D I S S E R T A T I O N

RAND

Does the Medicare Principal Inpatient Diagnostic Cost Group Model Adequately Adjust for Selection Bias?

Hongjun Kan

RAND Graduate School

The RAND Graduate School dissertation series reproduces dissertations that have been approved by the student's dissertation committee.

RAND is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND® is a registered trademark. RAND's publications do not necessarily reflect the opinions or policies of its research sponsors.

Published 2002 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 102, Pittsburgh, PA 15213
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information,
contact Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

PREFACE

This study examines bias in the Medicare Principal Inpatient Diagnostic Cost Group (PIP-DCG) model due to unobserved selection using HMO and FFS hospital use data. It found that unobserved selection is systematically different in the FFS and HMO populations, with HMO enrollees healthier and FFS beneficiaries sicker in ways not captured by the PIP-DCG model. As a result, the FFS-based model overestimates HMO enrollees' health care resource use compared to their use if they had been served in FFS.

This research should be of interest to researchers and policymakers who are interested in risk adjustment methodologies and are concerned with Medicare overpayments to Medicare+Choice health plans. Previous researchers have not been able to assess the PIP-DCG model using actual utilization data of HMO enrollees. The unique data set in this study comes from a RAND research project that linked California hospital discharge data to Medicare administrative data. The dissertation was completed in partial fulfillment of the requirements of the RAND Graduate School for the degree of Doctor of Philosophy in Policy Analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Medicare is traditionally a fee-for-service (FFS) health insurance program with an inherent incentive for overuse. Starting in the mid-1980s, the Centers for Medicare and Medicaid Services (CMS, formerly the Health Care Financing Administration) began to contract with risk-based managed care plans for the delivery of health care services to contain the growth of Medicare expenditures.

Medicare HMO enrollees are usually healthier than their FFS counterparts. To accurately prepay private health plans, Medicare has developed two risk adjustment models to predict underlying risk and health care costs of HMO enrollees. Medicare intended to pay private plans for health care costs of their enrollees if they had remained in FFS so that Medicare would be financially neutral as to whether its beneficiaries remain in FFS or join an HMO.

CMS first developed a risk adjustment method known as the Adjusted Average Per Capita Cost model (AAPCC), which risk adjusts for several demographic characteristics. The AAPCC model failed to adequately control for biased selection of HMO enrollees. The second model, which CMS started to phase in in 2000, is called the Principal Inpatient Diagnostic Cost Group (PIP-DCG). It risk adjusts for prior-year principal inpatient diagnoses as well as demographic characteristics. By risk adjusting for prior year hospitalizations, the new PIP-DCG model significantly improves upon the AAPCC.

However, there is some evidence indicating that this new model is still not up to the task of adequately adjusting for biased selection of HMO enrollees.

Medicare's failure to consistently estimate FFS costs of HMO enrollees is due to (1) the limited predictive power of its risk adjustment models and (2) the models' FFS basis. Biased selection of HMO enrollees does not necessarily lead to over or underestimation as long as it is fully adjusted for. A model that controls for all relevant risk factors for health care costs would be able to completely adjust away biased selection no matter how biased HMO enrollment is. The PIP-DCG model, with the necessity of being parsimonious and easy to implement, most likely has left out some significant predictors of health care costs. Given biased HMO enrollment, risk factors omitted in the model may be systematically different between the HMO and FFS populations. In this case, FFS-based model will only produce consistent estimates for a random FFS sample but not for an HMO sample.

The study aims to answer two specific questions: (1) Does the PIP-DCG model adequately adjust for HMO favorable selection? (2) If not, how biased is the FFS-based PIP-DCG model in predicting resource use of HMO enrollees if they had been served in FFS?

This study used a unique data set that contains demographic information and hospital discharge records for Medicare FFS and HMO beneficiaries in the state of California from 1994 to 1996. HMO enrollees stayed in FFS in Year 1 and joined an HMO at the

beginning of Year 2. FFS beneficiaries stayed in FFS in both Year 1 and Year 2. HMO enrollees display favorable selection in observed demographic characteristics and PIP-DCG groups compared to their FFS counterparts. While favorable demographic characteristics and PIP-DCG mix of HMO enrollees substantiate their lower level of hospital use in Year 2 compared to FFS beneficiaries, this study found that observed favorable selection of HMO enrollees cannot fully explain their lower hospital use. Their unobserved favorable selection also plays a significant role.

Since health status is an important factor for predicting HMO enrollment, as indicated by consistent favorable selection in Medicare HMOs, the PIP-DCG risk factors, which predict expected health status, are used to predict HMO enrollment as well as hospital use. It is hypothesized that the PIP-DCG risk factors only partially predict future health status. There are omitted health factors in the PIP-DCG model that affect both HMO enrollment and hospital use. In addition to unobserved health status, unobserved non-health factors such as different tastes for medical care and accessibility of providers may also affect hospital use and HMO enrollment simultaneously.

This study adopted a simultaneous equations model that jointly estimates an HMO enrollment model and two hospital-use models (having at least one admission in Year 2 and total length of stay given being admitted at least once in Year 2) to correct for the possible endogeneity of HMO enrollment. The hospital-use models include an HMO dummy to account for different HMO practice styles. Common and/or correlated unobserved variables omitted in the HMO enrollment and hospital-use models would

make the error terms in the three equations correlated to each other. The common and/or correlated unobserved variables are represented by a heterogeneity term. Unlike many other studies, this study avoided adopting a fully parametric specification for the distribution of the heterogeneity term. Instead, the heterogeneity term was approximated by a discrete distribution. This estimator was carefully studied by Mroz (1999) who proved that it was robust under various underlying distributions. A naïve 2-part model that estimates the two hospital use models independently ignoring the possible endogeneity of HMO enrollment was also tested and compared with the simultaneous equations model.

The robust simultaneous equations estimator showed that HMOs insignificantly reduce the admission probability by 1.8 percent, compared to 14.5 percent estimated in the naïve 2-part model. HMOs do reduce total length of stay given admission ($p=0.062$), but the magnitude of the reduction (10.3 percent) is much smaller than estimated in the naïve model (17.0 percent). Overall, HMOs shorten FFS hospital days of HMO enrollees by 11.9 percent. The estimated combined HMO effect represents the decline of hospital use that would be observed when moving a randomly selected Medicare beneficiary from FFS to an HMO. It is smaller than the 28.8 percent overall reduction estimated in the naïve 2-part model because the naïve model attributed HMO favorable selection to HMO efficiencies.

Based on the unbiased estimates of the simultaneous equations model, simulations were conducted to determine the magnitude of bias due to unobserved selection. First, I

decomposed the difference in mean observed inpatient use between FFS and HMO beneficiaries into three components: observed selection, FFS and HMO unobserved selection, and the HMO effect. The total difference in mean total hospital days between FFS beneficiaries and new HMO enrollees is 806 days, among which the PIP-DCG risk factors account for 315 days (39%). After adjustment for observed selection, new HMO enrollees still incur 490 (61%) more hospital days than FFS beneficiaries, among which 138 days (17%) is due to the HMO effect and 352 days (44%) is due to unobserved selection.

CMS uses a naïve PIP-DCG model developed on FFS cost data to estimate FFS costs of HMO enrollees. To simulate bias in the FFS-based model, a naïve 2-part model for hospital use was estimated over the FFS sample only and applied to predict hospital use of HMO enrollees. The bias in the FFS-based 2-part model is due to unobserved selection and the HMO effect.

HMO enrollees' hospital use predicted by the FFS-based naïve model averages 1,488 days. It is lower than mean observed use of FFS beneficiaries (1,828 days) because HMO enrollees have favorable observed characteristics compared to FFS beneficiaries. However, even after controlling for observed selection, the naïve model still overestimated HMO enrollees' observed use by 466 days (1,488-1,022), among which 138 days is due to the HMO effect. The remaining 328 days is attributable to ignoring unobserved selection in model development. FFS hospital use of HMO enrollees is the sum of their mean observed use (1,022 days) and the HMO effect (138 days). The

resulting 1,161 days is the counterfactual hospital use of HMO enrollees in the FFS setting given their unobserved favorable selection. Therefore, for the Medicare program, unobserved selection results in a 28 percent (328/1,161) overestimation of FFS hospital use of new HMO enrollees.

The PIP-DCG model will remain in use until 2007. This study found that the PIP-DCG model does not adequately adjust for selection bias. Unobserved selection is systematically different in the FFS and HMO populations, with HMO enrollees healthier and FFS beneficiaries sicker in ways not captured by the PIP-DCG model. While a naïve model developed on FFS data can consistently predict costs for a random FFS sample, its predictions would not be consistent for an HMO sample that is identical to a random FFS sample in observed characteristics. Bias was built into the model when it was developed over an FFS sample with unobserved selection different from that of the HMO population. In fact, the FFS-based naïve model overestimated hospital use of new HMO enrollees by 28 percent compared to their use if they had been served in FFS. Note that this study examines unobserved selection of new HMO enrollees, which is usually much stronger than enrollees of long tenure.

There are two problems associated with the FFS basis of the PIP-DCG methodology. First, the FFS basis of CMS risk adjustment models is not appropriate in the presence of systematically different unobserved selection between the HMO and FFS populations. Second, if HMOs reduce admissions or alter the mix of inpatient diagnoses, the FFS basis of a diagnosis-based model like the PIP-DCG may lead to inaccurate payment predictions

even if the model itself is actuarially impeccable. This problem is inherent in any risk adjuster that includes risk factors susceptible to manipulation. However, it does not seem to be a big problem for the PIP-DCG model. The HMO industry claims that the PIP-DCG methodology may penalize Medicare+Choice organizations that efficiently substitute less expensive care for hospital inpatient care. This is not empirically observed in 1995 California data. HMOs had little impact on admission rates. Given that, it is not unreasonable to suspect that HMOs may not have shifted the mix of inpatient diagnoses either. That is, a given group of individuals is likely to receive the same set of inpatient diagnoses whether they are in the HMO or FFS setting.

The bias in the PIP-DCG model is ultimately due to its limited predictive power rather than its FFS basis. The FFS basis is a deliberate policy choice, but the model's limited predictive power is a methodological flaw that needs to be fixed.

Overestimation of FFS costs of HMO enrollees by the PIP-DCG model, and to a lesser extent, the HMO efficiencies will continue to enable plans to offer extra benefits while charging no or low premiums. Excess payments to health plans essentially represent a government subsidy to health plans and their enrollees at the expense of taxpayers and FFS beneficiaries who are usually older and sicker than their Medicare+Choice counterparts. The substantial bias in the PIP-DCG model calls for a better risk adjustment model. However, managed care provides additional values such as less cost-sharing, coordinated care and preventive services. To the extent that these added values are worth the extra dollars, a better risk adjustment model that cuts further down on

payments should be balanced with its potential impact on the growth of the Medicare+Choice program.

ACKNOWLEDGEMENTS

I am deeply grateful to Dana Goldman, chair of my dissertation committee, for his consistent patience, support and encouragement. His advice on modeling and econometric issues is critical for the success of this dissertation work. I am also very grateful to Emmett Keeler and Glenn Melnick, my two other committee members, for their guidance on Medicare risk adjustment models and managed care payment policies. In particular, Glenn Melnick brought me onto a RAND research project early on that assessed various aspects of Medicare risk adjustment methods, which eventually led me to this topic. This dissertation used the data from that project. Teh-wei Hu, my external reviewer from University of California, Berkeley, carefully read an early draft of this dissertation and provided many insightful comments. Without their time, thoughtful advice, and encouragement, this work would not be possible.

Several other colleagues offered their help, comments and insights throughout the process. Nasreen Dhanani tirelessly helped me on numerous data issues, and provided many insights and invaluable information and advice all along. Her help was on a daily basis. I do not know how to thank her for her generosity and patience. I also benefited from Anil Bamerzai's insightful comments on this work. Jay Bhattacharya provided advice on econometric modeling.

This work was supported by grant number R03 HS11403 from the Agency for Healthcare Research and Quality. Data for the study were made accessible by additional support from the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department

of Health and Human Services. I am indebted to Robert Reddick and Beate Danielson who linked and managed the data files. I am also grateful to the DxCG, Inc. for providing the DxCG software, and the Centers for Medicare and Medicaid Services and the California Office of Statewide Health Planning and Development for providing the data that made this study possible.

My final thanks go to professors at RAND Graduate School and RAND researchers who taught and worked with me inside and outside of the classroom. Their talents, enthusiasm and friendship made my five years at RAND Graduate School truly rewarding and exciting.

CHAPTER 1

INTRODUCTION

Medicare Risk Adjustment and HMO Favorable Selection

Medicare is a federal health insurance program primarily for the elderly. Traditionally, it is a fee-for-service (FFS) program with an inherent incentive for overuse. Starting in the mid-1980s, the Centers for Medicare and Medicaid Services (CMS, formerly the Health Care Financing Administration) began to contract with risk-based managed care plans for the delivery of health care services under the Tax Equity and Fiscal Responsibility Act of 1982 to contain the growth of Medicare expenditures. By the time, health maintenance organizations (HMOs) had demonstrated their success in controlling health care costs in the private sector. It was hoped that Medicare HMOs, through negotiation of favorable prices for provider services, better coordination of care, and cost-effective health care delivery, could also generate savings for the government while maintaining or even improving quality of care.

Before the Balanced Budget Act (BBA) of 1997, Medicare HMO payments were risk adjusted for demographic characteristics and were tied to local costs of providing Medicare-covered services to FFS beneficiaries. The BBA of 1997 severed much of the linkage between local FFS costs and HMO payments and replaced the Medicare risk program with the Medicare+Choice program, allowing more alternative delivery and financing mechanisms including HMOs to participate in the Medicare program. In this

study, the term "Medicare+Choice plans" primarily refers to HMOs as the study is based on a sample of HMO enrollees prior to the BBA.

Health plans that participate in the Medicare program receive a fixed monthly payment for each beneficiary they enroll and bear the full risk of treating their enrollees. Health plans are required to provide Medicare-covered benefits and may offer more benefits for which they can charge a premium. Many plans have managed to pay down their enrollees' Medicare premiums and coinsurance and provide extra benefits such as prescription drugs while charging low or no premiums.

The proportion of the Medicare population enrolled in managed care plans grew from 3 percent in the late 1980s to 16 percent in 1998. The number of participating plans more than tripled from 1993 to 1998 (GAO, 2000a). The expansion of Medicare managed care began to slow down and eventually decline after 1999. In 2001, about 5.6 million, or nearly 15 percent of all Medicare beneficiaries, were enrolled in a Medicare+Choice plan (Scully, 2001).

This study is based on a sample of Medicare FFS and HMO beneficiaries in the state of California from 1994 to 1996. California has always led the nation with the largest Medicare managed care market. In 1995, 1.2 million Medicare beneficiaries were enrolled in about 30 risk-contract plans in California, accounting for 37 percent of all risk plan enrollees in the nation. The majority of the plans were of the Independent Physicians Association (IPA) model type. The penetration rate of Medicare risk HMOs

(the ratio of HMO enrollees to the Medicare population) in California was as high as 32 percent in 1995, compared to 8 percent in the nation.[1]

Medicare beneficiaries enroll in HMOs on a voluntary basis. For those who opt to leave FFS, they can enroll anytime in any participating health plan in their county. The option to join HMOs has proven attractive mainly to younger and healthier beneficiaries. Biased selection occurs when a health plan draws a non-random or different risk pool from the Medicare population. Numerous studies have found that HMO enrollees are healthier than their FFS counterparts even after adjustment for certain risk factors, a phenomenon called HMO favorable selection (Eggers and Prihoda, 1982; Kasper, et al., 1988; Riley, Lubitz and Rabey, 1991; Lichtenstein et al., 1991; Hill and Brown, 1990; Riley, Rabey, and Kasper, 1989; Hellinger, 1995; PPRC, 1996; Cox and Hogan, 1997; Thiede et al., 1999). In addition to favorable self-selection of plan enrollees, health plans may also have contributed to HMO favorable selection since they have incentives under the capitation payment system to enroll healthy beneficiaries and disenroll unhealthy ones (Luft, 1998; PPRC, 1996; Kronick and Beyer, 1997; Morgan et al., 1997).

To ensure that managed care payments reflect underlying health status of HMO enrollees, CMS first developed a risk adjustment method known as the Adjusted Average Per Capita Cost model (AAPCC) to predict future health care costs of HMO enrollees. Medicare risk adjustment models are meant to predict Medicare payments for HMO enrollees if they had remained in the traditional FFS program in order to ensure that the Medicare program would be no worse off after introducing the managed care option. The

[1] Author's analysis of CMS Medicare managed care penetration quarterly data files.

AAPCC model was developed over a nationally representative sample of FFS beneficiaries. It risk adjusts for five demographic characteristics, namely, age, sex, Medicaid enrollment, residence in an institution, and working-aged status.

It is well established that the AAPCC model does not adequately account for biased selection of HMO enrollees (Lubitz, Beebe, and Riley, 1985; Beebe, Lubitz and Eggers, 1985). The model has quite limited predictive power given that its demographic risk factors can only explain about 1.5 percent of variation in individual health care expenditures. It overpredicts costs for healthy beneficiaries and underpredicts for less healthy ones. Since Medicare HMO enrollees are usually healthier than their FFS counterparts and their better health status is not fully captured by the AAPCC risk factors, the AAPCC model has substantially overestimated FFS costs of HMO enrollees.

Under the AAPCC formula, health plans were paid 95 percent of the local cost a demographically similar beneficiary would incur in traditional FFS Medicare. The underlying assumption was that Medicare HMOs would generate more savings than the 5 percent discount taken by the government so that they could retain part of their savings for providing more benefits and earning a profit. However, according to Hill et al. (1992), the AAPCC model overestimated FFS costs of plan enrollees by 11.3 percent. Thus, the 5 percent discount was not even enough to offset Medicare's overpayments to HMOs. As a result, instead of achieving savings, Medicare ended up spending 5.7 percent more than it would have if HMO enrollees had stayed in FFS. Since the 5

percent discount is taken after risk adjustment, it is not a relevant issue in this study of Medicare risk adjustment.

A good risk adjustment model is essential for facilitating competition among health plans on the basis of price, service, and quality, rather than avoidance of risk. To address this deficiency, the 1997 Balanced Budget Act mandated the implementation of a health-based risk adjustment methodology. Among the different risk adjustment models developed, CMS began to phase in a risk adjuster called the Principal Inpatient Diagnostic Cost Group (PIP-DCG) in 2000. The PIP-DCG went beyond the AAPCC by factoring in previous hospitalizations as well as demographic characteristics. Prior-year principal inpatient diagnoses are grouped into sixteen PIP-DCG groups based on clinical coherence and future cost implications of each diagnostic code (International Classification of Diseases, 9th Revision, Clinical Modification). The model risk adjusts for the sixteen PIP-DCG groups and a set of demographic risk factors that is slightly different than the AAPCC risk factors. Its demographic risk factors include age, sex, Medicaid status, disabled as original reason for entitlement, and working-aged. Beneficiaries who are not admitted or have inpatient diagnoses excluded by the PIP-DCG methodology still have their costs predicted based on their demographic characteristics. This new model--like the AAPCC--was developed and evaluated over a nationally representative sample of Medicare FFS beneficiaries in order to obtain estimates of FFS costs of HMO enrollees. The development sample was drawn from the Medicare FFS population in 1995 and 1996. Inpatient diagnoses and demographic characteristics from

1995 were used to predict 1996 Medicare payments (see Pope et al., 2000a for a detailed description of the PIP-DCG model).

By risk adjusting for prior year hospitalizations, the new PIP-DCG formula significantly improves upon the AAPCC (Gruenberg, Tompkins, and Porell, 1989; Ash et al., 1989; Schauffler, Howland, and Cobb, 1992; Ellis and Ash, 1995; Weiner et al., 1996; Ellis et al., 1996; Iezzoni, 1997; Greenwald et al., 1998; Ingber, 1998; HCFA, 1999; Pope et al., 2000a). It can explain 6.2 percent of variation in individual health care costs. Nevertheless, the model still leaves a large portion of health status unaccounted for. Theoretically, a prospective risk adjustment model can predict up to 20-25 percent of individual expenditure variation (Newhouse et al., 1989).

The PIP-DCG model's predictive power remains limited because it relies on inpatient diagnoses to differentiate underlying health status. About twenty percent of FFS beneficiaries are hospitalized at least once every year. And only two-thirds of them incur inpatient diagnoses that would affect predicted expenditures under the PIP-DCG methodology. Thus, the sixteen PIP-DCG groups do not differentiate health status of almost ninety percent of FFS beneficiaries, whose costs are still being predicted based on their demographic characteristics (Pope et al., 2000a).

Not surprisingly, the PIP-DCG methodology has not been successful in eliminating excess payments to health plans. MedPAC (2000) estimated that the PIP-DCG risk factors only accounted for eight percentage points of a twenty-three percentage-point

difference in average costs between new HMO enrollees and FFS beneficiaries prior to HMO enrollment. A recent GAO study (2000) imputed HMO enrollees' post-enrollment costs based on how FFS costs regressed towards the mean over time, and found that, due to inadequate risk adjustment of the AAPCC model, Medicare spent 13.2 percent more ($3.2 billions) on health plan enrollees in 1998 than had they been served in traditional FFS Medicare, of which the new PIP-DCG model could only eliminate less than half when fully implemented.

The PIP-DCG model is currently being phased in according to the schedule shown in Table 1. Over the period from 2000 to 2007, an increasing portion of Medicare payments will be based on the PIP-DCG risk adjustment method, with the remainder still being based on a demographic model. CMS plans to adopt a more comprehensive health status adjuster that uses encounter data from all settings rather than just the inpatient setting to estimate Medicare payments for Medicare+Choice enrollees after 2007.

TABLE 1
Transition Schedule for Implementation of the Risk Adjustment Method

Year	Demographic-only Method (%)	PIP-DCG Method (%)
CY 2000	90	10
CY 2001	90	10
CY 2002	90	10
CY 2003	90	10
CY 2004	70	30
CY 2005	50	50
CY 2006	25	75
CY 2007	0	100

Source: Medicare Managed Care Manual. Baltimore, MD: CMS, 2001.

Why do Medicare Risk Adjustment Models Overestimate FFS Costs of HMO Enrollees?

Medicare's failure to consistently estimate FFS costs of HMO enrollees is due to (1) the limited predictive power of its risk adjustment models and (2) the models' FFS basis. Biased selection of HMO enrollees does not necessarily lead to over or underestimation as long as it is fully adjusted for. A model that controls for all relevant risk factors for health care costs would be able to completely adjust away biased selection no matter how biased HMO enrollment is. The PIP-DCG model, with the necessity of being parsimonious and easy to implement, most likely has left out some significant predictors of health care costs. Given biased HMO enrollment, risk factors omitted in the model may be systematically different between the HMO and FFS populations. In this case, FFS-based model will only produce consistent estimates for a random FFS sample but not for an HMO sample.

A different way to consider the inconsistency of CMS payment models is to examine how they actually work. The PIP-DCG and the AAPCC are both simple additive predictive models. They group demographic characteristics and previous principal inpatient diagnoses into a small set of cells, each of which has a marginal risk score. Each beneficiary is assigned to one or several of the cells, and his or her risk is calculated by summing up the marginal risk scores of all the cells s/he falls into. The small set of demographic cells and PIP-DCG groups, however, is unlikely to divide Medicare beneficiaries into a sufficiently large number of groups such that each group is more or

less homogenous. It is especially true for the elderly people who typically have more diverse and complicated conditions than the younger population. In order for an FFS-based model to consistently estimate FFS costs of HMO enrollees, HMO and FFS beneficiaries must have similar health care costs within each cell. If HMO enrollees are healthier than FFS beneficiaries within most demographic cells, and, in the case of the PIP-DCG model, have fewer hospitalizations or less severe conditions within PIP-DCG groups, an FFS-based model will overestimate HMO enrollees' FFS costs. These systematic within-cell differences correspond to unobserved selection that is not controlled for by a risk adjustment model. The FFS basis of Medicare risk adjustment is not appropriate in the presence of systematically different unobserved selection between the HMO and FFS populations.

The latest empirical evidence based on the data from the late 1990s indicates that the rapid expansion of Medicare managed care until the end of 1990s did not dilute HMO favorable selection. MedPAC (2000) found that the average cost in 1997 of beneficiaries who enrolled in Medicare+Choice plans during 1998 was 23 percent lower than the average cost of beneficiaries who remained in traditional Medicare. GAO (2000b) investigated 210 Medicare+Choice plans that enrolled 87 percent of all Medicare+Choice enrollees in 1998, and concluded that the vast majority of them experienced favorable selection after adjusting for demographic characteristics. This runs counter to the industry claim that favorable selection may have evaporated in markets with substantial Medicare+Choice penetration. A 2000 GAO testimony before the Congress stated that although many plans were withdrawing from the Medicare+Choice program, payments to

plans continued to exceed the expected FFS costs of their enrollees (GAO, 2000c). The 1998 Medicare Health Outcomes Survey was still able to identify better health status of managed care beneficiaries across age and sex (Haffer, 2000). Using inpatient encounter data from July 1997-June 1998 submitted by Medicare+Choice plans, Greenwald, Levy and Ingber (2000) calculated actual PIP-DCG relative risk scores (an indicator of expected relative costliness predicted by the PIP-DCG model) for Medicare+Choice enrollees. They found that the risk scores were substantially lower for beneficiaries in the Medicare+Choice program compared to beneficiaries in FFS for most counties in their study, indicating considerable favorable selection in HMOs.[2]

Substantial HMO favorable selection is most likely to be explained by both included and omitted risk factors in the PIP-DCG model. When relevant factors excluded in the model are systematically different between the FFS and HMO populations, the FFS-based model will not produce consistent estimates of FFS costs of HMO enrollees.

In this study, unobserved selection is defined as selection bias that remains after risk adjustment (i.e., within-cell differences). Observed selection refers to cross-cell differences that are already controlled for by a risk adjuster. This research examines unobserved selection in the PIP-DCG model.

[2] They took that as evidence of HMO favorable selection based on an implicit assumption that the PIP-DCG does produce unbiased estimates of FFS costs for HMO enrollees. This study concluded that the model overestimates resource use of HMO enrollees because of their unobserved favorable selection. Therefore, real HMO favorable selection could be even greater than what the differences in PIP-DCG risk scores indicated in their study.

Adverse Consequences of Overpaying Medicare HMOs

Overpaying Medicare HMOs has several unintended, deleterious consequences. It obviously costs taxpayers' money as 75 percent of Part B services are paid for out of general revenues. GAO (2000b) estimated that of Medicare's 1998 total payments of $29.8 billion to Medicare+Choice plans, $5.2 billion (21 percent) could have been saved if plan enrollees had received care in the FFS program.

Excess payments to Medicare HMOs enable plans to offer extra benefits to their enrollees. As a result, FFS and HMO beneficiaries receive different benefits under the same Medicare program. This inequality is particularly troublesome given that HMO enrollees are usually younger and healthier than their FFS counterparts.

Overpaying Medicare HMOs also raises access and quality of care concerns. As the AAPCC and the PIP-DCG overpay for the healthier and underpay for the less healthy, health plans have little incentive to develop a system of quality care that attracts the chronically sick, the disabled, and other high-cost groups. The quality of care track record of Medicare HMOs has been mixed so far (Miller and Luft, 1994).

Biased selection tended to be self-reinforcing under the AAPCC payment methodology. Before the 1997 BBA, HMO payments were directly linked to local FFS spending. Capitation payments to health plans were calculated as a product of county base payment rate and AAPCC risk score for each enrollee. County base payment rates are the costs of

serving a nationally average Medicare beneficiary in the FFS setting in each county. With more and more healthy beneficiaries moving to HMOs, beneficiaries that remained in the FFS sector became increasingly more expensive. Greater FFS expenditures automatically led to even higher payments to health plans through increased county base payment rates. More generous HMO benefits attracted even more healthy beneficiaries to join HMOs. In addition, since sicker FFS beneficiaries tend to be more likely to purchase high-end Medigap policies, Medigap premiums had been rising rapidly, encouraging more healthy beneficiaries to enroll in HMOs (Alecxih et al., 1997). This resulted in a scissor-like movement of healthy and unhealthy beneficiaries to two different sectors and a self-reinforcing cycle of biased selection and over-reimbursement. Thus, in addition to the AAPCC model's overestimation of costs of HMO enrollees, divergently rising county base payment rates also contributed to excess payments to health plans before the 1997 BBA went into effect (GAO, 1999).

The1997 BBA fixed county base payment rates at the 1997 level. The baseline rates are adjusted every year according to a specific formula designed to curb managed care payment growth and reduce cross-county and cross-time payment fluctuations. Although base payment rates are no longer tied to local FFS spending and are under tighter control of policymakers, over-reimbursements will not disappear. Actually, base payment rates in the initial years after the BBA were still set too high because the 1997 baseline rates had been too high and were not adjusted down quickly enough (GAO, 2000b). More importantly, even when base payment rates are set accurately, inadequate risk adjustment of the new PIP-DCG model will continue to guarantee overpayments to health plans. In

fact, during the period before the BBA when the inadequacy of the AAPCC model and inflated FFS-based county payment rates were both contributing to excess payments, inadequate risk adjustment of the AAPCC model accounted for the majority of excess payments (GAO, 1997).

Research Questions and Policy Significance

The purpose of this research is to examine the bias in the PIP-DCG model that is due to unobserved selection. The study aims to answer two specific questions:

1. Does the PIP-DCG model adequately adjust for HMO favorable selection?

2. If not, how biased is the FFS-based PIP-DCG model in predicting resource use of HMO enrollees if they had been served in FFS?

A microeconometric model of HMO enrollment and subsequent hospital services use will be developed to test whether the PIP-DCG model adequately captures selection bias. If not, the magnitude of the model's bias in predicting hospital use of HMO enrollees will be quantified through simulations. Finally, policy implications of the findings will be discussed.

Since Medicare HMOs were not required to submit encounter data on health services utilization of their enrollees before the 1997 BBA, plan enrollees' utilization data were

rarely available on a large scale. Many previous studies on biased selection of Medicare HMO enrollees had to rely on plan enrollees' utilization data *prior* to their enrollment.

Even for HMO enrollees under 65, where data have shown lower use than in FFS, the amount of selection is not well known. Since plan enrollees are usually healthier than FFS beneficiaries, their lower health care use could be due to their better health status and/or different (and presumably more efficient) managed care practice styles. The two effects have proven difficult to disentangle. Only a handful of studies, among which the RAND Health Insurance Experiment in the 1970s is the most prominent, were based on randomized experiments.

This research contributes to the literature by using a unique data set that contains actual hospital utilization data of Medicare HMO enrollees to assess the PIP-DCG model. Given the observational nature of the data, this study adopts a microeconometric model that explicitly takes into account the possibility of endogenous HMO enrollment. Although many inpatient diagnoses do not occur frequently even in the elderly population, the large size of the data set makes the assessment possible

Since the 1997 BBA went into effect, county base payment rates have been progressively delinked from local FFS spending. As a result, the contribution of inflated county base payment rates to excess payments has declined. Inadequate risk adjustment is becoming the single cause for excess payments and, thus, deserves more focused examination. This study examines how unobserved selection bias in the PIP-DCG model leads to

overestimation of resource use of HMO enrollees if they had been served in FFS. How excessive county payment rates had contributed to excess payments and how the split of HMO savings between health plans and Medicare affects excess payments are not relevant issues and will only be briefly discussed.

Biased selection of HMO enrollees is due to the way the entire system is set up. It is to be explained by the complex interactions between traditional FFS Medicare, Medicare managed care, Medigap, and many other factors (Kronick and Beyer, 1997). There is no quick fix to eliminate HMO favorable selection overnight. Barring a system overhaul, improving Medicare risk adjustment remains an effective way to reduce excess payments to health plans and alleviate the negative effects of biased selection.

A good risk adjustment method can reduce excess payments and financial loss for the Medicare program, level the playing field for health plans by basing reimbursements on their true opportunity costs, and encourage them to enroll the sick, the disabled and other vulnerable groups who will benefit more from coordinated care.

The PIP-DCG risk adjustment model will remain in use until 2007. The Balanced Budget Act of 1997 originally scheduled the PIP-DCG risk adjustment formula to be phased in from 2000 to 2003, and required the implementation of a comprehensive risk adjustment model based on encounter data from all settings in 2004. The Balanced Budget Refinement Act of 1999, however, slowed down the phase-in schedule and called for additional studies on risk adjustment implementation issues. The Benefits

Improvement and Protection Act of 2000 further extended the phase-in of the PIP-DCG methodology until 2007 (see Table 1 for the latest phase-in schedule), and postponed the replacement of the PIP-DCG model by a comprehensive risk adjustment model until 2007.

In 2000 and 2001, 10 percent of Medicare managed care payments were based on the PIP-DCG risk adjustment methodology, and 90 percent were based on a demographic model. Even with this limited implementation of the PIP-DCG model, Medicare has saved over 0.5 percent in each year compared to basing payments on the AAPCC model alone (Tudor, 2001). With annual Medicare payments running over 200 billion dollars (13 percent of the Federal budget), the savings are in the order of billions of dollars. Note that the savings is due to the improvement of the PIP-DCG upon the AAPCC model. Similar significant savings can be achieved by eliminating unobserved selection bias in the PIP-DCG model itself (GAO, 2000b).

CHAPTER 2

DATA, SAMPLE SELECTION, AND DESCRIPTIVE ANALYSES

Data and Sample Selection

This study investigates bias in the PIP-DCG payment model attributable to systematically different unobserved selection between the HMO and FFS populations. It uses a unique data set from a RAND project that contains demographic information and hospital discharge records for Medicare FFS and HMO beneficiaries in the state of California from 1994 to 1996. Beneficiaries' personal characteristics such as age, sex, disability, HMO enrollment status, Medicaid status, and county of residence were obtained from Medicare Denominator files. Hospital discharge records were obtained from the California Office of Statewide Health Planning and Development, and were linked to the Denominator files. More information about the data can be found in Dhanani et al. (forthcoming).

The study sample consists of beneficiaries in California who joined a Medicare risk HMO and a comparison group of those who remained in FFS (the choice-based nature of the data will be discussed in Chapter 3). New HMO enrollees stayed in FFS in Year 1 and received an HMO "treatment" in Year 2. FFS beneficiaries stayed in FFS throughout the study period (see Figure 1). The "HMO treatment" effect will be consistently estimated using a microeconometric model that controls for both observed and unobserved factors.

FIGURE 1 Enrollment Status of FFS and HMO Beneficiaries

	FFS Comparison Sample	New HMO Enrollees
Year 1	In FFS	In FFS
Year 2	In FFS	In HMO

The HMO sample consists of California Medicare beneficiaries who were enrolled in a risk HMO plan during 1995. To be included in the sample, HMO enrollees need to be at least age 65 at the beginning of Year 1, alive in the first month of Year 2, and without end-stage renal disease (ESRD).[3] In addition, HMO enrollees should be in FFS for at least twelve months before enrollment (Year 1), and stay in an HMO for at least twelve months or until death after enrollment (Year 2). All HMO enrollees had both Medicare Part A and Part B coverage as required by law. A one-year window was created for each enrollee before and after his or her enrollment. For example, for the beneficiary who enrolled in an HMO in May 1995, Year 1 was defined as the period from May 1994 to April 1995, and Year 2, from May 1995 to April 1996. Thus, everyone in the HMO sample stayed in FFS in Year 1 and joined an HMO at the beginning of Year 2. Since they had at least one year of FFS exposure before enrollment, most HMO enrollees should be first-time enrollees. The sample does not include those age-in's who went directly into an HMO when they became eligible for Medicare. New HMO enrollees typically display the strongest favorable selection. After enrollment, their favorable health status deteriorates over time, a phenomenon called regression towards the mean. The HMO sample consists of 78,693 new plan enrollees.

[3] Beneficiaries with ESRD were excluded because Medicare uses a different formula to pay private plans for these beneficiaries.

The FFS comparison sample is a random sample of California Medicare beneficiaries who were eligible for both Part A and Part B and stayed in FFS from 1994 to 1996 (or until death). The inclusion criteria for HMO enrollees were also applied to FFS beneficiaries. A pseudo month of "FFS enrollment" in 1995 was randomly assigned to each FFS beneficiary following the distribution of HMO enrollees' enrollment months (see Table 2). Those who died before "FFS enrollment" were dropped. This aligned the FFS and HMO samples at their starting points of Year 2 while making certain that both FFS and HMO beneficiaries survived until "enrollment" at the beginning of Year 2. The FFS sample size was made roughly equal to that of HMO.

TABLE 2
Distribution of Enrollment Months of HMO and FFS Beneficiaries (%)

1995	New HMO Enrollees	FFS Comparison Sample
January	15.5	15.5
February	8.2	8.2
March	9.6	9.5
April	9.9	9.9
May	8.6	8.5
June	8.0	7.9
July	8.6	8.6
August	8.9	8.8
September	7.6	7.6
October	5.5	5.6
November	4.9	5.0
December	4.7	4.9
Total	100.0	100.0

Beneficiaries who resided in counties with fewer than 100 new risk HMO enrollees or with a Medicare HMO penetration rate lower than 10 percent in 1995 were excluded. As a result, half of the California counties were dropped. This exclusion ensured that

beneficiaries in both sample groups had the option to stay in FFS or join an HMO. Not surprisingly, the remaining twenty-nine counties cluster around San Francisco and Los Angeles. Some peripheral counties surrounding San Francisco and Los Angeles such as Butte, Sutter, Yuba, Santa Cruz, Fresno, Madera and Monterey had quite low penetration rates (around 10 percent) in 1994, but saw a sharp increase in enrollment in 1995. Most of the other counties included in the sample had Medicare HMO penetration rates in the range of 30-50 percent in 1995.

Dependent Variables

The PIP-DCG model is a prospective model that predicts future Medicare payments based on prior year information. This study reassesses the PIP-DCG model using hospital utilization as the outcome variable. Only acute care hospital discharge records are available to this study.[4] The two principal dependent variables are (1) having one or more admissions in Year 2 and (2) total hospital length of stay given being admitted at least once in Year 2. Total length of stay approximates costs of hospital services.

Table 3 reports observed hospital use of HMO and FFS beneficiaries in Year 1 and Year 2. Year 1 (Year 2) hospitalizations are defined as those with an admission date in Year 1 (Year 2). Year 1 admission rates are lower than Year 2 because beneficiaries in the sample must survive until the first month of Year 2 but can die after that. A very small number of beneficiaries (341) were admitted in Year 2 but had less than one day of total

[4] In the actual PIP-DCG model implemented by the CMS, qualified inpatient diagnoses can come from facilities eligible for Medicare's prospective payment system as well as non-PPS facilities and units including psychiatric, rehabilitation, long-term, children's and other specialty hospitals.

length of stay. One day was assigned to these cases. As expected, compared to FFS beneficiaries, HMO enrollees were less likely to have one or more admissions, and had shorter total length of stay given admission in both Year 1 and Year 2.

TABLE 3
Observed FFS and HMO Inpatient Use in Year 1 and Year 2

	FFS Comparison Sample	New HMO Enrollees
Number of Observations	79,933	78,693
Inpatient Use in Year 1		
One or more admissions (%)	16.5	11.6
Mean total LOS, given admitted (days)	8.99	6.78
Median total LOS, given admitted (days)	6	4
Total hospital days (days/1,000)	1,481	786
Inpatient Use in Year 2		
One or more admissions (%)	20.0	15.0
Mean total LOS, given admitted (days)	9.14	6.82
Median total LOS, given admitted (days)	6	4
Total hospital days (days/1,000)	1,828	1,022

Note: Observed total length of stay greater than 60 days was set to 60 and one day was assigned to those who were admitted in Year 2 with less than 1 day of total length of stay.

Independent Variables

The PIP-DCG risk factors that predict future Medicare payments are deemed appropriate for predicting future hospital use as well. Independent variables in this study are defined almost in the same way as in the original PIP-DCG model.[5] The same demographic variables, namely, age at the beginning of Year 2, sex, any Medicaid coverage in Year 1, and disabled as original reason for entitlement, are used to predict hospital use. Although

[5] The original PIP-DCG methodology multiplies predicted payments by a factor of 0.21 for the working-aged, for whom Medicare is a secondary payer. Since the working-aged variable is not a risk factor in the PIP-DCG model itself, it is not included in our model.

disability as original reason for entitlement is not a perfect measure for disability in the population, the PIP-DCG methodology still uses it as a risk factor because originally disabled beneficiaries continue to incur considerably higher costs after they turn 65. Year 1 Medicaid eligibility status is defined as having been eligible for Medicaid anytime during Year 1. Medicaid is included as a prospective adjuster and a proxy for poverty in the PIP-DCG model. No differentiation of Medicaid eligibility categories is made as they are not reported consistently by states. PIP-DCG categories are assigned from Year 1 acute care hospital admissions.

The favorable demographic and health characteristics of HMO enrollees can be seen in Table 4. Medicare HMOs attracted younger beneficiaries and more males relative to FFS. Those who were dually eligible for Medicare and Medicaid were less likely to enroll in an HMO since Medicaid pays for Medicare cost-sharing and many non-covered benefits including drugs, lessening the incentives to join an HMO. Those who were originally entitled to Medicare for disability reasons were slightly less likely to enroll in an HMO. Compared to FFS beneficiaries, significantly fewer HMO enrollees died in Year 2. All of the demographic differences between FFS and HMO beneficiaries are statistically significant at the 0.1 percent level.

TABLE 4
Distribution of FFS and HMO Beneficiaries by Demographic Characteristics (%)

	FFS Comparison Sample	New HMO Enrollees
Number of Observations	79,933	78,693
Age (beginning of Year 2)		
66-68	13.3	17.7
69-71	16.1	19.0
72-74	16.9	17.9
75-77	14.4	14.3
78-80	12.0	11.1
81-83	9.6	8.3
84-86	7.1	5.5
87-89	4.8	3.5
90-92	3.0	1.7
93+	2.8	1.2
Male	39.8	42.0
Medicaid eligibility (Year 1)[1]	18.9	6.5
Disabled[2]	7.0	6.5
Died (Year 2)	7.4	4.2

Note: [1] Beneficiaries who had at least one month of Medicaid eligibility in Year 1.
[2] Beneficiaries whose original reason for entitlement to Medicare was their disability.
All demographic differences between HMO and FFS beneficiaries are statistically significant (P<.001).

Death in different months in Year 2 may lead to different levels of Year 2 hospital use as decedents usually use more health care services in their last several months leading to death. For example, those who died in the first month of Year 2 would incur less Year 2 hospital use than decedents in the last month of Year 2. Death, however, is not used as a predictor in our hospital-use model since this study reassesses the PIP-DCG model, which, as a prospective predictive model, certainly cannot include a death variable. However, in the development of the PIP-DCG model, CMS did adjust part-year expenditures of decedents by annualizing their expenditures and then weighting these cases by the fraction of the year alive. Part-year hospital use of decedents is not adjusted

in our model. Summary statistics show that admission rates and total length of stay are visibly lower for those who died in the first month of Year 2 (and somewhat lower for those who died in the second month), compared to those who died in later months. While length of stay can be easily annualized, it is not obvious how to annualize admission. Given the elderly people have a higher rate of hospital use right before death, annualizing hospital use of decedents may introduce more bias. Given that only a small number of beneficiaries died in the first month of year 2 (0.6% in FFS and 0.3% in HMO), I decided not to annualize hospital use of decedents. Unobserved adverse health status of decedents will be captured by the simultaneous equations model adopted in this study.

The PIP-DCG payment methodology uses an intricate algorithm to assign prior year principal inpatient diagnoses to sixteen PIP-DCG groups based on clinical coherence and future cost implications of each diagnostic code (Pope et al., 2000a). When an individual has multiple hospitalizations, his or her PIP-DCG category is determined by the diagnosis with the greatest future cost implications. DCG 0 is assigned to those with no hospitalizations in the previous year. The number associated with each non-zero PIP-DCG category indicates the costliness of that category. The higher the number, the costlier the category.

In the actual PIP-DCG model implemented by CMS, discretionary hospitalizations and short stays (< 2 days) are excluded in the assignment of PIP-DCG groups to curb incentives for unnecessary hospitalizations. Discretionary hospitalizations are those with minor, transitory or non-specific diagnoses. Short hospital stays often can be

accommodated in the outpatient setting. It is important to note that these hospitalizations are excluded from the PIP-DCG model not because they are not significant predictors of future health care costs but because CMS believed that the negative effects of undesirable behavioral incentives associated with using those diagnoses would outweigh the gain of predictive power.

This study, however, retains discretionary hospitalizations and short stays. DCG 4 contains all discretionary hospitalizations.[6] Short stays with non-discretionary diagnoses are assigned to the PIP-DCG categories in the same way as other hospitalizations. They are not excluded because the data in this study is from a period when the PIP-DCG model was not in place, and thus there is no need to worry about possible confounding behavioral responses. Since our hospital-use models counts all Year 2 hospitalizations, it is only appropriate to include all Year 1 hospitalizations in assigning PIP-DCG groups as well.

The distribution of PIP-DCG groups based on Year 1 principal inpatient diagnoses of FFS and HMO beneficiaries is presented in Table 5.[7] In Year 1, HMO enrollees were less likely to be admitted. The PIP-DCG mix of hospital users clearly shows that inpatient use of new HMO enrollees prior to their enrollment is less costly than their FFS counterparts. To increase the cell size, in the actual analysis, the sixteen non-zero DCGs

[6] In the actual PIP-DCG model, DCG 4 also contains admissions with invalid or missing diagnoses and all short stays. For payment purposes, both DCG 4 and DCG 0 have zero cost weight in the actual model.
[7] DxCG 5.1 for SAS developed by Boston, MA-based DxCG, Inc. was used to assign the PIP-DCG categories.

were collapsed into four groups: DCG 4-7, DCG 8-12, DCG 14-16, and DCG 18+. Using the full set of sixteen PIP-DCGs yielded little difference in estimation results.

TABLE 5
Distribution of Year 1 PIP-DCGs of FFS and HMO Beneficiaries (%)

PIP-DCG (Year 1)	FFS Comparison Sample	New HMO Enrollees
Hospital Users	16.5	11.6
Non-users (DCG 0)	83.5	88.4
Total	100.0	100.0
Users by DCGs		
DCG 4	31.6	34.2
DCG 5	0.8	0.9
DCG 6	1.0	1.1
DCG 7	0.4	0.5
DCG 8	10.7	13.6
DCG 9	8.6	9.1
DCG 10	5.3	5.8
DCG 11	8.2	8.3
DCG 12	9.7	8.4
DCG 14	2.7	2.1
DCG 16	12.0	10.7
DCG 18	2.0	1.6
DCG 20	3.5	1.7
DCG 23	1.9	1.1
DCG 26	1.2	0.7
DCG 29	0.6	0.4
Total	100.0	100.0

Note: The differences in the distributions of PIP-DCGs between HMO and FFS are statistically significant (P<.001).

Within-cell Differences in Hospital Use between FFS and HMO Beneficiaries

While favorable demographic characteristics and PIP-DCG mix of HMO enrollees substantiate their lower level of hospital use in Year 2 compared to FFS beneficiaries, the modeling results presented in Chapter 4 show that observed favorable selection of HMO

enrollees cannot fully explain their lower hospital use. Their unobserved favorable selection also plays a significant role.

The lower level of hospital use of HMO enrollees can be directly seen by comparing FFS and HMO hospital use within cells formed by the PIP-DCG risk factors. Table 6 lists mean hospital days /1,000 in Year 2 for FFS and HMO beneficiaries within 80 cells formed by the ten age groups, sex, any Medicaid coverage in Year 1, and any hospital use in Year 1. Any hospital use in Year 1 is a crude substitution for PIP-DCG groups. Disability as original reason for entitlement is not controlled for since only a small number of beneficiaries fall into this category. The majority of the PIP-DCG risk factors are controlled for in Table 6. In almost all of the 80 cells, FFS beneficiaries incurred more hospital use than HMO enrollees. Only in two cells the ratio of FFS hospital days to HMO hospital days is less than 1.

The fact that HMO enrollees incurred less hospital use even after controlling for the PIP-DCG risk factors can be explained by (1) more efficient HMO practice styles and (2) possible favorable unobserved selection of HMO enrollees. These two factors cannot be disentangled simply by descriptive statistics. The econometric analysis presented later in this study finds that the increase of use associated with moving HMO enrollees to the FFS setting (i.e., the HMO effect), everything else being equal, is equal to 13.5percent of observed use of HMO enrollees. Anything above the 13.5 percent is due to unobserved selection. For example, for male non-hospital users with no Medicaid coverage in Year 1, their ratio of FFS days to HMO days is 1.24. Since the HMO effect is a constant

fraction of observed HMO use, among the 24 percent difference, 10.5 percent (24% - 13.5%) is due to unobserved selection. In most cells, the FFS/HMO use ratio is greater than 1.135, indicating the presence of unobserved selection.

Since unobserved selection equals FFS/HMO ratios minus 1.135, we can compare the magnitude of unobserved selection across groups by directly comparing the FFS/HMO ratios. Female HMO enrollees tend to show stronger unobserved favorable selection compared to male HMO enrollees since their FFS/HMO ratios are greater than those of males in the majority of the cells formed by risk factors other than sex (i.e., age, Medicaid and prior-year hospital use). HMO enrollees who were hospital users in Year 1 show even stronger favorable unobserved selection than HMO enrollees who were not, controlling for other factors. HMO enrollees with Medicaid coverage in Year 1 also have stronger favorable unobserved selection than those without coverage, controlling for other factors. There does not appear to be a clear pattern of how unobserved selection varies with age though.

In sum, HMO enrollees have unobserved favorable selection after controlling for the PIP-DCG risk factors. Females, hospital users in Year 1, and Medicaid beneficiaries in Year 1 show particularly strong unobserved favorable selection.

TABLE 6 Mean Total Hospital Days / 1,000 in Year 2 for FFS and HMO Beneficiaries, by Age, Sex, Any Medicaid Coverage in Year 1, and Any Hospital Use in Year 1

Age	With No Medicaid Coverage in Year 1											
	Non-user in Year 1						Hospital User in Year 1					
	Male			Female			Male			Female		
	FFS	HMO	FFS/ HMO	FFS	HMO	FFS/ HMO	FFS	HMO	FFS/ HMO	FFS	HMO	FFS/ HMO
66-68	887	714	1.24	671	429	1.56	3,866	2,274	1.70	3,503	2,217	1.58
	(3,402)	(5,379)		(4,337)	(6,511)		(469)	(529)		(451)	(465)	
69-71	1,073	787	1.36	782	493	1.59	3,683	2,127	1.73	3,955	1,922	2.06
	(4,101)	(5,642)		(5,253)	(7,137)		(649)	(637)		(647)	(603)	
72-74	1,080	961	1.12	904	594	1.52	3,800	2,084	1.82	3,384	2,261	1.50
	(4,303)	(5,196)		(5,596)	(6,701)		(740)	(687)		(701)	(613)	
75-77	1,324	1,085	1.22	958	694	1.38	3,732	2,410	1.55	3,850	2,686	1.43
	(3,622)	(4,072)		(4,559)	(5,273)		(772)	(614)		(726)	(579)	
78-80	1,602	1,157	1.38	1,182	780	1.52	4,157	2,310	1.80	3,989	2,220	1.80
	(2,783)	(2,848)		(3,863)	(4,266)		(613)	(496)		(717)	(605)	
81-83	1,888	1,355	1.39	1,345	873	1.54	3,856	2,701	1.43	3,174	2,393	1.33
	(1,997)	(2,009)		(3,110)	(3,181)		(527)	(401)		(642)	(491)	
84-86	1,844	1,774	1.04	1,785	1,115	1.60	4,535	2,579	1.76	4,101	2,735	1.50
	(1,272)	(1,153)		(2,229)	(2,208)		(376)	(271)		(517)	(393)	
87-89	2,507	1,769	1.42	1,791	1,126	1.59	3,806	2,703	1.41	3,947	2,004	1.97
	(683)	(672)		(1,546)	(1,384)		(206)	(182)		(397)	(266)	
90-92	3,251	1,564	2.08	1,669	1,368	1.22	4,921	2,636	1.87	3,711	3,201	1.16
	(331)	(282)		(981)	(723)		(114)	(77)		(253)	(144)	
93+	2,671	1,868	1.43	1,251	1,180	1.06	5,785	2,628	2.20	4,016	2,123	1.89
	(237)	(167)		(841)	(532)		(65)	(43)		(188)	(122)	

TABLE 6: continued

Age	With Some Medicaid Coverage in Year 1											
	Non-user in Year 1						Hospital User in Year 1					
	Male			Female			Male			Female		
	FFS	HMO	FFS/ HMO	FFS	HMO	FFS/ HMO	FFS	HMO	FFS/ HMO	FFS	HMO	FFS/ HMO
66-68	1,629	1,361	1.20	1,715	974	1.76	8,490	3,857	2.20	7,009	1,688	4.15
	(625)	(327)		(967)	(533)		(145)	(77)		(223)	(77)	
69-71	1,244	1,488	0.84	1,287	714	1.80	5,669	1,311	4.32	6,648	2,327	2.86
	(618)	(258)		(1,143)	(517)		(157)	(61)		(287)	(101)	
72-74	1,762	1, 494	1.18	1,443	744	1.94	7,477	2,870	2.61	5,389	3,354	1.61
	(646)	(269)		(1,127)	(461)		(153)	(46)		(270)	(79)	
75-77	1,853	1,951	0.95	1,977	985	2.01	6,949	4,286	1.62	6,226	1,797	3.46
	(449)	(164)		(970)	(396)		(117)	(42)		(301)	(79)	
78-80	2,035	1,566	1.30	1,955	1,575	1.24	7,294	3,520	2.07	7,591	3,383	2.24
	(317)	(122)		(894)	(294)		(136)	(25)		(276)	(60)	
81-83	2,861	1,752	1.63	2,072	1,649	1.26	7,082	3,923	1.81	6,593	4,239	1.56
	(267)	(105)		(806)	(245)		(97)	(39)		(248)	(71)	
84-86	2,540	2,346	1.08	1,849	1,558	1.19	6,088	3,818	1.59	5,961	2,796	2.13
	(215)	(52)		(727)	(163)		(102)	(22)		(233)	(49)	
87-89	3,404	3,053	1.12	2,021	1,536	1.32	6,754	5,125	1.32	5,843	864	6.77
	(146)	(38)		(574)	(140)		(69)	(16)		(197)	(22)	
90-92	3,900	1,720	2.27	2,298	1,271	1.81	6,571	3,000	2.19	4,015	2,357	1.70
	(90)	(25)		(426)	(59)		(35)	(7)		(136)	(14)	
93+	3,072	188	16.38	1,802	827	2.18	3,949	1,500	2.63	4,258	1,833	2.32
	(97)	(16)		(606)	(52)		(39)	(4)		(186)	(12)	

Note: Cell size shown in parentheses.
Observed total length of stay greater than 60 days was set to 60 and one day was assigned to those who were admitted in Year 2 with less than 1 day of total length of stay.

CHAPTER 3

MODELS

An Econometric Approach to Correcting Biased Selection

Rational economic agents choose among competing alternatives on the basis of expected marginal returns. As a result, a group of self-selected people often does not represent the distribution of the underlying population. In order to obtain consistent estimates of population parameters, biased selection needs to be addressed. Endogenous selection can be handled with the sample selection approach initially suggested by Heckman (1976; 1978). Sample selection models obtain consistent population estimates by explicitly modeling biased selection, and are widely applied in empirical studies over non-random samples.

This study is based on an observational rather than randomized sample. HMO enrollees self-selected themselves into HMOs just as FFS beneficiaries voluntarily chose to stay in FFS. This is a typical situation where selection bias may occur. Consider the following simple latent variable model, where z_i is an indicator for staying in FFS and y_i* is utilization/costs of health care services:

HMO enrollment: $z_i^* = \boldsymbol{\gamma}'\mathbf{w}_i + u_i$

$z_i = 1$ if $z_i^* > 0$

$z_i = 0$ otherwise

Health care utilization: $y_i^* = \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i$.

The first model describes an individual's likelihood to stay in FFS ($z_i = 1$); the second, his or her utilization of health care services. Assume that only FFS beneficiaries' utilization is observed. That is, $y_i = y_i^*$ if $z_i^* > 0$, where y_i is observed FFS utilization. The latent variables z_i^* and y_i^* are not observed and have unobserved zero-mean errors. A regression defined by $E(y_i \mid \mathbf{x}_i, z_i^* > 0)$ can be consistently estimated over the FFS subsample as long as appropriate model assumptions are met. However, our objective is to estimate $\boldsymbol{\beta}'$--population parameters defined by $E(y_i^* \mid \mathbf{x}_i)$ for the entire population rather than for the FFS subpopulation. If there exist unobservable/unobserved variables that affect both enrollment and utilization, then FFS utilization will be incidentally truncated by the selection process, operating through the correlation between the two error terms in the selection and utilization equations. When the unobserved variables are correlated with observed variables included in the utilization equation, $\mathbf{x}_i$ will no longer be exogenous. Consequently, ordinary least squares (OLS) estimates of $\boldsymbol{\beta}'$ based on the FFS subsample will be biased. Statistically, it is a problem of misspecification of the conditional mean, $\boldsymbol{\beta}'\mathbf{x}_i$, since $E(\varepsilon_i \mid z_i^* > 0)$ is no longer equal to zero (Greene, 2000). If ε_i and u_i follow a bivariate normal distribution, $E(\varepsilon_i \mid z_i^* > 0) = \rho\sigma_\varepsilon\lambda_i(\alpha_u)$, where $\alpha_u = \boldsymbol{\gamma}'\mathbf{w}_i/\sigma_u$, $\lambda(\alpha_u) = \phi(\alpha_u)/\Phi(\alpha_u)$, and $\phi(.)$ and $\Phi(.)$ are the standard normal density and cumulative distribution functions. λ is called the inverse Mills ratio. It effectively behaves like an inappropriately omitted variable in the model of primary interest.

Unobserved variables that only affect either enrollment or utilization pose no problems (there may be a specification error if they are correlated with included variables). If the selection model is perfectly specified in the sense that all relevant variables are observed and included, then selection bias could be fully adjusted for by including appropriate conditioning variables in the utilization model. If the utilization model is perfectly specified, biased selection would obviously not cause any bias. The above scenarios all imply statistical independence of the errors in the two models. Different observed characteristics of HMO and FFS beneficiaries are neither a sufficient nor a necessary condition for unobserved selection to become a problem.

Unobserved Variables Affecting Both HMO Enrollment and Hospital Use

Since health status is an important factor for predicting HMO enrollment, as indicated by consistent favorable selection in Medicare HMOs, the PIP-DCG risk factors, which predict expected health status, are used to predict HMO enrollment as well as hospital use. If a rich set of variables is available, a simple OLS model may be sufficient for controlling for selection bias (for an example, see Hill et al., 1992). However, the PIP-DCG model is a parsimonious model that only uses four demographic risk factors and sixteen PIP-DCG categories to predict future health care costs. It is unlikely that the model has included enough health status predictors. Previous studies on risk adjustment have found the following significant predictors of future Medicare expenditures: demographic characteristics, prior acute-care hospital and physician services, non-discretionary acute-care hospital use, disability, self-reported health status, activities of

daily living and instrumental activities of daily living (ADLs and IADLs), medical risk indicators, disease-specific mortality rates, etc. (Gruenberg et al., 1996).

It is hypothesized that the PIP-DCG risk factors only partially predict future health status. There are omitted health factors in the PIP-DCG model that affect both HMO enrollment and hospital use. In addition to unobserved health status, unobserved non-health factors such as different tastes for medical care and accessibility of providers may also affect hospital use and HMO enrollment simultaneously. Different tastes for medical services can lead to different levels of health care use and physician ties, which may subsequently influence HMO enrollment. In rural areas, beneficiaries may have less access to both hospitals and Medicare health plans. Most unobserved selection is due to unobserved health status, however. Hill et al. (1992) found that unobserved selection in the AAPCC model was due mostly to unobserved health status (83%). Differences in attitudes toward health and health care, socioeconomic factors, and access to care accounted for 17 percent of unobserved selection.

The PIP-DCG model was developed on FFS data with no attempt to correct for possible unobserved selection. If our hypothesis is true, it would only be consistent for the FFS subpopulation but not for the HMO subpopulation. When there are common and/or correlated unobserved factors affecting both HMO enrollment and hospital use, a hospital-use model that includes an HMO dummy cannot be consistently estimated without correcting for the endogeneity of HMO enrollment.

As far as I know, only two previous studies used a sample selection approach to investigate unobserved selection in Medicare risk adjustment models. Dowd et al. (1996) used a fully parametric selection-corrected Tobit model to examine unobserved selection in the AAPCC model. Their sample was drawn from five risk HMOs and the Medicare FFS program in the Twin Cities from 1988 to 1989. Only FFS expenditures were observed. They jointly estimated a model for FFS expenditures and a model for HMO enrollment. However, their fully parametric model indicated favorable rather than adverse FFS selection. Unlike in ordinary linear regression models, misspecification of the distribution of unobserved variables may lead to implausible results in sample selection models (Mroz, 1999). This study will adopt a more robust semi-parametric model.

In an earlier study, Hornbrook et al. (1989) applied a normal-based fully parametric selection model to correct for unobserved selection in the AAPCC model. Their sample was drawn from the elderly members of the Northwest Region of Kaiser Permanente who were enrolled in a prospective payment demonstration program in the early 1980s. With a relatively small sample (about 500 risk enrollees), two independent variables (age and sex), and no identifying variable, it is not surprising that their model produced unstable estimates.

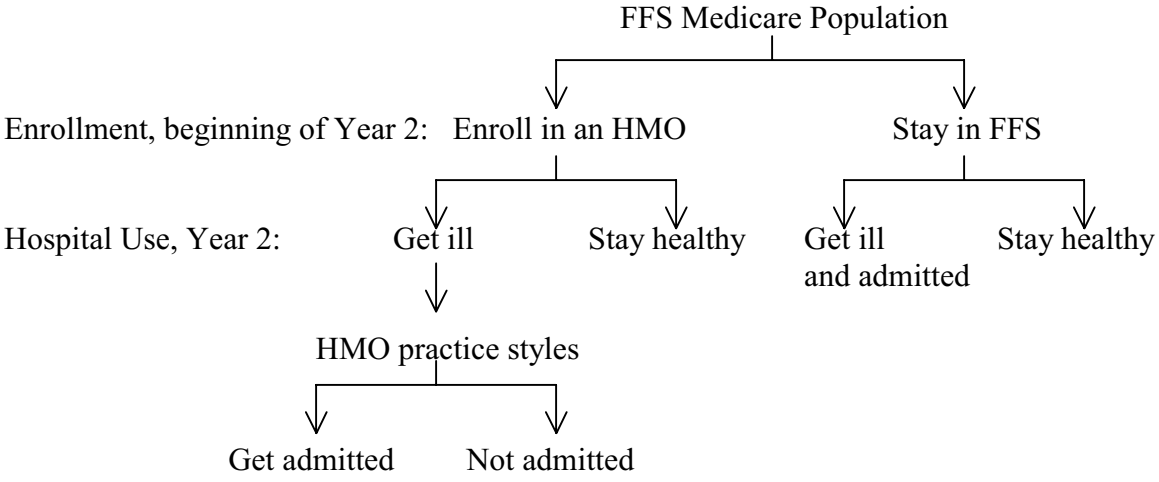
A few researchers studied the impact of non-Medicare HMOs on utilization and costs using the sample selection approach. Goldman et al. (1998) adopted a semi-parametric model to estimate childcare use in a Medicaid HMO from 1987 to 1988. After

controlling for selection, they found that the HMO reduced expenditures on children by 9.1 percent whereas a model assuming no selection found no HMO savings. They concluded that selection could substantially bias estimates of HMO savings. Dowd et al. (1991) used a parametric selection model to investigate the HMO effect on medical services use with a sample of employees from twenty Twin Cities firms in 1984. They found no evidence of omitted variable bias for HMO enrollees' health care utilization when the utilization equation included measures of chronic illness and demographic characteristics. Welch et al. (1984) studied a non-Medicare HMO offered free of charge to near-poor persons under age 65 in the early 1970s in Seattle and found no significant selection bias. Due to the specific way the demonstration HMO was set up, observed health status turned out to be not significant in predicting HMO enrollment, which they believed might explain why no unobserved selection was detected.

A Simultaneous Equations Model with Discrete Factor Approximations

This study adopts a more robust semi-parametric model to consistently estimate population parameters. It is different from traditional sample selection models in that it uses a discrete distribution to approximate unobserved heterogeneity in the population. The simultaneous equations model can be described by the following sequential process of HMO enrollment and hospital use. Decision to join an HMO took place at the beginning of Year 2. Then Year 2 hospital use of HMO enrollees and FFS stayers is observed.

FIGURE 2 HMO Enrollment and Hospital Use in Year 2



A conceptual economic model of HMO enrollment is briefly discussed below. A Medicare beneficiary chooses to join an HMO or stay in FFS at the beginning of Year 2 under the uncertainty about his/her health status in Year 2. S/he estimates future health status based on the information s/he currently has. An HMO is usually less costly than FFS but the downside is that it may restrict choice and utilization. The HMO effect is defined, in the broadest sense, as differences in resource use of a specific individual or group treated in the HMO and FFS settings due to different conditions of supply (incentives, efficiency, practice style, etc.) and demand (price, etc.). In Year 2, health status becomes known to the beneficiary, and s/he accordingly allocates his/her income between consumption goods and out-of-pocket medical care spending, given the FFS/HMO choice made before. His/her choice between HMOs and FFS can be derived by utility maximization in the presence of uncertainty of future health status.

This study adopts a robust semi-parametric microeconometric model that explicitly recognizes possible endogenous HMO enrollment. With a limited set of independent variables, only a reduced-form model is specified. The simultaneous equations model is defined below (see Goldman et al., 1998 and Mroz, 1999 for further discussions of the model):

(1) $y_{0i}^* = \boldsymbol{\gamma}'\mathbf{x}_i + \varepsilon_{0i}$

$y_{0i} = 1$ if $y_{0i}^* > 0$

$y_{0i} = 0$ otherwise

(2) $y_{1i}^* = \boldsymbol{\beta}_1'\mathbf{x}_i + \alpha_1 hmo_i + \varepsilon_{1i}$

$y_{1i} = 1$ if $y_{1i}^* > 0$

$y_{1i} = 0$ otherwise

(3) $(y_{2i} \mid y_{1i} = 1) = \boldsymbol{\beta}_2'\mathbf{x}_i + \alpha_2 hmo_i + \varepsilon_{2i}$

where y_{0i} = HMO enrollment at the beginning of Year 2

y_{1i} = having one or more admissions in Year 2

y_{2i} = total hospital length of stay given admission in Year 2 (logged)

$\mathbf{x}_i$ = PIP-DCG risk factors

hmo_i = HMO enrollment status

HMO enrollment (y_0), hospital admission (y_1) and total length of stay (y_2) in Year 2 are all modeled as a linear function of expected health status and other relevant predictors. Expected health status is partially predicted by the PIP-DCG risk factors $\mathbf{x}_i$, namely, age, sex, any Medicaid coverage in Year 1, disability as the original reason for entitlement, and PIP-DCG groups assigned from Year 1 inpatient diagnoses. The observed PIP-DCG risk variables are considered exogenous in the models.

Model (1) is a selection-corrected probit model that predicts the probability of HMO enrollment. One enrolls in an HMO (y_0=1) when y_0*>0. HMO enrollment is a function of expected health status and other factors. An individual uses more information than represented by the PIP-DCG risk factors to make his/her HMO enrollment decision. The extra information is not observed to the researcher. Unobserved health and non-health factors are reflected in the error term ε_{0i}, and those that also influence hospital use will be picked up by the simultaneous equations model.

The variables of primary interest in the model are (1) having one or more hospital admission in Year 2 (y_1) and (2) total length of stay given admission in Year 2 (y_2). Model (2) is a selection-corrected probit model that predicts the probability of having one or more admissions. Model (3) is a selection-corrected OLS model that predicts total length of stay given admitted at least once. Underlying health risk is partially predicted by the PIP-DCG risk factors ($\mathbf{x}_i$). Unobserved factors are again reflected in the error terms ε_{1i} and ε_{2i}. The hospital-use models resemble the original PIP-DCG model in that they include the same set of risk factors. Age and sex interaction terms are not

significant and are not included. The HMO dummy in hospital-use models (2) and (3) accounts for possible different HMO practice styles. Regional dummies will also be tested to control for possible regional differences in hospital use.

On its raw scale, total length of stay is quite skewed. An OLS model estimated directly on raw data may be subject to the undue influence of outliers, and the error term in the model may not be homoscedastic. Since all predictors in the model are categorical, a method suggested by Blough, Madden and Hornbrook (1999) was used to test if a logarithmic transformation is appropriate. Variance and mean of total length of stay were computed within each cell formed by the PIP-DCG risk factors. Then, logged variance was regressed on logged mean weighted by the degrees of freedom associated with variance in each cell. The slope coefficient is about 2.4 in both the HMO and FFS samples (see Figures 3 and 4), indicating that standard deviation of total length of stay is approximately proportional to conditional mean on the raw scale. In this case, a logarithmic transformation is most appropriate. It effectively stabilizes the variance of the error term. Logged total length of stay of HMO hospital users is not exactly normal but is quite close (skewness=0.13 and Kurtosis=2.52). For FFS users, skewness=0.09 and Kurtosis=2.56.

FIGURE 3 Logged Mean vs. Logged Variance of Total Length of Stay for HMO Enrollees

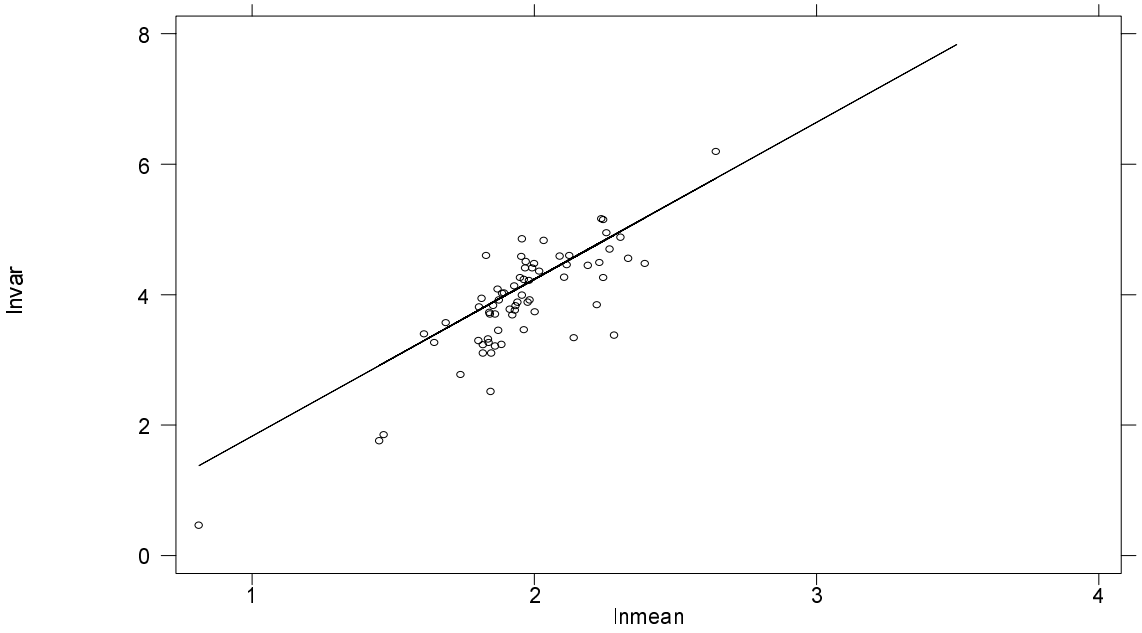
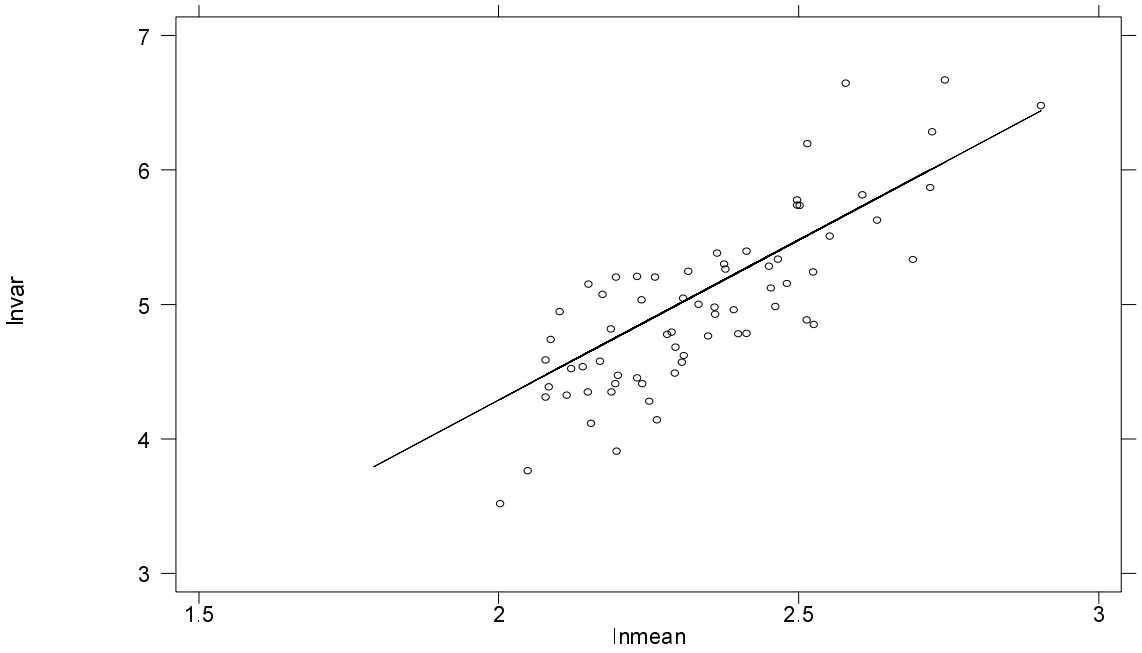


FIGURE 4 Logged Mean vs. Logged Variance of Total Length of Stay for FFS Beneficiaries



To reduce the influence of extreme values, in addition to the logarithmic transformation, total length of stay greater than 60 days was set to 60. The censoring point is the 99.77 percentile of the total length of stay distribution for the HMO sample, and the 99.11 percentile for the FFS sample. Censoring points other than 60 days yielded little changes in estimation results.

A sample selection model that includes an endogenous treatment dummy is called an endogenous treatment model. It imposes the restriction that all covariates have the same effects in the HMO and FFS settings. Indeed, most interactions between the HMO dummy and other covariates in the hospital-use models are not significant, which was also found to be true in Hill et al. (1992).

When the HMO enrollment and hospital-use models contain common and/or correlated unobserved variables, the error terms ε_0, ε_1, and ε_2 become interdependent. Unobserved variables are assumed to be orthogonal to the PIP-DCG risk factors. However, since the error terms in the HMO enrollment and hospital use equations are interdependent, the HMO dummy is not exogenous. A naïve 2-part model for hospital use, in which admission model (2) and total length of stay model (3) are independently estimated, will produce biased estimates.

Our simultaneous equations model produces consistent estimates by explicitly modeling endogenous HMO enrollment. Models (1)-(3) are jointly estimated. In the selection-corrected hospital-use models, α. and $\boldsymbol{\beta}.'$ represent the effects of HMO membership and

the PIP-DCG risk factors on hospital use in a Medicare population in which HMO enrollment is random. That is, the conditional means, $E(y_1 \mid \mathbf{x}_i)$ and $E(y_2 \mid \mathbf{x}_i)$, are expected hospital use of an individual randomly selected from the Medicare population. By contrast, observed hospital use of a self-selected FFS or HMO beneficiary includes the effects of unobserved variables as well as observed characteristics.

The effect of variation of unobserved/unobservable factors over the population is represented by a heterogeneity term θ. It represents common and/or correlated unobserved variables omitted in the HMO enrollment and hospital-use models. The zero-mean error terms in Models (1)-(3) thus can be decomposed into correlated and uncorrelated components:

(4) $\varepsilon_{0i} = \theta + \nu_{0i}$

(5) $\varepsilon_{1i} = \rho_2\theta + \nu_{1i}$

(6) $\varepsilon_{2i} = \rho_3\theta + \nu_{2i}$.

The heterogeneity term θ reflects the correlated component of unobserved selection that affects both HMO enrollment and hospital use. The extra information represented by θ is weighed by a beneficiary in his/her HMO enrollment decision, and also influences his/her subsequent hospital use. θ is assumed to be independently and identically distributed in the Medicare population. In addition to unobserved health status, θ may also represent differences in non-health factors such as different tastes for medical care and accessibility

of providers. As discussed above, the heterogeneity term θ largely represents unobserved health status.

The white noise terms, ν_{0i}, ν_{1i} and ν_{2i}, represent the uncorrelated component of unobserved selection, and are assumed to be identically distributed and independent from each other, θ, and the observed variables in the models. As Models (1) and (2) are both probit models, ν_{0i} and ν_{1i} are assumed to be normally distributed as N(0, 1) after correcting for common/correlated unobserved variables. The white noise term ν_{2i} in the logged total length of stay model is assumed to follow a normal distribution N(0, σ). Note that ν_{2i} represents the error term in a counterfactual Medicare population in which HMO enrollment is randomly assigned. It is different from the error term in the observed sample, ε_{2i}, which includes the effect of unobserved selection. Since unobserved selection is hypothesized to be different between HMO and FFS beneficiaries, ε_{2i} may not be identically distributed in the observed HMO and FFS populations.

With the common θ in ε_0, ε_1, and ε_2, HMO enrollment and hospital use become a function of jointly distributed random variables ε_0, ε_1, and ε_2. Our specification of the error terms explicitly recognizes the endogenous nature of HMO enrollment and formally models the selection process and its impact on hospital use.

A priori knowledge of the distribution of θ is seldom available. Distributional assumptions in simultaneous equations models are often made on the basis of familiarity and ease of computation. A multivariate normal distribution is most commonly assumed.

If θ follows a normal distribution, the distribution of the error terms ε_0, ε_1, and ε_2 will reduce to joint normal. However, misspecification of the joint distribution of the error terms can lead to inconsistent estimates in limited dependent variable models (for the models used in this study, see Mroz, 1999 and Goldberg, 1983). This study avoids a fully parametric distributional specification such as normal. Instead, the heterogeneity term θ is approximated by a discrete distribution with K points of support $\{\eta_k\}$, each of which has a probability π_k:

$$\text{Prob}(\theta = \eta_k) = \pi_k \qquad k = 1, \ldots, K,$$

where $\pi_k > 0$, $\Sigma\, \pi_k = 1$, and $\Sigma\, \eta_k \pi_k = 0$. Note that the mean of θ is set to zero. However, the scale of the zero-mean heterogeneity term θ is not restricted, so one of the factor loadings, ρ_1, is arbitrarily set to 1. Factor loadings ρ_2 and ρ_3 in the hospital use equations are identified. The discrete distribution is identified up to location and scale.

The covariance matrix of the error terms ε_0, ε_1 and ε_2 is

$$\begin{bmatrix} V_\theta + 1 & \rho_2 V_\theta & \rho_3 V_\theta \\ & \rho_2{}^2 V_\theta + 1 & \rho_2 \rho_3 V_\theta \\ & & \rho_3{}^2 V_\theta + \sigma^2 \end{bmatrix}$$

where $V_\theta = \Sigma\, \pi_k \eta_k{}^2$ is the variance of the heterogeneity term.[8]

[8] In a simultaneous equations system with three freely correlated heterogeneity terms, this specification would place inappropriate restrictions on the covariance matrix. As discussed above, θ primarily represents

The discrete factor, quasi-likelihood function for the model is

$$\prod_{i=1}^{N}\sum_{k=1}^{K}\pi_k\{[\int_{-\gamma' x_i-\eta_k}^{\infty}\phi(u)du]^{hmo_i}[\int_{-\infty}^{-\lambda' x_i-\eta_k}\phi(u)du]^{1-hmo_i}$$
$$[\int_{-\beta_1' x_i-\alpha_1 hmo_i-\rho_2\eta_k}^{\infty}\phi(u)du]^{adm_i}[\int_{-\infty}^{-\beta_1' x_i-\alpha_1 hmo_i-\rho_2\eta_k}\phi(u)du]^{1-adm_i}$$
$$[\frac{1}{\sigma}\phi(\frac{y_{2i}-\beta_2' x_i-\alpha_2 hmo_i-\rho_3\eta_k}{\sigma})]^{adm_i}\}$$

where N is the sample size, $\phi(.)$ is the standard normal density function, and adm_i is observed admission. The model parameters γ', β_1', β_2', α_1, α_2, σ, ρ_2, ρ_3, $\{\eta_k\}$, and $\{\pi_k\}$ are jointly estimated subject to the trivial normalizations discussed above.

This semi-parametric maximum likelihood estimator imposes minimum restrictions on the distribution of the error terms. As long as identification conditions are met, the semi-parametric estimator generally produces consistent estimates even without exclusion restrictions. Kiefer and Wolfowitz (1956) proved that, under usual regularity conditions, parameters of a common underlying distribution could be consistently estimated along with those of structural variables. Heckman and Singer (1994) and Cosslett (1983), who applied discrete factor approximations in a continuous time duration model and a binary choice model respectively, verified the assumptions of Kiefer and Wolfowitz (1956). As

---

unobserved health status. If there are other important unobserved factors affecting both HMO enrollment and inpatient use, adding a second heterogeneity term would be warranted. It would place no substantive restrictions on the covariance matrix. However, Wald tests based on estimation results of a model with two heterogeneity terms did not reject the restrictions imposed in the one-heterogeneity term model. It is consistent with the finding by Hill et al. (1992) that most unobserved selection is unobserved health status.

shown by Cameron and Taber (1998), consistency for the model used in this study is a straightforward extension of their work.

A Monte Carlo study demonstrated that this estimator compares favorably to a normal-based parametric maximum likelihood estimator in terms of precision and bias when the true distribution of the error terms is joint normal, and dominates other estimators including the normal-based MLE in a variety of situations when it was not (Mroz, 1999). In fact, a fully parametric maximum likelihood estimator may produce implausible estimates when the distribution of the error terms is misspecified. Even more impressively, the estimator produces quite accurate estimates of standard errors, allowing us to make conventional statistical inferences (Cameron and Taber, 1998).

Estimation Issues for the Simultaneous Equations Model

To identify the simultaneous equations model, a different set of independent variable are generally needed for the enrollment model (Manski, 1989). An identifying variable should affect enrollment but not utilization. Although having exclusion restrictions is preferable, Mroz' Monte Carlo study (1999) showed that the semi-parametric model could be effectively identified through functional form and distributional assumptions. Goldman et al. (1996) applied the model to estimate childcare use in a Medicaid HMO, and successfully identified the model using the same set of independent variables.

For a normal-based parametric model, Leung and Yu (1996) reported that with identical independent variables in two stages, a parametric estimator could be effective in identifying the model as long as at least one of the variables in the selection equation displays sufficient variation. This is because the transformation of predicted conditional mean into the inverse Mills ratio in a probit model is nonlinear at extreme values. The drawback of relying entirely on different functional form of the first stage model for identification is that it may produce unstable estimates when predictors do not vary over a wide range. Mroz (1999) demonstrated that the semi-parametric estimator outperformed other estimators including the normal-based MLE by a wide margin in model specifications with skewed error distributions and no exclusion restrictions.

The sample in this study was assembled from administrative data. The parameters (except the constant term) in the HMO enrollment model can be consistently estimated over the choice-based sample if the white noise term ν_0 in Equation (4) has a logistic distribution (see Manski and Lerman, 1977). Cameron and Taber (1998) and Mroz (1999) show that the discrete factor model works whether the stochastic term is normally or logistically distributed. In our model, ν_0 in Equation (4) is assumed to be normally distributed. There is no closed-form solution for maximizing the quasi-likelihood function, and derivatives are obtained numerically in the maximization process. Given that a logistic distribution is very close to a normal one, it makes little difference in practice whether the model is estimated by assuming ν_0 to be normally or logistically distributed.

The likelihood function of the simultaneous equations model (1)-(3) may have multiple local maxima. To ensure that the model converges to a meaningful point, a step-by-step estimation procedure is applied following the suggestion of Cameron and Taber (1998). First, a naïve model with a one-support point discrete distribution is estimated, and its coefficient estimates are used as initial values for estimating a 2-point model. Then, coefficient estimates of the 2-point model are used as initial points for estimating a 3-point model. This procedure is repeated until the likelihood stops improving appreciably. There is no consensus in the literature about how to choose the number of support points. Mroz (1999) suggested that one should liberally add points of support. With a sample size as large as 160,000, estimating the three equations jointly is time-consuming. Numeric problems are often encountered. This study stops at a 3-point discrete distribution. The 3-support point model performs well in terms of accuracy and consistency (Mroz, 1999).

Test Models

A naïve hospital-use model that ignores the possible endogeneity of HMO enrollment and a fully parametric simultaneous equations model will be tested, and their estimation results will be compared to those of the robust semi-parametric estimator. The naïve 2-part model for hospital admission and total length of stay given admission treats the HMO dummy as exogenous and estimates the admission equation (2) and the total length of stay equation (3) independently. Parameter estimates of this naïve model will be biased when the HMO dummy is correlated with the error terms. In the presence of

unobserved favorable selection of HMO enrollees, HMOs will appear more effective in reducing hospital use than they really are since the naïve model attributes HMO favorable selection to HMO efficiencies. The model is defined below.

(7) $y_{1i}^* = \boldsymbol{\beta_1}'\mathbf{x_i} + \alpha_1 hmo_i + \varepsilon_{1i}$

$y_{1i} = 1$ if $y_{1i}^* > 0$

$y_{1i} = 0$ otherwise

(8) $(y_{2i} \mid y_{1i} = 1) = \boldsymbol{\beta_2}'\mathbf{x_i} + \alpha_2 hmo_i + \varepsilon_{2i}$

where y_{1i} = having one or more admissions in Year 2

y_{2i} = total hospital length of stay given admission in Year 2 (logged)

$\mathbf{x_i}$ = PIP-DCG risk factors

hmo_i = HMO enrollment status

The error terms ε_{1i} and ε_{2i} in the above model are assumed to be uncorrelated.

A fully parametric normal-based simultaneous equations model for having at least one admission will also be tested. The model is specified as follows:

(9) $y_{0i}^* = \boldsymbol{\gamma_0}'\mathbf{x_i} + u_i$

$y_{0i} = 1$ if $y_{0i}^* > 0$

$y_{0i} = 0$ otherwise

(10) $y_{1i}^* = \boldsymbol{\gamma_1}'\mathbf{x_i} + \lambda hmo_i + \varepsilon_i$

$y_{1i} = 1$ if $y_{1i}^* > 0$

$y_{1i} = 0$ otherwise

where y_{0i} = HMO enrollment at the beginning of Year 2

y_{1i} = having one or more admissions in Year 2

$\mathbf{x_i}$ = PIP-DCG risk factors

hmo_i = HMO enrollment status

The model of main interest is a selection-corrected probit model for having one or more admissions in Year 2. Both admission and enrollment are binary variables. Common/correlated unobserved factors make u_i and ε_i correlated, and (u_i, ε_i) is assumed to follow a bivariate normal distribution with a correlation coefficient ρ. A negative ρ indicates favorable unobserved selection of HMO enrollees. That is, a high draw of u_i increases the probability of joining an HMO but decreases the probability of being admitted. This model is in the same spirit as the semi-parametric estimator except that it is now fully parametric and only models admission. This selection-corrected bivariate probit model was previously applied by Van de Ven and Van Praag (1981) and Boyes et al. (1989) to examine health insurance choices and credit scoring respectively.

All models in this study are estimated using STATA 7.0 (Stata, 2001). Standard errors are estimated robustly to protect against possible heteroscadasticity.

CHAPTER 4

ESTIMATION RESULTS

Naïve 2-part Model Results

Table 7 reports estimation results of the naïve 2-part model (7)-(8) estimated over the FFS and HMO sample ignoring the possible endogeneity of HMO enrollment. Models (7) and (8) were estimated independently. In that naïve model, HMO membership is highly significant, both statistically and practically, in reducing the probability of having one or more admissions as well as total length of stay given admission. To arrive at an estimate of the mean effect of HMOs on admission, I calculated predicted probability of admission for HMO enrollees in the HMO and FFS settings by switching the HMO dummy. On average, HMOs reduce the probability of having one or more admissions by 0.03 (or 14.5 percent) compared to mean predicted probability of admission for HMO enrollees if they had remained in FFS. HMOs shorten total length of stay given admission by 17.0 percent (1 - exp(-0.187)). Overall, HMOs reduce unconditional hospital days by 28.8 percent compared to what HMO enrollees would incur in FFS.

TABLE 7
Estimation Results of the Naïve 2-part Hospital-use Model using Total Sample

	Equation	
Variable	Any Admission in Year 2	Log Total LOS in Year 2
Age		
69-71	0.061*	-0.015
	(0.014)	(0.024)
72-74	0.117*	0.022
	(0.014)	(0.023)
75-77	0.205*	0.061*
	(0.014)	(0.023)
78-80	0.278*	0.083*
	(0.015)	(0.024)
81-83	0.362*	0.066*
	(0.016)	(0.024)
84-86	0.430*	0.142*
	(0.018)	(0.026)
87-89	0.501*	0.129*
	(0.020)	(0.028)
90-92	0.545*	0.091*
	(0.025)	(0.033)
93+	0.447*	0.077*
	(0.027)	(0.035)
Male	0.141*	0.029*
	(0.008)	(0.011)
Medicaid	0.153*	0.157*
	(0.011)	(0.016)
Disabled	0.272*	0.112*
	(0.014)	(0.020)
DCG 4-7	0.415*	0.096*
	(0.016)	(0.021)
DCG 8-12	0.588*	0.152*
	(0.014)	(0.018)
DCG 14-16	0.997*	0.425*
	(0.023)	(0.025)
DCG 18+	1.023*	0.526*
	(0.031)	(0.035)
HMO	-0.109*	-0.187*
	(0.008)	(0.012)
Constant	-1.303*	1.543*
	(0.012)	(0.020)

Note: Standard errors shown in parentheses.
* significant at 95% level;
** significant at 90% level.

Simultaneous Equations Model Results

Tables 8 and 9 report estimation results of the simultaneous equations model (1)-(3) with θ approximated by a 3-support point discrete distribution. The discrete factor model was successfully identified with no exclusion restrictions. The model's coefficient estimates are consistent because it explicitly models heterogeneity in the population.

Comparing Tables 7 and 8, except for the HMO effect, the coefficient estimates in the simultaneous equations model are quite similar to those in the naïve 2-part model. As expected, the probability of having at least one admission and total length of stay given admission increase with age. Males tend to use more hospital services. Medicaid entitlement, disability status, and costlier PIP-DCGs are all associated with more hospital use. Almost all observed PIP-DCG risk factors are statistically significant.

However, the HMO effect changed dramatically after correcting for the endogeneity of HMO enrollment. HMOs now have a much smaller and insignificant effect in reducing the probability of having at least one admission. HMOs only reduce the admission probability by 1.8 percent, compared to 14.5 percent estimated in the naïve model. HMOs do reduce total length of stay given admission ($p=0.062$), but the magnitude of the reduction (-0.109) is much smaller than estimated in the naïve model (-0.187). On the raw scale, HMOs shorten total length of stay given admission by 10.3 percent ($1 - \exp(-0.109)$).

Overall, HMOs shorten FFS hospital days of HMO enrollees by 11.9 percent (see Chapter 6). The combined HMO effect is larger than the 10.3 percent HMO reduction of total length of stay given admission because HMOs also insignificantly reduce the probability of admission by 1.8 percent. The estimated combined HMO effect represents the decline of hospital use that would be observed when moving a randomly selected Medicare beneficiary from FFS to an HMO. It is smaller than the 28.8 percent overall reduction estimated in the naïve 2-part model because the naïve model attributes HMO favorable selection to HMO efficiencies.

TABLE 8
Estimation Results of the Simultaneous Equations Model with θ Approximated by a 3-support point Discrete Distribution

Variable	Equation		
	HMO Enrollment at beginning of Year 2	Any Admission in Year 2	Log Total LOS in Year 2
Age			
69-71	-0.153*	0.064*	-0.012
	(0.024)	(0.014)	(0.024)
72-74	-0.321*	0.123*	0.027
	(0.027)	(0.015)	(0.024)
75-77	-0.406*	0.212*	0.067*
	(0.031)	(0.016)	(0.024)
78-80	-0.520*	0.287*	0.090*
	(0.040)	(0.017)	(0.024)
81-83	-0.589*	0.373*	0.074*
	(0.045)	(0.018)	(0.025)
84-86	-0.782*	0.442*	0.152*
	(0.069)	(0.021)	(0.027)
87-89	-0.866*	0.514*	0.140*
	(0.081)	(0.024)	(0.029)
90-92	-1.386*	0.563*	0.106*
	(0.190)	(0.030)	(0.035)
93+	-1.878*	0.467*	0.095*
	(0.289)	(0.033)	(0.038)
male	0.015	0.142*	0.030*
	(0.017)	(0.008)	(0.011)
Medicaid	-2.557*	0.179*	0.178*
	(0.471)	(0.025)	(0.022)
disabled	0.177*	0.270*	0.110*
	(0.031)	(0.015)	(0.020)
DCG 4-7	-0.439*	0.422*	0.102*
	(0.049)	(0.017)	(0.021)
DCG 8-12	-0.409*	0.595*	0.158*
	(0.048)	(0.015)	(0.019)
DCG 14-16	-0.607*	1.007*	0.433*
	(0.084)	(0.025)	(0.026)
DCG 18+	-1.602*	1.042*	0.540*
	(0.412)	(0.035)	(0.037)
HMO	--	-0.013	-0.109**
	--	(0.083)	(0.062)
constant	1.880	-1.365*	1.490*
	(1.215)	(0.055)	(0.044)
σ	--	--	0.932*
	--	--	(0.004)

Note: Standard errors in parentheses. * significant at 5%; ** significant at 10%.

The estimated HMO effect is consistent with the literature. HMOs have achieved cost reduction mainly by controlling inpatient use, especially, length of stay. Hill et al. (1992) found that risk plans had no effect on admission rates but reduced length of stay by 16.8 percent. The authors suggested that the lack of HMO effect on admission might be because discretionary hospitalizations had largely disappeared in the 1990s. Since the California health care market in 1995 was probably more mature and competitive than the national market in 1990, it is not surprising that this study found a smaller HMO effect on length of stay (10.3 percent). A literature synthesis by Miller and Luft (1994) confirmed that HMOs (including Medicare HMOs) had a much more significant impact on shortening length of stay than reducing admission rates and physician visits. More recently, Christensen and Shinogle (1997) found that the Medicare HMO membership was associated with a statistically significant decline in length of stay and an insignificant increase of the probability of being admitted. Dhanani et al. (forthcoming) estimated the HMO effect with a 2-part model, using the same linked data as this study from 1991 to 1995, and found that the HMO membership had no effect on the probability of being admitted at least once but was associated with a decrease of 13 percent in length of stay.

Table 9 presents the parameter estimates of the 3-support point discrete distribution. The estimated locations of the three support points are –2.364, 1.309 and 14.215 with a probability of 0.577, 0.360, and 0.063 respectively, suggesting that the underlying distribution of the heterogeneity term is not normal.

TABLE 9
Parameter Estimates for the 3-Support Point Discrete Distribution

Factor Locations	
η_1	-2.364*
	(1.208)
η_2	1.309
	(1.050)
η_3	14.215
	(p=.350[1])
Factor Weights	
π_1	0.577*
	(p=.000[1])
π_2	0.360*
	(p=.000[1])
π_3	0.063*
	(p=.007[1])
Factor Loadings	
ρ_1	set to 1
	--
ρ_2	-0.022
	(0.029)
ρ_3	-0.018
	(0.022)
Calculated Corr.	
corr(ε_0, ε_1)	-0.085
	(p=.120[1])
corr(ε_0, ε_2)	-0.078*
	(p=.001[1])
corr(ε_1, ε_2)	0.007
	(p=.251[1])

Note: Standard errors shown in parentheses.
[1] p values of Wald statistics.
* significant at 95% level;
** significant at 90% level.

The estimated correlation matrix of ε_0, ε_1 and ε_2 is:

$$\begin{bmatrix} 4.189 & -0.085 & -0.078 \\ & 1.003 & 0.007 \\ & & 0.935 \end{bmatrix}$$

The error term ε_0 in the HMO enrollment equation is negatively correlated with ε_1 and ε_2 in the hospital use equations (corr(ε_0, ε_1)= -.085, corr(ε_0, ε_2)= -.078). That is, a high draw of ε_0 would increase the probability of HMO enrollment in Year 1 but decrease the probability of admission and total length of stay in year 2, indicating unobserved favorable selection in the HMO population and unobserved adverse selection in the FFS population even after adjustment for the PIP-DCG risk factors. Wald tests show that the correlation is highly significant in the total length of stay model (p=0.001), but it is not significant in the admission model (p=0.120). Nevertheless, the fact that the estimated HMO effect on admission changed dramatically compared to the naïve model validates that controlling for the endogeneity of HMO enrollment is important. It can be concluded that overall the PIP-DCG model does not fully adjust for selection bias in hospital use.

The error terms ε_1 and ε_2 in the admission and total length of stay equations are virtually uncorrelated, indicating that admission and the extent of use after admission are two different processes that are not affected by common unobserved factors. The two equations constitute another sample selection model with correlated error terms. This is in contrast to a standard 2-part model that assumes no correlation between the error terms

(Duan et al., 1984). In their study of the effects of HMO and other Medicare insurance supplements on health care utilization, Christensen and Shinogle (1997) also found insignificant correlation between the error terms in models of probability of any hospital use and length of stay.

The counties included in the sample cluster around San Francisco and Los Angeles. A regional dummy for Southern and Northern California was added to the hospital-use models to test if there exist regional differences in hospital use. A complete set of county dummies was also tested. Regional differences in admission and total length of stay turned out to be very small, and coefficient estimates changed little. Given that estimation results did not change significantly, I decided not to include any regional dummies in the model. Another consideration for not doing that is that the original PIP-DCG model does not include any regional dummies.

Bivariate Probit Model Results

The bivariate probit model (9)-(10) failed to produce plausible estimates. It produced a positive estimate of ρ (.446), indicating excessive FFS favorable selection, an insignificant Medicaid effect (-.052) on admission, and a big reduction of admission by HMOs (-.823). Other coefficient estimates also appeared implausible.

Mroz (1999) found that a fully parametric maximum likelihood estimator could generate absurd estimates when the distribution of error terms was misspecified. The estimated 3-

point discrete distribution suggests that the heterogeneity term is not normally distributed. Thus, it is not surprising that the normal-based bivariate probit model failed to produce meaningful estimates. As mentioned above, Dowd et al. (1996), using a fully parametric selection-corrected Tobit model, also found favorable FFS selection in the AAPCC model. Our results demonstrate again the importance of correctly and robustly specifying the distribution of error terms in a limited dependent variable model.

A normal-based simultaneous equations model can presumably be used to model HMO enrollment, admission, and total length of stay simultaneously. Since the bivariate probit model did not work out, it was not pursued.

CHAPTER 5

SIMULATIONS

A Framework for Decomposing Total Bias

Using the estimation results of the simultaneous equations model, simulations are conducted over the same sample to quantify and decompose bias in the PIP-DCG model. The sample was not split into two subsamples for development and evaluation because hospitalizations for many diagnoses do not occur frequently. The simulations are conducted under the scenario of moving HMO enrollees to the FFS setting since CMS risk adjustment models are meant to estimate FFS costs of HMO enrollees. The purpose of the simulations is not to evaluate the PIP-DCG model's theoretical performance which has been thoroughly studied (Pope et al., 2000a).

The simulations focus on hospital days per 1,000, a better measure of inpatient resource use than admission rates. Given that the HMO effect on admission is essentially zero, the difference between observed FFS and HMO admission rates is attributable to observed and unobserved selection only. However, since HMOs do shorten total length of stay, the difference between observed FFS and HMO hospital days is attributable to the HMO effect as well as observed and unobserved selection. Thus, hospital days per 1,000 is also more interesting than admission rates.

Our model estimates are consistent for a Medicare population in which HMO enrollment is random (i.e., there is no unobserved selection). Expected hospital days for an individual randomly assigned to the FFS or HMO sector is estimated as follows:

(11) Any admission: $\hat{y}_{1i}^{*} = \Phi(\hat{\beta}_1' \mathbf{x}_i + \hat{\alpha}_1 \text{hmo}_i)$

(12) Logged total LOS given admission: $\hat{y}_{2i} = \hat{\beta}_2' \mathbf{x}_i + \hat{\alpha}_2 \text{hmo}_i$

(13) Hospital days: $\hat{y}_i = \hat{y}_{1i}^{*} \times \exp(\hat{y}_{2i}) \times \exp(\hat{\sigma}^2 / 2)$,

where $\Phi(.)$ is the standard normal cumulative distribution function. Predicted hospital use in Equation (13) has the effect of unobserved selection purged out. It represents the counterfactual use of FFS and HMO beneficiaries if HMO enrollment had been random. Observed hospital use, however, includes the effect of unobserved selection.

Equation (12) predicts total length of stay on the logarithmic scale. The retransformation of predicted total length of stay from the log scale to the raw scale is based on the log normality assumption in Model (3). Note that ν_{2i} represents the error if there had been no unobserved selection. In contrast, ε_{2i} $(= \rho_3\theta + \nu_{2i})$ includes the effect of unobserved selection. Since HMO and FFS display different unobserved selection, the variance of ε_{2i} does not need to be equal in the two populations. In fact, the variance of the predicted residual, $\hat{\varepsilon}_{2i}$, is greater in the FFS sample than in the HMO sample, indicating that unobserved factors in the FFS sample vary more than in the HMO sample. The expectation of exponentiated ν_{2i} may also be estimated by a robust estimator developed

by Duan (1983) as the average of the exponentiated residuals. But since the residuals ν_{2i} are not observed and the observed residuals, $\hat{\varepsilon}_{2i} = y_{2i} - (\hat{\beta}_2' \mathbf{x_i} + \hat{\alpha}_2 \text{hmo}_i)$, are contaminated by unobserved selection, Duan's estimate is not applicable.

Let $E(y_f \mid \mathbf{x}, \text{FFS})$ and $E(y_h \mid \mathbf{x}, \text{HMO})$ be expected hospital use conditional on FFS and HMO enrollment for self-selected FFS and HMO beneficiaries. In the simulations, we use observed use instead of predicted use conditional on enrollment since they differ little. The difference between FFS and HMO observed hospital use,

(14) $$E(y_f \mid \mathbf{x}, \text{FFS}) - E(y_h \mid \mathbf{x}, \text{HMO}),$$

can be decomposed into the (1) the HMO effect, (2) HMO unobserved selection, and (3) FFS unobserved selection. First, HMOs practice a different style of medicine and achieve hospital cost savings primarily by shortening length of stay. Second, HMO enrollees are generally healthier than FFS beneficiaries, and their better health status cannot be fully captured by the PIP-DCG model. Unobserved selection makes HMO observed use lower than predicted use and FFS observed use higher than predicted use. Unbiased population estimates in the simultaneous equations model enable us to parse total bias (14) into these three sources.

Let $E(y_f^* \mid \mathbf{x})$ and $E(y_h^* \mid \mathbf{x})$ be counterfactual use predicted by Equation (13) for beneficiaries randomly assigned to the HMO and FFS sectors. The asterisk indicates that

these are counterfactual use with unobserved selection purged out. The difference between them,

(15) $E(y_f^* \mid \mathbf{x}) - E(y_h^* \mid \mathbf{x})$,

represents the pure HMO effect that would be observed when moving a randomly picked beneficiary from the FFS sector to the HMO sector. Since the HMO dummy does not interact with observed PIP-DCG risk factors in the model, the HMO effect does not vary with **x**.

Observed hospital use, $E(y_f \mid \mathbf{x}, \text{FFS})$ and $E(y_h \mid \mathbf{x}, \text{HMO})$, includes the effect of unobserved selection as well as the HMO effect. Predicted hospital use, $E(y_f^* \mid \mathbf{x})$ and $E(y_h^* \mid \mathbf{x})$, only includes the HMO effect. So the difference between predicted and observed use represents the effect of unobserved selection. For example, since HMO enrollees have unobserved favorable health status, their predicted hospital use, with unobserved selection purged out, is expected to be greater than their observed use. The effect of HMO unobserved selection equals

(16) $E(y_h^* \mid \mathbf{x}) - E(y_h \mid \mathbf{x}, \text{HMO})$.

$E(y_h^* \mid \mathbf{x})$ is a beneficiary's expected utilization had s/he been randomly enrolled in an HMO, and $E(y_h \mid \mathbf{x}, \text{HMO})$, his/her expected utilization conditional on self-selected HMO

enrollment. $E(y_h^* \mid \mathbf{x})$ is expected to be greater than $E(y_h \mid \mathbf{x}, HMO)$ because the latter includes the effect of unobserved better health status of self-selected HMO enrollees.

Similarly, the effect of FFS unobserved selection is

(17) $E(y_f \mid \mathbf{x}, FFS) - E(y_f^* \mid \mathbf{x})$.

Given unobserved adverse selection of FFS beneficiaries, $E(y_f^* \mid \mathbf{x})$ is expected to be less than $E(y_f \mid \mathbf{x}, FFS)$.

Therefore, the difference in observed hospital use between an FFS beneficiary and an HMO enrollee with identical observed characteristics, $E(y_f \mid \mathbf{x}, FFS) - E(y_h \mid \mathbf{x}, HMO)$, can be decomposed into the HMO effect, HMO unobserved selection, and FFS unobserved selection:

(14) = (15) + (16) + (17)

The magnitude of prediction bias is obviously a function of **x**.

Since the FFS and HMO samples are not matched, the difference in mean observed hospital use in the two samples will also reflect the differences in observed characteristics (**x**) between the two samples. Therefore, the difference in mean observed hospital days between FFS and HMO beneficiaries can be decomposed into four components: observed

selection, HMO effect, HMO unobserved selection, and FFS unobserved selection. To evaluate group-level average effects, predictive ratios will be calculated. The predictive ratio is defined as the ratio of the sum of predicted utilization of all individuals in a group to the sum of their actual utilization.

Predicting Hospital Use with Unobserved Selection Purged Out

The following two sections decompose the difference in mean observed inpatient use between FFS and HMO beneficiaries into observed selection, FFS and HMO unobserved selection, and the HMO effect. Table 10 compares predicted and observed means of inpatient use for FFS and HMO beneficiaries. Mean predicted admission rate is greater than mean observed rate for HMO enrollees, but less than mean observed rate for FFS beneficiaries. Observed hospital days per 1,000 averages 1,022 days for HMO enrollees and 1,828 days for FFS beneficiaries, whereas predicted days averages 1,161 and 1,633 days respectively. Consistent with the estimated direction of unobserved selection, HMO enrollees would have incurred more hospital use if they had been randomly enrolled whereas FFS beneficiaries would have incurred less hospital use if there had been no unobserved adverse selection.

TABLE 10
FFS and HMO Inpatient Use in Year 2, Observed and Predicted by the Simultaneous Equations Model

			HMO		FFS	
	HMO	FFS	Year 1 non-users	Year 1 users	Year 1 non-users	Year 1 users
Number of obs	78,693	79,933	69,572	9,121	66,756	13,177
Having one or more admissions in Year 2						
Observed	0.150	0.200	0.130	0.304	0.161	0.398
Predicted	0.160	0.188	0.137	0.339	0.152	0.373
Obsd/pred-1 (%)	-6.4	6.4	-5.1	-10.3	6.3	6.8
Total hospital days in Year 2 (days/1,000)						
Observed	1,022	1,828	843	2,389	1,303	4,487
Predicted	1,161	1,633	923	2,972	1,183	3,910
Obsd/pred-1 (%)	-11.9	12.0	-8.7	-19.6	10.2	14.8

Note: Observed total LOS greater than 60 days was set to 60 and one day was assigned to those who were admitted in Year 2 with less than 1 day of total length of stay.

According to Equations (16) and (17), mean unobserved adverse selection of FFS beneficiaries is equal to 195 days (1,828-1,633), whereas mean unobserved HMO favorable selection is 138 days (1,161-1,022 with rounding errors).[9]

An overall measure of unobserved selection is calculated as:

Unobserved selection =

(observed days per thousand / predicted days per thousand - 1) * 100%.

[9] The decompositions are all in terms of mean effects. For example, FFS observed use is the average use of the beneficiaries in the FFS sample. Therefore, the simulation results do not depend on the proportion of HMO and FFS beneficiaries in our sample. Our sample has a roughly equal number of FFS and HMO beneficiaries. Risk HMO plans enrolled less than 10 percent of FFS beneficiaries in California in 1995.

It measures the extent to which observed hospital days deviate from unbiased predicted days. Unobserved favorable selection of HMO enrollees makes their observed hospital days 11.9 percent lower than predicted days. In contrast, unobserved adverse selection of FFS beneficiaries makes their observed hospital days 12.0 percent higher than predicted days. Unobserved selection is particularly large among those who used hospital services in Year 1, consistent with the descriptive analyses shown in Table 6.

Decomposing the Difference in Hospital Use between FFS and HMO Beneficiaries

The difference in mean observed hospital days between FFS and HMO, 806 days (1,828-1,022), can be disaggregated into four components--the HMO effect, HMO unobserved selection, FFS unobserved selection, and observed selection as discussed above. The last section quantified the magnitude of unobserved selection. This section will quantify the HMO effect and observed selection.

While the difference between observed and predicted hospital days is attributable to unobserved selection, the difference between FFS and HMO mean predicted hospital days, 1,633-1,161=472 days, reflects the HMO effect as well as observed selection. For a given group of individuals, the difference between their predicted hospital use in the FFS and HMO settings would only reflect the HMO effect. However, since our HMO and FFS samples are not matched, the difference also reflects observed selection.

To estimate the combined HMO effect on hospital days per 1,000, we need to estimate hospital use of HMO enrollees in the FFS setting. Since our HMO and FFS samples are not matched, moving FFS beneficiaries to the HMO setting is not equivalent to moving HMO enrollees to FFS. The following simulations are conducted under the scenario of moving HMO enrollees to the FFS setting because CMS risk adjustment models are meant to predict FFS costs of plan enrollees.

HMO enrollees' predicted hospital days per 1,000 averages 1,161 days whereas their observed hospital use averages 1,022 days. Given their favorable unobserved selection, HMO enrollees would have incurred more hospital use (1,161-1,022 = 138 days) if they had been randomly enrolled. The 138 days represents their unobserved favorable selection.

If the same plan enrollees had been randomly enrolled but treated in FFS (switch the HMO dummy in Equation (13) from HMO to FFS), their predicted hospital use would average 1,317 days. The difference between predicted days of HMO enrollees treated in the FFS and HMO settings, 157 days (1,317-1,161), represents the mean HMO effect (the increase of use) that would be observed when moving randomly selected HMO enrollees from HMOs to FFS. Thus, HMOs shorten the counterfactual FFS hospital days of HMO enrollees by 11.9 percent (1-1,161/1,317). Equivalently, HMO enrollees would incur an additional amount of use that is equal to 13.5 percent (1/(1-11.9%)-1) times HMO days when moved to the FFS setting.

The 1,317 days is the counterfactual use of HMO enrollees treated in the FFS setting had they been randomly chosen from the entire population. FFS predicted use (1,633) and this counterfactual use both have unobserved selection purged out. The only difference is that FFS and HMO beneficiaries have different observed PIP-DCG characteristics. Thus, the difference between them, 315 days (1,633-1,317), represents observed selection only.

The HMO effect calculated above (157 days) is based on unbiased predicted days. It is the HMO effect that would have been observed had HMO enrollees been randomly enrolled. To calculate the actual increase in use when moving self-selected HMO enrollees to the FFS setting, HMO enrollees' unobserved favorable health status needs to be taken into account. The HMO effect should be based on observed days, which reflects HMO enrollees' unobserved better health, rather than predicted days. The question is how much more hospital use self-selected HMO enrollees may incur, given their unobserved favorable selection, if they had been treated in FFS. Given their unobserved selection, HMO enrollees' use could have increased by 1,022*13.5%=138 days to 1,161 days if they had remained in FFS. The HMO effect based on observed use is 138 days. Coincidently, HMO unobserved favorable selection also equals 138 days. HMO unobserved selection (138 days) times 13.5%, which equals 19 days, represents the difference when the HMO effect is calculated based on observed use instead of predicted use. To estimate the HMO effect when moving self-selected HMO enrollees to FFS, the 19 days must be deducted from the HMO effect calculated based on predicted use (157), which overestimates the HMO effect because of HMO unobserved favorable selection.

The 19 days should be added back to HMO unobserved favorable selection. Adjusted HMO unobserved selection is 138+19=157 days. HMO unobserved selection is larger in the FFS setting because the HMO effect has compressed everything including unobserved selection in the HMO setting.

Therefore, the total difference in observed hospital days between HMO and FFS (806 days) can be disaggregated into the following components:

- HMO effect: 138 days (17% of 806 days)
- Observed selection: 315 days (39%)
- Total unobserved selection: 352 days (44%)

  -Unobserved FFS adverse selection: 195 days (24%)

  -Unobserved HMO favorable selection: 157 days (19%)

MedPAC (2000) found that the PIP-DCG model explained 35 percent of total difference in Medicare payments between new Medicare+Choice enrollees and FFS beneficiaries in 1997, which is quite close to our estimate (39 percent). Their sample was selected in the exact same way as ours, but they compared pre-enrollment Medicare payments as opposed to post-enrollment hospital days.

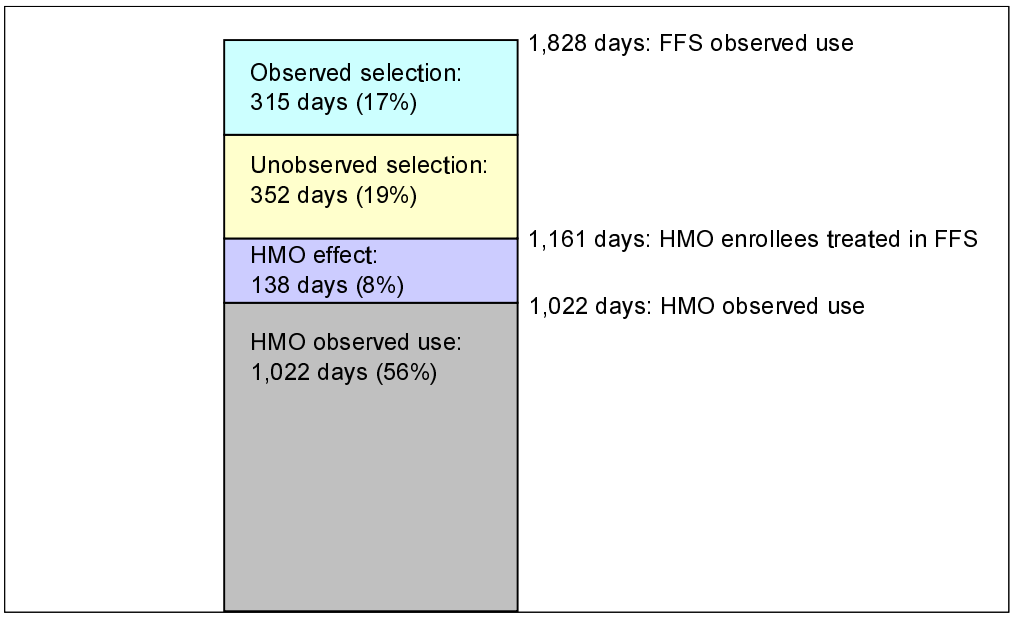
Hill et al. (1992) evaluated the Medicare risk HMO program based on a detailed survey. Their nationally representative HMO sample consists of those who had been enrolled at

or before the beginning of 1990 whereas our sample only includes new plan enrollees. But since Medicare only started to contract with risk-based managed care plans in 1985, many of their enrollees should be quite new at the time. Their sample has an equal number of FFS and HMO beneficiaries but does not include any decedents. They tested for unobserved selection and concluded that their OLS model adequately controlled for selection bias. They found that the AAPCC factors explained about 25 percent of the raw enrollee-nonenrollee difference in means of hospital days. Adding other covariates, which is equivalent to unobserved variables in our model since selection was fully captured by their model, accounted for another 40 percent of the enrollee-nonenrollee difference. The remaining difference (35%) was due to the HMO effect. Their estimate of unobserved selection (40%) is quite close to ours (44%), which is not surprising given that HMO favorable selection did not wane in the 1990s. As expected, the PIP-DCG risk factors explain more of the HMO-FFS difference in hospital days (39%) than the AAPCC factors (25%). This study finds a smaller HMO effect (17%) than theirs (35%) possibly because the health care market in California was more mature and competitive in 1995 than the national market in 1990.

Using the 1,828 observed hospital days of FFS beneficiaries as the denominator, the HMO effect accounts for about 8 percent of FFS observed use (138/1,828). Unobserved selection, adverse in FFS and favorable in HMO, accounts for 19 percent (352/1,828). Observed selection accounts for 17 percent (315/1,828). The decomposition of the total difference between FFS and HMO observed use is shown in Figure 5.

After adjustment for the PIP-DCG risk factors, HMO enrollees still incur about 27 percent fewer hospital days ((352+138)/1,828) than FFS beneficiaries. The HMO effect accounts for about 30 percent of it with the remainder due to unobserved selection.

FIGURE 5 Disaggregating the Difference between FFS and HMO Observed Hospital Days



Finally, let us take a quick look at admission rates. Given that HMOs have an insignificant small effect on admission, the HMO effect is ignored. The difference in mean admission rates between FFS and HMO, 0.20-0.15=0.05, can be decomposed into two components: unobserved and observed selection. The difference in predicted admission rates, 0.188-0.160=0.028, is attributable to observed selection. The rest of it, 0.05-0.028=0.022, is attributable to unobserved selection. Therefore, the difference in

mean observed admission rates between FFS and HMO is roughly equally explained by unobserved and observed selection.

Decomposing Bias in an FFS-based Naïve Model for Hospital Use

CMS uses a naïve PIP-DCG model developed on FFS cost data to estimate FFS costs of HMO enrollees with no attempt to correct for possible unobserved selection. To simulate bias in the FFS-based model, a naïve 2-part model for hospital use is estimated over the FFS sample only and applied to predict hospital use of HMO enrollees. The bias in the FFS-based 2-part model comes from unobserved selection and the HMO effect. Estimation results of the naïve 2-part model developed over the FFS sample are presented in Table 11.

TABLE 11
A Naïve 2-part Hospital-use Model Estimated over the FFS Sample

	Equation	
Variable	Any Admission in Year 2	Log Total LOS in Year 2
69-71	0.083*	-0.024
	(0.020)	(0.034)
72-74	0.107*	0.014
	(0.020)	(0.034)
75-77	0.215*	0.044
	(0.021)	(0.033)
78-80	0.284*	0.092*
	(0.021)	(0.034)
81-83	0.370*	0.077*
	(0.022)	(0.034)
84-86	0.405*	0.147*
	(0.024)	(0.036)
87-89	0.493*	0.147*
	(0.027)	(0.038)
90-92	0.513*	0.096*
	(0.032)	(0.045)
93+	0.410*	0.077**
	(0.033)	(0.045)
male	0.117*	0.004
	(0.011)	(0.016)
Medicaid	0.150*	0.179*
	(0.013)	(0.018)
disabled	0.266*	0.122*
	(0.020)	(0.027)
DCG 4-7	0.454*	0.103*
	(0.021)	(0.027)
DCG 8-12	0.632*	0.179*
	(0.018)	(0.024)
DCG 14-16	1.037*	0.481*
	(0.030)	(0.032)
DCG 18+	1.072*	0.548*
	(0.037)	(0.041)
constant	-1.301*	1.539*
	(0.017)	(0.029)

Note: * significant at 5%;
** significant at 10%.

Table 12 compares mean observed hospital use of HMO enrollees with mean HMO use predicted by the FFS-based naïve model as well as the simultaneous equations model. HMO enrollees' hospital use predicted by the FFS-based naïve model averages 1,488 days. It is lower than mean observed use of FFS beneficiaries (1,828 days) because HMO enrollees have favorable observed characteristics compared to FFS beneficiaries.

However, even after controlling for observed selection, the naïve model still overestimates HMO enrollees' observed use by 46 percent (1,488/1,022 - 1). Overprediction is measured by predictive ratio (predicted/observed – 1). This occurs because the FFS-based naïve model ignored the HMO effect and unobserved selection when developed on the FFS sample. Overprediction is particularly large for those who had hospital use in Year 1. The simultaneous equations model's overprediction is much smaller than the FFS-based naïve model because its overprediction is due only to HMO unobserved selection.

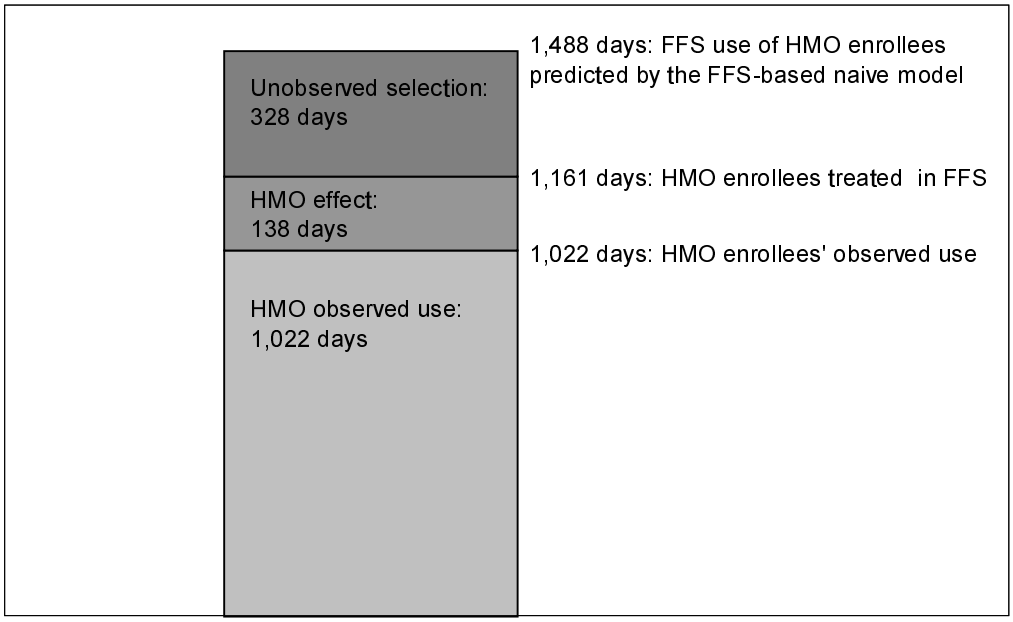
TABLE 12
HMO Inpatient Use Predicted by the Simultaneous Equations Model and the FFS-based Naïve Model

		HMO Sample	
	HMO Sample	Year 1 non-users	Year 1 users
Number of Observations	78,693	69,572	9,121
Having one or more admissions in Year 2			
Observed	0.150	0.130	0.304
Predicted by the simultaneous eq. model	0.160	0.137	0.339
Predicted by the FFS-based naïve model	0.175	0.150	0.371
Overprediction by the FFS-based model (%)	17.0	15.5	21.9
Total hospital days in Year 2 (days/1,000)			
Observed	1,022	843	2,389
Predicted by the simultaneous eq. model	1,161	923	2,972
Predicted by the FFS-based naïve model	1,488	1,178	3,855
Overprediction by the FFS-based model (%)	45.6	39.7	61.4

Note: Observed total LOS greater than 60 days was set to 60 and one day was assigned to those who were admitted in Year 2 with less than 1 day of total length of stay.

The FFS-based naïve model overpredicts hospital use of HMO enrollees because it ignored the HMO effect and FFS and HMO unobserved selection. Among the total overestimation of the FFS-based model (466 days), 138 days is the HMO effect if HMO enrollees had remained in FFS (see the previous section). Thus, about 30 percent (138/466) of the model's overprediction is attributable to the HMO effect, and the remaining 70 percent (328/466) is attributable to ignored unobserved selection. The total bias in the FFS-based naïve model is disaggregated in Figure 6.

FIGURE 6 Disaggregating Hospital Days for HMO Enrollees Predicted by the FFS-based Naïve Model



From the perspective of CMS, the Medicare+Choice program should provide beneficiaries more choices and additional values without costing Medicare more than it would cost to provide the basic Medicare package to enrollees through the FFS program (MedPAC, 2000). Thus, the CMS risk adjustment models try to estimate costs of plan enrollees if they had remained in the traditional FFS program. If CMS risk adjustment models did produce unbiased estimates of FFS costs of HMO enrollees, Medicare would be financially neutral as to whether beneficiaries choose Medicare+Choice or stay in the traditional FFS program.

Unfortunately, unobserved selection ignored by the FFS-based naïve model leads to a substantial overestimation of HMO enrollees' resource use. FFS hospital use of HMO

enrollees is the sum of their mean observed use (1,022 days) and the HMO effect (138 days) if they had been treated in FFS. The resulting 1,161 days is the counterfactual hospital use of HMO enrollees in the FFS setting given their unobserved favorable selection. Ignoring unobserved selection results in an overestimation of 328 days. Therefore, for the Medicare program, unobserved selection leads to a 28 percent (328/1,161) overestimation of FFS hospital use of new HMO enrollees.

CHAPTER 6

POLICY IMPLICATIONS

The PIP-DCG model will remain in use until 2007. This study found that the PIP-DCG model does not adequately adjust for selection bias. Unobserved selection is systematically different in the FFS and HMO populations, with HMO enrollees healthier and FFS beneficiaries sicker in ways not captured by the PIP-DCG model. While a naïve model developed on FFS data can consistently predict costs for a random FFS sample, its predictions would not be consistent for an HMO sample that is identical to a random FFS sample in observed characteristics. Bias was built into the model when it was developed over an FFS sample with unobserved selection different from that of the HMO population. In fact, the FFS-based naïve model overestimated hospital use of new HMO enrollees by 28 percent compared to their use if they had been served in FFS. It is noteworthy that the overestimation is not only caused by unobserved HMO favorable selection as commonly stated. Rather, both FFS and HMO unobserved selection are responsible for the overestimation. Unobserved FFS adverse selection may well be responsible for as much bias as unobserved HMO favorable selection.

The 28 percent overestimation of FFS hospital days of HMO enrollees is substantial in dollar terms as hospital days are expensive. Medicare spending is roughly evenly split between Part A and Part B. Assuming that Part A expenditures are proportional to hospital days and that the PIP-DCG model does not overestimate Part B spending much, the 28 percent overestimation would translate into an approximately 15 percent

overestimation of Medicare spending on new HMO enrollees (actual excess payments shall be smaller since Medicare discounts estimated payments by 5 percent). However, the two assumptions need to be modified. First, since marginal costs of hospital days decline over time, given that HMOs do not have much impact on admission rates, the 28 percent overpredicted days may not be as expensive as initial hospital days. Thus, the model's overprediction of payments may be less than 15 percent. Second, according to Hill et al. (1992), the AAPCC overestimates Part B spending by 7.5 percent and Part A spending by 12.4 percent before the 5 percent discount. Their finding is based on a representative sample of the entire Medicare HMO population. Overestimation of Part B spending should be greater in our sample that contains new HMO enrollees only, but it may be mitigated by the improved predictive power of the PIP-DCG model. On balance, the PIP-DCG is likely to overestimate Part B spending as well. Therefore, the PIP-DCG's overestimation of overall Medicare spending on new HMO enrollees could be more or less than 15 percent if the above two effects do not offset each other.

Note that this study examines unobserved selection of new HMO enrollees only, which is usually much stronger than enrollees of long tenure. Net biased selection experienced by a particular health plan is the result of a dynamic process of enrollment, disenrollment, and regression toward the mean during the interim (Welch, 1985; Halvorson and Stix, 1988). Regression toward the mean of HMO enrollees of long tenure will bring predictions of the FFS-based model more in line with true costs. MedPAC (2000) investigated the change of mortality rates of HMO and FFS beneficiaries over time, and found that after controlling for age, sex and Medicaid, beneficiaries enrolled in

Medicare+Choice for less than one year had a 21 percent lower mortality rate than traditional Medicare beneficiaries, but those enrolled for more than 5 years only had a rate 11 percent lower. Similarly, GAO (2000b) found that although beneficiaries who had been plan members for several years continue to use fewer health care services than FFS beneficiaries, the gap between them grew much narrower over time. Since favorable health status of plan enrollees never completely regresses to the mean, overprediction of the PIP-DCG model will remain quite significant.

Overestimation of FFS costs of HMO enrollees by the PIP-DCG model, and to a lesser extent, the HMO efficiencies will continue to enable plans to offer extra benefits while charging no or low premiums. Excess payments to health plans essentially represent a government subsidy to health plans and their enrollees at the expense of taxpayers and FFS beneficiaries who are usually older and sicker than their Medicare+Choice counterparts. Better risk adjustment can reduce excess payments and financial loss for the Medicare program, level the playing field for private plans by basing reimbursements on their true opportunity costs, and encourage them to enroll the sick, the disabled and other vulnerable groups who will in principle benefit more from coordinated care.

There are two problems associated with the FFS basis of the PIP-DCG methodology. First, the FFS basis of CMS risk adjustment models is not appropriate in the presence of systematically different unobserved selection between the HMO and FFS populations. It is true that in order to forestall profitable cream-skimming, a Medicare risk adjustment model does not need to be perfect. CMS only needs to make sure private health plans do

not outperform the government in risk adjustment (Van de Ven and Ellis, 2000). However, an imperfect model developed on FFS data does not consistently predict FFS costs of HMO enrollees when unobserved selection left out by the model is systematically different between the HMO and FFS populations.

The bias cannot be easily eliminated because other than abandoning the FFS basis, the only solution is to develop an actuarially impeccable model that captures all selection bias. Although it produces consistent estimates for randomly selected beneficiaries, the simultaneous equations model itself is not that useful to CMS for payment purposes because CMS needs to take into account unobserved favorable selection of HMO enrollees in its payments to Medicare+Choice plans.

One may think that a model developed on HMO data should be free of bias in predicting costs of HMO enrollees in the HMO setting. However, it would deprive health plans of all the savings they may achieve. In addition, costs of individual enrollees must be derived from utilization or other data sources, which may not be easy to do. Extensive data collection and reporting burdens may discourage prospective health plans from participating in the Medicare+Choice program. It may also encourage inefficient practice styles on the part of health plans.

Second, the FFS basis of a diagnosis-based model like the PIP-DCG has an inherent problem that may hamper its ability to accurately predict FFS costs of HMO enrollees even when the model itself is unbiased. Suppose that HMO enrollment were random, or

that the PIP-DCG could perfectly capture selection bias, would the model be able to accurately predict FFS costs of HMO enrollees? The answer is maybe. If a group of individuals randomly assigned to the HMO setting receives a different set of inpatient diagnoses than they would in FFS, their expected costs predicted by a diagnosis-based model would obviously be different than if they had remained in FFS. Unlike demographic risk factors, inpatient diagnoses are subject to change for a variety of reasons including HMO practice syles and possible gaming. If HMOs increase or decrease admissions and/or shift the mix of inpatient diagnoses, a perfect risk adjustment model may still over or underestimate FFS costs of HMO enrollees.

This problem is inherent in any risk adjuster that includes risk factors susceptible to manipulation. However, it does not seem to be a big problem for the PIP-DCG model. The HMO industry claims that the PIP-DCG methodology may penalize Medicare+Choice organizations that efficiently substitute less expensive care for hospital inpatient care. This is not empirically observed in our 1995 California data. HMOs had little impact on admission rates. Given that, it is not unreasonable to suspect that HMOs may not have shifted the mix of inpatient diagnoses either. That is, a given group of individuals is likely to receive the same set of inpatient diagnoses whether they are served in the HMO or FFS setting.

At least in California, given that historically HMOs did not reduce admission rates, it is highly unlikely that they would do so at present with the PIP-DCG risk adjustment in place. If anything, it is only likely that they might be upcoding inpatient diagnoses.

However, the PIP-DCG model was designed with the need to reduce undesirable incentives in mind. For example, the exclusion of short stays and discretionary hospitalizations and no reward for multiple hospitalizations all mitigate possible gaming. Since hospitalizations are expensive and higher future payments from previous admissions are uncertain as enrollees may die or disenroll, plans do not have strong incentives to admit enrollees unnecessarily. Furthermore, upcoding can be relatively easily checked through auditing and model improvement.

The bias in the PIP-DCG model is ultimately due to its limited predictive power rather than its FFS basis. The FFS basis is a deliberate policy choice, but the model's limited predictive power is a methodological flaw that needs to be fixed. The FFS development basis would not cause any bias for an actuarially perfect model unless HMO practice styles alter some of its risk factors.

The substantial bias in the PIP-DCG model calls for a better risk adjustment model. CMS plans to introduce a comprehensive payment model based on diagnoses from hospital inpatient, hospital outpatient, and physician encounters in 2007. Less than 20 percent of Medicare beneficiaries receive one or more diagnoses in the inpatient setting in one year, but over 85 percent of them receive one or more diagnoses in the outpatient setting (Weiner et al., 1996). If health plans could further reduce hospital admissions and substitute less expensive care for inpatient care, under the proposed all-encounter model, they would be financially better off than under the PIP-DCG methodology that gives no credit for these efficiency improvements. However, the model will place even more data

collection burdens on health plans than the PIP-DCG. An all-encounter risk adjustment model within the DCG family has been developed (Pope et al., 2000b). It can explain 11.2 percent of individual expenditure variation, two times greater than the PIP-DCG. Even though the all-encounter model was still developed on FFS data, it would be less biased in predicting FFS costs of HMO enrollees since it captures more unobserved selection than the PIP-DCG. This means that health plans could get paid even less than they currently are.

Health plans started to withdraw from the Medicare+Choice program in the late 1990s for a variety of reasons including allegedly insufficient Medicare reimbursements. This study suggests otherwise: the PIP-DCG model still overpays. If the primary goal of the Medicare+Choice program is to control costs, a more actuarially sound risk adjuster such as the all-encounter model should be introduced as soon as possible. Plan withdrawals are not necessarily bad news if they mean decrease in excess payments. However, managed care provides additional values such as less cost-sharing, coordinated care and preventive services. To the extent that these added values are worth the extra dollars, a better risk adjustment model that cuts further down on payments should be balanced with its potential impact on the growth of the Medicare+Choice program.

CHAPTER 7

CONCLUSIONS AND LIMITATIONS

Conclusions

This study found that the PIP-DCG model cannot adequately control for selection bias. Unobserved selection is systematically different between the HMO and FFS populations, with HMO enrollees healthier and FFS beneficiaries sicker in ways not captured by the PIP-DCG model. Thus, a naïve model developed on FFS data ignoring unobserved selection considerably overestimates hospital use of new HMO enrollees even after controlling for the PIP-DCG risk factors.

HMOs reduce the probability of having at least one admission per year by 1.8 percent. But the effect is highly insignificant. HMOs do shorten total length of stay given admission per year by 10.3 percent. In total, HMOs shorten hospital days/1,000 per year for HMO enrollees by 11.9 percent compared to what they would incur in the FFS setting.

The difference in mean total hospital days between FFS beneficiaries and new HMO enrollees is 806 days, among which the PIP-DCG risk factors account for 315 days (39%). After adjustment for observed selection, new HMO enrollees still incur 490 (61%) more hospital days than FFS beneficiaries, among which 138 days (17%) is due to the HMO effect and 352 days (44%) is due to unobserved selection.

A naïve 2-part model for hospital use developed on FFS data ignoring unobserved selection overestimates hospital days of new HMO enrollees by 46 percent. After subtracting the HMO effect, the naïve FFS-based model overestimates hospital use of new HMO enrollees by 28 percent compared to their use if they had been treated in FFS.

The overestimation bias is ultimately due to the limited predictive power of the PIP-DCG model. Unobserved selection not captured by the model is adverse in the FFS population but favorable in the HMO population. As a result, the limited FFS-based model overestimates FFS use of HMO enrollees. To reduce excess payments, CMS needs a more comprehensive model that can better capture underlying health risk. A risk adjustment model that includes many predictors, however, may be cumbersome to implement. Such a model will further reduce excess payments to Medicare+Choice plans that have been experiencing financial difficulties since the late 1990s. More plan withdrawals could further reduce Medicare beneficiaries' access to managed care. Thus, CMS faces a trade-off between the need for reducing excess payments, the administrative burden of its risk adjustment models, and the negative impacts of reduced payments on health plan participation in the Medicare+Choice program.

As for modeling, this study demonstrates that the endogeneity of HMO enrollment may substantially bias model estimates. Controlling for unobserved selection is necessary for any serious endeavor to measure the true HMO effect. When the underlying distribution of error terms is not normal, a more robust specification like the one adopted in this study

works much better than a normal-based maximum likelihood estimator. Wrongly specified distribution of error terms may lead to implausible estimates.

Limitations

The study has some limitations. FFS beneficiaries may not have any supplemental health insurance coverage. Among those who have supplemental coverage, they may have Medicaid, Medigap, and/or employer-sponsored coverage. However, only Medicaid coverage is known in our data set. To the extent that beneficiaries with these supplemental health insurance did not display much different moral hazards in hospital services use (i.e., hospital use did reflect underlying health status), not controlling for them would not pose a big problem. Hospitalizations were largely non-discretionary at least in California. Different model types of HMO (e.g., staff, group, network, or independent practice association) are not controlled for either. Most of the HMOs in our sample are of the IPA model type. It is not clear whether unobserved selection of HMO enrollees differs across HMO model types. If HMO enrollees did differentiate HMO model types, our HMO enrollment model that only distinguished between enrollment and non-enrollment would be too simplistic.

The dependent variable in this study is hospital use rather than Medicare payments. Ultimately we are concerned about how much the PIP-DCG overpays health plans in dollar terms. However, it is not easy to derive Medicare payments from hospital days as

discussed above. Given the model's 28 percent overprediction of hospital days, overpayments should remain substantial under the PIP-DCG.

The study is based on a sample of new HMO enrollees in California. The HMO population in California may not represent the national population. In addition, the California health care market is more competitive than many other markets. Thus, we should be cautious in extrapolating the findings to markets other than California. Also, the study uses historical data from a period when the PIP-DCG risk adjustment methodology was not in place. The findings may not remain valid in the current environment due to possible behavioral responses to PIP-DCG risk adjustment. However, undesirable behavioral responses should be curbed by various built-in features of the PIP-DCG methodology.

Appendix: STATA Code for Estimating the Simultaneous Equations Model

```
************************************************************
* ESTIMATING THE SIMULTANEOUS EQUATIONS MODEL WITH θ        *
* APPROXIMATED BY A 3-SUPPORT POINT DISCRETE DISTRIBUTION   *
*                                                           *
************************************************************

capture log close
clear
cd "C:\WINDOWS\Desktop"
log using dfm3.log, replace
set more off
set mem 70m
set matsize 700

************************
* Get needed variables *
************************

global rhs3  "age71 age74 age77 age80 age83 age86 age89 age92 age00
male medicaid disabled dcga-dcgd"

use $rhs3 hmo user totlos using C:\WINDOWS\Desktop\sample95

replace totlos=1 if totlos==0
replace totlos=min(totlos,60) if totlos~=.
gen lnlos=ln(totlos) if totlos~=.


****** Model Specification ******
* $ML_y1: hmo
* $ML_y2: user
* $ML_y3: logged total los
* xb1: hmo enrollment
* xb2: any admission (including an hmo dummy)
* xb3: logged total los (including an hmo dummy)
* sigma: sigma (>0) of the white noise term in the LOS equation
*
* Factor distribution:
* d1: 1st point of support
* d2: 2nd point of support
* d3: 3rd point of support
* d1*p1+d2*p2+d3*p3=0
* p1: 0<p1<1
* p2: 0<p2<1
* p3: 0<p3<1
* p1+p2+p3=1
* rho2: factor loading in the user equation
* rho3: factor loading in the total LOS equation
*********************************
```

```
#delimit  ;

capture program drop discrete;
program define discrete;

      args lnf xb1 xb2 xb3 sigma d1 d2 p1_t p2_t rho2 rho3;

      tempname p1 p2 p3 d3;
      scalar `p1'=normprob(`p1_t');
      scalar `p2'=normprob(`p2_t');
      scalar `p3'=1-`p1'-`p2';
      scalar `d3'=(-`p1'*`d1'-`p2'*`d2')/`p3';

      * lnf of hmo enrollees with use;
      quietly replace `lnf' = ln(
      `p1'  *normprob(`xb1'+`d1')
            *normprob(`xb2'+`rho2'*`d1')
            *normd((($ML_y3-`xb3'-`rho3'*`d1')/`sigma')/`sigma'

      +`p2' *normprob(`xb1'+`d2')
            *normprob(`xb2'+`rho2'*`d2')
            *normd((($ML_y3-`xb3'-`rho3'*`d2')/`sigma')/`sigma'

      +`p3' *normprob(`xb1'+`d3')
            *normprob(`xb2'+`rho2'*`d3')
            *normd((($ML_y3-`xb3'-`rho3'*`d3')/`sigma')/`sigma'
            )
      if $ML_y1==1 & $ML_y2==1;

      * lnf of hmo enrollees with no use;
      quietly replace `lnf' = ln(
      `p1'  *normprob(`xb1'+`d1')
            *normprob(-`xb2'-`rho2'*`d1')

      +`p2' *normprob(`xb1'+`d2')
            *normprob(-`xb2'-`rho2'*`d2')

      +`p3' *normprob(`xb1'+`d3')
            *normprob(-`xb2'-`rho2'*`d3')
            )
      if $ML_y1==1 & $ML_y2==0;

      * lnf of ffs stayers with use;
      quietly replace `lnf' = ln(
      `p1'  *normprob(-`xb1'-`d1')
            *normprob(`xb2'+`rho2'*`d1')
            *normd((($ML_y3-`xb3'-`rho3'*`d1')/`sigma')/`sigma'

      +`p2' *normprob(-`xb1'-`d2')
            *normprob(`xb2'+`rho2'*`d2')
            *normd((($ML_y3-`xb3'-`rho3'*`d2')/`sigma')/`sigma'

      +`p3' *normprob(-`xb1'-`d3')
            *normprob(`xb2'+`rho2'*`d3')
            *normd((($ML_y3-`xb3'-`rho3'*`d3')/`sigma')/`sigma'
            )
      if $ML_y1==0 & $ML_y2==1;
```

```
      * lnf of ffs stayers with no use;
      quietly replace `lnf' = ln(
      `p1'  *normprob(-`xb1'-`d1')
            *normprob(-`xb2'-`rho2'*`d1')

      +`p2' *normprob(-`xb1'-`d2')
            *normprob(-`xb2'-`rho2'*`d2')

      +`p3' *normprob(-`xb1'-`d3')
            *normprob(-`xb2'-`rho2'*`d3')
            )
      if $ML_y1==0 & $ML_y2==0;
end;

ml model lf discrete
      (hmo: hmo= $rhs3)
      (user: user= $rhs3 hmo)
      (lnlos: lnlos=$rhs3 hmo)
      (sigma: )

      (dis1: )
      (dis2: )

      (p1_t: )
      (p2_t: )

      (rho2: )
      (rho3: )
      ,
      collinear missing robust
;

* The following initials are from the 2-point discrete factor model;
ml init

hmo:age71=-0.1451083
hmo:age74=-0.2980528
hmo:age77=-0.3737578
hmo:age80=-0.4722116
hmo:age83=-0.5283571
hmo:age86=-0.6802937
hmo:age89=-0.7402008
hmo:age92=-1.106411
hmo:age00=-1.436496
hmo:male=0.0146907
hmo:medicaid=-1.964826
hmo:disabled=0.158928
hmo:dcga=-0.3787442
hmo:dcgb=-0.3427379
hmo:dcgc=-0.5059057
hmo:dcgd=-1.069566
hmo:_cons=0.8703299

user:age71=0.0647852
user:age74=0.1253552
user:age77=0.2149543
```

```
user:age80=0.2903561
user:age83=0.3764425
user:age86=0.4475521
user:age89=0.519636
user:age92=0.5705091
user:age00=0.4776512
user:male=0.1412207
user:medicaid=0.1897212
user:disabled=0.2693069
user:dcga=0.4241311
user:dcgb=0.5967085
user:dcgc=1.009417
user:dcgd=1.044879
user:hmo=0.0295529
user:_cons=-1.389035

lnlos:age71=-0.0107824
lnlos:age74=0.0302122
lnlos:age77=0.070319
lnlos:age80=0.0947908
lnlos:age83=0.0789811
lnlos:age86=0.158088
lnlos:age89=0.1467759
lnlos:age92=0.1139843
lnlos:age00=0.1054534
lnlos:male=0.0291414
lnlos:medicaid=0.1883947
lnlos:disabled=0.1096801
lnlos:dcga=0.1046068
lnlos:dcgb=0.1602996
lnlos:dcgc=0.4362221
lnlos:dcgd=0.5456479
lnlos:hmo=-0.063837
lnlos:_cons=1.463821

sigma:_cons=.93

dis1:_cons=-3.7
dis2:_cons=.02

p1_t:_cons=0.2
p2_t:_cons=-.4

rho2:_cons=-0.02
rho3:_cons=-0.02
;

* tolerance defaults for internal Stata programs: 0 and -4 respectively
(defaults for ml programs: -7 and -6);

ml max, ltolerance (0) tolerance (1e-4) difficult;

#delimit cr

***************
log close
exit
```

BIBLIOGRAPHY

Alecxih, L.M.B, Lutzky, S., Sevak, P., and Claxton, G.: *Key Issues Affecting Accessibility to Medigap Insurance.* New York, NY: The Commonwealth Fund, August 1997.

Ash, A., Porell, F, Gruenberg, L., Sawitz, E., and Beiser, A.: Adjusting Medicare capitation payments using prior hospitalization data. *Health Care Financing Review* 10(4):17-29, 1989.

Beebe, J., Lubitz, J., and Eggers, P.: Using prior utilization to determine payments Medicare enrollees in health maintenance organizations. *Health Care Financing Review* 6(3):27-38, 1985

Blough, D., Madden, C., and Hornbrook, M.: Modeling risk using generalized linear models. *Journal of Health Economics* 18: 153-171, 1999

Boyes, W., Hoffman, D., and Low, S.: An econometric analysis of the bank credit scoring problem. *Journal of Econometrics* 40:3-14, 1989.

Christensen, S., and Shinogle, J.: Effects of supplemental coverage on use of services by Medicare enrollees. *Health Care Financing Review* 19(1):5-17, 1997.

Cameron, S., and Taber, C.: *Evaluation and identification of semiparametric maximum likelihood models of dynamic discrete choice*. Working paper, Department of Economics, Columbia University. 1998

Cosslett, S.: Distribution-free maximum likelihood estimator of the binary choice model. Econometrica 51(3):765-782, 1983.

Cox, D.F., and Hogan, C.: Biased selected and Medicare HMOs: analysis of the 1989-1994 experience. *Medicare Care Research and Review* 54(3):259-74, 1997

Dhanani, N., O'leary, J., Keeler, E. et al.: The Effect of Medicare Risk HMOs on Inpatient Utilization. *Health Services Research*, forthcoming.

Dowd, B., Feldman, R., Cassou, S. et al.: Health plan choice and the utilization of health care services. *Review of Economics and Statistics* 73(1): 85-93, 1991

Dowd, B., Feldman, R., Moscovice, I. et al.: An analysis of unobserved selection in the Medicare AAPCC (adjusted average per capita cost). *Health Care Financing Review* 17(3):35-57, 1996

Duan, N., Manning, W., Morris, C. et al.: Choosing between the sample-selection model and the multi-part model. *Journal of Business and Economic Statistics* 283-289, July 1984.

Duan, N.: Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association* 78(383):605-610, 1983.

Eggers, P.W., and Prihoda, R: Pre-enrollment reimbursement patterns of Medicare beneficiaries enrolled in "at-risk" HMOs. *Health Care Financing Review* 4(2), 1982

Ellis, R., and Ash, A.: Refinements to the Diagnostic Cost Group (DCG) Model. *Inquiry* 32:418-29, 1995.

Ellis, R., Pope, G., Lisa, M. et al.: Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financing Review* 17(3):101-127, 1996

General Accounting Office: *Medicare HMOs: HCFA can promptly eliminate hundreds of millions in excess payments*. GAO/HEHS-97-16, 1997.

General Accounting Office: *Medicare managed care: better risk adjustment expected to reduce excess payments overall while making them fairer to individual plans.* Statement of William Scanlon, Director, Health Financing and Systems, Health, Education, and Human Services Division. Testimony Before the Subcommittee on Health and Environment, Committee on Commerce, House of Representatives. GAO/HEHS-99-72, 1999.

General Accounting Office: *Medicare+Choice: Plan withdrawals indicate difficulty of providing choice while achieving savings*. GAO/HEHS-00-183, 2000a.

General Accounting Office: *Medicare+Choice: Payments Exceed Cost of Fee-for-Service Benefits, Adding Billions to Spending*. GAO/HEHS-00-161, 2000b.

General Accounting Office: *Medicare: Refinements should continue to improve appropriateness of provider payments*. Statement of William Scanlon, Director, Health Financing and Systems, Health, Education, and Human Services Division. Testimony Before the Subcommittee on Health and Environment, Committee on Commerce, House of Representatives. GAO/T-HEHS-00-160, 2000c

Goldberger, A: Abnormal selection bias. In *Studies in Econometrics, Time Series and Multivariate Statistics*, ed. Karlin, S., Amemiya, T., and Goodman, L. New York, NY: Academic Press, 1983.

Goldman, D., Leibowitz, A., and Buchanan, J.L.: Cost-containment and adverse selection in Medicaid HMOs. *Journal of American Statistical Association* 93(441):54-62, March 1998.

Greene, W.H.: *Econometric Analysis*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2000.

Greenwald, L., Esposito, A., Ingber, M. et al.: Risk adjustment for the Medicare program: lessons learned from research and demonstrations. *Inquiry* 35:193-209, 1998

Greenwald, L., Levy, J., and Ingber, M.: Favorable selection in the Medicare+Choice program: new evidence. *Health Care Financing Review* 21(3):127-35, 2000.

Gruenberg, L., Tompkins, C., and Porell, F.: The health status patterns of the elderly: implications for setting Medicare payments to HMOs. *Advances in Health Economics and Health Services Research* 10:41-73, 1989.

Gruenberg, L., Kaganova, E., Hornbrook M.: Improving the AAPCC with health-status measures from the MCBS. *Health Care Financing Review* 17(3):59-75, 1996.

Haffer, S.C.: *Comparing the Health Status of Medicare Beneficiaries in Fee-For-Service and Managed Care: Baseline Results from the 1998 Medicare Health Outcomes Survey*. A presentation at the 2000 AHSR Annual Meeting, Los Angeles: CA, June 2000

Halvorson, G., and Stix, P.: Hospital use by new and continuing HMO enrollees. *HMO Practice* 2(3):97-100, 1988.

Health Care Financing Administration, Office of Strategic Planning: *Report to Congress: Proposed Method of Incorporating Health Status Risk Adjusters into Medicare+Choice Payments*. Washington, DC, 1999.

Heckman, J.: The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5:475-92, 1976.

Heckman, J.: Dummy endogenous variables in a simultaneous equations system. *Econometrica* 46:931-959, 1978.

Heckman, J., and Singer, B.: A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52(2): 271-320, 1984.

Hellinger, F.: Selection bias in HMOs and PPOs: a review of the evidence. *Inquiry* 32:135-42, 1995

Hill, J.W., and Brown, R.S.: *Biased Selection in the TEFRA HMO/CMP Program.* Princeton, NJ: Mathematica Policy Research, Inc., 1990.

Hill, J., Brown, R., Chu, D., and Bergeron, J.: *The Impact of the Medicare Risk Program on the Use of Services and Costs to Medicare.* Princeton, NJ: Mathematica Policy Research, Inc., 1992.

Hornbrook MC, Bennett MD, Greenlick MR: Adjusting the AAPCC (adjusted average per capita cost) for selectivity and selection bias under Medicare risk contracts. Advances in *Health Economics and Health Services Research* 10:111-49. Greenwich, CT: JAI, 1989

Iezzoni, L., ed.: *Risk Adjustment for Measuring Health Care Outcomes*, 2nd ed. Ann Arbor, Michigan: Health Administration Press, 1997.

Ingber, M.: The current state of risk adjustment technology for capitation. *Journal of Ambulatory Care Manage* 21(4):1-28, 1998.

Kasper, J., Riley, G., McCombs, J. et al.: Beneficiary selection, use, and charges in two Medicare capitation demonstrations. *Health Care Financing Review* 19(1):37-49, 1988

Kiefer, J., and Wolfowitz, J.: Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Annals of Mathematical Statistics* 27:363-366, 1956.

Kronick, R., and Beyer, J.: Risk Adjustment is Not Enough: Strategies to Limit Risk Selection in the Medicare Program. New York, NY: The Commonwealth Fund, 1997.

Langwell, K.L., Rossiter, R., Brown, L. et al.: Early experience of health maintenance organizations under Medicare competition demonstrations. *Health Care Financing Review* 8:37-55, 1987.

Leung, S.F., and Yu,S.: On the choice between sample selection and two-part models. *Journal of Econometrics* 72(1-2):197-230, 1996

Lichtenstein, R., Thomas, J.W., Adams-Watson et al.: Selection bias in TEFRA at-risk HMOs. *Medicare Care* 29(4):318-31, 1991

Lubitz, J., Beebe, J., and Riley, G.: Improving Medicare HMO payment formula to deal with biased selection. In Scheffler, R., and Rossiter, L., eds.: *Advances in Health Economics and Health Services Research* 6:101-123. Greenwich, CT: JAI, 1985.

Luft, H.S. *Health maintenance organizations: dimensions of performance*. New York, NY: John Wiley and Sons, 1981.

Luft, H.S. Medicare and managed care. *Annual Review of Public Health* 19:169-175, 1998.

Manning, W., Leibowitz, A, Goldberg, G. et al.: A controlled trial of the effect of a prepaid group practice on use of services. *New England Journal of Medicine* 310(23): 1505-1510, 1984.

Manning, W.: The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics* 17:283-295, 1998

Manning, W.G., and Mullahy, J.: Estimating Log models to transform or not to transform? Technical Working Paper 246. National Bureau of Economic Research, 1999.

Manski, C. and Lerman, S.: The estimation of choice probabilities from choice based samples. *Econometrica* 45(8): 1977-1988, 1977

Manski, C.: Anatomy of the selection problem. *Journal of Human Resources* 24:343-60, 1989.

Medicare Payment Advisory Commission: *Report to the Congress: improving risk adjustment in Medicare.* Washington, DC: MedPAC, November 2000

Miller, R., Luft, H. Managed care plan performance since 1980. *JAMA* 271(19):1512-19, 1994

Morgan, R.O., Virnig, B.A., DeVito, C.A. et al.: The Medicare-HMO revolving door – the healthy go in and the sick go out. *New England Journal of Medicine* 337:169-75, 1997

Mroz, T.: Discrete factor approximations in simultaneous equation models: estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics* 92:233-274, 1999.

Newhouse, J., Manning, W., Keeler, E. et al.: Adjusting capitation rates using objective health measurers and prior utilization. *Health Care Financing Review* 10(3):41-54, 1989

Physician Payment Review Commission: *Prior Research on Risk Selection in Medicare HMOs.* Washington, DC: PPRC, 1996.

Pope, G., Ellis, R., Ash, A. et al.: Principal inpatient diagnostic cost group model for Medicare risk adjustment. *Health Care Financing Review* 21(3):127-35, 2000a

Pope, G., Ellis, R., Ash, A. et al.: *Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment*. Final report prepared for Health Care Financing Administration, December, 2000b.

Rice, T., and Bernstein, J.: *Supplemental Health Insurance for Medicare Beneficiaries. Medicare Brief #6*. Washington, DC: National Academy of Social Insurance, November 1999.

Riley, G., Rabey, E., and Kasper, J.: Biased selection and regression toward the mean in three Medicare HMO demonstrations: a survival analysis of enrollees and disenrollees. *Medicare Care* 27(4):337-51, 1989

Riley, G., Lubitz, J., and Rabey, E.: Enrollee health status under Medicare risk contracts: an analysis of mortality rates. *Health Services Research* 26(2):137-63, 1991

Schauffler, H., Howland, J., and Cobb, J.: Using chronic disease risk factors to adjust Medicare capitation payments. *Health Care Financing Review* 14(1):79-89, 1992.

Scully, T.: *Testimony of Thomas A. Scully, Administrator of Centers for Medicare & Medicaid Services, before the House Ways & Means Health Subcommittee on the Medicare+Choice Program, December 4, 2001*. Accessed on http://cms.hhs.gov in June 2002.

Stata Corporation. *Stata 7.0.* College Station, TX. 2001.

Thiede, C., Dowd, B., Feldman, R. et al.: Selection experiences in Medicare HMOs: pre-enrollment expenditures. *Health Care Financing Review* 20(4):197-209, 1999

Tudor, C.: *Risk adjustment implementation for Medicare*. A presentation by Cynthia G. Tudor with Health Care Financing Administration. March 2001.

Van de Ven, W., and Van Praag, B.: The demand for deductibles in private health insurance. *Journal of Econometrics* 17:229-252, 1981.

Van de Ven, W., and Ellis, R.: Risk adjustment in competitive health plan markets. In *Handbook of Health Economics*, Volume 1A, edited by Culyer, A., and Newhouse, J.: 755-847. New York, NY: Elsevier, 2000.

Weiner, J., Dobson, A., Maxwell, S. et al.: Risk-adjusted Medicare capitation rates using ambulatory and inpatient diagnoses. *Health Care Financing Review* 17(3):77-99, 1996

Welch, W.P., Frank, R., and Diehr, P.: Health care costs in health maintenance organizations: correcting for self-selection. In Scheffler, R., and Rossiter, L., eds.: *Advances in Health Economics and Health Services Research* 5:95-128. Greenwich, CT: JAI, 1984.

Welch, W.P.: Medicare capitation payments to HMOs in light of regression toward the mean in health care costs. In Scheffler, R., and Rossiter, L., eds.: *Advances in Health Economics and Health Services Research* 6:75-96. Greenwich, CT: JAI, 1985.