



An Analysis of Unobserved Selection in an Inpatient Diagnostic Cost Group Model

HONGJUN KAN*
Eli Lilly and Company

kanho@lilly.com

DANA GOLDMAN
EMMETT KEELER
RAND Corporation

dgoldman@rand.org
emmett@rand.org

NASREEN DHANANI
GLENN MELNICK
School of Policy, Planning and Development, University of Southern California; RAND Corporation

dhanani@rand.org
gmelnick@usc.edu

Received June 5, 2003; Revised November 14, 2003; Accepted January 28, 2004

Abstract. The study assesses unobserved selection bias in an inpatient diagnostic cost group (DCG) model similar to Medicare's Principal Inpatient Diagnostic Cost Group (PIP-DCG) risk adjustment model using a unique data set that contains hospital discharge records for both FFS and HMO Medicare beneficiaries in California from 1994 to 1996. We use a simultaneous equations model that jointly estimates HMO enrollment and subsequent hospital use to test the existence of unobserved selection and estimate the true HMO effect. It is found that the inpatient DCG model does not adequately adjust for biased selection into Medicare HMOs. New HMO enrollees are healthier than FFS beneficiaries even after adjustment for the included PIP-DCG risk factors. A model developed over an FFS sample ignoring unobserved selection overestimates hospital use of new HMO enrollees by 28 percent compared to their use if they had remained in FFS. Models that better captures selection bias are needed to reduce overestimation of Medicare HMO enrollees' resource use.

Keywords: risk adjustment, PIP-DCG, Medicare HMOs

1. Introduction

This study assesses unobserved selection bias in a type of prospective risk adjustment model that predicts individual health care costs using previous inpatient diagnoses and demographic risk factors. The inpatient diagnostic cost group (DCG) model was first developed for the Medicare program to estimate health care costs of beneficiaries enrolled in the Medicare + Choice program [7]. It is also used by health plans that enroll commercial populations for payment, profiling, and care management purposes. This paper evaluates the consistency of this inpatient DCG model in the Medicare context.

*Author to whom correspondence should be addressed.

Medicare is traditionally a fee-for-service (FFS) government insurance program. To contain costs and make managed care available to Medicare beneficiaries, the Health Care Financing Administration (now called the Centers for Medicare and Medicaid Services or CMS) began to contract with risk-based health management organizations (HMOs) to deliver health care services in 1985. Under the Medicare risk program, beneficiaries were allowed to enroll in an HMO or return to FFS Medicare in any month. Numerous studies have found that Medicare HMO enrollees are healthier than their FFS counterparts even after adjustment for certain risk factors, a phenomenon called HMO favorable selection [6, 8, 12, 21, 24]. The Medicare + Choice program, the latest version of Medicare's alternative delivery and financing mechanisms established by the Balanced Budget Act of 1997, enrolled about 5.6 million, or nearly 15 percent of the Medicare population in 2001 [26].

To try to make managed care payments reflect the underlying health status of HMO enrollees, Medicare first developed a risk adjustment model called the Adjusted Average Per Capita Cost (AAPCC) model that used a small number of demographic characteristics to risk-adjust capitation payments to HMOs. The model failed to adequately account for HMO favorable selection and led to overpayments to HMOs [13]. Under the requirement of the Balanced Budget Act of 1997, CMS began to phase in a new health-based risk adjustment model called the Principal Inpatient Diagnostic Cost Group (PIP-DCG) in 2000 that explicitly incorporates health status information in predicting future health care costs. The inpatient DCG model classifies principal inpatient diagnoses from previous hospitalizations into sixteen distinct PIP-DCG groups (DCG 4–DCG 29) based on clinical coherence and future cost implications. The number of each PIP-DCG group approximates future costliness (in \$1,000 units) of treating an individual with a diagnosis included in that group. Each individual is assigned one PIP-DCG. Patients with multiple previous hospitalizations have their PIP-DCG determined by the inpatient diagnosis with the greatest future cost implications. DCG 0 is assigned to those who incurred no hospital use in the past. The new risk adjustment model uses the PIP-DCG groups as well as demographic risk factors including age, sex, prior-year Medicaid coverage, disability as the original reason for Medicare entitlement, and working-aged status to predict future health care costs (see Pope et al. [22] for a detailed description of the PIP-DCG model). The PIP-DCG formula can explain 6.2 percent of variation in individual health care expenditures [22], significantly improving upon the AAPCC model which only explains 1.5 percent of individual expenditure variation. However, its predictive power still falls short of the theoretical limit of 14.5 percent that a prospective risk adjustment model can possibly achieve [20].

The PIP-DCG model was developed over a representative sample of FFS Medicare beneficiaries and estimates FFS costs of HMO enrollees if they had received treatment through the traditional FFS program. The model groups demographic characteristics and inpatient diagnoses into a limited number of cells, each with an associated marginal cost factor. An HMO enrollee's expected cost in the FFS setting is calculated by summing up the marginal costs of the cells he or she falls in. In order for the FFS-based model to consistently estimate costs of HMO enrollees in the FFS setting, HMO and FFS beneficiaries must have similar health status within each cell. In that case, biased selection would be fully captured by the included risk factors. However, to avoid putting excessive data collection burdens on HMOs, Medicare risk adjustment models such as the PIP-DCG are parsimonious. The small

number of cells may not be “narrow” enough to make individuals in each cell homogenous. When HMO enrollees are healthier than FFS beneficiaries within the demographic cells and/or have fewer hospitalizations or less severe conditions within the PIP-DCG cells, the model will overestimate HMO enrollees’ FFS costs. These systematic within-cell differences represent unobserved selection that is not captured by the risk adjustment model.

Several studies based on indirect measures of health care costs of HMO enrollees indicated that the PIP-DCG model is not adequate in controlling for HMO favorable selection. MedPAC [16] estimated that the PIP-DCG risk factors only accounted for 8 percentage points of a 23 percentage-point difference in average costs between HMO enrollees and FFS beneficiaries before they enrolled in HMOs. GAO [8] imputed HMO enrollees’ post-enrollment costs based on the pattern of changes in FFS costs over time and found that due to inadequate risk adjustment of the AAPCC, Medicare spent 13.2 percent more (\$3.2 billions) on plan enrollees in 1998 than if they had remained in FFS, of which the PIP-DCG model could only eliminate less than half when fully implemented. Excess payments essentially represent a government subsidy to health plans and HMO enrollees. Since HMOs tend to attract good risks from the Medicare population, beneficiaries left in the FFS sector are often older and sicker than HMO enrollees and yet they receive less money spent on them from the Medicare program. Better risk adjustment can reduce overpayments to HMOs and encourage health plans to compete on the basis of price, service, and quality, rather than on avoidance of risk.

The large observed differences in resource use between HMO and FFS beneficiaries reflect differences in observed risk factors, differences in unobserved risk factors, and more efficient managed care practice styles. We will call the impact of these styles, as quantified by the average HMO reduction in spending from FFS on identical populations, the “HMO effect”. Many studies have difficulty disentangling the effect of selection bias from the HMO effect. The RAND Health Insurance Experiment in the 1970s randomly assigned subjects to different FFS and managed care settings and found substantial cost-saving effects of HMOs [15]. However, their study focused on working-age people. A few other studies [5, 14, 17] rely on econometric techniques to control for selection bias and separate out the Medicare HMO effect from selection bias. For example, Dowd et al. [5] adopted a selection-corrected Tobit model to test the existence of unobserved selection bias in the AAPCC. Their sample was drawn from five risk HMOs and the FFS Medicare program in the Twin Cities from 1988 to 1989. Only FFS expenditures were observed. Their fully parametric model indicated favorable rather than adverse FFS selection.

This study investigates the consistency of the new PIP-DCG formula. It contributes to the risk adjustment literature by combining rarely available observed utilization data of individual HMO enrollees with a relatively new econometric method to assess unobserved selection in a diagnosis-based risk adjustment model. It aims to answer three specific questions: (1) Does a model of inpatient use that risk adjusts for PIP-DCGs and demographic characteristics adequately account for HMO favorable selection? (2) If not, how biased is an FFS-based model in predicting HMO enrollees’ inpatient use? (3) How much do Medicare HMOs reduce or increase inpatient use?

We use observed utilization data of HMO enrollees and a robust simultaneous equations model to estimate the HMO effect consistently and test unobserved selection bias in the

inpatient DCG model. The specific model assessed in this study is different from Medicare's original PIP-DCG model. Due to the lack of individual cost data on HMO enrollees, we used inpatient utilization instead of total health care costs as the outcome measure and a simplified set of PIP-DCG risk factors to predict inpatient utilization.

2. Data and sample selection

This study is based on a unique data set that contains acute care hospital discharge records for both FFS and HMO Medicare beneficiaries in the state of California from 1994 to 1996. Patient-level hospital discharge records were obtained from the California Office of Statewide Health Planning and Development. Since Medicare HMO enrollment status is not always accurately coded in hospital discharge records, the discharge records were linked to demographic and HMO enrollment data obtained from CMS. A third-party contractor performed the linkage and forwarded us the resulting file after removing all identifiers. Additional information about the data can be found in Dhanani et al. (forthcoming).

California has the largest Medicare managed care market in the nation. In 1995, 1.2 million Medicare beneficiaries were enrolled in about thirty risk HMOs in California, accounting for 37 percent of the entire Medicare HMO population in the nation. Its Medicare HMO penetration rate (percentage of Medicare beneficiaries enrolled in HMOs) was as high as 32 percent, compared to 8 percent in the nation.¹

The analysis sample consists of elderly beneficiaries in California who joined a Medicare risk HMO in 1995 and a comparison group of beneficiaries who remained in FFS. Medicare HMO enrollees are required by law to have both Part A and Part B coverage. The HMO subsample includes beneficiaries who enrolled in a risk HMO during the year of 1995, stayed enrolled for at least twelve months or until death, were at least age 65 at the beginning of Year 1, alive at the beginning of Year 2, and did not have end-stage renal disease (ESRD).² A one-year window was created for each HMO enrollee before and after enrollment. For example, for a 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 subsample stayed in FFS in Year 1 and joined an HMO at the beginning of Year 2.

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 same selection criteria were applied in selecting FFS beneficiaries. A pseudo month of "FFS enrollment" in 1995 was randomly assigned to each FFS beneficiary following the distribution of enrollment month of HMO enrollees in 1995. Those who died before "FFS enrollment" were dropped. This matching operation aligned the starting points of Year 2 for HMO and FFS beneficiaries in the sample and ensured that everyone survived until the beginning of Year 2.

Beneficiaries who resided in counties with a Medicare HMO penetration rate lower than 10 percent or fewer than 100 new risk HMO enrollees in 1995 were excluded. As a result, about half of the California counties were dropped. This ensured that all beneficiaries in our sample did have an option to enroll in an HMO.

3. Analytic methods

This study assessed a similar model to Medicare's original PIP-DCG model in terms of health risk factors included. But our outcome measure is inpatient utilization rather than total health care costs. Inpatient utilization is measured by (1) having one or more hospital admissions in Year 2, and (2) total hospital length of stay conditional on having at least one admission in Year 2.

Risk factors included in the study model are similar to those in the original PIP-DCG model. They are age at the beginning of Year 2, sex, any Medicaid coverage in Year 1, disability as the original reason for Medicare entitlement, and PIP-DCGs assigned from Year 1 hospitalizations.³ Although they are meant to predict health care costs, these risk factors are considered appropriate for predicting future hospital use too. We did not include any interaction terms between the demographic variables because most of them are not significant.

Consistently observed favorable selection into Medicare HMOs indicates that HMO enrollment, like hospital use, is also a function of health status. Thus, the same risk factors that predict future health status were also used to predict HMO enrollment. It is hypothesized that the included PIP-DCG risk factors do not adequately predict future health status (see Gruenberg et al. [10] for a list of significant predictors of health status for the Medicare population). HMO enrollment becomes endogenous in a simple regression of hospital use over the included risk factors and an HMO enrollment indicator when there are omitted risk factors affecting both hospital use and HMO enrollment. Omitted risk factors may include unobserved health status as well as non-health related factors. For example, different tastes for medical services may result in different levels of health care use and physician ties, which may subsequently influence HMO enrollment. However, most unobserved selection, at least in the AAPCC model, is due to unobserved health status. Hill et al. [13] developed a comprehensive model to predict future health care costs based on a detailed survey of Medicare HMO enrollees. They found no evidence of selection bias in their model. By parsing out the effects of the independent variables included in the model, they concluded that unobserved selection in the AAPCC was due mostly to unobserved health status (83%), with the rest of it due to differences in attitudes toward health and health care, socioeconomic factors, and access to care.

With HMO enrollment being potentially endogenous, this study employed a simultaneous equations model that jointly models HMO enrollment and inpatient use to consistently estimate the HMO effect. Based on the consistent estimate of the HMO effect, unobserved selection bias in an FFS-based model that predicts HMO enrollees' inpatient use will be quantified. Instead of making a fully parametric distributional assumption, we approximated the distribution of unobserved factors by a discrete distribution. The discrete factor approximation allows greater flexibility in characterizing the underlying distribution of unobserved selection and is more robust to misspecifications compared to traditional parametric distributional assumptions. A sample selection model that assumes the distribution of error terms to be joint normal may produce implausible results if the assumption is not correct [19]. The discrete factor approach has been employed in recent studies to analyze the effect of managed care on health care costs and use [9, 17]. Heckman and Singer [11]

applied a similar approach to correcting for heterogeneity for continuous-time duration models.

The simultaneous equations model describes the sequential events of HMO enrollment and hospital use. All beneficiaries in the sample were enrolled in the traditional FFS program in Year 1. At the beginning of Year 2, they made a decision to either remain in FFS or join an HMO, and then received their Year 2 health care in the setting they chose. Both HMO enrollment and Year 2 hospital use are a function of expected health status in Year 2, which can only be partially predicted by the included Year 1 PIP-DCG risk factors. Given the limited set of independent variables, a reduced-form model is specified. The simultaneous equations model is defined as follows:

$$\begin{aligned} y_{0i}^* &= \gamma' \mathbf{x}_i + \varepsilon_{0i} \\ y_{0i} &= 1 \quad \text{if } y_{0i}^* > 0 \\ y_{0i} &= 0 \quad \text{otherwise} \end{aligned} \quad (1)$$

$$\begin{aligned} y_{1i}^* &= \beta_1' \mathbf{x}_i + \alpha_1 y_{0i} + \varepsilon_{1i} \\ y_{1i} &= 1 \quad \text{if } y_{1i}^* > 0 \\ y_{1i} &= 0 \quad \text{otherwise} \end{aligned} \quad (2)$$

$$(y_{2i} | y_{1i} = 1) = \beta_2' \mathbf{x}_i + \alpha_2 y_{0i} + \varepsilon_{2i} \quad (3)$$

where

y_{0i} : 1 if individual i enrolled in an HMO at the beginning of Year 2 and 0 otherwise

y_{1i} : 1 if admitted at least once in Year 2 and 0 otherwise

y_{2i} : logged total length of stay conditional on admission in Year 2

\mathbf{x}_i : observed risk factors in Year 1 (age, sex, Medicaid coverage, originally disabled, and PIP-DCGs).

Model (1) is a probit model that predicts HMO enrollment status in Year 2. Model (2) is also a probit model that predicts having at least one hospital admission in Year 2. Model (3) is a log linear model for total length of stay conditional on admission in Year 2. The hospital use models (2) and (3) include an HMO membership indicator (y_{0i}). The three models are estimated jointly to allow correlation among the error terms that may arise if there are common or correlated unobserved variables affecting both HMO enrollment and hospital use.

On its raw scale, total length of stay is quite skewed. An ordinary least square model directly estimated on raw data may be subject to undue influence of outliers. Since all predictors in our model are categorical, a method suggested by Blough et al. [2] was used to test if a logarithmic transformation is appropriate. Variance and mean of total length of stay were calculated for each cell formed by the included risk factors. Logged variance then was regressed on logged mean weighted by the degrees of freedom associated with variance in each cell. The slope coefficient is about 2.4 for both the HMO and FFS subsamples, 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. To further reduce the influence of outliers, total length of stay greater than 60 days was set to 60. Censoring points other than 60 yielded little change in estimation results. We also assigned one day to those who were admitted in Year 2 but had a zero total length of stay ($n = 341$). The resulting logged total length of stay of HMO hospital users is close to being normal (skewness = 0.13 and Kurtosis = 2.52). For FFS users, skewness = 0.09 and Kurtosis = 2.56.

It is assumed that the observed risk factors in Year 1, \mathbf{x}_i , are exogenous and only partially predict expected health status in Year 2. Unobserved risk factors are reflected in the error terms. The error terms ε_0 , ε_1 , and ε_2 become interdependent when unobserved risk factors influence both HMO enrollment and hospital use. Significance tests for correlation among the error terms constitute tests for the existence of unobserved selection. Unobserved risk factors affecting both HMO enrollment and hospital use is represented by a heterogeneity term θ . The error terms ε_0 , ε_1 , and ε_2 thus can be decomposed into correlated and uncorrelated components:

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

$$\varepsilon_{1i} = \rho_1\theta + v_{1i} \quad (5)$$

$$\varepsilon_{2i} = \rho_2\theta + v_{2i}. \quad (6)$$

The heterogeneity term θ represents the permanent part of health risk known to beneficiaries but not captured by the included risk factors. Its relationship with the dependent variables is assumed to be linear. The uncorrelated components, v_{0i} , v_{1i} and v_{2i} , are independent from each other, θ , and the included variables. v_{0i} and v_{1i} are assumed to follow a standard normal distribution $N(0, 1)$, and v_{2i} is assumed to follow $N(0, \sigma)$. Instead of specifying a parametric distribution for the heterogeneity term θ such as a normal one (which would effectively make the three error terms joint normal), we adopted a more robust semi-parametric specification that approximates θ by a discrete distribution with K points of support $\{\eta_k\}$:

$$\text{Prob}(\theta = \eta_k) = \pi_k = \Phi(\xi_k) \quad k = 1, \dots, K, \quad (7)$$

where $0 < \pi_k < 1$, $\sum \pi_k = 1$, and $\sum \eta_k \pi_k = 0$. $\Phi(\cdot)$ is the standard normal cumulative distribution function. The probability π_k associated with η_k is formulated as $\Phi(\xi_k)$ to make sure $0 < \pi_k < 1$. Since the scale of the zero-mean heterogeneity term θ is not restricted, one of the factor loadings, ρ_0 , is arbitrarily set to 1. The covariance matrix of the error terms ε_0 , ε_1 and ε_2 is

$$\begin{bmatrix} V_\theta + 1 & \rho_1 V_\theta & \rho_2 V_\theta \\ \rho_1^2 V_\theta + 1 & \rho_1 \rho_2 V_\theta & \\ & \rho_2^2 V_\theta + \sigma^2 & \end{bmatrix} \quad (8)$$

where $V_\theta = \sum \pi_k \eta_k^2$ is the variance of the heterogeneity term.

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

$$\prod_{i=1}^N \sum_{k=1}^K \pi_k \left\{ \left[\int_{-\gamma'x_i - \eta_k}^{\infty} \phi(u) du \right]^{y_{0i}} \left[\int_{-\infty}^{-\gamma'x_i - \eta_k} \phi(u) du \right]^{1-y_{0i}} \right. \\ \times \left[\int_{-\beta_1'x_i - \alpha_1 y_{0i} - \rho_2 \eta_k}^{\infty} \phi(u) du \right]^{y_{1i}} \left[\int_{-\infty}^{-\beta_1'x_i - \alpha_1 y_{0i} - \rho_2 \eta_k} \phi(u) du \right]^{1-y_{1i}} \\ \left. \times \left[\frac{1}{\sigma} \phi \left(\frac{y_{2i} - \beta_2'x_i - \alpha_2 y_{0i} - \rho_3 \eta_k}{\sigma} \right) \right]^{y_{1i}} \right\} \quad (9)$$

where N is the sample size, $\phi(\cdot)$ is the standard normal density function. The model parameters γ' , β_1' , β_2' , α_1 , α_2 , σ , ρ_1 , ρ_2 , $\{\eta_k\}$, and $\{\pi_k\}$ are jointly estimated (see Mroz [19] for a detailed discussion of the estimator).

A Monte Carlo study conducted by Mroz [19] showed that this estimator compares favorably to a normal-based parametric maximum likelihood estimator in terms of consistency and accuracy when the true distribution of the error terms is joint normal, and dominates other estimators in a variety of situations when it is not. In fact, when the distribution of the error terms is misspecified, a fully parametric maximum likelihood estimator may lead to implausible estimates. Although having exclusion restrictions is preferable, the Monte Carlo study demonstrated that the semi-parametric model can be effectively identified without exclusion restrictions. In particular, the estimator outperforms other estimators by a wide margin in model specifications with skewed error distributions and no exclusion restrictions.⁴ In their study of childcare use in a Medicaid HMO, Goldman et al. [9] applied this estimator with the same set of independent variables. The Monte Carlo study also showed that a 3-support point discrete distribution is often adequate in providing good estimates in terms of consistency and accuracy. The heterogeneity term θ in this study was approximated by a 3-point discrete distribution.

Hospital admission and total length of stay models (2) and (3) were also estimated independently ignoring possible endogeneity of HMO enrollment. The 2-part model serves as a baseline model and is expected to overestimate the HMO effect in the presence of HMO favorable selection as it would attribute the effect of unobserved favorable health status of HMO enrollees to HMOs. All models in this study were estimated using STATA 7.0 [27]. Standard errors were adjusted for possible heteroskedasticity using the Huber/White correction.

4. Results

Table 1 compares unadjusted measures of Year 2 hospital use of FFS and HMO beneficiaries, which are the outcome variables in the hospital use model. Twenty percent of FFS beneficiaries and 15 percent of HMO enrollees incurred at least one hospital admission in Year 2. Conditional on admission, FFS beneficiaries' mean total length of hospital stay is 9.14 days, 2.32 days shorter than that of HMO enrollees. Total hospital days per 1,000 is 1,022 days for HMO enrollees, much lower than 1,828 days for FFS beneficiaries.⁵

Table 1. Observed HMO and FFS hospital use in Year 2.

	FFS comparison sample	New HMO enrollees
Number of observations	79,933	78,693
One or more admissions (%)	20.0	15.0
Mean total LOS, given admission (days)	9.14	6.82
Median total LOS, given admission (days)	6	4
Total hospital days per 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 but had less than 1 day of total length of stay.

Table 2. Distribution of HMO and FFS beneficiaries by demographic characteristics (%).

	FFS comparison sample	New HMO enrollees
Number of observations	79,933	78,693
Age (beginning of Yr. 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 (Yr. 1) ^a	18.9	6.5
Disabled ^b	7.0	6.5
Died (Yr. 2)	7.4	4.2

Notes: ^aBeneficiaries who had at least one month of Medicaid eligibility in Year 1.

^bBeneficiaries whose original reason for entitlement to Medicare was their disability. All demographic differences between HMO and FFS beneficiaries are statistically significant ($P < .001$).

The favorable demographic and health characteristics of HMO enrollees can be seen in Table 2. Medicare HMOs enrolled younger beneficiaries and more males than traditional FFS Medicare. Beneficiaries dually eligible for Medicare and Medicaid were less likely to enroll in an HMO as Medicaid pays for cost-sharing and many benefits not covered by Medicare such as prescription drugs. Those who were originally entitled to Medicare for disability reasons were slightly less likely to enroll in an HMO. Compared with FFS

Table 3. Distribution of HMO and FFS beneficiaries by PIP-DCGs assigned from Year 1 principal inpatient diagnoses (%).

PIP-DCG (Yr. 1)	FFS comparison sample	New HMO enrollees
Users (DCGs 4 and above)	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 higher the DCG number, the costlier the PIP-DCG group is. The differences in the PIP-DCG mix between HMO and FFS beneficiaries are all statistically significant ($P < .001$).

beneficiaries, fewer HMO enrollees died in Year 2. These differences are all statistically significant ($P < .001$).

The distribution of PIP-DCGs assigned from Year 1 hospitalizations is presented in Table 3.⁶ The higher the PIP-DCG number, the costlier the DCG group is. Fewer HMO enrollees were hospitalized in Year 1. Conditional on hospital use, the PIP-DCG mix of HMO hospital users prior to HMO enrollment is clearly less costly than that of FFS hospital users.

It is not surprising to observe less hospital use of HMO enrollees given their favorable demographic and PIP-DCG mix. However, we hypothesize that the observed favorable health status of HMO enrollees does not fully explain their less hospital use. Unobserved favorable selection also plays an important role.

Table 4 shows estimation results of the 2-part model (2) and (3). The 2-part model is biased because it ignores possible endogeneity of HMO enrollment. HMO membership in that model is highly significant, both statistically ($p = 0.000$) and practically, in reducing

Table 4. Estimation results of the 2-part model using total sample.

Variable	Equation			
	Any admission in Yr. 2		Log total length of stay in Yr. 2	
	Coeff.	Robust S.E.	Coeff.	Robust S.E.
Constant	-1.304**	(0.012)	1.544**	(0.020)
HMO	-0.109**	(0.008)	-0.186**	(0.012)
Age				
69-71	0.061**	(0.014)	-0.015	(0.024)
72-74	0.117**	(0.014)	0.022	(0.023)
75-77	0.205**	(0.014)	0.061**	(0.023)
78-80	0.278**	(0.015)	0.083**	(0.024)
81-83	0.363**	(0.016)	0.065**	(0.024)
84-86	0.430**	(0.018)	0.139**	(0.026)
87-89	0.503**	(0.020)	0.125**	(0.028)
90-92	0.547**	(0.025)	0.087**	(0.033)
93+	0.450**	(0.027)	0.076*	(0.035)
Male	0.141**	(0.008)	0.031**	(0.011)
Medicaid	0.152**	(0.011)	0.153**	(0.016)
Disabled	0.271**	(0.014)	0.111**	(0.020)
PIP-DCGs				
DCG 4	0.425**	(0.016)	0.101**	(0.021)
DCG 5	0.185	(0.110)	0.194	(0.188)
DCG 6	0.294**	(0.091)	-0.166	(0.126)
DCG 7	0.296*	(0.145)	0.173	(0.170)
DCG 8	0.544**	(0.026)	-0.004	(0.034)
DCG 9	0.421**	(0.031)	0.085*	(0.038)
DCG 10	0.651**	(0.037)	0.170**	(0.046)
DCG 11	0.703**	(0.030)	0.293**	(0.037)
DCG 12	0.656**	(0.029)	0.230**	(0.037)
DCG 14	0.705**	(0.055)	0.290**	(0.062)
DCG 16	1.060**	(0.026)	0.448**	(0.027)
DCG 18	0.919**	(0.063)	0.558**	(0.069)
DCG 20	1.016**	(0.052)	0.523**	(0.058)
DCG 23	1.017**	(0.069)	0.559**	(0.079)
DCG 26	1.166**	(0.088)	0.352**	(0.097)
DCG 29	1.167**	(0.118)	0.671**	(0.133)

Note: *significant at the 95% level; **significant at the 1% level.

admission rates as well as total length of stay conditional on admission. We calculated marginal HMO effects on the probability of admission and unconditional hospital days of HMO enrollees by setting the HMO indicator, y_{0i} , to 0 (FFS) and 1 (HMO) and then taking the difference between mean predicted values of HMO enrollees treated in the two settings and dividing it by mean predicted values in FFS. On average, HMOs reduce the probability of admission by 0.025 (or 14.5 percent), and shorten total length of stay conditional on admission by 17.0 percent ($=1 - \exp(-0.186)$) for HMO enrollees. Overall, HMOs reduce hospital days per 1,000 of HMO enrollees by 28.7 percent compared to their use if they had remained in FFS (see Table 7).

Tables 5 and 6 present estimation results of the simultaneous equations model (1), (2) and (3). The simultaneous equations model produces consistent estimates by explicitly modeling unobserved selection. As expected, both hospital use measures (probability of admission and total length of stay conditional on admission) generally increase with age except for the oldest. Males use more hospital services. Medicaid entitlement, previous disability, and costlier PIP-DCGs are all associated with more hospital use.

According to the simultaneous equations model, HMOs only reduce the probability of admission of HMO enrollees by 0.3 percent, compared to the 14.5 percent reduction estimated by the 2-part model. The coefficient of the HMO membership in the admission model is highly insignificant ($p = .983$). HMOs do shorten total length of stay conditional on admission by 11.5 percent ($=1 - \exp(-0.122)$) ($p = 0.072$), compared to the 17.0 percent reduction estimated by the 2-part model. Overall, Medicare HMOs reduce hospital days per 1,000 of HMO enrollees by 11.7 percent, much smaller than the overall effect (-28.7 percent) estimated by the 2-part model. HMOs reduce total hospital days primarily through shortening length of stay after admission. The estimated HMO effect, after correcting for unobserved selection, represents the true HMO effect we would observe in a randomized experiment with the HMO membership as the treatment. It is smaller than the 2-part model's estimate because the 2-part model attributed unobserved HMO favorable selection to the HMO effect. Table 7 summarizes the marginal HMO effects on hospital use of HMO enrollees estimated by the simultaneous equations model and the 2-part model.

The error term ε_0 in the HMO enrollment model is negatively correlated with ε_1 and ε_2 in the two hospital use models ($\text{corr}(\varepsilon_0, \varepsilon_1) = -.103$, $\text{corr}(\varepsilon_0, \varepsilon_2) = -.074$), meaning that a high draw of ε_0 would increase the probability of HMO enrollment but reduce the probability of admission and total length of stay conditional on admission. That is, relative to the entire Medicare population, HMO enrollees are healthier and FFS beneficiaries are sicker in ways not captured by the included risk factors. Wald tests show that the correlation between ε_0 and ε_2 in the total length of stay model is highly significant ($p = 0.002$). Although the correlation between ε_0 and ε_1 in the admission model is not significant ($p = 0.262$), the estimated HMO effect on admission did change dramatically after correction for unobserved selection. These results indicate that the included PIP-DCG risk factors cannot adequately control for biased selection. Unobserved favorable selection of HMO enrollees contributes in a significant way to their less hospital use, especially, total length of stay conditional on admission. It is also noteworthy that the error terms ε_1 and ε_2 are virtually uncorrelated, indicating that admission and the extent of use after admission are two different processes governed by unrelated unobserved factors.

Table 5. Estimation results of the simultaneous equations model.

Variable	Equation					
	HMO enrollment at beginning of Yr. 2		Any admission in Yr. 2		Log total length of stay in Yr. 2	
	Coeff.	Robust S.E.	Coeff.	Robust S.E.	Coeff.	Robust S.E.
Constant	1.731	(1.195)	-1.373**	(0.068)	1.497**	(0.045)
HMO	-	-	-0.002	(0.099)	-0.122	(0.068)
Age						
69-71	-0.157**	(0.024)	0.064**	(0.015)	-0.012	(0.024)
72-74	-0.325**	(0.029)	0.123**	(0.016)	0.026	(0.024)
75-77	-0.411**	(0.033)	0.214**	(0.017)	0.066**	(0.024)
78-80	-0.523**	(0.042)	0.288**	(0.018)	0.090**	(0.025)
81-83	-0.588**	(0.049)	0.375**	(0.020)	0.072**	(0.025)
84-86	-0.770**	(0.079)	0.444**	(0.022)	0.148**	(0.027)
87-89	-0.846**	(0.096)	0.518**	(0.025)	0.135**	(0.030)
90-92	-1.298**	(0.236)	0.567**	(0.032)	0.100**	(0.036)
93+	-1.704**	(0.343)	0.474**	(0.036)	0.091*	(0.038)
Male	0.016	(0.017)	0.142**	(0.008)	0.032**	(0.011)
Medicaid	-2.253**	(0.562)	0.181**	(0.029)	0.171**	(0.023)
Disabled	0.173**	(0.032)	0.269**	(0.015)	0.110**	(0.020)
PIP-DCGs						
DCG 4	-0.433**	(0.058)	0.433**	(0.018)	0.106**	(0.022)
DCG 5	-0.534	(0.285)	0.194	(0.110)	0.201	(0.188)
DCG 6	-0.509*	(0.232)	0.303**	(0.091)	-0.160	(0.126)
DCG 7	-0.035	(0.255)	0.298*	(0.146)	0.176	(0.170)
DCG 8	-0.249**	(0.071)	0.550**	(0.027)	0.000	(0.035)
DCG 9	-0.321**	(0.091)	0.427**	(0.031)	0.090*	(0.038)
DCG 10	-0.366**	(0.112)	0.659**	(0.038)	0.174**	(0.046)
DCG 11	-0.399**	(0.119)	0.711**	(0.032)	0.299**	(0.037)
DCG 12	-0.728**	(0.135)	0.668**	(0.032)	0.238**	(0.038)
DCG 14	-0.727**	(0.209)	0.716**	(0.056)	0.299**	(0.062)
DCG 16	-0.555**	(0.101)	1.072**	(0.029)	0.456**	(0.028)
DCG 18	-0.796*	(0.316)	0.934**	(0.065)	0.567**	(0.069)
DCG 20	-1.630**	(0.435)	1.040**	(0.058)	0.536**	(0.059)
DCG 23	-1.362**	(0.474)	1.038**	(0.073)	0.572**	(0.080)
DCG 26	-2.037*	(0.867)	1.197**	(0.092)	0.372**	(0.098)
DCG 29	-1.680*	(0.751)	1.195**	(0.122)	0.688**	(0.134)
σ	-	-	-	-	0.931**	(0.004)

Note: *significant at the 95% level; **significant at the 1% level.

Table 6. Parameter estimates for the 3-support point discrete distribution.

Factor locations		
η_1	-2.24	(-1.181)
η_2	1.15	(-1.110)
η_3^a	16.018	($p = .443$)
Factor weights		
π_1^b	0.570**	($p = .000$)
π_2^b	0.377**	($p = .000$)
π_3^c	0.053	($p = .128$)
Factor loadings		
ρ_0	Normalized to 1	-
ρ_1	-0.026	(-0.037)
ρ_2	-0.017	(-0.023)
Correlation between errors ^d		
$\text{corr}(\varepsilon_0, \varepsilon_1)$	-0.103	($p = .262$)
$\text{corr}(\varepsilon_0, \varepsilon_2)$	-0.074**	($p = .002$)
$\text{corr}(\varepsilon_1, \varepsilon_2)$	0.008	($p = .351$)

Notes: Standard errors or p values of Wald statistics shown in parentheses.

*significant at 95% level; **significant at the 1% level.

^a $\eta_3 = (-\pi_1\eta_1 - \pi_2\eta_2)/\pi_3$. Wald tests were performed on all calculated values.

^b $\pi_{1,2} = \Phi(\xi_{1,2})$. $\Phi(\cdot)$ is the standard normal cumulative distribution function.

^c $\pi_3 = 1 - \pi_1 - \pi_2$.

^dCorrelation coefficients are calculated according to the covariance matrix (8).

Table 7. Marginal HMO effects on hospital use of HMO enrollees compared to their use if they had remained in FFS.

	Marginal HMO effect on probability of admission	Marginal HMO effect on days conditional on admission ($= 1 - \exp(\hat{\alpha}_2)$)	Overall marginal HMO effect
2-part model	-14.5%	-17.0%	-28.7%
Simultaneous equations model	-0.3%	-11.5%	-11.7%

Note: Marginal HMO effects on the probability of admission and hospital days per 1,000 of HMO enrollees were calculated by setting the HMO indicator, y_{0i} , to 0 (FFS) and 1 (HMO) in the HMO subsample and then dividing the difference between mean predicted hospital use of HMO enrollees treated in the two settings (mean predicted use | FFS—mean predicted use | HMO) by (mean predicted use | FFS).

5. Simulations

Based on the estimation results of the simultaneous equations model, simulations were conducted over the same sample to quantify the magnitude of unobserved selection bias in a hospital use model estimated over the FFS subsample and used to predict HMO enrollees' use. The FFS-based model is expected to overestimate HMO enrollee's hospital use because of (1) the HMO reduction of hospital use and (2) unobserved favorable

selection of HMO enrollees. Total bias in the FFS-based model thus can be disaggregated into the two factors. We focused on the model's bias in predicting use of HMO enrollees because the actual PIP-DCG model was developed over Medicare FFS data to predict HMO enrollees' costs. We did not split our sample into two separate ones for development and evaluation purposes since hospitalizations for some diagnoses do not occur frequently.

The selection-corrected simultaneous equations model produces estimates of counterfactual hospital use of HMO enrollees as if they were not self-selected but selected randomly from the Medicare population and therefore had no unobserved health status. Expected hospital days of HMO enrollees if they had been randomly assigned to the HMO setting was calculated by multiplying estimated admission probability and total length of stay conditional on admission:

$$\hat{y}_i = \Phi(\hat{\beta}'_1 \mathbf{x}_i + \hat{\alpha}_1 y_{0i}) * \exp(\hat{\beta}'_2 \mathbf{x}_i + \hat{\alpha}_2 y_{0i}) * \exp(\hat{\sigma}^2/2), \quad (10)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Note that the retransformation of predicted total length of stay conditional on admission from the log scale to the raw scale is based on the log normality assumption in Model (3).

If they had been randomly enrolled and possessed no unobserved selection, HMO enrollees would have incurred more hospital use than their observed use which reflects their unobserved favorable health status. That is, predicted hospital days, \hat{y}_i , with unobserved selection purged out, should be greater than observed days for HMO enrollees. Indeed, mean predicted hospital days per 1,000 for HMO enrollees is 1,158 days, 136 days greater than their observed use of 1,022 days.

Counterfactual use of the same HMO enrollees if they had been randomly selected but treated in the FFS setting is 1,312 days, which was calculated by setting the HMO indicator in Equation (10) to 0. HMOs thus slash hospital days by 11.7 percent ($= (1,312 - 1,158)/1,312$) (see Table 7). This is the true HMO effect we would observe moving a randomly selected beneficiary from the FFS to HMO setting.

The PIP-DCG model being implemented by CMS was developed on FFS data and produces estimates of FFS costs of HMO enrollees. No attempt was made to correct for possible unobserved selection. To simulate unobserved selection bias in the FFS-based model, a naïve 2-part model for hospital use was estimated over the FFS subsample and then applied to HMO enrollees. Estimation results of the FFS-based hospital use model are shown in Table 8. The model is consistent only in the FFS population.

The FFS-based model predicts a mean of 1,487 hospital days for HMO enrollees, much less than observed use of FFS beneficiaries (1,828 days) because of observed favorable demographic characteristics and PIP-DCG mix of HMO enrollees (see Tables 2 and 3). After controlling for observed favorable selection, the FFS-based model still overestimates HMO enrollees' use by 465 days or 45.5 percent ($= (1,487 - 1,022)/1,022$) because of (1) the HMO effect and (2) unobserved favorable selection of HMO enrollees.

HMO enrollees' observed use 1,022 days reflects their observed as well as unobserved favorable selection. If the self-selected HMO enrollees had remained in FFS, their use would have increased by 136 days ($= 1,022 * (1/(1 - 11.7\%) - 1)$) to 1,158 days after removing

Table 8. Estimation results of the 2-part model over the FFS subsample.

Variable	Equation			
	Any admission in Yr. 2		Log total length of stay in Yr. 2	
	Coeff.	Robust S.E.	Coeff.	Robust S.E.
Constant	-1.302**	(0.017)	1.542**	(0.029)
Age				
69-71	0.083**	(0.020)	-0.025	(0.034)
72-74	0.107**	(0.020)	0.012	(0.034)
75-77	0.216**	(0.021)	0.043	(0.033)
78-80	0.284**	(0.021)	0.090**	(0.034)
81-83	0.371**	(0.022)	0.073*	(0.034)
84-86	0.406**	(0.024)	0.142**	(0.036)
87-89	0.496**	(0.027)	0.140**	(0.038)
90-92	0.515**	(0.032)	0.089*	(0.045)
93+	0.415**	(0.033)	0.076	(0.045)
Male	0.118**	(0.011)	0.006	(0.016)
Medicaid	0.149**	(0.013)	0.172**	(0.018)
Disabled	0.264**	(0.020)	0.121**	(0.027)
PIP-DCGs				
DCG 4	0.471**	(0.021)	0.105**	(0.027)
DCG 5	0.124	(0.147)	0.344	(0.265)
DCG 6	0.156	(0.124)	-0.099	(0.186)
DCG 7	0.387*	(0.197)	0.066	(0.233)
DCG 8	0.595**	(0.035)	0.015	(0.045)
DCG 9	0.453**	(0.040)	0.059	(0.048)
DCG 10	0.723**	(0.049)	0.131*	(0.058)
DCG 11	0.737**	(0.039)	0.350**	(0.048)
DCG 12	0.687**	(0.036)	0.295**	(0.047)
DCG 14	0.718**	(0.068)	0.270**	(0.076)
DCG 16	1.110**	(0.033)	0.519**	(0.034)
DCG 18	1.000**	(0.078)	0.661**	(0.082)
DCG 20	1.074**	(0.060)	0.534**	(0.065)
DCG 23	1.064**	(0.083)	0.632**	(0.093)
DCG 26	1.128**	(0.103)	0.202	(0.116)
DCG 29	1.209**	(0.141)	0.687**	(0.158)

Note: *significant at the 95% level; **significant at the 1% level.

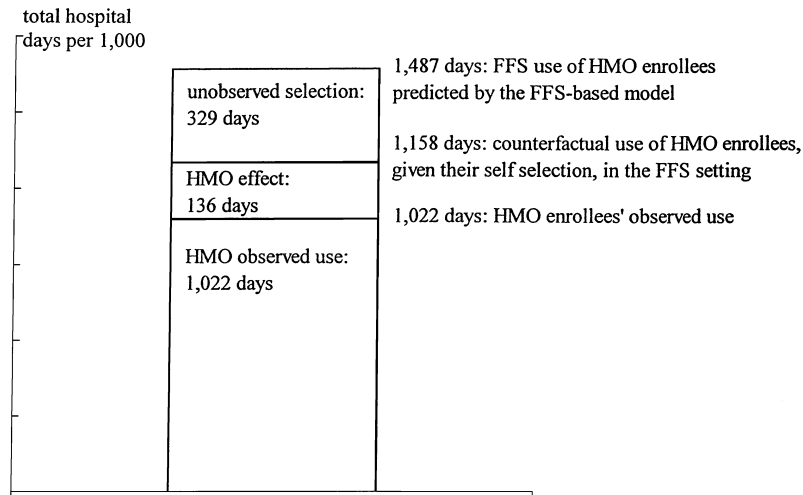


Figure 1. Disaggregating hospital use of HMO enrollees predicted by the FFS-based model.

the HMO effect. The 1,158 days represents the counterfactual use of HMO enrollees, given their unobserved favorable selection, in the FFS setting. The remaining overestimation, 329 days ($= 465 - 136$ days), is due to unobserved selection ignored in model development. Thus, unobserved HMO favorable selection results in a 28.4 percent ($= 329/1,158$) overestimation of FFS hospital use of HMO enrollees. The total bias in the FFS-based model is disaggregated in figure 1.

6. Discussion

The HMO effect estimated by the simultaneous equations model is consistent with the recent literature, which shows that HMOs achieve cost reduction mainly by controlling inpatient use and, especially, length of stay. Hill et al. [13] conducted a nationally representative survey of Medicare HMO enrollees in 1990 and found that risk HMOs had no effect on admission rates but did reduce total length of stay by 16.8 percent. The authors suspected that the lack of HMO effect on admission may have occurred because of disappearance of discretionary hospitalizations in increasingly competitive health care markets. This study found a smaller HMO effect on total length of stay (-11.5 percent) possibly because the California health care market in 1995 was more competitive than the national market in 1990. A literature synthesis by Miller and Luft [18] confirmed that HMOs (including Medicare HMOs) had a much more significant impact on shortening length of stay than reducing admission rates and physician visits. Christensen and Shinogle [3] analyzed data from a 1994 national survey and found that the Medicare HMO membership was associated with a statistically significant decline in length of stay and an insignificant increase in hospital admission. Using the same data (the time period is slightly different) but a different analytic method, Dhanani et al. (forthcoming) found HMO effects on hospital use similar to the findings of this study. One limitation of the simultaneous

equations model is that the heterogeneity term θ was used to index both the marginal distribution of unobserved factors in the entire population in Models (1) and (2) and the conditional distribution of those who were hospitalized in Model (3).⁷ However, the fact that the HMO effect estimated by the model is consistent with the recent literature and Dhanani et al. (forthcoming) suggests that the restrictions we imposed may represent reasonable simplifications.

This study concludes that an inpatient DCG model that predicts inpatient utilization using PIP-DCGs and demographic risk factors does not fully capture biased selection of Medicare HMO enrollees. Unobserved selection is significantly different between the FFS and HMO populations, with HMO enrollees healthier and FFS beneficiaries sicker in ways not captured by the model. Although a model developed on FFS data is consistent for the FFS population, it is not consistent for HMO enrollees when there is unobserved selection. Unobserved selection leads the FFS-based model to overestimate HMO enrollees' hospital days by 28.4 percent compared to their use if they had stayed in FFS.

One needs to be cautious in interpreting the reduction in hospital days in dollars terms because marginal costs of hospital days may decline over time during a hospital stay. Given that HMOs do not have much impact on admission rates, HMOs achieve cost savings primarily by making average length of stay shorter, but initial days of a hospital stay may be more intensive and expensive than the overpredicted days. It should also be noted that this study only examined hospital use of first-year HMO enrollees, whose unobserved selection is usually stronger than enrollees of longer 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 [29].

Biased selection does not necessarily lead to over or underestimation as long as it is fully captured. A perfect risk adjuster that includes all relevant risk factors and fully controls for selection can always consistently predict costs no matter how biased the underlying sample is. Over the period from 2000 to 2003, Medicare uses the PIP-DCG model to risk adjust 10 percent of its payments to Medicare + Choice organizations, with the remaining payments continuing to be adjusted by a demographic-only model. CMS will phase in a new model within the DCG family that incorporates both inpatient and ambulatory diagnoses starting from 2004. The comprehensive risk adjustment model has a stronger predictive power than the PIP-DCG [22], and is expected to be less biased in its estimation of Medicare + Choice costs.

It is true that in order to forestall profitable cream-skimming, Medicare does not need a perfect risk adjustment model but only needs to make sure that private health plans do not outperform government risk adjustment [28]. But even if HMOs do not outsmart Medicare in risk adjustment, a model that is developed on FFS data and applied on the HMO population can still be biased when unobserved risk factors are different in the two populations. This bias cannot be eliminated unless the model perfectly captures biased selection or unobserved selection diminishes.

Although a perfect model is always consistent, it still may not produce accurate estimates in reality. This is because even if there is no unobserved selection left, one may not receive the same diagnoses in the HMO and FFS settings for a variety of reasons such as

different practice styles and incentives and differences in the quality of reporting of diagnosis codes. If HMOs reduce hospital admissions, a perfect inpatient DCG model may underpay HMOs that efficiently shift patients to less expensive sites of care. This inaccuracy has nothing to do with the inconsistency of the FFS-based model. Although this study found a virtually zero HMO effect on hospital admission, this may not hold true in other circumstances. As a general matter, using risk factors that may be easily altered for good or bad reasons or may not be reported in uniform quality across different settings may lead to undesirable over or underestimation even when the risk adjuster itself is actuarially impeccable.

The substantial unobserved selection bias in the inpatient DCG model calls for a better risk adjustment model. A more actuarially sound risk adjuster such as the comprehensive DCG model mentioned above may further reduce overpayments to Medicare + Choice organizations. However, Medicare managed care provides extra values such as coordinated care and preventive services that are usually less available in the FFS sector. To the extent that these added values are worth the extra dollars, a better risk adjuster that cuts further down on payments needs to be balanced with its possible negative impacts on health plan participation. In any case, overpayments should be made explicit rather than as a result of incomplete risk adjustment. Future research on unobserved selection bias in the comprehensive DCG model that incorporates diagnoses from all settings will be particularly interesting.

Since this study is based on a sample of Medicare beneficiaries in California, its findings may not be generalizable to other populations. For example, biased selection in the employer-sponsored commercial health insurance market may have different dynamics. One general lesson that can be drawn, however, is that when a risk adjustment model is developed on one population and applied on another, we need to seriously consider the possibility that the model may not be powerful enough to sufficiently control for biased selection and as a result may produce biased prediction. Unobserved selection is much more difficult to deal with than observed selection. The possibility of incomplete control for selection is especially difficult to rule out in retrospective data that contain limited information. This study shows that unobserved selection can substantially bias model estimate of the true effect of a treatment and lead to biased estimation of outcomes.

Acknowledgments

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. We are grateful for the programming support provided by Robert Reddick of RAND Corporation and by Beate Danielson of Health Information Solutions who linked the data files. We are also grateful to the DxCG, Inc. for providing the DCG 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. We also would like to thank three anonymous reviewers for their insightful comments.

Notes

1. Authors' analysis of CMS Medicare managed care penetration quarterly data files.
2. Beneficiaries with ESRD were excluded because Medicare uses a different formula to pay private plans for these beneficiaries.
3. The original PIP-DCG methodology multiplies predicted costs by a factor of 0.21 for the working-aged, for whom Medicare is a secondary payer. Since the working-aged variable is not one of the risk factors included in the PIP-DCG model itself, it was not included in our study model.
4. Although including an instrumental variable in the HMO selection model is preferable, our administrative data provide no information that may serve as a valid instrument. One potential instrumental variable is county-specific system-wide HMO penetration or Medicare HMO penetration rates. However, Baker [1] found that higher system-wide HMO penetration reduced health care expenditures in the FFS Medicare sector. In particular, in California, acute care hospital discharges, length of stay and inpatient days were found to decline more rapidly in areas with a higher overall HMO presence from 1984 to 1994 [25]. Medicare HMO penetration rates, on the other hand, may also be correlated with utilization when favorable health status of new HMO enrollees diminishes in areas well penetrated by Medicare HMOs where many healthy beneficiaries have already moved to the HMO sector. Evidence indicates that Medicare HMO market penetration is a significant predictor of hospital days conditional on admission [17].
5. HMO and FFS beneficiaries all survived until the beginning of Year 2. Death in different months in Year 2 may lead to different levels of Year 2 hospital use as decedents tend to use more health care services in their last months leading to death. For example, those who died in the first month of Year 2 may incur less hospital use in Year 2 than those who died in the last month of Year 2. However, we did not adjust part-year hospital use of decedents. Summary statistics show that those who died in the first month of Year 2 did incur less hospital use, compared to those who died in later months. Only a small number of people died in the first month of Year 2 (0.6% in FFS and 0.3% in HMO). While length of stay may be easily annualized, it is not obvious how to annualize admission. Also, since the elderly people have a higher rate of hospital use right before death, annualizing hospital use of decedents may introduce more bias.
6. This study only examined acute care hospital discharge records. In the actual PIP-DCG model implemented by CMS, inpatient diagnoses may come from facilities eligible for Medicare's prospective payment system as well as non-PPS facilities and units such as psychiatric, rehabilitation, long-term, children's and other specialty hospitals. Also, this study retained discretionary hospitalizations (DCG 4) and short stays (<2 days), which are excluded by the original PIP-DCG model. These hospitalizations are excluded not because they do not predict future health care costs, but because they may provide incentives for undesirable behavior on the part of health plans. However, this is not a concern to this study as the data is from a period when the PIP-DCG was not in use. DxCG 5.1 for SAS developed by DxCG, Inc. was used for mapping inpatient diagnoses to the PIP-DCG groups (see Pope et al. [22] for a description of how principal inpatient diagnoses are grouped into PIP-DCGs).
7. We thank one anonymous reviewer for pointing this out.

References

1. Baker, L., "Association of managed care market share and health expenditures for fee-for-service Medicare patients," *JAMA* 281(5), 432–437, 1999.
2. Blough, D., Madden, C., and Hornbrook, M., "Modeling risk using generalized linear models," *Journal of Health Economics* 18, 153–171, 1999.
3. 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.
4. Dhanani, N., O'leary, J., Keeler, E., et al., "The effect of HMOs on the inpatient utilization of Medicare beneficiaries," *Health Services Research*, forthcoming.
5. 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.
6. 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.

7. Ellis, R., Pope, L., Iezzoni, J., et al., "Diagnosis-based risk adjustment for Medicare capitation payments," *Health Care Financing Review* 18(1), 1996.
8. General Accounting Office, "Medicare + Choice: Payments exceed cost of fee-for-service benefits, adding billions to spending," GAO/HEHS-00-161, 2000.
9. 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, 1998.
10. Gruenberg, L., Kaganova, E., and Hornbrook, M., "Improving the AAPCC with health-status measures from the MCBS," *Health Care Financing Review* 17(3), 59–75, 1996.
11. Heckman, J. and Singer, B., "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica* 52, 271–320, 1984.
12. Hellinger, F., "Selection bias in HMOs and PPOs: A review of the evidence," *Inquiry* 32, 135–142, 1995.
13. 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," Mathematica Policy Research, Inc., Princeton, NJ, 1992.
14. Hornbrook, M., Bennett, M., and Greenlick, M., "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, JAI, Greenwich, CT, 1989.
15. 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, 1505–1910, 1984.
16. Medicare Payment Advisory Commission, "Report to the Congress: Improving risk adjustment in Medicare," MedPAC, Washington, DC, Nov. 2000.
17. Melllo, M., Stearns, S., and Norton, E., "Do Medicare HMOs still reduce health services use after controlling for selection bias?" *Health Econ.* 11, 323–340, 2002.
18. Miller, R. and Luft, H., "Managed care plan performance since 1980," *Journal of the American Medical Association* 271(19), 1512–1519, 1994.
19. 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.
20. Newhouse, J., Manning, W., Keeler, E., et al., "Adjusting capitation rates using objective health measures and prior utilization," *Health Care Financing Review* 10(3), 41–54, 1989.
21. Physician Payment Review Commission, "Prior research on risk selection in Medicare HMOs," PPRC, Washington, DC, 1996.
22. Pope, G., Ellis, R., Ash, A., et al., "Principal inpatient diagnostic cost group model for Medicare risk adjustment," *Health Care Financing Review* 21(3), 127–135, 2000.
23. 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, Dec., 2000.
24. Riley, G., Lubitz, J., and Rabey, E., "Enrollee health status under Medicare risk contracts: An analysis of mortality rates," *Health Services Research* 26(2), 137–163, 1991.
25. Robinson, J., "Decline in hospital utilization and cost inflation under managed care in California," *JAMA* 276, 1060–1064, 1996.
26. 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, Dec. 4, 2001, Accessed on <http://cms.hhs.gov> in June 2002.
27. Stata Corporation, Stata 7.0. College Station, TX, 2001.
28. Van de Ven, W. and Ellis, R., "Risk adjustment in competitive health plan markets," in *Handbook of health economics*. (A. Culyer and J. Newhouse, eds.), Elsevier, New York, NY, vol. 1A, pp. 755–847, 2000.
29. Welch, W., "Medicare capitation payments to HMOs in light of regression toward the mean in health care costs," *Advances in Health Economics and Health Services Research* 6, 75–96, JAI, Greenwich, CT, 1985.