Comparing Patient Outcomes across Payer Types: Implications for Using Hospital Discharge Records to Assess Quality

Daniel D. Maeng and Grant R. Martsolf

Objective. To explain observed differences in patient outcomes across payer types using hospital discharge records. Specifically, we address two mechanisms: hospital-payer matching versus unobserved patient heterogeneity.

Data Source. Florida's hospital discharge records (1996–2000) of major surgery patients with private health insurance between the ages of 18 and 65, Health Maintenance Organization (HMO) market penetration data, hospital systems data, and the Area Resource File.

Study Design. The dependent variable is occurrence of one or more in-hospital complications as identified by the Complication Screening Program. The key independent variable is patients' primary-payer type (HMO, Preferred Provider Organization, and fee-for-service). We estimate five different logistic regression models, each representing a different assumption about the underlying factors that confound the causal relationship between the payer type and the likelihood of experiencing complications.

Principal Finding. We find that the observed differences in complication rates across payer types are largely driven by unobserved differences in patient health, even after adjusting for case mix using available data elements in the discharge records.

Conclusion. Because of the limitations inherent to hospital discharge records, making quality comparisons in terms of patient outcomes is challenging. As such, any efforts to assess quality in such a manner must be carried out cautiously.

Key Words. Acute inpatient care, quality of care/patient safety (measurement), quality improvement/report cards (interventions)

Because quality of care became an important health policy issue in the United States (Institute of Medicine 2001), public reporting of comparative quality information has received attention as a way to improve quality (Marshall et al. 2000). Consequently, the recognition of the fact that this approach to quality improvement requires accurate measurements and reliable data sources has

led to an increased interest in using readily available data sources such as hospital discharge records to assess quality. For instance, the Healthcare Cost & Utilization Project and the Patient Safety Indicators (PSIs) developed by Agency for Healthcare Research and Quality illustrate such efforts.

Administrative data such as hospital discharge records offer the potential to be an objective yet inexpensive source of quality data. On the other hand, the disadvantage lies precisely in the fact that such data are not collected for the explicit purpose of quality measurement. As a result, hospital discharge records typically lack key clinical information (e.g., test results) and individual characteristics (e.g., income, marital status) that would be important for case-mix adjustments (GAO 1994; Iezzoni et al. 1996).

In this paper, we attempt to explain the observed differences in patient outcomes among different payer types using hospital discharge records and, in the process, identify the difficulties of doing so. Specifically, we test two separate hypotheses. First, managed care organizations (MCO) may selectively contract with either low- or high-quality hospitals. Alternatively, there may be a systematic but unobserved difference in health status between the patients covered by MCO and those by fee-for-service (FFS) health insurers. We find that the observed differences in patient outcome are driven largely by the unobserved differences in patient health status not captured by the existing data elements within the hospital discharge records. We thus conclude that there is no significant difference in the inpatient quality of care received by MCO and FFS patients, and that hospital discharge records as currently available are unlikely to be an adequate data source for assessing quality of care in terms of inpatient treatment outcomes.

CONCEPTUAL FRAMEWORK

We propose three possible hypotheses that explain any observed differences in patient outcomes across payer types. First, there may be selective contracting between providers and payer types based on quality. For instance, lower-quality providers may have an incentive to seek out contracts with certain payers to attract patients; likewise, certain payers may have an incentive to

Address correspondence to Daniel D. Maeng, Ph.D., Geisinger Center for Health Research, 100 N. Academy Avenue M.C. 44-00, Danville, PA 17822; e-mail: ddmaeng@geisinger.edu. Grant R. Martsof, M.P.H., R.N., is with the Department of Health Policy and Administration and Center for Health Care and Policy Research, The Pennsylvania State University, University Park, PA.

contract with either high-quality providers (to ensure high quality of care to their enrollees) or low-quality providers (to secure lower costs at the expense of quality). The previous literature on this issue suggests mixed results. Chernen, Hayward, and Scanlon (1996) have found that MCO are more likely to contract with high-volume hospitals, which indicate higher quality, whereas Escarce, Shea, and Chen (1997) have found essentially the opposite. Given that these two studies had focused on different markets (California and Florida, respectively), these findings may reflect different MCO behaviors in different markets.

The second hypothesis is that there may be unobserved health status differences among the patients covered by different payer types. While the conventional wisdom suggests that MCO selectively enroll healthier enrollees to reduce cost, there is evidence that MCO enrollees may in fact be sicker than their non-MCO counterparts (Schaefer and Reschovsky 2002). Huesch (2010) also suggests sicker Medicare patients may selectively enroll into Medicare-managed care plans, although the author is unable to find any direct evidence supporting that conclusion. Thus, it is possible that MCO patients may in fact be more likely to experience worse treatment outcomes than FFS patients.

There are several possible explanations for this: First, because MCO typically require lower cost-sharing than FFS plans, those who are in worse health conditions and are therefore more likely to face cost sharing may choose MCO over FFS, while MCO are unable to adequately screen out such patients. Second, it may be that because the MCO enrollees had been previously denied of needed care, they are more prone to complications. Moreover, once their enrollees are admitted, MCO may also impose stricter restrictions on who receives a surgery, which may lead to a selection of patients who are more prone to complications at the time of surgery. Unfortunately, the current literature offers little guidance as to which one of these might be the most likely explanation.

Lastly, providers may respond to MCO financial incentives by reducing both the quantity and the quality of treatments rendered, resulting in worse outcomes for their enrollees (Berenson 1986; Landon et al. 2008). If providers are able to easily alter their treatment patterns by each patient's payment source, this may explain the differences in patient outcomes across payer types within a hospital. However, in hospital settings where processes and procedures are increasingly standardized, it is difficult to see how such discrimination by payer types can actually occur. Therefore, in our analysis, we ignore this last hypothesis and focus on the first two as the most likely explanations.

The prior work that most closely resembles this paper is the paper by Haile and Stein (2002). Using the hospital discharge records from 39 California hospitals in 1996 and 1997, they report a significant association between payer types and in-hospital complication rates, suggesting that MCO patients appeared to be more likely to experience in-hospital complications than FFS patients. They conclude that this association is attributable to the variation in treatment patterns across hospitals rather than within each hospital. Their finding is therefore consistent with the selective contracting hypothesis between MCO and low-quality hospitals. However, their empirical model does not explicitly account for the unobserved differences in patient health status. In this paper, we build upon this prior analysis by explicitly controlling for any unobserved patient heterogeneity in our models.

DATA

The main source of data for this analysis was the Florida hospital inpatient discharge records collected by Florida Agency for Health Care Administration. The data set consisted of patient-level data containing all the discharge records from approximately 250 of Florida's major hospitals. Of the available years of data, those from 1996 through 2000 were chosen and pooled. This yielded 1,712,873 discharges for which one or more major surgeries had taken place during the period. From this, we retained only those who had private Health Maintenance Organization (HMO), Preferred Provider Organization (PPO), or FFS as the primary payer and were between the ages of 18 and 64 to obtain a more homogeneous population in terms of health status (1,195,797 excluded).

To reduce the computational burden due to the large sample size, we took a 50 percent random sample in each hospital (258,538 excluded). This random sampling was done before dropping those observations with missing variables (19,890 excluded) to ensure that the random sampling would not be subject to any bias resulting from potential nonrandom missingness in the variables. The final sample size included 238,648 discharges from 175 hospitals located throughout the state of Florida.

Private Health Insurance Market in Florida 1996–2000

From 1996 through 2000, Florida's private managed care market experienced a period of change that closely reflects what had occurred in other states during the same period (Baumgarten 2003). Following a rapid period of

growth between 1996 and 1998, the HMO enrollment in Florida began to shrink beginning in 2000. This was attributable to the increases in the HMO premium, which prompted employers to switch to other managed care types such as PPO. Therefore, this was the period in which there were perceivable distinctions between HMO and PPO plans, at least from the purchasers' perspective. Furthermore, our study period captured the time during which HMO as well as PPO had consisted a significant portion of the private health insurance market.

INDEPENDENT VARIABLES

The key independent variable of interest was the primary payer indicator variable for each patient. In particular, those patients whose primary payer was identified as either HMO or PPO were considered to be covered by MCO. Those patients whose primary payer was listed as "commercial charge-based" insurance were considered to be covered by private FFS insurance. The data also contained information on the type of admission (e.g., emergency, urgent, or elective) and the source of admission (e.g., transfer from another hospital, long-term care facility, or emergency room) on each discharge, as well as up to 10 diagnosis and procedure codes.

We also had detailed information on hospital characteristics such as the bed size, ownership type (for-profit or nonprofit), teaching status, as well as the total number of admissions. In addition, using the hospital zip code, we merged in a set of geographical and market-level variables that have been known to be associated with hospital quality (Sari 2002; Dranove, Lindrooth, and White 2008), which included the county-level yearly managed care penetration data compiled by Baker (1995) as well as the Florida hospital systems data compiled by Dranove, Lindrooth, and White (2008) to calculate the Herfindahl-Hirschman index that captures the local hospital market concentration. For the purposes of this study, market was defined as county.

In particular, the managed care penetration data provide an important control variable in our empirical model that captures the "spill over" effect of MCO. Baker (2003) suggests that an increased MCO enrollment within a market may induce overall changes in the treatment patterns and resource utilization across all patients in the area, affecting outcomes even among those not covered by MCO. This means that those hospitals located within the high-MCO penetration markets may be expected to have different patient

outcomes than those located in the low-MCO penetration markets. Thus, our market-level managed care penetration rate variable allows us to control for this effect.

Our discharge data set contained limited information on patient demographics (age, gender, and race), including the zip code of each patient's residence. Most notably, the data set did not contain information on patient income, which may be problematic if patient income is a significant predictor of patients' health status and treatment outcomes. In prior studies, researchers used each patient's zip code to obtain the median income of that zip code to proxy for the patient's income (Shapiro et al. 1994; Haile and Stein 2002; Encinosa and Bernard 2005). We recognized two issues with this approach: First, this assumed that the median income of each zip code corresponds to the income of a particular patient, which may or may not be true. Second, given that patients' proximity to hospitals is a significant predictor of patients' choice of hospitals (Gowrisankaran and Town 1999), this approach was essentially analogous to obtaining the local market conditions of each hospital. As such, we merged in the relevant market condition variables—unemployment rate and percentage of population who are 65 years old or older—from the Area Resource File.

DEPENDENT VARIABLES

This study utilizes in-hospital complication rates among patients who had undergone major surgeries computed by the Complication Screening Program (CSP) as the primary outcome variable (Iezzoni et al. 1999b). CSP is similar to PSI in that it relies on commonly available hospital discharge data to identify cases in which one or more preventable complications may have potentially occurred due to medical errors caused by the provider. In this study, we use CSP instead of PSI in order to build upon the earlier work by Haile and Stein, who also utilized CSP to construct their outcome variables. CSP includes 28 complication categories based on ICD-9-CM codes. In our data set, we find that about 13 percent of the major-surgery patients had experienced at least one complication. Refer to Iezzoni et al. (1999a, b) for more detailed descriptions and the logic of CSP (see Appendix SA2 for the descriptive statistics of all the variables used in this analysis). Table 1 lists the types of complications considered under CSP and their frequencies in our data.

Table 1: CSP Complication Categories

<i>Complication Description</i>	<i>Frequency (%)</i>
Iatrogenic complications	6.13
Postoperative pulmonary compromise	2.10
Miscellaneous complications	1.29
Postoperative pneumonia	1.22
Mechanical complications	1.06
Postoperative myocardial infarction	0.50
Venous thrombosis/pulmonary embolism	0.37
Other postoperative cardiac abnormalities	0.35
Postoperative urinary tract complications	0.29
Postoperative GI hemorrhage/ulceration	0.28
Postprocedure hemorrhage/hematoma	0.21
Technical difficulty with medical care	0.20
Postoperative nervous system complications	0.16
Procedure-related perforation/laceration	0.15
Postoperative stroke	0.15
Wound infection	0.14
Postoperative coma/stupor	0.13
Cellulitis/decubitus ulcer	0.12
Postoperative physiologic and metabolic derangements	0.08
Reopening of surgical site	0.07
Septicemia	0.05
Postoperative infections (not pneumonia, wound)	0.05
Poisoning due to medication	0.02
In-hospital hip fracture/fall	0.01
Aspiration pneumonia	0.00
Postoperative shock due to anesthesia	0.00
Complication due to anesthetic agents	0.00
Shock/cardiorespiratory arrest in hospital	0.00

Case-Mix Adjustment

CSP provides its own algorithm for case-mix adjustment. First, CSP classifies patients into different risk pools. For the purposes of this study, we have selected those patients who had undergone major surgeries (Risk Pool A) based on the assumptions that (1) they are likely to be more homogeneous in terms of the severity of their conditions and that (2) they are more likely to yield meaningful variation in the complication rates relative to those in other, less severe, CSP risk pools. Furthermore, CSP suggests controlling for the Major Diagnostic Categories and presence of 13 chronic conditions when making comparisons.

In addition to patient age, gender, and race, we have included the numbers of diagnosis and procedure codes (up to 10 each) as covariates in our

empirical models to further capture the severity of illness of each patient, even though CSP does not explicitly call for their inclusion. Clearly, these are endogenously determined—that is, experiencing complications would increase the number of diagnoses and procedure codes. We have estimated alternative models without these variables and found that our results did not change significantly.

We also include the admission types and sources as control variables that capture any differences in patient severity at the time of admission. We recognize that these variables may also capture the MCO influences and therefore are likely to be correlated with our key independent variable, patient's payer type. However, since the prior studies similar to this one (e.g., Haile and Stein 2002; Encinosa and Bernard 2005) have included them as case-mix control variables, we maintain the interpretation of these variables as case-mix control variables and include them in our empirical models.

METHODOLOGY

To minimize the presence of confounding factors in our data, our analysis is restricted to major surgery patients between the ages of 18 and 64 covered by private health insurance, which was defined as private FFS, HMO, or PPO. We begin with the following basic specification:

$$C_{ijt} = \alpha_0 + \alpha_1 MC_{ijt} + \alpha_2 P_{ijt} + \alpha_3 H_{jt} + \alpha_4 MK_{jt} + \omega_{ijt} + \lambda_{jt} + \varepsilon_{ijt} \quad (1)$$

where C_{ijt} equals 1 if patient i undergoing a major surgery in hospital j at time t experienced one or more complications and 0 otherwise. MC_{ijt} is the categorical variable for the patient i 's payer type, with the private FFS insurance being the referent category. P_{ijt} is a vector of the patient characteristics, and H_{jt} represents the vector of hospital characteristics. MK_{jt} represents the market condition for hospital j at time t . Also, ω_{ijt} represents the unobserved patient heterogeneity that affects his or her complication in hospital j and time t , while λ_{jt} represents the unobserved quality of hospital j at time t . Because of the discrete nature of the dependent variable, (1) is estimated via binary logistic regression model, with the assumption that the error term ε_{ijt} is logistically distributed with mean 0. Our basic approach is to see how α_1 changes under different model specifications.

To test whether the association between patients' payer types and complication rates are driven by selective contracting between hospitals and

MCO based on hospital quality, we use hospital fixed effects (i.e., include dummy variables for 174 hospitals if the unobserved hospital quality λ_{jt} is time invariant) and hospital-year fixed effects (i.e., include dummy variables for 174 hospitals $\times$ 4 years if λ_{jt} is time variant) to capture λ_{jt} , while assuming that the unobserved patient heterogeneity, ω_{ijt} , is 0. Compared with a naïve model in which both ω_{ijt} and λ_{jt} are assumed 0, these fixed-effects models should yield an estimate of α_1 that is closer to 0. That is, if MCO selectively contract with lower-quality hospitals, controlling for the unobserved hospital quality via hospital fixed effects should eliminate the positive and significant effect of the payer type variables.

It is important to note that this constitutes only an indirect method of testing the selective contracting hypothesis. In other words, we are unable to conclude from this part of our analysis whether selective contracting based on hospital quality does or does not occur. Rather, the purpose of this is to show whether selective contracting is a plausible explanation for the observed association between payer types and complication rates.

If the significant and positive effect still remains even after controlling for the unobserved hospital quality via fixed effects, we then move on to testing whether the association is driven by the unobserved patient heterogeneity, ω_{ijt} . To do so, we take a more structural approach. More specifically, we specify a two-equation model as the following:

$$MC_{ijt} = \beta_0 + \beta_1 P_{ijt} + \beta_2 H_{jt} + \beta_3 MK_{jt} + \varepsilon_1 \quad (2)$$

$$C_{ijt} = \tau_0 + \tau_1 MC_{ijt} + \tau_2 P_{ijt} + \tau_3 H_{jt} + \tau_4 MK_{jt} + \varepsilon_2 \quad (3)$$

The positive and significant estimate of β_1 arises from the cross-equation correlation between ε_1 and ε_2 ; that is, if MCO patients are in worse health conditions than their FFS counterparts, then they may be more likely to experience complication. To account for this correlation, the error terms are decomposed as the following:

$$\varepsilon_2 = \rho_1 \omega_{ijt} + u_1 \quad (4)$$

$$\varepsilon_1 = \rho_2 \omega_{ijt} + u_2 \quad (5)$$

Thus, the cross-equation correlation of the error terms arises from the presence of ω_{ijt} in both (2) and (3).

Put differently, the difference between the naïve model as shown in Equation (1) and the two-equation model is the explicit recognition that MCO patients are different from FFS patients in terms of the unobserved

characteristics that contribute to one's likelihood of experiencing complications. The naïve model implicitly assumes that patients across all payer types are essentially homogeneous in terms of their unobserved characteristics, whereas the two-equation model allows MCO patients to be systematically different from FFS patients.

Discrete Factor Approximation Method (DFM)

To remove the effects of ω_{ijt} in our model, the DFM is used (Mroz 1999). Rather than imposing a stringent distributional assumption on ω_{ijb} , DFM assumes that the distribution may be approximated using a finite discrete set of mass points (Heckman and Singer 1984), or "factors." Mechanically, DFM estimates (2) and (3) simultaneously via maximum likelihood estimation (MLE), conditional on each mass point representing a value of the random variable ω_{ijt} . Then, the conditional likelihood values are summed over all the mass points to obtain the unconditional likelihood values.

Intuitively, the mass points represent the unobserved "types" of individuals in the data set. That is, we assume that there are finite types of patients in our data set and that those who belong to the same type are similar to one another in terms of the unobserved health status. Thus, rather than assuming an arbitrary distribution of ω_{ijb} which is unknown, we approximate the empirical distribution of the types in our data with discrete mass points. Conceptually, once we know the approximate empirical distribution of the types, we can then "integrate out" the unobserved ω_{ijt} from our equations by summing over all discrete values of ω_{ijt} .

Although our DFM model is technically identified by its functional form, the identification is strengthened via exclusion restriction. In our model, complication rates among the publicly insured patients of each hospital serve as the exclusion restriction. That is, it appears as a right-hand-side variable in (3) to predict whether or not a privately insured patient experiences a complication, but it does not predict the payer type in (2). Publicly insured patients are defined as those covered by Medicare or Medicaid (mostly Medicare, as Medicaid patients accounted for <8 percent of the total hospital admissions in Florida during our study period [Encinosa and Bernard 2005]). To the extent that how a hospital treats its publicly insured patient is strongly correlated with how it treats its privately insured patients, this is a reasonable variable to include in our complication model. There is ample literature to suggest that this may indeed be a reasonable assumption (Needleman et al. 2003; Wennberg et al. 2004; Baker, Fisher, and Wennberg 2008).

The complication rate among the publicly insured patients in each hospital thus represents the underlying risk of complication experienced by all patients treated in that hospital—that is, unobserved hospital quality. To help identify our model, we further assume that the unobserved hospital quality is uncorrelated with the payer types of the privately insured patients. Therefore, this exclusion restriction would not hold if the hospital-MCO selective contracting hypothesis were true. That is, if MCO are able to observe the hospital quality and thus “steer” their enrollees to particular hospitals, the complication rates among the publicly insured patients would be correlated with the privately insured patients’ payer types as well. However, we believe this is unlikely based on our findings presented below. In short, we find little evidence of MCO-hospital matching based on unobserved hospital quality, suggesting that our exclusion restriction is warranted.

We use DFM rather than the traditional instrumental variable (IV) method because the functional form of DFM helps the identification of the parameters of interest and thus avoids the restrictive properties of IV. Using our current data source, it is difficult to identify a plausible IV that predicts patients’ payer types but not directly the patients’ complication rates. Instead, we have identified a variable that is correlated with the complication rates but not with the payer types. Thus, we are able to achieve identification using an exclusion restriction and the functional form of DFM rather than resorting to an implausible IV.

DFM Model Specification

Let η represent one of M mass points approximating the distribution of ω_{ijt} . Combining equations (2) through (6), the two-equation model can be expressed as the following:

$$MC_{ijt} = \alpha_0 + \alpha_1 P_{ijt} + \alpha_2 H_{jt} + \alpha_3 MK_{jt} + \rho_1 \eta_m + u_1 \quad (6)$$

$$C_{ijt} = \beta_0 + \beta_1 MC_{ijt} + \beta_2 P_{ijt} + \beta_3 H_{jt} + \beta_4 MK_{jt} + \rho_2 \eta_m + u_2 \quad (7)$$

Since η represents a mass point in a discrete probability distribution, each η_m has a probability—denoted by π_m —associated with it such that $\sum \pi_m = 1$. Thus, the probability corresponding to each η_m is given by the following logit transformation:

$$\pi_m = \frac{\varepsilon^{\eta_j}}{1 + \sum_{j=1}^{M-1} \varepsilon^{\eta_j}} \quad (8)$$

Thus, we estimate φ_m instead of π_m , which can be calculated from (8). Following Mroz's suggestion, ρ_1 is constrained to one to determine the scale of the discrete factors. The estimated magnitude and sign of ρ_2 , therefore, indicate the extent to which the unobserved patient heterogeneity leads to the biased estimates of β_1 . The number of mass points is arbitrary. Deb (2001) suggests that three or four mass points are usually sufficient to approximate the discrete empirical distribution of ω_{ijt} . Thus, our strategy is to obtain separate model estimates with two, three, and four mass points and choose the one that yields the highest likelihood value—in this case, four mass points. For more details on DFM and the derivation of the likelihood function, see Appendix SA2.

RESULTS

Table 2 summarizes the estimated coefficient estimates obtained from estimating the 1-equation model as described in (1).

The “Naïve Model” column shows the parameter estimates as obtained under the assumption that there are no unobserved patient or hospital characteristics. On the other hand, the hospital FE and hospital-year FE columns show the estimates obtained by controlling for the unobserved hospital quality. In these models, we omit those independent variables that do not vary over time (in the hospital FE model) and those that do not vary within each hospital in a given year (in the hospital-year FE model) because they are perfectly collinear with the fixed effect terms.

The naïve model estimates suggest that those patients covered by HMO and PPO are significantly more likely to experience in-hospital complications than their FFS counterparts. Interestingly, however, even after controlling for the unobserved hospital quality via the fixed effects, the coefficient estimates on the payer type variables remain positive and statistically significant, albeit the magnitudes get smaller toward zero. This suggests that the hospital-HMO selective contracting is unlikely to be the primary driver of the observed relationship between payer type and treatment outcome. Thus, we move on to test the alternative hypothesis that MCO patients may be more prone to complication than FFS patients.

For this step, we use DFM as described above. Before implementing DFM, however, we collapse the HMO and PPO indicator variables into a single binary variable that equals one if the patient is covered by HMO or PPO and zero if covered by FFS. This is done because (1) we fail to reject that the coefficient estimates on the HMO and PPO dummy variables are equivalent

Table 2: Single Equation Logistic Regression Results

<i>Patient Experienced at Least 1 Complication</i>	<i>Coefficient (SE)</i>		
	<i>Naïve Model</i>	<i>Hospital FE</i>	<i>Hospital-Year FE</i>
Payer type			
Primary payer is HMO	0.0976 (0.0347**)	0.0813 (0.0279**)	0.0757 (0.0288**)
Primary payer is PPO	0.1035 (0.037**)	0.0727 (0.0231**)	0.057 (0.0231*)
Chronic conditions			
Cancer with a poor prognosis	− 0.1071 (0.0741)	− 0.1139 (0.0724)	− 0.1076 (0.0675)
Metastatic cancer	− 0.4051 (0.0399**)	− 0.4506 (0.04**)	− 0.4814 (0.0408**)
AIDS	0.528 (0.2415*)	0.5742 (0.2415*)	0.5888 (0.2187**)
Chronic pulmonary disease	− 0.2241 (0.0293**)	− 0.2281 (0.029**)	− 0.2209 (0.0279**)
Coronary artery disease	− 0.6041 (0.0527**)	− 0.6139 (0.0536**)	− 0.6179 (0.0531**)
Congestive heart failure	0.0904 (0.0395*)	0.0778 (0.0393*)	0.0645 (0.0373)
Peripheral vascular disease	− 0.6585 (0.063**)	− 0.6654 (0.0644**)	− 0.6776 (0.0628**)
Severe chronic liver disease	− 0.4559 (0.072**)	− 0.4813 (0.0716**)	− 0.4975 (0.0686**)
Diabetes with end organ damage	− 0.9166 (0.0505**)	− 0.9286 (0.0491**)	− 0.9592 (0.0476**)
Chronic renal failure	− 0.0141 (0.0928)	− 0.0236 (0.0926)	− 0.0064 (0.0888)
Nutritional deficiencies	0.2507 (0.0573**)	0.2584 (0.0611**)	0.2357 (0.0614**)
Dementia	− 0.715 (0.3596*)	− 0.7959 (0.3981*)	− 0.7005 (0.4)
Functional impairment	0.0093 (0.0974)	− 0.0029 (0.0954)	0.0068 (0.0966)
Major diagnostic categories			
Nervous system	0.2232 (0.058**)	0.2431 (0.0553**)	0.2615 (0.054**)
Respiratory system	0.2083 (0.0762**)	0.2101 (0.0743**)	0.2199 (0.0707**)
Circulatory system	0.1827 (0.0835*)	0.1729 (0.0882*)	0.1988 (0.0842*)
Digestive system	0.7736 (0.0442**)	0.8023 (0.0427**)	0.861 (0.0448**)
Hepatobiliary system and pancreas	0.185 (0.0489**)	0.2088 (0.0472**)	0.2417 (0.0481**)
Musculoskeletal system and connective tissue	− 0.0265 (0.0461)	− 0.0085 (0.0447)	0.0109 (0.0429)
Kidney and urinary tract	0.2222 (0.0562**)	0.2358 (0.0547**)	0.227 (0.0553**)
Male reproductive system	0.1281 (0.0757)	0.1426 (0.0778)	0.1522 (0.0734*)
Number of diagnostic and procedure codes			
Number of diagnosis codes	0.4381 (0.0085**)	0.4482 (0.0085**)	0.4563 (0.0083**)
Number of procedure codes	0.1595 (0.0092**)	0.1653 (0.0085**)	0.1683 (0.0084**)
Demographics			
Woman	− 0.2653 (0.0234**)	− 0.272 (0.0231**)	− 0.2753 (0.0221**)
Age	− 0.0162 (0.0065*)	− 0.0175 (0.0073*)	− 0.016 (0.0074*)
Age squared	0.0002 (0.0001**)	0.0002 (0.0001**)	0.0002 (0.0001**)

continued

Table 2. *Continued*

<i>Patient Experienced at Least 1 Complication</i>	<i>Coefficient (SE)</i>		
	<i>Naïve Model</i>	<i>Hospital FE</i>	<i>Hospital-Year FE</i>
Black	0.112 (0.0363**)	0.0976 (0.0293**)	0.0919 (0.0283**)
Hispanic	− 0.0052 (0.0446)	− 0.0417 (0.0407)	− 0.0288 (0.0358)
Asian	0.0865 (0.1111)	0.0783 (0.1104)	0.0935 (0.1067)
Other race	0.1244 (0.0581*)	0.0329 (0.0537)	0.0507 (0.0496)
<i>Hospital characteristics</i>			
Number of beds available	0.0386 (0.0157*)	(Omitted)	(Omitted)
For-profit status	− 0.0706 (0.0688)	(Omitted)	(Omitted)
Teaching hospital	− 0.0665 (0.1508)	(Omitted)	(Omitted)
% Patients covered by Medicare	0.0127 (0.0064*)	0.019 (0.0095*)	(Omitted)
% Patients covered by Medicaid	− 0.0189 (0.0425)	− 0.0886 (0.0582)	(Omitted)
<i>Market characteristics</i>			
HMO penetration	0.0077 (0.0053)	− 0.0085 (0.0113)	(Omitted)
HH index	0.1521 (0.2403)	− 0.453 (0.4952)	(Omitted)
County unemployment rate	− 0.0462 (0.0269)	− 0.0488 (0.064)	(Omitted)
% County population over 65	− 0.0138 (0.0067*)	− 0.0104 (0.0588)	(Omitted)
<i>Year dummies</i>			
Year 1997	− 0.2645 (0.0436**)	− 0.2037 (0.0542**)	(Omitted)
Year 1998	− 0.3179 (0.0651**)	− 0.2221 (0.0936*)	(Omitted)
Year 1999	− 0.3971 (0.069**)	− 0.2882 (0.1209*)	(Omitted)
Year 2000	− 0.441 (0.0825**)	− 0.3075 (0.161)	(Omitted)

Note. Also includes (but not shown) indicator variables for admission sources (emergency room, hospital, long-term care facility, and Other) and admission types (emergency, urgent, elective, and Other).

*Significant at 5% level;

**Significant at 1% level.

(p -value = .35 in the Hospital-Year FE model), and (2) this greatly simplifies the DFM likelihood function to be estimated via MLE. As a comparison, we also estimate the two-equation model without DFM, which corresponds to the “Naïve Model” in Table 2. The results are shown in Tables 3 and 4 below.

In Table 3, we find that under DFM, the estimated coefficient on the payer type dummy variable is -0.17 and statistically insignificant. Compared with the estimated coefficient of about 0.1 under the naïve model (no DFM), the sign of the coefficient estimate under DFM is now reversed and statistically insignificant. Thus, we conclude that those covered by HMO and PPO are no more or less likely to experience complications than their FFS counterparts. This is in contrast to what the estimates from the one-equation logistic

Table 3: Simultaneous Equation Logistic Regression Results for Complication Rates

<i>Patient Experienced at Least 1 Complication</i>	<i>Coefficient (SE)</i>	
	<i>No DFM</i>	<i>DFM</i>
Payer type		
Primary payer is HMO/PPO	0.0999 (0.0264**)	− 0.1668 (0.1719)
Treatment outcome of other patients		
% Comp for public pay patients	0.0418 (0.004**)	0.0911 (0.0085**)
Chronic conditions		
Cancer with a poor prognosis	− 0.1038 (0.0727)	− 0.1856 (0.1868)
Metastatic cancer	− 0.4365 (0.0389**)	− 1.0248 (0.1103**)
AIDS	0.5629 (0.2519*)	1.386 (0.6967*)
Chronic pulmonary disease	− 0.2265 (0.0286**)	− 0.5418 (0.0822**)
Coronary artery disease	− 0.6232 (0.0531**)	− 1.5569 (0.207**)
Congestive heart failure	0.088 (0.0383*)	0.3535 (0.1075**)
Peripheral vascular disease	− 0.6581 (0.0621**)	− 1.6728 (0.1909**)
Severe chronic liver disease	− 0.4442 (0.0729**)	− 1.0161 (0.1802**)
Diabetes with end organ damage	− 0.9018 (0.0503**)	− 2.102 (0.1819**)
Chronic renal failure	− 0.014 (0.0914)	0.0265 (0.2288)
Nutritional deficiencies	0.2442 (0.0615**)	0.7777 (0.1901**)
Dementia	− 0.7335 (0.3646*)	− 1.8165 (1.0273)
Functional impairment	0.0013 (0.0972)	0.0146 (0.2392)
Major diagnostic categories		
Nervous system	0.2246 (0.0566**)	0.4813 (0.1419**)
Respiratory system	0.1899 (0.075*)	0.3635 (0.1687*)
Circulatory system	0.1683 (0.0828*)	0.4724 (0.1978*)
Digestive system	0.7817 (0.0424**)	1.6906 (0.1324**)
Hepatobiliary system and pancreas	0.1863 (0.0464**)	0.4589 (0.0996**)
Musculoskeletal system and connective tissue	− 0.025 (0.0446)	0.068 (0.0994)
Kidney and urinary tract	0.2183 (0.0551**)	0.4251 (0.1277**)
Male reproductive system	0.1158 (0.0757)	0.144 (0.1476)
Number of diagnosis and procedure codes		
Number of diagnosis codes	0.4318 (0.0085**)	1.066 (0.0696**)
Number of procedure codes	0.1647 (0.0085**)	0.4072 (0.0325**)
Demographics		
Woman	− 0.2694 (0.0236**)	− 0.6066 (0.0629**)
Age	− 0.0175 (0.007*)	− 0.0471 (0.0128**)
Age squared	0.0002 (0.0001**)	0.0006 (0.0001**)
Black	0.1134 (0.0318**)	0.1849 (0.0689**)
Hispanic	− 0.0321 (0.0382)	− 0.128 (0.0848)
Asian	0.0937 (0.1041)	0.1575 (0.2198)
Other race	0.0744 (0.0511)	0.1581 (0.1138)
Hospital characteristics		
Number of beds available	0.0004 (0.0073)	0.0101 (0.0184)
For-profit status	− 0.0178 (0.0529)	0.0462 (0.109)

continued

Table 3. *Continued*

<i>Patient Experienced at Least 1 Complication</i>	<i>Coefficient (SE)</i>	
	<i>No DFM</i>	<i>DFM</i>
% Patients covered by Medicare	0.0129 (0.0054*)	0.0244 (0.0109*)
% Patients covered by Medicaid	− 0.0218 (0.0305)	− 0.0717 (0.0718)
Teaching hospital	− 0.0744 (0.0735)	− 0.1532 (0.1923)
Market characteristics		
HMO penetration	0.0041 (0.0032)	0.0065 (0.0072)
HH index	0.0649 (0.1207)	− 0.0604 (0.2795)
% County population over 65	− 0.0104 (0.0053*)	− 0.0243 (0.0101*)
County unemployment rate	− 0.0225 (0.0181)	− 0.0563 (0.0378)
Year dummies		
Year 1997	− 0.1658 (0.0349**)	− 0.3422 (0.0781**)
Year 1998	− 0.2177 (0.0516**)	− 0.4325 (0.111**)
Year 1999	− 0.2847 (0.0536**)	− 0.5477 (0.1199**)
Year 2000	− 0.3162 (0.0653**)	− 0.6451 (0.142**)

Note. Also includes (but not shown) indicator variables for admission sources (emergency room, hospital, long-term care facility, and Other) and admission types (emergency, urgent, elective, and Other).

*Significant at 5% level;

**Significant at 1% level.

regression in Table 2 implies and is consistent with the hypothesis that those patients covered by HMO and PPO tend to be more prone to complications than their FFS counterparts.

Table 4 reinforces this conclusion. It suggests that MCO patients tend to be more likely to be chronically ill—metastatic cancer, chronic pulmonary disease, coronary artery disease, congestive heart failure, peripheral vascular disease, diabetes with end organ damage, and nutritional deficiencies—than FFS patients. Thus, MCO patients may have been in worse health conditions than their FFS counterparts even before they were admitted. The results also indicate that those who are covered by MCO are more likely to have fewer diagnosis codes and undergo fewer procedures. One possible explanation for this is that through the use of gatekeepers and practice guidelines, MCO may reduce the use of certain procedures by providers, without impacting the treatment outcomes of their enrollees.

The estimates of the DFM parameters as shown in Table 4 indicate the presence of the unobserved patient heterogeneity affecting the coefficient estimate on the payer type variable. In particular, the ρ_2 parameter is positive and significant, suggesting that there is an upward bias on the payer type coefficient in the naïve model estimates. Again, this is consistent with the

Table 4: Two-Equation Logistic Regression Results for Payer Type

<i>Primary payer is HMO/PPO</i>	<i>Coefficient (SE)</i>	
	<i>No DFM</i>	<i>DFM</i>
Chronic conditions		
Cancer with a poor prognosis	0.0655 (0.0745)	0.0658 (0.0746)
Metastatic cancer	0.2084 (0.0461**)	0.2074 (0.0453**)
AIDS	0.4692 (0.2971)	0.4696 (0.2988)
Chronic pulmonary disease	0.092 (0.0279**)	0.092 (0.028**)
Coronary artery disease	0.1007 (0.043*)	0.1017 (0.0431*)
Congestive heart failure	0.2085 (0.0523**)	0.2077 (0.0522**)
Peripheral vascular disease	0.1463 (0.05**)	0.1482 (0.0502**)
Severe chronic liver disease	0.0864 (0.1323)	0.084 (0.1319)
Diabetes with end organ damage	0.2267 (0.0607**)	0.2278 (0.0608**)
Chronic renal failure	0.1591 (0.1004)	0.1562 (0.0993)
Nutritional deficiencies	0.1898 (0.078*)	0.189 (0.0781*)
Dementia	− 0.4812 (0.2274*)	− 0.4799 (0.2287*)
Functional impairment	0.0302 (0.0919)	0.0318 (0.0926)
Major diagnostic categories		
Nervous system	− 0.0972 (0.0484*)	− 0.0909 (0.0551)
Respiratory system	0.1145 (0.0667)	0.1214 (0.0694)
Circulatory system	0.4111 (0.0876**)	0.4183 (0.0867**)
Digestive system	0.1697 (0.0593**)	0.177 (0.0636**)
Hepatobiliary system and pancreas	0.2403 (0.0657**)	0.246 (0.0697**)
Musculoskeletal system and connective tissue	− 0.4137 (0.0406**)	− 0.406 (0.0369**)
Kidney and urinary tract	0.1861 (0.0702**)	0.1934 (0.0759*)
Male reproductive system	0.046 (0.0714)	0.0576 (0.0641)
Number of diagnosis and procedure codes		
Number of diagnosis codes	− 0.0244 (0.01*)	− 0.0238 (0.0097*)
Number of procedure codes	− 0.0894 (0.0119**)	− 0.0894 (0.0124**)
Demographics		
Woman	0.1121 (0.0284**)	0.119 (0.0246**)
Age	0.0862 (0.0177**)	0.0922 (0.0079**)
Age squared	− 0.001 (0.0002**)	− 0.001 (0.0001**)
Black	0.2086 (0.0487**)	0.2091 (0.048**)
Hispanic	− 0.0048 (0.1623)	− 0.007 (0.1629)
Asian	0.0572 (0.1165)	0.0579 (0.114)
Other race	− 0.0932 (0.1175)	− 0.092 (0.1159)
Hospital characteristics		
Number of beds available	0.1113 (0.0284**)	0.1125 (0.0301**)
For-profit status	0.3415 (0.1492*)	0.3449 (0.1537*)
% Patients covered by Medicare	− 0.0049 (0.0184)	− 0.0043 (0.0183)
% Patients covered by Medicaid	− 0.1498 (0.0896)	− 0.1464 (0.0939)
Teaching hospital	− 0.9474 (0.3059**)	− 0.9538 (0.3075**)
Market characteristics		
HMO penetration	− 0.0119 (0.0081)	− 0.0113 (0.0092)
HH index	− 1.5241 (0.3546**)	− 1.5107 (0.3583**)

continued

Table 4. *Continued*

<i>Primary payer is HMO/PPO</i>	<i>Coefficient (SE)</i>	
	<i>No DFM</i>	<i>DFM</i>
% County population over 65	− 0.0414 (0.0135**)	− 0.0409 (0.0144**)
County unemployment rate	− 0.045 (0.0467)	− 0.0425 (0.0522)
Year dummies		
Year 1997	0.3775 (0.0866**)	0.3791 (0.0872**)
Year 1998	0.4711 (0.1115**)	0.4729 (0.112**)
Year 1999	0.6186 (0.1178**)	0.6219 (0.1192**)
Year 2000	0.7903 (0.1215**)	0.793 (0.1228**)
DFM parameters		
η_1		− 0.209 (0.7143)
η_2		− 0.328 (0.7058)
η_3		− 0.0351 (0.7339)
η_4		− 0.1239 (0.7212)
φ_2		− 0.5386 (0.2951)
φ_3		− 1.2734 (0.2614**)
φ_4		− 1.073 (0.3853**)
ρ_2		41.51 (21.15*)

Note. Also includes (but not shown) indicator variables for admission sources (emergency room, hospital, long-term care facility, and Other) and admission types (emergency, urgent, elective, and Other).

*Significant at 5% level;

**Significant at 1% level.

hypothesis that there are unobserved differences in the underlying health status between MCO and FFS patients.

Our results may be summarized as the following: MCO patients appeared to be more prone to complications than FFS patients after a risk adjustment based only on the observed characteristics. This association remained even after controlling for the unobserved hospital quality that may be correlated with both the patients' payer types and complications (i.e., selective contracting). However, in a model in which unobserved patient heterogeneity (e.g., health status) was explicitly accounted for, the association disappeared, suggesting that the observed association between payer types and complications was largely driven by the unobserved differences in patient health status.

DISCUSSION

There are two implications to our study: First, further study is necessary to examine why MCO patients appear to be in worse health condition than their

FFS counterparts. One possible explanation is that the MCO patients are more prone to complications because utilization restrictions make it more difficult for them to access necessary care before hospital admission. Alternatively, patients with chronic conditions may be attracted to MCO because of the lower cost sharing. Our results appear to be more consistent with the latter hypothesis, as we have found that certain chronic conditions appear to be more prevalent among the MCO patients than among the FFS patients. This is in contrast to the commonly held perception that MCO “cherry pick” the healthier individuals. Thus, if this adverse selection of sicker individuals by MCO is true, then it suggests a further area to explore in terms of how well private insurers are able to assess the risk of each potential enrollee.

Second, our results illustrate the limitations of using hospital discharge data to make inferences on quality of care, particularly when quality is measured in terms of patient outcome. Even after extensively adjusting for patient case mix using the available data elements in the discharge records, we find that there is likely to be substantial unobserved differences in patient health status that confound the relationship between payer types and outcomes. However, as other researchers have also pointed out (Wray et al. 1997), adequately capturing case mix from discharge records is difficult. Researchers who are aware of this issue have offered several methods of addressing it (e.g., Elixhauser et al. 1998; Geweke, Gowrisankaran, and Town 2003), but there has not been one universally accepted case-mix adjustment method. Our paper demonstrates that health care administrators and researchers must be cautious when using such data to make quality comparisons and, if feasible, should consider augmenting the administrative data with relevant clinical data that can help reduce the unobserved differences in patients’ underlying health conditions.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: None.

Disclosures: None.

Disclaimers: None.

REFERENCES

- Baker, L. C. 1995. *County-Level Measures of HMO Enrollment and Market Share*. Unpublished manuscript, Stanford University, Stanford, CA.

- . 2003. "Managed Care Spillover Effects." *Annual Review of Public Health* 24: 435–56.
- Baker, L. C., S. Fisher, and J. E. Wennberg. 2008. "Variations in Hospital Resource Use for Medicare and Privately Insured Populations in California." *Health Affairs* 27 (2): w123–34.
- Baumgarten, A. 2003. "Trend Note: HMO Enrollment Continues to Decrease in 2001–2002" [accessed on November 30, 2010]. Available at http://www.allan.baumgarten.com/index.cfm?fuseaction=dsp_news_detail&id=39
- Berenson, R. A. 1986. "Capitation and Conflict of Interest." *Health Affairs* 5 (1): 141–46.
- Chernew, M., R. Hayward, and D. Scanlon. 1996. "Managed Care and Open-Heart Surgery Facilities in California." *Health Affairs* 15 (1): 191–201.
- Deb, P. 2001. "A Discrete Factor Random Effects Probit Model with Application to the Demand for Preventative Care." *Health Economics* 10: 371–83.
- Dranove, D., R. Lindrooth, and W. D. White. 2008. "Is the Impact of Managed Care on Hospital Prices Decreasing?" *Journal of Health Economics* 27 (2): 362–76.
- Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey. 1998. "Comorbidity Measures for Use with Administrative Data." *Medical Care* 36: 8–27.
- Encinosa, W. E., and D. M. Bernard. 2005. "Hospital Finances and Patient Safety Outcomes." *Inquiry* 42: 60–72.
- Escarce, J. J., J. A. Shea, and W. Chen. 1997. "Segmentation of Hospital Markets: Where Do HMO Enrollees Get Care?" *Health Affairs* 16 (6): 181–92.
- Geweke, J., G. Gowrisankaran, and R. Town. (2003). "Bayesian Inference for Hospital Quality in a Selection Model." *Econometrica* 71 (4): 1215–38.
- Gowrisankaran, G., and R. J. Town. 1999. "Estimating the Quality of Care in Hospitals Using Instrumental Variables." *Journal of Health Economics* 18: 747–67.
- Haile, P. A., and R. M. Stein. 2002. "Managed Care Incentives and Inpatient Complications." *Journal of Economics and Management Strategy* 11 (1): 37–79.
- Heckman, J., and B. Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52 (2): 271–320.
- Huesch, M. D. 2010. "Managing Care? Medicare Managed Care and Patient Use of Cardiologists." *Health Services Research* 45 (2): 329–54.
- Iezzoni, L. I., R. B. Davis, R. H. Palmer, M. Cahalane, M. B. Hamel, K. Mukamal, R. S. Phillips, N. J. Banks, and D. T. Davis. 1999a. "Does the Complications Screening Program Flag Cases with Process of Care Problems? Using Explicit Criteria to Judge Processes." *International Journal of Quality Health Care* 11 (2): 107–18.
- Iezzoni, L. I., A. Lawthers, R. B. Davis, E. P. McCarthy, S. N. Weingart, R. H. Palmer, M. Cahalane, M. B. Hamel, K. Mukamal, R. S. Phillips, D. T. Davies, and N. J. Banks. 1999b. *Screening Quality of Care Using Administrative Data Final Report Appendix A Logic of the Complication Screening Program*. Washington, DC: Agency for Health Care Policy and Research.
- Iezzoni, L. I., M. Shwartz, A. S. Ash, J. S. Hughes, J. Daley, and Y. D. Mackiernan. 1996. "Severity Measurement Methods and Judging Hospital Death Rates for Pneumonia." *Medical Care* 34 (1): 11–28.
- Institute of Medicine. 2001. *Crossing the Quality Chasm*. Washington, DC: National Academies Press.

- Landon, B. E., M. B. Rosenthal, S. T. Normand, R. G. Frank, and A. M. Epstein. 2008. "Quality Monitoring and Management in Commercial Health Plans." *American Journal of Managed Care* 14: 344–86.
- Marshall, M. N., P. G. Shekelle, S. Leatherman, and R. Brook. 2000. "The Public Release of Performance Data: What Do We Expect to Gain? A Review of the Evidence." *Journal of the American Medical Association* 283: 1866–74.
- Mroz, T. A. 1999. "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome." *Journal of Econometrics* 92: 233–74.
- Needleman, J., P. I. Buerhaus, S. Matke, M. Stewart, and K. Zelevinsky. 2003. "Measuring Hospital Quality: Can Medicare Data Substitute for All-Payer Data?" *Health Services Research* 38 (6): 1487–508.
- Sari, N. 2002. "Do Competition and Managed Care Improve Quality?" *Health Economics* 11 (7): 571–84.
- Schaefer, E., and J. D. Reschovsky. 2002. "Are HMO Enrollees Healthier Than Others? Results from the Community Tracking Study." *Health Affairs* 21 (3): 249–58.
- Shapiro, M. F., R. E. Park, J. Keeseey, and R. H. Brook. 1994. "The Effect of Alternative Case-Mix Adjustments on Mortality Differences between Municipal and Voluntary Hospitals in New York City." *Health Services Research* 29 (1): 95–112.
- United States General Accounting Office. 1994. *Report Cards: Are Useful But Significant Issues Need to be Addressed*. Washington, DC: United States General Accounting Office.
- Wennberg, J. E., E. S. Fisher, T. A. Stukel, and S. M. Sharp. 2004. "Use of Medicare Claims Data to Monitor Provider-Specific Performance among Patients with Severe Chronic Illness." *Health Affairs* 23 (variation suppl): VAR5–18.
- Wray, N., J. Hollingsworth, N. Petersen, and C. Ashton. 1997. "Case-Mix Adjustment Using Administrative Databases: A Paradigm to Guide Future Research." *Medical Care Research and Review* 54: 326–56.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Appendix SA2: Derivation of DFM Likelihood Function.

Table S1: Descriptive Statistics—Variable Mean and Standard Deviation.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.