} . This marginal cost is the sum of three effects.²⁶

First, increasing *generosity* changes the pool of enrollees that jk attracts. Increasing g attracts $\frac{dMS_{jk}(v|g, p, \Theta)}{dg_{jk}} dF(v|\sigma)$ new enrollees with preferences v . These new enrollees change the insurer's costs by $(\gamma_0 + \gamma_1 g_{jk}) v_g \frac{dMS_{jk}(v|g, p, \Theta)}{dg_{jk}} dF(v|\sigma)$. Integrating this indirect cost over v yields the first term in the right hand side of (9). It captures the overall effect on the insurer's costs from changes in his pool of enrollees. If increasing g_{jk} attracts many *new* high v_g enrollees, and adverse selection is present, then costs will increase. If there is no advantageous or adverse selection ($\gamma_0 = \gamma_1 = 0$), then changes in jk 's pool of enrollees do not affect costs.

Next, increasing g_{jk} has a direct effect on costs. It becomes more costly to insure each of jk 's enrollees, even those enrolled in jk prior to the increase. If adverse selection is present, it is particularly costly to provide high v_g enrollees with additional g . This effect is captured by the second term in (9). If adverse and advantageous selection do not exist, this direct effect on costs does not depend on jk 's pool of enrollees. The third effect on marginal costs is determined by X_m , Z_j , ψ_{jk}^0 and ψ_{jk}^1 .

The terms attached to γ_0 and γ_1 in (9) ($\int_v v_g * \frac{dMS_{jk}(v)}{dg_{jk}}$ and $\int_v v_g * MS_{jk}(v)$) are functions of p_{jk} and g_{jk} . Therefore, they are correlated with ψ_{jk}^0 and ψ_{jk}^1 since j chooses p_{jk} and g_{jk} after observing ψ_{jk}^0 and ψ_{jk}^1 . To identify γ_0 and γ_1 , I use the number of insurers and plans in each market (J and JK), ξ , and insurers' opponents' characteristics, Z_{-j} as instruments. The four stage game in section (4) implies ξ , J and JK are independent of ψ_{jk} . I assume Z_{-j} is uncorrelated with ψ_{jk} .²⁷ The instruments are correlated with the endogenous terms because $\frac{dMS_{jk}(v)}{dg_{jk}}$ and $MS_{jk}(v)$ depend on the entire choice set, not just g_{jk} and p_{jk} . The instruments affect market structure and competitors' choices, and therefore, the choice set.

There is a relationship between the endogenous term, $\int_v v_g * MS_{jk}(v)$, and the instrument, JK , because consumers' sorting depends on how many plans are available. $\int_v v_g * MS_{jk}(v)$, the expected strength of jk 's enrollees' preferences, is increasing in *generosity*, at a rate that is increasing in JK . In markets with many plans, sorting is more efficient, and there is less pooling within plans among consumers with different

²⁶In (9), $\tilde{\psi}_{jk}^z = \pi_z + \delta'_z X_m + \lambda'_z Z_j + \psi_{jk}^z$ for $z = 0, 1$.

²⁷ ψ_{jk}^0 and ψ_{jk}^1 are realized conditional on Z_j and X_m , and therefore uncorrelated with Z_j and X_m . Therefore, it seems reasonable to assume Z_{-j} is also uncorrelated with ψ_{jk}^0 and ψ_{jk}^1 .

v_g . Low v_g consumers are likely to have low coverage, inexpensive plans to choose from, and therefore, will not enroll in a high g plan. Similarly, high v_g consumers are unlikely to enroll in low g plans when the product space is saturated. A similar relationship between JK and $\int_v v_g * \frac{dMS_{jk}(v)}{dg_{jk}}$ can be derived for any given g_{jk} .

Intuition also relates $\int_v v_g * MS_{jk}(v)$ to j 's opponents' characteristics, Z_{-j} . Z_{-j} affects g_{-j} , and therefore, the pool of enrollees that j attracts. Suppose j faces one competitor who offers one plan with $g_{jk} > g_{-jk}$ or $g_{jk} < g_{-jk}$. If $g_{jk} > g_{-jk}$, j will attract more high v_g consumers than when $g_{jk} < g_{-jk}$. If, for example, $Total_Assets_{-j}$ is positively correlated with g_{-jk} , then $Total_Assets_{-j}$ will be negatively correlated with $\int_v v_g * MS_{jk}(v)$.²⁸ An insurer whose competitor has high assets (and therefore, is likely to offer a generous plan), will tend to draw low v_g consumers. Similar relationships can be derived between Z_{-j} and $\int_v v_g * \frac{dMS_{jk}(v)}{dg_{jk}}$.

The moment conditions used to identify γ_0 and γ_1 are listed below.

$$\begin{aligned}
E[f(Z'_{-jm}) * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[JK_m * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[J_m * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[f(Z'_{-jm}) * JK_m * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[f(Z'_{-jm}) * J_m * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[\xi_j * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[f(\xi_{-j}) * \psi_{jk}^n] &= 0 \quad \forall j, k, n
\end{aligned}$$

Sample analogues are constructed by summing across plans, insurers, and markets. Any function f can be used in these moment conditions. I use $Mean(Z_{-j})$, $Min(Z_{-j})$, and $Max(Z_{-j})$.

γ_0 and γ_1 are identified only if the sample moments are different from zero at incorrect guesses of γ_0 and γ_1 . The mechanical relationships between g and p and the cost shocks ψ , provide intuition for this condition. Consider an increase in γ_1 away from its true value. Costs increase, particularly for plans that are attractive to high v_g consumers. The implied cost shocks, ψ , of plans attractive to high v_g consumers, must go down so insurers' first order conditions continue to hold. If g_{jk} is positively correlated with $Total_Assets_j$, then $f_{mean}(Total_Assets_{-j}) * \psi_{jk}^1$ will, on average, be affected by the increase in γ_1 . Consumers with high v_g tend not to enroll in plans offered by insurers with low assets (who offer low g plans). Rather, they are likely to choose plans offered by competitors of low asset insurers. Thus, increasing γ_1 will increase costs and decrease ψ most for insurers competing against low asset insurers. This causes the correlation between $f_{mean}(Total_Assets_{-j})$ and ψ_{jk}^1 to increase away from zero. Similar intuition relates the other moment

²⁸Reduced form regressions of Z_j on g_{jk} and p_{jk} are provided in section (8).

conditions to γ_0 and γ_1 .

5.2.2 Identifying π , δ , and λ

π_0 and π_1 are attached to constant terms that enter $FC(v_g, \Theta^{Supply})$ and $MC'_{jk}(v_g, \Theta^{Supply})$. δ_0 and δ_1 are attached to the firm controls, Z_j , and λ_0 and λ_1 are attached to market controls, X_m . The following moment conditions are used to identify π , δ and λ :

$$\begin{aligned} E[\psi_{jk}^n] &= 0 \quad \forall jkn \\ E[X_m * \psi_{jk}^n] &= 0 \quad \forall jkn \\ E[Z_j * \psi_{jk}^n] &= 0 \quad \forall jkn \end{aligned}$$

For any choice of γ , a π , δ , and λ that satisfy the sample versions of these moment conditions can be retrieved using an ordinary least squares regression. Let $\tilde{\psi}_{jk}^0$ and $\tilde{\psi}_{jk}^1$ be defined by:

$$\begin{aligned} \tilde{\psi}_{jk}^0 &= \pi_0 + \delta'_0 X_m + \lambda'_0 Z_j + \psi_{jk}^0 \\ \tilde{\psi}_{jk}^1 &= \pi_1 + \delta'_1 X_m + \lambda'_1 Z_j + \psi_{jk}^1 \end{aligned}$$

I invert insurers' first order conditions to recover $\tilde{\psi}_{jk}$. OLS regressions of $\tilde{\psi}_{jk}$ on a constant, X_m , and Z_j generate estimates of π , δ , and λ .²⁹

5.3 Summary of Estimation Routine

The estimation routine is simulated method of moments. It searches over $(\alpha, \beta, \theta, \sigma, \gamma)$ for the minimum of a criterion function, $\Psi(\alpha, \beta, \theta, \sigma, \gamma)$:

$$\begin{aligned} \Psi(\alpha, \beta, \theta, \sigma, \gamma) &= \\ & \left[\begin{array}{c} Demand_Moments(\alpha, \beta, \theta, \sigma, \gamma) \\ Supply_Moments(\alpha, \beta, \theta, \sigma, \gamma) \end{array} \right]' * W \\ & * \left[\begin{array}{c} Demand_Moments(\alpha, \beta, \theta, \sigma, \gamma) \\ Supply_Moments(\alpha, \beta, \theta, \sigma, \gamma) \end{array} \right] \end{aligned}$$

Evaluating Ψ can be broken into four steps for each choice of $(\alpha, \beta, \theta, \sigma, \gamma)$:

1. Use the Berry contraction mapping to generate $\hat{\xi}(\alpha, \beta, \theta, \sigma, \gamma)$
2. Plug $\hat{\xi}$ into the insurers' first order conditions and back out $\tilde{\psi}$.

²⁹This method is not asymptotically efficient. A weighted IV regression using X_m , Z_j , and Z_{-j} is implied by the GMM criterion function. In practice, this regression has little effect on parameter estimates and the standard errors.

Table 1: Total Number of Markets

<u># Insurers</u>	<u>2000</u>	<u>2001</u>	<u>2002</u>	<u>2003</u>	<u>Total</u>
1	607	517	581	722	2427
2	560	522	333	434	1849
3	784	405	244	351	1784
4	373	131	81	127	712
5	266	107	27	30	430
6	195	43	13	42	293
>6	136	108	83	112	439
Total	2921	1833	1362	1818	

3. Regress $\tilde{\psi}$ on a constant, X_m and Z_{jm} to find $(\hat{\pi}_0, \hat{\pi}_1, \hat{\delta}'_0, \hat{\delta}'_1, \hat{\lambda}'_0, \hat{\lambda}'_1)$ and ψ_j
4. Evaluate the remaining supply and demand moment conditions and calculate the value of Ψ at $(\alpha, \beta, \theta, \sigma, \gamma)$.

The market share probabilities and insurers' cost shocks ψ are continuous in the parameters. In the text above, intuition for identification was given. Assuming standard regularity conditions yields consistency of the parameter estimates and their asymptotic distribution (Newey & McFadden 1994).

6 Descriptive Results

The tables in this section describe insurers, the markets insurers participate in, and the plans that insurers offer. Tables 1 displays the distribution of insurers across markets and across time. The majority of markets have at least two insurers, but the overall trend in insurer participation is negative. HMO participation was at its strongest in 2000, when payment rate reductions put in place by the BBA began to take effect.³⁰ In 2003, the Medicare Modernization Act was passed, paving the way for expansions in privatized Medicare.

Table 2 displays the characteristics of markets that have different numbers of insurers. There are strong correlations between the number of insurers in a market and the government's payment rate, as well as between the number of insurers and market size.³¹

Tables 3 and 4 describe the types of insurers participating in Medicare, and their characteristics. Most are Health Maintenance Organizations, but Preferred Provider Organizations and Private Fee for Service insurers are becoming increasingly common.³²

Table 5 displays means and standard deviations of plan characteristics included in utility. It is through dental, vision, and prescription drug coverage that private insurers distinguish themselves from traditional

³⁰The BBA reduced payments on average to insurers by 6% (Congressional Budget Office 1999).

³¹Maruyama (2006) finds that CMS subsidy rates are an important determinant of entry by Medicare HMOs.

³²PPOs share characteristics with traditional Indemnity Plans and HMOs. HMOs tend to be restrictive about where enrollees can receive medical care and traditional Indemnity plans are not. Within PPOs, medical care can be received from all providers, but small penalties are imposed for out of network care. PFFS insurers provide coverage identical to traditional Medicare.

Table 2: Market Characteristics

<u># Insurers</u>	<u>Payment Rate</u>	<u>Outside Mkt. Sh.</u>	<u># Plans in Mkt</u>	<u>Mkt Population</u>
1	513.70	92.3%	2.06	121.92
2	526.78	85.5%	3.79	294.40
3	545.49	77.2%	6.96	442.00
4	537.97	73.5%	6.36	567.10
5	582.84	71.4%	8.95	663.70
6	603.54	72.8%	11.71	1395.16
>6	628.69	65.4%	12.70	3128.84

Notes: Payment rate is equal to the base subsidy paid per month/enrollee to HMOs by the CMS. Outside Mkt Share equal to the percentage of enrollees remaining in traditional Medicare. # Plans is equal to the total number of plans offered by insurers. Population is in 1000s.

Table 2 Ctd.: Market Characteristics

<u># Insurers</u>	<u>% White</u>	<u>Income</u>	<u>% BA Degree</u>	<u># Hospitals</u>
1	0.91	19.00	18.7%	2.70
2	0.88	21.23	21.9%	5.26
3	0.84	23.04	26.0%	7.52
4	0.85	22.80	25.2%	8.96
5	0.78	24.43	27.6%	10.76
6	0.72	23.60	26.9%	19.26
>6	0.71	20.38	23.0%	35.38

Note: Per Capita Income in 1000's. % BA Degree is the percentage of a market's population with a college degree.

Table 3: Insurer Types

<u>Year</u>	<u>HMO</u>		<u>PFFS</u>		<u>PPO</u>	
	<u>%</u>	<u>#</u>	<u>%</u>	<u>#</u>	<u>%</u>	<u>#</u>
2000	92.85%	2200	0.00%	0	7.15%	134
2001	93.62%	1273	1.93%	16	4.45%	63
2002	87.17%	675	1.83%	11	11.00%	116
2003	80.52%	868	0.96%	7	18.52%	259

Notes: Table gives distribution of firm types in those markets used in estimation.

Table 4: Mean Insurer Characteristics

	<u>Mean</u>	<u>Std. Dev</u>
Total Assets	\$374.63	791.78
Net Income	\$10.82	42.12
# Physicians / Enrollee	47.97	67.01
% Administrative Expenses	9.88%	0.94
Medical Expenditures / Enrollee	\$140.49	189.25
% Business Medicare	22.38%	2.17

Notes: All data reported at the insurer level, not the insurer-market level. Assets and Net Income reported in \$Millions. #Physicians/Enrollee is equal to the total number of member physicians divided by the total number of enrollees. Administrative expenses are given as a percentage of total premium income. Medical Expenditure/Enrollee the average amount spent on each enrollee per month. %Business Medicare is percent of enrollees who are Medicare beneficiaries. All variables lagged one year.

Table 5: Plan Characteristics

	<u>Mean</u>	<u>Std. Dev</u>
Premium	\$56.72	49.68
\$ Specialist	\$13.12	9.44
Vision	83.10%	3.70
\$ Inpatient Hospital Admission	\$90.55	214.53
Dental	19.78%	3.98
Generic PD	61.32%	4.87
Brand PD	50.05%	5.00
Generic PD Unlimited Coverage	34.17%	4.74
Brand PD Unlimited Coverage	6.44%	2.45
\$ Copay Brand PD	\$10.56	15.01

Notes: Premium is paid monthly. \$Specialist and \$Inpatient Hospital are paid when an enrollee visits a specialist or is admitted to an inpatient hospital. Vision, Dental, and Prescription Drug (PD) are dummy variables, indicating whether any coverage is offered. Unlimited coverage variables indicate if there are limits to prescription drug coverage. \$ Copay Brand PD is enrollee's cost for a 30 day supply of a brand prescription drug. Payments all in 2000 dollars

Table 6: Mean Plan Characteristics by # Insurers

<u># Insurers</u>	<u>Premium</u>	<u>Dental</u>	<u>Vision</u>	<u>\$ Specialist</u>
1	\$68.64	11.21%	71.98%	\$12.88
2	\$57.55	20.23%	83.23%	\$13.62
3	\$62.18	20.96%	88.34%	\$14.26
4	\$42.91	22.47%	90.73%	\$12.25
5	\$45.19	24.19%	94.19%	\$12.30
6	\$32.56	30.03%	91.47%	\$11.77
>6	\$14.03	45.45%	94.17%	\$10.82

Medicare.³³ The remaining variables proxy for cost-sharing requirements (i.e. copayments) in each plan.³⁴ In particular, there is extensive variation in premiums and the types of prescription drug coverage that are offered.

Table 6 displays a clear correlation between the number of insurers in a market and the amount of coverage that insurers offer. This reflects competition, payment rates, and perhaps costs. In markets with one insurer, the mean monthly premium is \$68.64. In markets with six or more insurers, the mean premium is \$14 per month. Similarly, in markets with more insurers, prescription drug, dental, and vision coverage become more common.

³³Town & Liu (2003) find that the main source of surplus gains from HMO participation in Medicare are utility gains from prescription drug insurance which was absent from traditional Medicare until 2006.

³⁴Clearly, there are more plan characteristics in a typical insurance plan than listed. Unfortunately, reporting methods make comparison of many plans' characteristics difficult. These variables should serve as a reasonable proxy for each plan's amount of coverage.

Table 6 ctd.

<u># Insurers</u>	<u>\$ Inpatient Hospital</u>	<u>Generic PD</u>	<u>Brand PD</u>
1	\$100.77	46.48%	36.59%
2	\$86.91	57.00%	46.02%
3	\$82.09	67.21%	53.53%
4	\$104.46	74.86%	62.64%
5	\$31.04	82.79%	73.26%
6	\$81.97	83.28%	74.74%
>6	\$126.02	80.42%	68.07%

Notes: Same as Table 5

7 Estimation Results

7.1 Utility

The main specification for utility is given in equation (1). Table 7 displays estimates of the parameters entering this specification.

The parameter estimates have the expected signs and are very precise. In particular, Medicare beneficiaries receive significant utility from prescription drug coverage. The parameters on *Generic PD* and *Brand PD Unlimited* are significant at the 0.1% level. The parameter attached to *Generic PD Unlimited* is significant at the 1% level. Brand prescription drug coverage provides little utility beyond what is provided by generic coverage. This is probably the result of a strong correlation between *Generic PD* and *Brand PD*.

Dental and vision coverage also provide positive utility. The payment variables are negative, as expected. The parameter attached to $\$Specialist$ is not precisely estimated, but utility decreases by a significant amount when a plan's inpatient hospital and prescription drug cost-sharing increase.

Utility is decreasing in premiums, although less so in high income markets. To determine beneficiaries' sensitivity to premium increases, I calculate the mean semi-elasticity, $\frac{\partial MS_{jk}}{\partial P_{jk}} * \frac{1}{MS_{jk}}$. A \$1 increase in a monthly premium reduces enrollment, on average, by .46%. Town & Liu (2003) estimate that a one dollar increase in all of an insurer's plans decrease the insurer's market share by .9%. This elasticity is comparable to mine since the average insurer offers 1.64 plans.³⁵ I also calculate mean market share elasticities with respect to *generosity* and ξ ($\frac{\partial MS_{jk}}{\partial g_{jk}} * \frac{g_{jk}}{MS_{jk}}$ and $\frac{1}{n_j} \frac{\partial MS_j}{\partial \xi_j} * \frac{\xi_j}{MS_j}$). A one percentage increase in a plan's *generosity* generates, on average, a .98% increase in market share. Increasing unobserved quality by one percent generates, on average, a .72% increase in market share for each plan.

³⁵Buchmueller (2000) find a premium semi-elasticity of -.7%. Atherly, Dowd, and Feldman (2003) find -.7% and -.4%. Both of these studies also examine Medicare beneficiaries.

Table 7: Utility Parameter Estimates

	(1)
<u>Plan Characteristics</u>	
Dental	0.6386 (.1182)***
Vision	0.269 (.1284)**
\$ Specialist	-0.0924 (.1316)
\$ Inpatient Hosp Admission	-0.3626 (.1202)**
Generic PD	1.1796 (.1311)***
Brand PD	0.1022 (.1878)
\$ Brand PD Copay	-0.0636 (.0041)***
Generic PD Unlimited	0.1984 (.118)**
Brand PD Unlimited	1.1578 (.1794)***
Premium	-3.1885 (.1013)***
<u>Vertical Preferences</u>	
Sigma^2 Generosity	0.2911 (.0348)***
Sigma^2 Premium	0.0000 (4.4312)
Market Income _{Premium}	-0.0094 (.003)***

Notes: Standard errors in parentheses. Variable definitions given in Table 5 and appendix.
 * Significant at 5% level, ** at 1% level, *** at 0.1% level

7.2 Willingness to Pay

The bottom portion of Table 6 describes estimated variation in consumers' preferences for *generosity* and *premiums*. σ_g^2 is positive and estimated very precisely, suggesting there is heterogeneity in preferences for insurance coverage. Conditional on income, however, the model does not find any variation in disutility from premiums. The strong correlation between g_{jk} and p_{jk} may prevent separate identification of both parameters.³⁶

To interpret the economic importance of this heterogeneity, I calculate the willingness to pay for plan characteristics by consumers with different v_g . After adjusting for normalizations, a Medicare eligible is willing to pay $\frac{-(1+v_g)\beta_n}{(1+\theta I_m+v_p)\alpha}$ for one unit of x_n . Table 8 provides estimates of the willingness to pay for each component of *generosity* at the 10th, 50th, and 90th percentiles of the distribution of v_g .

Table 8: Willingness to Pay Estimates

	10th Percentile	Median	90th Percentile
Dental	\$10.81 (1.89)	\$13.78 (2.41)	\$19.05 (3.33)
Vision	\$4.58 (2.12)	\$5.85 (2.71)	\$8.08 (3.74)
\$ Specialist	\$0.07 (.11)	\$0.09 (.14)	\$0.12 (.19)
\$ Inpatient Hospital Admission	\$0.01 (.004)	\$0.02 (.01)	\$0.02 (.01)
PD Generic	\$19.57 (2.04)	\$24.97 (2.64)	\$34.52 (3.78)
PD Brand	\$1.96 (3.10)	\$2.49 (3.96)	\$3.45 (5.47)
\$ Copay Brand	\$1.07 (.08)	\$1.36 (.01)	\$1.88 (.14)
PD Generic Unlimited	\$3.23 (1.00)	\$4.11 (2.45)	\$5.67 (3.49)
PD Brand Unlimited	\$19.42 (2.86)	\$24.78 (3.64)	\$34.24 (5.01)

Notes: WTP estimates give monthly value (in 2000 \$) to consumer from \$1 reductions in payment variables and full coverage of indicator variables. It is assumed for each percentile, consumers have median preferences for premiums. Standard errors in parentheses. They are constructed after taking parameter draws (except for Sigma² Premium) from their estimated distributions.

The median consumer is willing to pay up to \$13.78 per month for dental coverage. An individual at the 90th percentile of the v_g distribution is willing to pay \$5.27 more per month. Medicare beneficiaries are not sensitive to changes in doctor visit and inpatient hospital admission copayments. With respect to prescription drug insurance, a consumer at the 10th percentile is willing to pay \$44.18 per month for prescription drug coverage that includes unlimited generic and branded drugs. An individual at the 90th

³⁶The residuals in the utility function are equal to $\sigma_g |v_g| * g + \sigma_p |v_p| * \alpha p$ where v_g and v_p are drawn from a $|N(0,1)|$ distribution. With perfect correlation ($g = \lambda p$), the residuals are equal to $(\sigma_g |v_g| + \sigma_p |v_p|) * (\alpha + \lambda)p$, in which case σ_g and σ_p are not separately identified. The estimation routine chooses a $\widehat{\sigma}_g > 0$ and $\widehat{\sigma}_p \approx 0$, through the supply side. If adverse selection is an important determinant of firm behavior, $\widehat{\sigma}_g = 0$, would not allow my model to capture it.

Table 9: Consumer Surplus Estimates

# Insurers	CS/Medicare Beneficiary	CS/M+C Enrollee
1	\$19.85 (1.248)	\$251.98 (12.117)
2	\$45.29 (1.908)	\$265.33 (12.577)
3	\$70.17 (3.108)	\$285.89 (13.356)
4	\$82.12 (3.636)	\$290.53 (13.407)
5	\$88.44 (4.260)	\$307.96 (13.981)
6	\$90.40 (3.792)	\$293.12 (13.162)
>6	\$109.40 (4.824)	\$293.12 (13.456)
Overall	\$59.48 (2.604)	\$276.45 (10.659)

Notes: Column 1 gives yearly value in year 2000 dollars to a representative consumer from living in market with N insurers instead of zero insurers. Column 2 gives the total surplus per consumer actually enrolling in a private insurer. Values are population weighted means. Standard errors of means, calculated by taking draws from estimated distribution of parameters, are in parentheses.

percentile is willing to pay \$77.88 per month for the most generous prescription drug package.

7.2.1 Consumer Surplus Estimates

Table 6 shows a relationship between the number of insurers in a market, generous coverage, and lower premiums. Through improved coverage, lower premiums and plan variety, Medicare beneficiaries will extract more surplus from the M + C program in markets with more insurers. Table 9 reflects this. Following Train (2003), the expected consumer surplus a representative Medicare beneficiary receives in a market m from the presence of M + C insurers is $\int \frac{1}{(1+\theta_p I_m + v_p)^{\alpha}} \log \left(1 + \sum_{jk \in \Omega_m} e^{u_{jk}(v)} \right)$.³⁷

A representative Medicare beneficiary residing in a market with one insurer receives \$19.85 in surplus per year from the M + C program. In a market with more than six private insurers, \$109.40 in consumer surplus is generated per Medicare beneficiary per year. The second column of Table 9 calculates the total surplus in each market per beneficiary enrolled in a M + C plan. Across all markets and years, the total contribution to consumer surplus is estimated to be \$5.911 billion.³⁸ A breakdown by year is given in Table 10.

³⁷Relative to a market with only the outside good.

³⁸Town & Liu (2003) estimate national consumer surplus from Medicare HMOs to \$4.061 billion in 2000. Maruyama obtains similar estimates of \$3.881 and \$3.963 billion in 2003 and 2004.

Table 10: Total Consumer Surplus Estimates

<u>Year</u>	<u>Total Consumer Surplus</u>
2000	\$1.808
2001	\$1.482
2002	\$1.242
2003	\$1.379
Overall	\$5.911

Note: Estimates are in year 2000 billions of dollars. Full set of markets used in calculations

Table 11 Descriptive Evidence of Adverse Selection:
Premiums and Preferences

	<u>Premium</u>
Constant	41.55 (8.30)***
Mean v_g	73.32 (10.19)***
Generosity	-5.71 (1.95)**
# Insurers	-3.7 (.466)***
Unobserved Insurer Quality	0.259 (.043)***
N	5622
R-Squared	0.629

Notes: Insurer controls, market controls, and year dummies omitted from table. Mean v_g and generosity constructed from estimated demand parameters. Standard Errors are in parentheses. *, **, and *** indicate significance at the 5%, 1% and 0.1% levels

7.3 Descriptive Evidence of Adverse Selection

If adverse selection is present, insurers that attract enrollees with strong preferences for insurance have high costs, all else equal. These insurers will need to charge high premiums. Descriptive evidence consistent with adverse selection is given in table 11. It displays regression results that relate each premium p_{jk} , to the mean v_g of jk 's enrollees, g_{jk} , ξ_j , the number of insurers in jk 's market, and the full set of market and insurer controls. There is a strong, positive relationship between p_{jk} and jk 's mean v_g .^{39, 40, 41}

These regression results, however, are only suggestive of adverse selection because they do not adequately control for the nature of competition in privatized Medicare or differences in price sensitivities across consumer types. Estimates of insurers' costs are presented in the next section.

³⁹ Mean v_g is calculated as $\frac{\int v_g * MS_{jk}(v) dF(v|\hat{\sigma})}{MS_{jk}}$

⁴⁰ The standard errors in Table 11 have not yet been corrected for simulation error. However, running the specification several times using different draws of preferences from $N(0, \hat{\sigma}^2)$ had only slight effects on the parameter estimates.

⁴¹ When $Mean v_g$ is removed from the specification, the coefficient on *Generosity* is positive.

7.4 Costs

The costs to an insurer from offering plan jk to a consumer with preferences for generosity, v_g , is given in equation (2). Table 12 displays estimates of this specification.^{42, 43}

The second row describes parameter estimates relating to adverse selection ($\widehat{\gamma}_0$ and $\widehat{\gamma}_1$). Both parameter estimates strongly suggest the presence of adverse selection: beneficiaries with strong preferences for insurance are more costly to insure. A one unit increase in an beneficiary's type v_g causes a \$22.29 increase in insurers' fixed costs of providing care and a \$5.58 increase in insurers' marginal costs of *generosity*. Both parameter estimates are significant at the 0.1% level.

Across plans, the mean expected marginal cost to insurers from providing an additional unit of *generosity* to a consumer at the 10th percentile of $F_g(v|\widehat{\sigma})$ is \$22.60 per month. For a consumer at the 90th percentile of $F_g(v|\widehat{\sigma})$, this marginal cost is \$27.17 per month, 20.3% higher. Fixed costs per enrollee (i.e. they do not depend on coverage) also vary by v_g . For enrollees at the 10th and 90th percentiles of $F_g(v|\widehat{\sigma})$, mean fixed costs are \$5,588 and \$5,807 per year. These results suggest that a consumer's preferences for generosity have an economically larger impact on the marginal costs of *generosity* than the fixed costs of enrollment. This conforms with economic intuition. Unhealthy consumers don't impose large burdens on insurers who offer little coverage.

Dental coverage is equivalent to .64 units of generosity. This implies that net of cost-sharing, dental coverage typically costs an insurer \$173.15 and \$208.81 per year for enrollees at the 10th and 90th percentiles of $F_g(v|\widehat{\sigma})$. Similarly, providing a prescription drug package with unlimited coverage costs insurers \$524.32 and \$630.49 per year for enrollees at the 10th and 90th percentiles of $F_g(v|\widehat{\sigma})$.

7.4.1 Firm and Market Controls

The estimated coefficients attached to insurer and market controls in table 12 vary in statistical significance. Also, some conform with economic intuition and others do not. HMOs appear to have higher fixed and marginal costs than non-HMOs, and overall costs are lower for large insurers. Insurers employing more physicians per enrollee behave as if they pay higher costs per enrollee.

These controls' ability to explain costs, however, is unimportant relative to their ability to explain observed levels of insurance coverage and premiums. It must be true that the elements of Z_j are correlated with j 's plans' characteristics, conditional on X_m . If so, Z_{-j} can be used to generate exogenous variation in q_{-j} and p_{-j} that identifies γ_0 and γ_1 . Table 13 displays results from a regression of q_{jk} and p_{jk} on Z_j and X_m . Insurer characteristics do explain their premiums and chosen levels of insurance coverage.

⁴²At the estimated parameters, I evaluate insurers' second order conditions for g and p to determine whether insurers are in fact behaving optimally. All of insurers' choices maximize profits locally, and in most cases, over a wide range of feasible (g, p) combinations.

⁴³Estimates of alternative specifications are provided in the appendix. Specifically, equation (2) assumes costs are linear in *generosity*. If costs are in fact convex, my model may attribute convexity to adverse selection. If costs are concave in *generosity*, my model may understate the severity of adverse selection. The appendix provides estimates of a cost function that allows for nonlinearity in *generosity*. The estimates suggest that $c(v_g, g, \Theta^{Supply})$ is approximately linear in g .

Table 12: Cost Parameter Estimates

	<u>Fixed Cost</u>	<u>Marginal Cost of Generosity</u>
Constant	13.51 (6.669)**	13.34 (.08)***
<u>Adverse Selection</u>		
v_g	22.29 (4.43)***	5.58 (.38)***
<u>Market Controls</u>		
AB Rate / 10	9.16 (.15)***	-0.001 (.002)
AB Rate 1997 / 10	-0.46 (.12)***	0.004 (.002)**
Population / 1000	-0.01 (.02)	-0.002 (.0002)***
Per Capita Income / 1000	-6.08 (1.29)***	3.84 (.02)***
# Hospitals	-0.3 (.14)**	0.008 (.002)***
2001 Dummy	21.85 (1.77)***	0.015 (.022)
2002 Dummy	22.47 (2.16)***	0.008 (.027)
2003 Dummy	31.78 (1.89)***	0.005 (.024)
<u>Insurer Controls</u>		
HMO Dummy	4.118 (2.83)	-0.001 (.035)
Total Assets (\$million)	-2.88 (1.01)**	-0.042 (.013)***
Net Income (\$million)	-0.03 (.02)	0.0009 (.0002)***
# Network Physicians / Enrollee	0.69 (.16)***	0.0014 (.002)
% Administrative Expenses	-9.3 (1.16)***	-0.005 (.014)
Medical Expenditures / Enrollee	0.0025 (.0061)	0 (.0001)
% Business Medicare	-2.33 (.40)***	0.06 (.005)***

Notes: Insurer controls lagged one year. AB Rate and AB Rate 1997 are current and 1997 government payment rates to insurers. Other variables defined in tables 2 and 4. Estimates of the parameters attached to market and insurer controls obtained from an OLS regression using implied costs as dependent variables. *,**,*** indicate significance at 5%,1%, and 0.1% levels. Standard Errors are in parentheses.

Table 13: Reduced Form Supply Side Regressions

	<u>Generosity</u>	<u>Premium</u>
Constant	1.22 (.13)***	90.91 (7.68)***
<u>Market Controls</u>		
AB Rate / 10	0.009 (.003)***	0.469 (.17)**
AB Rate 1997 / 10	-0.009 (.002)***	-1.34 (.134)***
Population / 1000	0.001 (.0003)***	0.012 (.021)
Per Capita Income / 1000	0.016 (.025)	11.38 (1.48)**
# Hospitals	-0.007 (.003)**	-0.429 (.163)**
2001 Dummy	-0.59 (.034)***	2.76 (2.04)
2002 Dummy	-0.71 (.041)***	-1.13 (2.49)
2003 Dummy	-0.62 (.036)***	14.86 (2.18)***
<u>Insurer Controls</u>		
HMO Dummy	-0.25 (.054)***	-1.99 (3.26)
Total Assets (\$ million)	0.052 (.019)**	-1.5 (1.16)
Net Income (\$ million)	0.0002 (.0004)	-0.026 (.023)
# Network Physicians / Enrollee	-0.002 (.003)	0.788 (.188)***
% Administrative Expenses	-0.096 (.022)***	-11.86 (1.33)***
Medical Expenditures / Enrollee	0.0002 (.0001)	-0.004 (.007)
% Business Medicare	-0.045 (.008)***	-2.9909 (.4601)***
N	5622	5622
R-Squared	0.331	0.292

Notes: Firm controls lagged one year. Standard Errors are in parentheses. *, **, and *** indicate significance at the 5%, 1% and 0.1% levels

Table 14: Profit Margins with Adverse Selection

# Insurers	All Insurers	0-25th % Generosity	75-100th % Generosity
1	\$22.24 (.019)	\$22.34 (.019)	\$20.43 (.018)
2	\$22.95 (.018)	\$23.44 (.018)	\$22.12 (.019)
3	\$23.97 (.019)	\$23.56 (.019)	\$23.03 (.019)
4	\$23.27 (.019)	\$23.67 (.018)	\$22.85 (.019)
5	\$23.07 (.018)	\$23.43 (.019)	\$21.70 (.019)
6	\$22.04 (.018)	\$22.09 (.018)	\$20.96 (.018)
>6	\$21.50 (.018)	\$20.82 (.017)	\$21.37 (.018)

Notes: Estimates of profit margins are per enrollee and per month at the plan level. Generosity quartiles are formed using the distribution of generosity across all markets. Reported values are market share and population weighted means. Standard deviations of means in parentheses

8 Counterfactual Simulations

8.1 Adverse Selection and Expected Profit Margins

Insurers that offer different amounts of coverage are affected differently by adverse selection for two reasons. First, the costs of insurers who offer generous insurance are more sensitive to adverse selection than insurers who offer little coverage (in $c(v_g, g)$, $\gamma_1 v_g g$ is increasing in g). Second, generous insurance plans attract more enrollees with strong preferences for insurance.

To investigate the economic importance of adverse selection, I first examine implied profit margins in the presence and absence of adverse selection costs. To remove insurers' costs from adverse selection, I adjust government payments to insurers to reflect the differences in costs across enrollees. Let \bar{v}_g denote the median draw of v_g . Payments for consumers with $v_g < \bar{v}_g$ are reduced and payments for consumers with $v_g > \bar{v}_g$ are increased so that $c(v_g, g) = c(\bar{v}_g, g)$ for all v_g and g . After this policy is implemented, insurers are no longer penalized for attracting high v_g consumers (the second effect).

Table 14 displays mean profit margins in markets with different numbers of insurers before adverse selection is removed from insurers' costs. Column 1 displays mean profit margins across all insurers. Columns 2 and 3 display profit margins for plans that offer coverage in the bottom and top quartiles of the distribution of *generosity*. Table 15 holds premiums and plan characteristics fixed and duplicates table 14 after adverse selection is removed from insurers' costs. All profit margins are monthly and per enrollee.

In markets with different numbers of insurers, profit margins are similar at each level of *generosity*, with and without the effects of adverse selection. This is because CMS payment rates are higher in markets with more insurers. Otherwise (holding premiums fixed), margins would decline in the number of insurers, as

Table 15: Profit Margins without Adverse Selection

# Insurers	All Insurers	0-25th % Generosity	75-100th % Generosity
1	\$26.77 (.018)	\$23.47 (.019)	\$34.61 (.023)
2	\$27.71 (.018)	\$23.90 (.019)	\$33.49 (.022)
3	\$29.45 (.019)	\$23.94 (.019)	\$34.08 (.021)
4	\$28.81 (.019)	\$24.07 (.019)	\$34.56 (.021)
5	\$29.48 (.018)	\$23.59 (.021)	\$34.94 (.021)
6	\$28.02 (.018)	\$22.26 (.020)	\$33.72 (.021)
>6	\$28.89 (.017)	\$20.49 (.018)	\$33.34 (.019)

Notes: Estimates of profit margins are per enrollee and per month at the plan level. Generosity quartiles are formed using the distribution of generosity across all markets. Reported values are market share and population weighted means. Standard deviations of means in parentheses

expected. Profit margins of insurers offering low levels of *generosity* increase slightly when adverse selection costs are removed. Generous plans, however, experience large increases in profit margins.

Estimates of insurers' costs and the changes in profit margins in tables 14 and 15 suggest that insurers do have incentives to distort their coverage to deter enrollment by consumers with high v_g . These findings may explain the low levels of coverage offered by M + C insurers. For example, the median consumer is willing to pay \$76 per month for an insurance plan that does not charge doctor or inpatient hospital copayments and offers dental, vision, and unlimited prescription drug coverage (this corresponds to a *generosity*, $g^{Max} = 3.55$). After receiving the government's payment, 38% of insurance plans could offer g^{Max} to the median consumer at a cost less than \$76. Yet this plan is only observed in less than 1% of markets at an average premium of \$268 per month. Among the set of insurers able to profitably offer g^{Max} to \bar{v}_g , the mean level of *generosity* offered is 1.1. I allow each of these insurers to unilaterally increase their coverage to g^{Max} , whereupon their costs increase, on average, by \$63 per enrollee. A large portion of these cost increases can be attributed to changes in the pools of consumers that each insurer attracts. If the government adjusts payments to remove the effects of adverse selection on costs, insurers' unilateral increases in *generosity* still increase costs, but only by \$49.60 per enrollee.

Below, I remove the distortionary effects adverse selection has on insurers' incentives and costs. I then allow insurers to *reoptimize* their choices of (g, p) . Tables 14 and 15 suggest that initially generous plans will be most affected in a new equilibrium without adverse selection. This exercise is necessary to quantify distortions to behavior caused by adverse selection, and to measure welfare losses due to its presence.

8.2 Endogenous Plan Characteristics and Adverse Selection

Below, I rearrange an insurer's cost function to make the costs from adverse selection explicit. As before, the median consumer type is denoted by \bar{v}_g .

$$\begin{aligned} c_{jk}(g_{jk}, v_g) &= (\gamma_o v_g + \psi_{jk}^0) + (\gamma_1 v_g + \psi_{jk}^1) * g_{jk} = \\ &(\gamma_o \bar{v}_g + \gamma_o (v_g - \bar{v}_g) + \psi_{jk}^0) + (\gamma_1 \bar{v}_g + \gamma_1 (v_g - \bar{v}_g) + \psi_{jk}^1) * g_{jk} \end{aligned}$$

I assume that in a world without adverse selection, $c_{jk}(g_{jk}, v_g) = c_{jk}(g_{jk}, \bar{v}_g)$ for all v_g . The costs imposed on jk attributable to adverse selection are given by $c_{jk}(v, g_{jk}) - c_{jk}(\bar{v}_g, g_{jk}) = \gamma_o(v_g - \bar{v}_g) + \gamma_1(v_g - \bar{v}_g) * g_{jk}$. Suppose outside policy makers remove a percentage P of these adverse selection costs. Insurers' expected costs per enrollee, for a given g_{jk} , are adjusted by $P * \int_v MS(g_{jk}, v) * [\gamma_o(v_g - \bar{v}_g) + \gamma_1(v_g - \bar{v}_g) * g_{jk}] dF(v|\sigma)$. When $P = 1$, adverse selection has no effect on insurers' costs and insurers do not have incentives to distort their coverage to deter enrollment by high v_g consumers.⁴⁴

8.2.1 Recalculating Equilibria

Under a policy P , insurers' costs are given by:

$$\begin{aligned} c_{jk}(g, v_g, P) &= [\gamma_o \bar{v} + (1 - P) * \gamma_o (v_g - \bar{v}) + \psi_{jk}^0] \\ &+ [\gamma_1 \bar{v} + (1 - P) * \gamma_1 (v_g - \bar{v}) + \psi_{jk}^1] * g \end{aligned}$$

Holding the number of insurers and plans fixed, a Nash equilibrium can be characterized as a set

$$(Premium^*, Generosity^* | P) \equiv \left\{ \left[(g_{jkm}^*, p_{jkm}^*)_{k=1 \dots n_j} \right]_{j=1 \dots m_j} \right\}_{m=1 \dots M}$$

such that for all jkm , $\frac{\partial Profit_{jm}(Premium^*, Generosity^* | P)}{\partial g_{jkm}} = \frac{\partial Profit_{jm}(Premium^*, Generosity^* | P)}{\partial p_{jkm}} = 0$.

In the data, I observe an equilibrium at $P = 0$. To determine how $(Premium^*, Generosity^*)$ will change in response to a small change in P , I implicitly differentiate the full set of first order conditions with respect to P :⁴⁵

$$\begin{bmatrix} \frac{\partial Premium^*}{\partial P} \\ \frac{\partial Generosity^*}{\partial P} \end{bmatrix} = - \begin{bmatrix} \frac{\partial FOC_{Premium}}{\partial Premium} & \frac{\partial FOC_{Generosity}}{\partial Premium} \\ \frac{\partial FOC_{Generosity}}{\partial Premium} & \frac{\partial FOC_{Generosity}}{\partial Generosity} \end{bmatrix}^{-1} * \begin{bmatrix} \frac{\partial FOC_{Premium}}{\partial P} \\ \frac{\partial FOC_{Generosity}}{\partial P} \end{bmatrix} \quad (10)$$

⁴⁴There are other ways to remove adverse selection. I could allow insurers to charge individuals with different v_g different premiums or I could remove v_g from consumers' preferences. The counterfactual policy I employ seems most appropriate because the CMS has experimented with risk adjustment and it is feasible. In principle, the CMS could implement perfect risk adjustment by compensating insurers ex-post.

⁴⁵Implicit differentiation yields the changes in $(Premium^*, Generosity^* | P)$ in an open neighborhood surrounding P . Conditions for its use are that each first order condition and its derivatives with respect to P and $(Premium, Generosity)$ are continuous. These conditions are satisfied.

Using (10) and $(Premium^*, Generosity^*|P = 0)$, I can find $(Premium^*, Generosity^*|P = \omega)$ for a small ω . I then find $(Premium^*, Generosity^*|P = 2\omega)$. I slowly continue along this path until $P = 1$.⁴⁶ I interpret $(Premium^*, Generosity^*|P = 1)$ as the result of a set of policy changes that slowly (but entirely) eliminate adverse selection.⁴⁷

8.2.2 Equilibrium changes in g and p ⁴⁸

Table 16 regresses equilibrium changes in *generosity* and *premiums* on initial plan and market characteristics. As expected, the simulated equilibria exhibit expanded insurance coverage, particularly for insurers who offer the most generous plans in their market and for those initially enrolling many high v_g enrollees (see column one).

When fifty percent of adverse selection costs are removed ($P=.50$), the mean level of offered *generosity* increases by .09 units. After all adverse selection costs are removed ($P=1$), the mean level of offered *generosity* increases by .23 units. In almost all markets (88%), the maximum amount of offered coverage increases; on average, by .33 units. To provide meaning to these unit changes, recall that dental coverage is equivalent to .64 units of *generosity* and is worth \$13.78 per month to the median consumer.

In the simulated equilibria, I also observe reduced premiums. Mean premiums fall by \$.72 and \$1.19 per month after 50% and all adverse selection costs are removed. These effects are particularly strong for generous plans (see column two of table 16). When adverse selection is present, generous insurers charge premiums that reflect the direct costs of providing generous coverage and the indirect costs of attracting an unhealthy pool of enrollees. After adverse selection is removed, the latter costs no longer impact insurers' incentives. The correlation between *generosity* and *premiums* is equal to .23 in the original equilibrium. After adverse selection is removed, the correlation between *generosity* and *premiums* falls to .10.⁴⁹

When the premium schedule flattens, Medicare beneficiaries enrolled in private plans *consume* more generous insurance. The expected amount of *generosity* consumed increases by .18 units after 50% of adverse selection is removed, and by .42 units when it is fully removed.

8.2.3 Equilibrium Changes in Welfare

The changes in *generosity* and *premiums* will benefit consumers. Medicare beneficiaries initially enrolled in M + C plans receive increased utility from expanded coverage and lower premiums. Beneficiaries originally in traditional Medicare also benefit, since the improved menu of plans makes opting out of traditional

⁴⁶This path does not necessarily exist. If it does, it is not necessarily unique. After each small change in P , *generosity* and *premiums* are adjusted if necessary to make each first order condition hold within a small tolerance. In practice, this process was time consuming, but not difficult.

⁴⁷In this experiment, I do not allow plans to exit the market or new plans to enter. The latter seems more likely to occur. I only allow existing plans to change their characteristics. In future work, I hope to endogenize entry.

⁴⁸The results below are preliminary and may change. As of this version, I have simulated new equilibrium in a randomly chosen 790 markets in my data-set. The aggregate calculations below extrapolate from these 790 markets to the full set of markets.

⁴⁹I can not attribute the mean reductions in premiums to adverse selection, because total government payments to insurers increase. In the future, I plan to implement revenue neutral policies. Adverse selection does explain why generous plans' premiums fall the most in the new equilibrium.

Table 16: Premium and Generosity Changes

	Change in Generosity	Change in Premium
Constant	0.643 (.064)***	0.753 (.40)*
# Insurers	0.012 (.003)***	0.075 (.104)
Initial Generosity	-0.012 (.006)*	-3.03 (.18)***
Initial Mean v_g	0.491 (.067)***	
Max G	0.103 (.028)***	-0.344 (.88)
Max G * Generosity	-0.003 (.009)	-0.423 (.28)
N	3277	3277
R ²	0.446	0.15

Note: Results from an OLS regression. Dependent variables are changes in coverage and premiums after all costs of adverse selection are removed. Initial Generosity and Mean v_g are levels of coverage and enrollees' preferences in original equilibrium. Max G is equal to one if plan was the most generous in its market in original equilibrium. Standard errors in parentheses. *, **, and *** indicate significance at the 5%, 1%, and 0.1% levels

Medicare more attractive. Many exercise this option. After adverse selection is eliminated, the percentage of Medicare beneficiaries enrolled in a M + C plan increases from 23.9% to 35.2%.⁵⁰

Insurers' profits also increase. The risk adjustment policy increases payments to all insurers except those that only attract consumers with weak preferences for insurance. The payment increases offset changes in *generosity* and *premiums* that increase costs and decrease revenues. Let $(Generosity^*, Premium^*)$ and $(Generosity_{No\ Adv\ Sel}^*, Premium_{No\ Adv\ Sel}^*)$ denote the equilibrium distributions of *generosity* and *premiums* before and after adverse selection is removed. The change in total welfare per beneficiary is calculated as follows:^{51,52}

⁵⁰The federal government could potentially save money from these enrollment changes if the cost of enrolling beneficiaries in traditional Medicare is greater than the payment rates made to insurers, even after accounting for risk adjustment. The welfare calculations below assume the cost of treating a beneficiary in traditional Medicare is equal to the payment rate made to insurers. Thus, these calculations may understate welfare changes. On the other hand, some studies have shown that the least costly Medicare beneficiaries are most likely to enroll in a M + C plan (Hellinger & Wong 2000). In this case, my calculations may understate increases in government expenditures, and therefore, overstate welfare changes.

⁵¹In the simulated equilibria, government expenditures increase. In my welfare calculations, I assume there are no welfare losses from increases in taxation. This assumption suggests another reason for me to consider revenue neutral risk adjustment policies.

⁵² $CS(Generosity, Premium)$ is equal to consumer surplus per beneficiary given $(Generosity, Premium)$. $\pi(Generosity, Premium, P)$ is equal to the sum of insurer profits per beneficiary given $(Generosity, Premium)$ and a risk adjustment policy P .

Table 17: Changes in Yearly Total Surplus / Medicare Beneficiary

# Insurers	Original Equilibrium	No Adv. Sel. Equilibrium	Change
1	\$36.49 (5.12)	\$42.95 (5.54)	\$6.46 (.64)
2	\$81.14 (2.82)	\$93.31 (3.05)	\$12.17 (.45)
3	\$109.24 (4.09)	\$125.38 (4.49)	\$16.14 (.74)
4	\$122.39 (7.34)	\$141.84 (7.91)	\$19.45 (1.33)
5	\$112.62 (6.89)	\$132.68 (7.86)	\$20.06 (1.71)
6	\$137.08 (11.84)	\$156.91 (11.85)	\$19.83 (1.75)
>6	\$155.20 (13.04)	\$184.39 (13.21)	\$29.19 (2.61)

Note: Calculations are per Medicare beneficiary and per year. Means are weighted by number of beneficiaries in each market. Values are in 2000 dollars. Standard deviations of means are in parentheses.

$$\Delta TW = \Delta CS + \Delta Profits + \Delta Govt Payments$$

where

$$\Delta CS = CS(Generosity_{No Adv Sel}^*, Premium_{No Adv Sel}^*) - CS(Generosity^*, Premium^*)$$

$$\Delta Profits = \pi(Generosity_{No Adv Sel}^*, Premium_{No Adv Sel}^*, P = 1) - \pi(Generosity^*, Premium^*, P = 0)$$

$$\Delta Govt Payments = \pi(Generosity_{No Adv Sel}^*, Premium_{No Adv Sel}^*, P = 1) - \pi(Generosity_{No Adv Sel}^*, Premium_{No Adv Sel}^*, P = 0)$$

Table 17 displays yearly changes in total surplus after adverse selection is eliminated from insurers' costs by perfect risk adjustment. Total surplus increases most in markets with many insurers. This suggests that welfare losses from adverse selection are larger in markets with more insurers. The mean change in surplus in markets with one insurer is \$6.46 per beneficiary. In markets with six or more insurers, surplus increases by \$29.19 per beneficiary. All surplus changes are per year and per beneficiary.

Table 18 decomposes the surplus changes from table 17 into changes in consumer surplus, total insurer profits, and government expenditures. In each market, government expenditures increase to fund the risk adjustment policy. To remove adverse selection, government expenditures on privatized Medicare increase by 1.73%.⁵³ Consumer surplus and total insurer profits, however, increase by amounts that offset this

⁵³Initial government expenditures on M + C are equal to the number of M+C enrollees multiplied by the payment rate.

Table 18: Decomposing Changes in Total Surplus / Medicare Beneficiary

# Insurers	Change in CS	Change in Profits	Change in Govt Exp
1	\$10.21 (1.39)	\$4.30 (.99)	\$8.06 (1.48)
2	\$19.39 (.71)	\$8.11 (.44)	\$15.34 (.62)
3	\$27.49 (1.17)	\$10.81 (.65)	\$22.16 (.94)
4	\$34.42 (2.58)	\$9.96 (1.27)	\$24.92 (1.49)
5	\$34.10 (3.19)	\$9.41 (1.56)	\$23.45 (1.90)
6	\$37.88 (3.13)	\$14.99 (2.11)	\$33.04 (2.21)
>6	\$54.35 (4.89)	\$15.00 (3.97)	\$40.16 (3.67)

Note: Calculations are per Medicare beneficiary and per year. Values for profits sum across all participating insurers. All means weighted by number of beneficiaries in each market. Values are in 2000 dollars. Standard Deviations of means are in parentheses.

increase in expenditures, and total surplus increases.

Most of the realized surplus gains are enjoyed by Medicare beneficiaries. In markets with one insurer, the average beneficiary realizes an \$10.21 increase in consumer surplus, whereas total insurer profits increase by \$4.30 per beneficiary. In markets with more than one insurer, the relative surplus gains are qualitatively similar. Consumers consistently realize more gains in surplus than insurers after adverse selection is removed.

Summing across months, markets and the number of Medicare beneficiaries in each market, aggregate changes in total surplus are given in table 19. Eliminating the costs of adverse selection increases the total surplus associated with the M + C between 2000 and 2003 by \$1.50 billion. This corresponds to a 16.2% increase. Equivalently, total surplus increases by an amount equal to 1.1% of total payments made by the government to M + C insurers before adverse selection was eliminated.

9 Conclusion

In this paper, I use observed insurer behavior to investigate welfare losses caused by adverse selection in privatized Medicare. I estimate a model of privatized Medicare in which insurers choose how much coverage to offer and what premiums to charge, and consumers vary in their preferences for insurance. The model allows consumers with different preferences to impose different costs on their insurers. To measure adverse selection, I use exogenous variation in market structure to identify a causal relationship between consumers' preferences and insurers' costs. This strategy allows me to infer whether consumers' preferences

After eliminating adverse selection, expenditures increase by the per beneficiary change in profits (from Table 18) multiplied by the number of beneficiaries enrolled in M + C in the new equilibrium. This calculation assumes that the cost of treating a beneficiary in traditional Medicare is equal to the payment rate made to M + C insurers. Thus, when beneficiaries switch between M + C and traditional Medicare, the government's costs do not change.

Table 19: Aggregate Changes in Total Surplus

# Insurers	Original Equilibrium	No Adv. Sel. Equilibrium	Change
1	\$998.908 (140.16)	\$1,175.750 (151.66)	\$176.841 (17.52)
2	\$1,896.539 (65.91)	\$2,180.996 (71.29)	\$284.457 (10.52)
3	\$2,262.368 (84.70)	\$2,596.629 (92.99)	\$334.261 (1,532.55)
4	\$1,140.445 (68.40)	\$1,321.683 (73.71)	\$181.238 12.39
5	\$458.571 (28.06)	\$540.253 (32.00)	\$81.681 (6.96)
6	\$687.105 (59.35)	\$786.502 (59.40)	\$99.397 (8.77)
>6	\$1,811.251 (152.18)	\$2,151.911 (154.17)	\$340.660 (30.46)
Total	\$9,255.188 (251.62)	\$10,753.723 (265.48)	\$1,498.535 (1,533.08)

Note: Changes in TS take changes in TS per beneficiary and multiply by the number of Medicare Beneficiaries in each market. Values in table are constructed by summing over markets and years. Results are in year 2000 millions of \$. Standard deviations of means are in parentheses.

for insurance, which determine how much insurance they consume, contain information about their expected health. My measure is analogous to previous tests of adverse selection in the sense that I relate consumers' choices to insurers' costs. But unlike tests that rely only on consumer behavior, my model of insurer behavior allows me to quantify the welfare losses caused by adverse selection. Market failures caused by adverse selection begin with distortions in insurer behavior. Measurement of the resulting welfare losses, therefore, requires a model of insurer behavior.

My empirical results confirm that consumers with strong preferences for generous insurance are more costly to insure. Using counterfactual simulations, I find that this adverse selection has substantial effects on welfare. I implement policies of perfect risk adjustment that are extreme versions of policies the government has experimented with in privatized Medicare in recent years. These policies adjust government payments to insurers according to the types of consumers they enroll, thereby eliminating the effects of adverse selection on insurers' costs and incentives. The simulated equilibria exhibit expanded insurance coverage and lower premiums. Between 2000 and 2003, total surplus associated with privatized Medicare increases by \$1.50 billion (16.2%), or equivalently, by an amount equal to 1.1% of the total payments made to insurers by the government. These results suggest that expanded efforts by the government to implement risk adjustment in privatized Medicare would increase welfare.

Welfare losses from adverse selection are larger in markets with more insurers. In markets with one insurer, total surplus increases by \$6.46 per Medicare beneficiary in each year after adverse selection is removed. In markets with six or more insurers, surplus increases by \$29.19 per Medicare beneficiary. These results suggest a tradeoff between the welfare losses from adverse selection and the benefits of competition.

Welfare in privatized Medicare could potentially be enhanced if the government limits the number of insurers allowed to enter each market, but requires each insurer to offer more plans. In this paper, I am limited in my ability to consider this possibility because I took market structure as exogenous. In future research, I plan to generalize my model so I can further investigate the tradeoff between welfare losses from adverse selection and the benefits of competition.

It should also be pointed out that this analysis assumes away consequences from consumers having multiple dimensions of private information. This omission may be particularly important in scenarios where consumers with the same preferences, but different characteristics, impose different costs on insurers. If insurers are able to screen along these separate dimensions, the welfare implications may differ from those predicted above. In principle, my approach can be used to study the consequences of such heterogeneity. Thus, this is also a likely avenue for future research.

References

- [1] Achman, L. and Gold, M. "Trends in Medicare + Choice Benefits and Premiums, 1999-2002." Mathematica Policy Research, Inc. November 2002.
- [2] Akerlof, G. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84(3), pp. 488-500
- [3] Atherly, A., Dowd, B., Feldman, R. "The Effects of Benefits, Premiums, and Health Risk on Health Plan Choice in the Medicare Program." Mimeo, Division of Health Services Research and Policy, University of Minnesota, 2003.
- [4] Berry, S., "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics*, Vol. 25 (1994), pp. 242-262.
- [5] Berry, S., Levinsohn, J., and Pakes, A. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Vehicle Market," *Journal of Political Economy*, 112(1), 68-104.
- [6] Buchmueller, T.C. (2000) "The Health Plan Choices of Retirees Under Managed Competition." *Health Services Research*, Vol. 35, pp. 949-976.
- [7] Cardon, J. and Hendel, I. (2001) "Asymmetric Information in Health Care and Health insurance Markets: Evidence from the National Medical Expenditure Survey." *Rand Journal of Economics*, 32(3), 408-427.
- [8] Chiappori, P., and Salanie, B. (2000), "Testing for asymmetric information in Insurance Markets," *Journal of Political Economy* 108(1), 56-78.
- [9] Chiappori, P., Jullien, B., Salanie, B., and Salanie, F. (2006), "Asymmetric information in insurance: General Testable implications," *Rand journal of Economics* 37(4).
- [10] Cutler, D.M. and Reber, S.J. 1998 "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection." *Quarterly Journal of Economics*, Vol. 113 (1998), pp. 433-466.
- [11] Cutler, D.M. and Zeckhauser, R.J. "The Anatomy of Health Insurance.", *Handbook of Health Economics*, Vol. 1 (2000), pp. 564-643.
- [12] Cutler, D.M. and Zeckhauser, R.J. 1997 "Adverse Selection in Health Insurance." NBER Working Paper 6107.
- [13] Fang, H., Keane, M., and Silverman, D. (2006), "Sources of advantageous Selection: Evidence from the Medigap insurance market," mimeo, Yale University.
- [14] Finkelstein, A., and McGarry, K. (2006) "Multiple dimensions of private information: Evidence from the long-term Care Insurance Market," *American Economic Review* 96(4), 938-958.

- [15] Finkelstein, A., and Poterba, J. (2004), "Adverse Selection in insurance Markets; Policyholder evidence from the U.K. Annuity Market," *Journal of Political Economy*, 112(1), 193-208.
- [16] Finkelstein, A. and Poterba, J. (2006), "Testing for Adverse Selection with 'Unused Observables'." NBER Working Paper No. 12112
- [17] Hellinger, F. and Wong, H. (2000) "'Selection Bias in HMOs: A Review of the Evidence.'" *Medical Care Research and Review* 57(4) 405-438.
- [18] Maruyama, S., (2006) "Welfare analysis incorporating a Structural Entry-Exit model: A Case Study of Medicare HMOs", Working Paper, Northwestern University.
- [19] McFadden, D., and Newey, W. (1994) "Large Sample Estimation and Hypothesis Testing", *Handbook of Econometrics*. Volume 4, Ch. 36.
- [20] Medicare Payment Advisory Commission, (2000). "Report to the Congress: Improving Risk Adjustment in Medicare." Washington, DC: MedPac.
- [21] Medicare Payment Advisory Commission, (2001). "Report to the Congress: Medicare Payment Policy." Washington, DC: MedPac.
- [22] Medicare Payment Advisory Commission, (2004). "Report to the Congress: Medicare Payment Policy." Washington, DC: Medpac.
- [23] Moon, M., (2006) "Medicare: A Policy Primer". Urban Institute Press.
- [24] Rothschild, M., and J.E. Stiglitz. (1976) "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect information.", *Quarterly Journal of Economics*, 90(1), 630-649.
- [25] Town, R. and Liu, S., (2000) "The welfare impact of Medicare HMOs," *Rand journal of Economics*, 34(4), 719-736.
- [26] Wooldridge, J. "Econometric Analysis of Cross Section and Panel Data." MIT Press: 2001.

10 Appendix

10.1 Data

The main sources of data are acquired from the Center for Medicare and Medicaid Services. The data primarily comes from three sources: the yearly Medicare Compare database, the December version of the CMS's State-County-Plan files, and the MCBS survey. In the CMS data files, there is a clear distinction between organizations, contracts, and products. The organization is the broadest of the three and is the company responsible for overseeing each contract. For example, Pacificare, Aetna, and Humana are all organizations often observed within the data. In the Medicare program, each organization enters into one or more contracts with the CMS for the purpose of providing health care to Medicare beneficiaries. Each contract is assigned a unique contract number by the government. Often times, organizations enter into multiple contracts with the government, with different contracts serving different markets. Other times, the same contract extends across markets. Within each contract, organizations offer, and Medicare beneficiaries can enroll in, multiple products. Products within a given contract can vary along multiple dimensions, including premiums, copayments, coverage types, etc. In the main text, I assume organizations' behavior is independent across markets. I refer to a contract as a firm or insurer and a product as a plan.

10.1.1 Identifying Market Entrants

To identify insurers, I focus on the contract level. The State-County-Plan files disaggregate enrollment data at the contract level into enrollment data at the county level. For each Medicare HMO contract, I observe which counties have a positive number of contract enrollees. If this positive number is greater than ten, I observe actual enrollment information. Using the SCP data alone, however, greatly overstates entry by Medicare HMOs. When initially enrolling in a Medicare HMO, beneficiaries are only able to choose among those plans offered within their county. But once enrolled, beneficiaries can remain in a particular plan even after changing residences. Thus, within the SCP files, there are many counties that contain information on the enrollment by Medicare beneficiaries in a particular contract, despite that contract not being available within the county. To enter the final sample, a contract must be listed as being available in the Medicare Compare database, and have a market share of at least one percent.

10.1.2 Omitted Plan Types

Under the Medicare + Choice program, several types of private health insurance plans other than Medicare HMOs are eligible to contract with the CMS to provide coverage to Medicare beneficiaries. Other types of organizations include Preferred Provider Organizations (PPO), Provider Sponsored Organizations (PSO), Private Fee For Service Plans (PFFS), and cost contract HMOs. From a beneficiary's point of view, these plans are full-fledged alternatives to Medicare HMOs as long as they are available in the beneficiary's market. They are included in the model.

Other types of plans, however, are excluded from supply and demand. These types include Health Care Prepayment Plans (HCPP), which cover only outpatient services, Programs of All-Inclusive Care for the Elderly (PACE), which are combination programs with Medicaid that provide comprehensive community and medical services (to be enrolled in a PACE plan a Medicare beneficiary needs to be certified as eligible for nursing home care by the appropriate state agency), and demonstration plans (DEMO) which are designed to evaluate the effects and impacts of various health care initiatives.

10.1.3 Plan Characteristics

Definitions of variables entering Medicare beneficiaries utility are provided below. All product characteristics are extracted from the Medicare Compare database.

Premium: Premium information is provided at the monthly level. They do not include the required payment for Medicare Part B, which is charged to all Medicare HMO enrollees. For a few plans, premium information was unavailable. These plans were dropped from the final sample.

\$InpatientHospital: Required payments for inpatient hospital stays varied across plans in structure. Some described required copayments for days 0 – 20 and 20 – 100, others had additional tiers. This variable is equal to the charged copayment upon admission to an inpatient hospital. For some plans, this information is explicitly provided in the Medicare Compare Database. For others, the day 1 copayment is used.

\$Specialist: Required copayment for each visit to a specialist. Across plans, the format of this data is constant.

Dental: This is a dummy variable indicating whether any supplementary dental coverage is described in the Medicare compare database. Traditional Medicare offers no dental coverage. For some plans, additional information on dental coverage was provided, such as annual deductibles, dentist copayments, etc. But much of this information was incomplete and its structure varied extensively across plans. If no mention of Dental coverage was made, I assumed there was none.

Vision: This is a dummy variable indicating whether any supplementary vision coverage is described in the Medicare compare database. Traditional Medicare offers very little coverage of vision services. The formats of the data provided was very similar in nature to *Dental*.

PrescriptionDrugs: Five prescription drug variables were used in demand. First, *PDBrand* and *PDGeneric* are dummy variables indicating whether each plan offered any form of generic and branded prescription drug coverage. Between 2000 and 2003, traditional Medicare offered no such coverage. *\$BrandPD* is equal to the copayment for a 31 day supply of a brand prescription drug. Some plans described this copayment using different lengths of time. For these, all copayments were prorated to 31 days. The generosity of provided prescription drug coverage is captured by *UL_BrandPD* and *UL_GenericPD*. Both are dummy variables indicating whether generic and brand coverage are unlimited. Most plans offered only limited drug coverage, for example \$1000 annually. Because of variation in the format of coverage data across plans, it was difficult to construct additional variables detailing the generosity of any provided drug coverage. If for

each plan, any mention is made of unlimited generic or brand drug coverage these indicator variables are set to one.

10.2 Firms' Choice Variables

The model assumes insurers choose generosity levels, g , and not the determinants of generosity, x . Note that in insurers' costs and in utility the individual elements of x_{jk} are perfect substitutes. This restriction implies that it is without loss of generality to assume that insurers choose g . Suppose that insurers did choose x to maximize:

$$\int_v MS(g(x), p, v) * [p - c(v, x)] dF(v)$$

where:

$$\begin{aligned} g(x) &= \beta'x \\ c(v, x) &= (\gamma_0 v + \psi_0) * [\lambda'x] \end{aligned}$$

Claim 1 *If insurers are behaving optimally and for all m and n there exists plans k and k' such that $\frac{x_{kn}}{x_{km}} \neq \frac{x_{k'n}}{x_{k'm}}$ then $\frac{\beta_n}{\beta_m} = \frac{\lambda_n}{\lambda_m}$ for all n and m .*

Proof. Assume not. Then for some n and m , $\frac{\beta_n}{\beta_m} > \frac{\lambda_n}{\lambda_m}$. By assumption, there exist plans k and k' such that $\frac{x_{kn}}{x_{km}} < \frac{x_{k'n}}{x_{k'm}}$ (i.e. plans k and k' offer x_m and x_n in different ratios). If the characteristics in k' are optimally chosen, then the characteristics in k are not. Plan k can earn higher profits by increasing x_{kn} and decreasing x_{km} . Consider a one unit increase in x_{kn} and a $\frac{\beta_n}{\beta_m}$ decrease in x_{km} . This shift in characteristics leaves g_k unchanged. Therefore $MS(g(x_k), p_k, v)$ is unchanged for all v . But profit margins will increase. The effect on costs is negative since:

$$\begin{aligned} \lambda_n \Delta x_{kn} + \lambda_m \Delta x_{km} &= \\ \lambda_n - \lambda_m \frac{\beta_n}{\beta_m} < 0 &\iff \\ \frac{\beta_n}{\beta_m} > \frac{\lambda_n}{\lambda_m} & \end{aligned}$$

The last line holds by assumption. ■

Thus, after normalization, insurers' cost functions are equivalent to $c(v, g) = (\gamma_0 v + \psi_0) * g$. Firms differentiate themselves via choices of overall insurance coverage, not the individual elements of x . It can be verified for each n and m that sufficient variation in product characteristics exists in the data so that the

claim's assumption is satisfied.

10.3 Robustness to Multi-dimensional Types

Suppose that consumers' preferences for generosity are determined by their type $t_i \equiv (t_{i1}, t_{i2}, \dots, t_{in})$ so that (after omitting irrelevant terms) utility is:

$$u_{ijk} = v_{gi}(t_{i1}, t_{i2}, \dots, t_{in}) * g_{jk} + \alpha p_{jk} + \xi_j^m + \varepsilon_{ijk}^m$$

Consumers' preferences for insurance now explicitly depend on their characteristics such as risk type, risk aversion, cognitive ability, etc. The restriction implicit in this specification is that insurers are unable to screen between consumers with the same v but different t . This assumption is admittedly strong. Once it is made, however, it is without loss of generality under certain assumptions to consider the more general, and realistic, cost function below:

$$\begin{aligned} c_{jk}(t_1, t_2, \dots, t_n, g, \Theta^{Supply}) &= (\rho_{01}t_1 + \dots + \rho_{0n}t_n + \psi_{jk}^0) + \\ &(\rho_{11}t_1 + \dots + \rho_{1n}t_n + \psi_{jk}^1) * g \end{aligned}$$

This allows individuals with the same preferences, but different characteristics, to impose different costs on insurers. Because an insurer can not discriminate between consumers with distinct t , but the same v , their cost function reduces to $E[c_{jk}(t_1, t_2, \dots, t_n, g, \Theta^{Supply})|v]$. This can be seen by simply rearranging the insurer's profit function:

$$\begin{aligned} \int_{t_1} \dots \int_{t_n} \sum_{k=1}^{n_j} MS_{jk}(g, p, v(t_1, \dots, t_n) | G, P, \Theta^{Demand}) * [S_m + p_{jk} - c_{jk}(t, g, \Theta^{Supply})] dH(t_1, \dots, t_n) &= \\ \int_{\tilde{v}} \int_{t:v(t)=\tilde{v}} \sum_{k=1}^{n_j} MS_{jk}(g, p, \tilde{v} | G, P, \Theta^{Demand}) * [S_m + p_{jk} - c_{jk}(t, g, \Theta^{Supply})] dH(t_1, \dots, t_n | v(t) = \tilde{v}) &= \tilde{v} dF(\tilde{v}) = \\ \int_{\tilde{v}} \sum_{k=1}^{n_j} MS_{jk}(g, p, \tilde{v} | G, P, \Theta^{Demand}) * [S_m + p_{jk} - E[c_{jk}(t, g, \Theta^{Supply})|v(t) = \tilde{v}]] dF(\tilde{v}) &= \end{aligned}$$

Thus, ignoring the determinants of preferences for insurance can be made without loss of generality if:

$$\begin{aligned} E[c_{jk}(t, g, \Theta^{Supply})|v(t) = \tilde{v}] &= c_{jk}(v, g, \Theta^{Supply}) = \\ &(\gamma_0 v + \psi_{jk}^0) + (\gamma_1 v + \psi_{jk}^1) * g \end{aligned}$$

where

$$E[c_{jk}(t_1, t_2, ..t_n, g, \Theta^{Supply})|v] = (\rho_{01}E[t_1|v] + \dots + \rho_{0n}E[t_n|v] + \psi_{jk}^0) + (\rho_{11}E[t_1|v] + \dots + \rho_{1n}E[t_n|v] + \psi_{jk}^1) * g$$

Under specific conditions on the CDF of t , $H(t_1, ..t_n)$, this statement is true. Note that $c_{jk}(v, g, \Theta^{Supply})$ is linear in v and has an intercept $\psi_{jk}^0 + \psi_{jk}^1 * g$ when $v = 0$. Assume that $v(t) = 0$ if and only if $t_m = 0$ for each m . Thus, $E[c_{jk}(t_1, t_2, ..t_n, g, \Theta^{Supply})|v]$ also has an intercept $\psi_{jk}^0 + \psi_{jk}^1 * g$ when $v = 0$. Now, note that $c_{jk}(v, g, \Theta^{Supply})$ is linear in v with slope $\gamma_0 + \gamma_1 * g$. $E[c_{jk}(t_1, t_2, ..t_n, g, \Theta^{Supply})|v]$ is also linear in v if $\frac{dE[t_m|v]}{dv} = \Delta_m$ for all v and each m . For example, if $v_g(t_1, .., t_n) = t_1 + .. + t_n$ and $t_n \sim |N(0, \sigma_{tn}^2)|$ for each n and $t_{n'}$ is uncorrelated with each t_n , this linearity condition holds. The slope of $E[c_{jk}(t_1, t_2, ..t_n, g, \Theta^{Supply})|v]$ in v is now equal to $(\rho_{01}\Delta_1 + \dots + \rho_{0n}\Delta_n) + (\rho_{11}\Delta_1 + \dots + \rho_{1n}\Delta_n) * g$. Thus, the restricted model, with $c_{jk}(v, g, \Theta^{Supply})$ can duplicate the more general model if $\gamma_0 = \rho_{01}\Delta_1 + \dots + \rho_{0n}\Delta_n$ and $\gamma_1 = \rho_{11}\Delta_1 + \dots + \rho_{1n}\Delta_n$.

10.4 Estimation

10.4.1 Demand

Construction of Demand Moments Under the assumption that the horizontal shock to preferences, $\varepsilon_{ij'k}^m$, is a type-1 logit error, the probability that individual i enrolls in $j'k'$ in market m , conditional on i 's unobserved type, v_i , can be written as:

$$\begin{aligned} Prob(j'k'|v_i, m, \Theta) &= \frac{e^{(1+\nu_{1i})*g_{j'k'} - (\alpha + \theta_p I_m + \nu_{pi})*p_{j'k'} + \xi_{j'}}}{1 + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{(1+\nu_{1i})*g_{jk} - (\alpha + \theta_p I_m + \nu_{pi})*p_{jk} + \xi_j}} \\ &= \frac{e^{u_{ij'k'} + \xi_{j'}}}{1 + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}} \end{aligned}$$

where Ω_m is the set of firms that have entered market m and Ω_{jm} is the set of products offered by each firm j that is participating in market m .

Three other probability functions are used in estimation. The probability that consumer i enrolls in a product offered by firm j' in market m can be written by summing over $Prob(j'k'|v_i)$ for each $k' \in \Omega_{j'm}$:

$$Prob(j'|v_i, m, \Theta) = \frac{\sum_{k' \in \Omega_{j'm}} e^{u_{ij'k'} + \xi_{j'}}}{1 + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}}$$

The probability of i enrolling in product k' conditional on enrolling in firm j' can be written as:

$$Prob(k'|j', v_i, m, \Theta) = \frac{e^{u_{ij'k'}}}{\sum_{k \in \Omega_{mj'}} e^{u_{ijk}}}$$

Note that firm level $\xi_{j'}$ falls out of $Prob(k'|j', v_i, m, \Theta)$.

The probability of i enrolling in product $j'k'$ in market m' , and then product $\tilde{j}\tilde{k}$ one year later in market \tilde{m} can be written as:

$$Prob(j'k', \tilde{j}\tilde{k}|v_i, m', \tilde{m}, \Theta) = \frac{e^{u_{ij'k'} + \xi_{j'}}}{1 + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}} * \frac{e^{u_{i\tilde{j}\tilde{k}} + \xi_{\tilde{j}'}}}{1 + \sum_{j \in \Omega_{\tilde{m}}} \sum_{k \in \Omega_{\tilde{m}j}} e^{u_{ijk} + \xi_j}}$$

It is necessary to work with the unconditional share probabilities. These are obtained simply by integrating over $Prob(j'k'|v_i)$ and $Prob(j|v_i)$ with respect to v :

$$Prob(j'k'|m, \Theta) = \int_v \frac{e^{u_{ij'k'} + \xi_{j'}}}{e^{u_{i0}} + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}} dF(v|\sigma)$$

$$Prob(j'|m, \Theta) = \int_v \frac{\sum_{k' \in \Omega_{j'm}} e^{u_{ij'k'} + \xi_{j'}}}{e^{v_{0i}} + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}} dF(v|\sigma)$$

$$Prob(k'|j', m, \Theta) = \int_v \frac{e^{u_{ij'k'} + \xi_{j'}}}{\sum_{k \in \Omega_{mj'}} e^{u_{ijk} + \xi_j}} dF(v|\sigma)$$

$$Prob(j'k', \tilde{j}\tilde{k}|m, \Theta) = \int_v \frac{e^{u_{ij'k'} + \xi_{j'}}}{1 + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}} * \frac{e^{u_{i\tilde{j}\tilde{k}} + \xi_{\tilde{j}'}}}{1 + \sum_{j \in \Omega_{\tilde{m}}} \sum_{k \in \Omega_{\tilde{m}j}} e^{u_{ijk} + \xi_j}} dF(v|\sigma)$$

With a set of L random draws from $f(v|\sigma)$ I simulate the above probabilities with $\widehat{Prob}(j'k'|m, \Theta) = \frac{1}{L} \sum_{l=1}^L \frac{e^{u_{ij'k'} + \xi_{j'}}}{e^{u_{i0}} + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}}$, $\widehat{Prob}(j'|m, \Theta) = \frac{1}{L} \sum_{l=1}^L \frac{\sum_{k' \in \Omega_{j'm}} e^{u_{ij'k'} + \xi_{j'}}}{e^{v_{0i}} + \sum_{j \in \Omega_m} \sum_{k \in \Omega_{mj}} e^{u_{ijk} + \xi_j}}$, and $\widehat{Prob}(k'|j', m, \Theta) = \frac{1}{L} \sum_{l=1}^L \frac{e^{u_{ij'k'} + \xi_{j'}}}{\sum_{k \in \Omega_{mj'}} e^{u_{ijk} + \xi_j}}$.

All moments are constructed using these simulated probabilities.

Three forms of moment conditions are used in estimation. The population moments, described in the main text are repeated below:

$$E[x_{jk} * (d_{ijk} - Prob(k|j, m, \Theta))|ij] = 0 \quad (11)$$

$$E[x_{jkl} * x'_{j'k'} * (d_{ijk, j'k'} - Prob(jk \& j'k'|m, m', \Theta))|i] = 0 \quad (12)$$

$$E[f(x_{-jk}|x_{jk}) * (d_{ijk} - Prob(k|j, m, \Theta))|ij] = 0 \quad (13)$$

The corresponding sample moments are:

$$\begin{aligned} & \frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{x_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}] = 0 \\ & \frac{1}{M} \sum_{i=1}^M \frac{1}{|\Omega_m| * |\Omega_{m'}|} \sum_{jk \in \Omega_m} \sum_{j'k' \in \Omega_{m'}} x_{jk} x'_{j'k'} * (Prob(jk, j'k'|m, \Theta) - d_{i,jk,j'k'}) = 0 \\ & \frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{f(x_{-jk}|x_{jk})}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}] = 0 \end{aligned}$$

Above, j sums over insurers, k over products and i over MCBS respondents. These summations are divided through by $\sqrt{n_{jm}}$ instead of n_{jm} to ensure that the variance of residuals interacted with instruments ($\frac{1}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}]$) does not depend on the number of survey respondents from each firm. This leads to more weight being placed on contracts for whom we observe more individuals enrolling.

Accounting for Simulation Error Consider a sample moment condition employed in demand:

$$\frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}] = 0$$

where $Prob(k|j, m, \Theta)$ is the true probability of enrolling in product k conditional on enrolling in firm j . Rewrite this moment condition, using $\widehat{Prob}_m(k|j, m, \Theta)$ to represent simulated probabilities:

$$\begin{aligned} & \frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) \\ & - \widehat{Prob}(k|j, m, \Theta) + \widehat{Prob}(k|j, m, \Theta) - d_{ijk}] = \\ & \frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - \widehat{Prob}(k|j, m, \Theta)] \\ & + \frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_{jk}}{\sqrt{n_{jm}}} \sum_i [\widehat{Prob}(k|j, m, \Theta) - d_{ijk}] = 0 \end{aligned}$$

Because simulation error is independent of the randomness inherent to $\sum_i [\widehat{Prob}(k|j, m, \Theta) - d_{ijk}]$, the variance of the sample moment is equal to:

$$\begin{aligned}
& Var \left[\frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - \widehat{Prob}(k|j, m, \Theta)] \right] + \\
& Var \left[\frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_{jk}}{\sqrt{n_{jm}}} \sum_i [\widehat{Prob}(k|j, m, \Theta) - d_{ijk}] \right]
\end{aligned}$$

The first term is variance due to simulation error. Note that $\widehat{Prob}_m(k|j, m, \Theta)$ is the sum of $nsim$ draws of $Prob(k|j, v_i, m, \Theta)$ where v_i is distributed according to $F(v|\sigma)$. Because $E(Prob(k|j, v_i, m, \Theta)) = Prob(k|j, m, \Theta)$, the delta method implies:

$$\widehat{Prob}_m(k|j, v_i, m, \Theta) - Prob(k|j, m, \Theta) \sim N\left(0, \frac{\partial \widehat{Prob}_m(k|j, v_i, m, \Theta)}{\partial v} \sigma \frac{\partial \widehat{Prob}_m(k|j, v_i, m, \Theta)}{\partial v'}\right)$$

Using this result, it is straightforward to construct the variance of

$\frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - \widehat{Prob}_m(k|j, m, \Theta)]$. Analogous procedures can be used for the remaining moments.

Selection in Demand Moments Consider a moment condition that interacts Z_{jk} with the random residuals, $\sum_i [Prob(k|j, m, \Theta) - d_{ijk}]$. Assume for a particular firm j that $Z_{jk} = Z_j$ for all k . That firm's contribution to the sample moment condition is equal to zero, regardless of the choice of parameters:

$$\begin{aligned}
& \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{Z_j}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}] = \\
& \frac{Z_j \sqrt{n_{jm}}}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{1}{n_{jm}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}] = \\
& \frac{Z_j}{|\Omega_{jm}|} \frac{1}{\sqrt{n_{jm}}} \sum_{k=1}^{|\Omega_{jm}|} (Prob(k|j, m, \Theta) - \frac{1}{n_{jm}} \sum_i d_{ijk}) = \\
& \frac{Z_j}{|\Omega_{jm}|} \frac{1}{\sqrt{n_{jm}}} * \left[\sum_{k=1}^{|\Omega_{jm}|} (Prob(k|j, m, \Theta) - \sum_{k=1}^{|\Omega_{jm}|} \frac{1}{n_{jm}} \sum_i d_{ijk}) \right] = \\
& \frac{Z_j}{|\Omega_{jm}|} \frac{1}{\sqrt{n_{jm}}} * [1 - 1] = 0
\end{aligned}$$

To model this selection issue, let s_{jZ} be an indicator variable that equals one if and only if there is variation in Z_{jk} across k . Interacting s_{jZ} with Z and yields the following moment condition:

$$\frac{1}{N_j} \sum_{j=1}^{N_j} \frac{1}{|\Omega_{jm}|} \sum_{k=1}^{|\Omega_{jm}|} \frac{s_{jZ} * Z_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}]$$

This sample moment condition should equal zero in expectation only if for each firm,

$$E \left[\sum_{k=1}^{|\Omega_{jm}|} \frac{s_{jZ} * Z_{jk}}{\sqrt{n_{jm}}} \sum_i [Prob(k|j, m, \Theta) - d_{ijk}] \right] = 0$$

The discussion in Wooldridge (pg 573), argues that the selection case here is of the most favorable kind. Note that s_{jZ} is a deterministic function of the instruments themselves, and therefore $E[s_{jZ}(Z) * Z_{jk} * \epsilon_{jk}] = 0$ if $E[\epsilon_{jk} | Z] = 0$ via iterated expectations. That is indeed the case, and therefore, the sample moment conditions still equal zero in expectation.

10.4.2 Supply

Recall that in each market the collection of plan characteristics, $\{g_{jk}, p_{jk}\}_{jk \in \Omega_m}$ satisfy a Nash Equilibrium. Therefore, all firms' first order conditions with respect premiums and insurance generosity must equal zero:

$$\frac{\partial Profit_j}{\partial g_{jk}} = \frac{\partial Profit_j}{\partial p_{jk}} = 0 \quad \forall j, k$$

Backing out cost shocks ψ Firms' first order conditions for p_{jk} and g_{jk} reduce to the following set of linear equations. First, for the first order condition on premium:

$$\begin{aligned} \frac{\partial Profit_{jkm}}{\partial p_{jk'}} &= c_{p_{jk'}}^0 + \sum_{k=1}^{m_j} \sum_{z=0}^2 c_{p_{jk'}}^{k,z} * \psi_{jk}^z = 0 \\ &\text{where} \\ c_{p_{jk'}}^0 &= \int_v \sum_{k=1}^{m_j} \frac{\partial MS_{jk}(g, p, v | G, P, \Theta^{Demand})}{\partial p_{jk'}} * [S_m + p_{jk} - \gamma'_0 * v] \\ &\quad - \gamma'_1 * v * g_{jk} dF(v, \vartheta) \\ &\quad + \int_v MS_{jk'}(g, p, v | G, P, \Theta^{Demand}) dF(v | \sigma) \\ c_{p_{jk'}}^{k,0} &= - \int_v \frac{\partial MS_{jk}(g, p, v | G, P, \Theta^{Demand})}{\partial p_{jk'}} dF(v | \sigma) \quad \forall k \\ c_{p_{jk'}}^{k,1} &= -g_{jk} \int_v \frac{\partial MS_{jk}(g, p, v | G, P, \Theta^{Demand})}{\partial p_{jk'}} dF(v | \sigma) \quad \forall k \end{aligned}$$

and below is the first order condition on quality:

$$\frac{\partial Profit_{jkm}}{\partial g_{jk'}} = c_{g_{jk'}}^0 + \sum_{k=1}^{m_j} \sum_{z=0}^1 c_{g_{jk'1}}^{k,z} * \psi_{jk'}^1 = 0$$

where

$$c_{g_{jk'1}}^0 = \int_v \sum_{k=1}^{m_j} \frac{\partial MS_{jk}(g, p, v | G, P, \Theta^{Demand})}{\partial g_{jk'1}} * [S_m + p_{jk} - \gamma'_0 * v - \gamma'_1 * v * g_{jk'}] dF(v, \vartheta) - \int_v MS_{jk'}(g, p, v | G, P, \Theta^{Demand}) * (\pi_1 + \gamma'_1 * v + \delta'_1 * [X_m Z_{jm}]) dF(v, \vartheta)$$

$$c_{g_{jk'1}}^{k,1} = -g_{jk} \int_v \frac{\partial MS_{jk}(g, p, v | G, P, \Theta^{Demand})}{\partial g_{jk'}} dF(v, \vartheta) \quad \forall k \neq k'$$

$$c_{g_{jk'1}}^{k',1} = -g_{jk'} \int_v \frac{\partial MS_{jk'}(g, p, v | G, P, \Theta^{Demand})}{\partial g_{jk'}} dF(v, \vartheta) - \int_v MS_{jk'}(g, p, v | G, P, \Theta^{Demand}) dF(v, \vartheta)$$

By inspection, these linear equations are non-redundant. Backing out unobserved cost shocks is done via the following matrix algebra:

$$\begin{pmatrix} \tilde{\psi}_{j1}^0 \\ \tilde{\psi}_{j1}^1 \\ \cdot \\ \cdot \\ \tilde{\psi}_{jm_j}^0 \\ \tilde{\psi}_{jm_j}^1 \end{pmatrix} = -1 * \begin{pmatrix} c_{p_{j1}}^{1,0} & c_{p_{j1}}^{1,1} & \cdot & \cdot & c_{p_{j1}}^{m_j,0} & c_{p_{j1}}^{m_j,0} \\ c_{g_{j1}}^{1,0} & c_{g_{j1}}^{1,1} & \cdot & \cdot & c_{g_{j1}}^{m_j,0} & c_{g_{j1}}^{m_j,1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{p_{jm_j}}^{1,0} & c_{p_{jm_j}}^{1,1} & \cdot & \cdot & c_{p_{jm_j}}^{m_j,0} & c_{p_{jm_j}}^{m_j,1} \\ c_{g_{jm_j}}^{1,0} & c_{g_{jm_j}}^{1,1} & \cdot & \cdot & c_{g_{jm_j}}^{m_j,0} & c_{g_{jm_j}}^{m_j,1} \end{pmatrix}^{-1} * \begin{pmatrix} c_{p_{j1}}^0 \\ c_{g_{j1}}^0 \\ \cdot \\ \cdot \\ c_{p_{jm_j}}^0 \\ c_{g_{jm_j}}^0 \end{pmatrix}$$

Construction of Supply Side Moments Supply side moments are constructed by exploiting assumed population moment conditions relating firms' unobserved costs shocks to instruments. These conditions are:

$$\begin{aligned} E[\psi_{jk}^n] &= 0 \quad \forall jkn \\ E[X_m * \psi_{jk}^n] &= 0 \quad \forall jkn \\ E[Z_j * \psi_{jk}^n] &= 0 \quad \forall jkn \end{aligned}$$

$$\begin{aligned}
E[f(Z'_{-jm}) * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[N_{Prod,m} * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[N_{Firm,m} * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[f(Z'_{-jm}) * N_{Prod,m} * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[f(Z'_{-jm}) * N_{Firm,m} * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[\xi_j * \psi_{jk}^n] &= 0 \quad \forall j, k, n \\
E[f(\xi_{-j}) * \psi_{jk}^n] &= 0 \quad \forall j, k, n
\end{aligned}$$

The corresponding sample moment conditions are below. Notation is the same as above. Within each firm, the residual is constructed as $\frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n$ so that the variance does not depend on the number of products offered by firm j , m_j .

$$\begin{aligned}
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(\frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(X'_m * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(Z'_{jm} * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(f(Z'_{-jm}) * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(N_{Pr od,m} * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(N_{Firm,m} * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(f(Z'_{-jm}) * N_{Pr od,m} * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(f(Z'_{-jm}) * N_{Firm,m} * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(\xi_j * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1 \\
\frac{1}{N_j} \sum_{j=1}^{N_j} \left(f(\xi_{-j}) * \frac{1}{\sqrt{m_j}} \sum_{k=1}^{m_j} \psi_{jk}^n \right) &= 0 \quad \forall j, k, n = 0, 1
\end{aligned}$$

Simulation Error in Supply Moments Simulation error is also present in the supply moments because they are constructed from cost shocks ψ which are non-linear functions of the simulated market share functions. Because the simulation error enters each moment condition in a non-fashion, biases may be present. Changing the number of simulation draws, however, does not have large effects on parameter estimates. To correct variance estimates for this simulation error, I repeatedly take samples of simulation draws to simulate the variation of each moment condition due to simulation. As above, this variation is independent of variation in the sample moments due to observable data, and in practice quite small. All reported standard errors are corrected for this simulation error.

10.5 Properties of $\hat{\Theta}$

Following Newey & McFadden (1994) the following conditions are necessary for $\hat{\Theta} \xrightarrow{p} \Theta_0$, where $\hat{\Theta}$ minimizes $\Omega_n(m, \Theta) \equiv (M_n(\Theta))' * W * M_n(m, \Theta)$.

1. $\Omega(m, \Theta)$ is uniquely minimized at Θ_0
2. The parameter space Θ is compact.
3. $\Omega(m, \Theta)$ is continuous
4. $\Omega_n(m, \Theta)$ converges uniformly in probability to $\Omega(m, \Theta)$.

The first two conditions are assumed. $M_n(\Theta)$ is highly non-linear, thus proving identification is a challenge. Intuition for identification, however, is provided throughout the main text. Condition 3 is satisfied by simple inspection of moment conditions. All market share functions entering demand moments are continuous functions over the entire parameter space Θ . Similarly, firms' first order conditions are all continuously differentiable at all Θ . Thus, the unobserved cost shocks, ψ are continuous functions of Θ since inversion preserves continuity. Because $M_n(\Theta)$ is continuous for all Θ , so is $M_0(m, \Theta)$ and $\Omega(m, \Theta)$. The fourth condition is true under standard regularity conditions.

10.6 Alternative Specifications

The parameter estimates in table 12 are for a cost function that is linear in *generosity*. Equation (14), below, allows costs to be convex in *generosity*.

$$\begin{aligned}
 c_{jk}(v_g, g, \Theta^{Supply}) &= FC_{jk}(v_g, \Theta^{Supply}) + MC_{jk}(v_g, \Theta^{Supply}) * g^\phi & (14) \\
 &= (\gamma_0 v_g + \tilde{\psi}_{jk}^0) + (\gamma_1 v_g + \tilde{\psi}_{jk}^1) * g^\phi
 \end{aligned}$$

If I mistakenly assume $\phi = 1$ when $\phi > 1$, my model may generate adverse selection when none exists. When $\phi = 1$, adverse selection makes $c_{jk}(v_g, g, \Theta^{Supply})$ convex in g because increases in g attract more

Table 20: Utility Parameter Estimates

	(2)
<u>Plan Characteristics</u>	
Dental	0.5291 (.1054)***
Vision	0.2204 (.1129)*
\$ Specialist	-0.0811 (.0778)
\$ Inpatient Hosp Admission	-0.3303 (.0909)***
Generic PD	1.1989 (.1128)***
Brand PD	(.1582) (.1526)
\$ Brand PD Copay	(-.0649) (0038)***
Generic PD Unlimited	0.1331 (.0708)*
Brand PD Unlimited	1.276 (.1664)***
Premium	-3.179 (.1199)***
<u>Vertical Preferences</u>	
Sigma ² Generosity	0.3922 (.0634)***
Sigma ² Premium	(.0108) (.003)***
Market Income _{Premium}	-0.0097 (.003)***

Notes: Standard errors in parentheses. * Significant at 5% level, ** at 1% level, *** at 0.1% level

costly pools of enrollees. Tables 20 and 21 provide estimates of the parameters that enter utility and costs when $c_{jk}(v_g, g, \Theta^{Supply})$ is given by (14). The parameter results suggest that independent of adverse selection, costs are approximately linear in *generosity* ($\hat{\phi} = 1.03$). The parameter estimates still imply that adverse selection is present.

Table 21: Cost Parameter Estimates

	<u>Fixed Cost</u>	<u>Marginal Cost of Generosity</u>
Constant	14.936 (6.508)*	12.52 (.13)**
Phi		1.034 (.025)**
<u>Adverse Selection</u>		
v_g	13.052 (5.4926)***	5.0288 (.6689)***
<u>Market Controls</u>		
AB Rate / 10	9.16 (.14)***	0.002 (.003)
AB Rate 1997 / 10	-0.392 (12)***	-0.004 (.002)*
Population / 1000	-0.02 (.02)	-0.001 (.0003)***
Per Capita Income / 1000	-4.04 (1.26)***	3.29 (.025)***
# Hospitals	-0.25 (.14)*	-0.0001 (.003)
2001 Dummy	21.61 (1.73)***	0.13 (.03)***
2002 Dummy	21.92 (2.11)***	0.225 (.041)***
2003 Dummy	31.21 (1.85)***	0.067 (.04)*
<u>Insurer Controls</u>		
HMO Dummy	3.78 (2.76)	0.032 (.05)
Total Assets (\$million)	-2.84 (.98)**	-0.01 (.02)
Net Income (\$million)	-0.02 (.02)	0.001 (.0004)***
# Network Physicians / Enrollee	0.63 (.16)***	0.015 (.003)***
% Administrative Expenses	-8.68 (1.23)***	-0.091 (.02)***
Medical Expenditures / Enrollee	0.003 (.01)	-0.0004 (.001)
% Business Medicare	-2.18 (.39)***	0.003 (.008)

Notes: Insurer controls lagged one year. AB Rate and AB Rate 1997 are current and 1997 government payment rates to insurers. Other variables defined in tables 2 and 4. Estimates of the parameters attached to market and insurer controls obtained from an OLS regression using implied costs as dependent variables. *, **, *** indicate significance at 5%, 1%, and 0.1% levels. Standard Errors are in parentheses.