Bayesian Inference for Hospital Quality in a Selection Model

John Geweke
Departments of Economics and Statistics
University of Iowa
john-geweke@uiowa.edu

Gautam Gowrisankaran
Department of Economics
University of Minnesota,
Federal Reserve Bank of San Francisco,
and NBER
gautam@econ.umn.edu

Robert J. Town
Graduate School of Management
University of California-Irvine
rjtown@uci.edu

August, 2001

Abstract

This paper develops new econometric methods to infer hospital quality in a model with discrete dependent variables and non-random selection. Mortality rates in patient discharge records are widely used to infer hospital quality. However, hospital admission is not random and some hospitals may attract patients with greater unobserved severity of illness than others. In this situation the assumption of random admission leads to spurious inference about hospital quality. This study controls for hospital selection using a model in which distance between the patient's residence and alternative hospitals are key exogenous variables. Bayesian inference in this model is feasible using a Markov chain Monte Carlo posterior simulator, and attaches posterior probabilities to quality comparisons between individual hospitals and groups of hospitals. The study uses data on 74,848 Medicare patients admitted to 114 hospitals in Los Angeles County from 1989 through 1992 with a diagnosis of pneumonia. It finds the smallest and largest hospitals to be of high quality and public hospitals to be of low quality. There is strong evidence of dependence between the unobserved severity of illness and the assignment of patients to hospitals. Consequently a conventional probit model leads to inferences about quality markedly different than those in this study's selection model.

1. Introduction

This paper develops new econometric methods to estimate hospital quality and other models with discrete dependent variables and non-random selection. Assessing the quality of care in hospitals is an important problem for public policy and a challenge for applied econometrics.[1] Policy changes in Medicare reimbursement rates and the rise of managed care as well as technological innovations have affected hospital incentives, and through that, hospital quality.[2] These quality changes have large welfare effects and hence the potential for large deadweight losses.[3]

Hospital patient discharge databases provide several indicators plausibly associated with hospital quality. Since they cover large numbers of patients and hospitals and are much less expensive to obtain and access than other sources of information, they have been widely used in comparisons of hospital quality. Mortality has been the most popular indicator of hospital quality in the literature: it is unambiguously defined and free of measurement error, and its link with quality of care is so strong as to be tautological.[4]

In this widely used framework, the conceptual experiment that reveals hospital quality is hospital-specific mortality rates following random assignment of a population of patients to hospitals. Patients, however, are not randomly assigned to hospitals. Patients or their physicians are likely to choose hospitals based on factors such as location, convenience and their severity of illness. The experiment implicit in the data is not random assignment. If assignment were nonrandom, but random conditional on observed characteristics, then conventional dichotomous outcome models could be used to infer the outcome of the conceptual experiment from the available data. However, discharge data contain only crude summaries of medically pertinent information and hence many aspects of the severity of illness are unobserved. Thus, the assumption of random conditional assignment is not tenable and patients with the same observed

---

[1] "As described by a leading study, "Quality of care is the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge…," Lohr (1990, p. 4).

[2] See Cutler (1995), Kessler and McClellan (2000), McClellan and Noguchi (1998) for studies on the effects of Medicare policy, the impact of managed care and the impacts of technological change on medical outcomes, respectively.

[3] For instance, if changes in Medicare policies cause hospitals to reduced their pneumonia mortality rates by one percentage point, this would translate to over 6,000 lives saved annually in the U.S.

[4] Strictly speaking mortality is an indicator of hospital mediocrity; mortality is an inverse indicator of quality. Subsequently we provide a precise definition of hospital quality in the context of the model developed in this study.

characteristics are not equally likely to be admitted to all hospitals. For instance, if patients with high unobserved severity of illness select high quality hospitals, then observed mortality rates for high quality hospitals will be inconsistent and upwardly biased measures of mortality from the conceptual experiment. This will be true even after controlling for observed measures of severity of illness. Conventional statistical methods that ignore unobserved severity will produce misleading inferences about hospital quality. This has led prominent medical experts to make a pessimistic assessment of the usefulness of discharge data in assessing hospital quality.[5]

Recent work by Gowrisankaran and Town (1999) developed a framework to control for the non-random assignment of patients. This work modeled mortality as a function of indicator variables for each hospital and patient discharge information. The authors treat mortality as continuous, and directly apply linear instrumental variables methods. The identifying assumption is that a patient's mortality is not affected by how far that patient's residence is from alternative hospitals. Combined with the demonstrable fact that patients are more likely to choose hospitals that are closer to home, other things equal, the conventional conditions for consistency of instrumental variables estimation in a linear model are satisfied. Conceptually, the estimator would predict hospital A to be of higher quality than hospital B if patients residing near hospital A have lower mortality than patients residing near hospital B, after controlling for their medical and demographic characteristics.

The difficulty with this approach is that because the outcome variable, mortality, is dichotomous, any internally consistent model of hospital quality and choice must be nonlinear. This paper develops a logically coherent model designed to infer the outcome of the conceptual experiment that randomly assigns patients to hospitals, given data that has non-random patient assignment.[6] Inference with this model is challenging because the amount of information per observation is small.[7] This paper develops an approach to inference in this model that is practical with the large data sets required to extract signal from noise in hospital patient discharge databases. This approach is potentially applicable to a wide range of policy evaluations of

---

[5] Leading medical researchers, including Iezzoni *et al.* (1996), and government studies (US GAO (1994)) have both argued that discharge databases are problematic, for this reason.

[6] Though the methods of Gowrisankaran and Town (1999) are much simpler than the ones developed in this paper, there is no formal statistical model that rationalizes their approach.

[7] Simple measures of fit always indicate that most variation in mortality cannot be ascribed to covariates. Even if all the difference in mortality rates were attributable to quality, the variation in these rates is small.

economic interest where the outcome variable is dichotomous.[8]

The model developed here incorporates hospital choice and mortality as endogenous variables, and fixed hospital and patient characteristics as exogenous variables. Hospital choice is described by a multinomial probit model, and mortality by a binary probit model. The mortality model includes indicator variables for each hospital to accommodate hospital specific differences in quality. It is structural, in the sense that it predicts outcomes given any hypothetical assignment of patients to hospitals, including random assignment. The multinomial probit model is a reduced form relationship that provides probabilities of hospital choice conditional on observed covariates that are a function of demographic characteristics and distance of the hospital from the patient's home. The random component in the binary probit model includes unobserved severity of illness, and is permitted to be correlated with the random component in the multinomial choice model. Thus, the model accommodates the possibility that the greater a patient's probability of mortality due to unobserved severity, the more likely it is that the patient is admitted to some hospitals rather than others.

Substantively the selection model works as follows. Suppose that patients with high unobserved severity of illness are more likely to be admitted to hospital A, after controlling for distance, than is the case on average; and conversely for low unobserved severity of illness. Three features of the selection model, in combination, accommodate this pattern. First, the shock to the mortality equation (whose covariates include hospital specific dummies as well as demographic and observed disease characteristics) is interpreted as unobserved severity of illness. Second, the shock to the hospital A choice equation in the multinomial hospital choice model captures the likelihood of being admitted to hospital A after controlling for distance. Finally, correlation between these two shocks (in this case positive) accommodates the systematic impact of unobserved disease severity on the probability of admission to hospital A.

The methodology developed here exploits the similarity between this model and the conventional linear simultaneous equation model. If the mortality probit were observed, the mortality probit equation would be a linear structural equation. Similarly the hospital choice equations would be conventional reduced form equations if their latent utility outcomes were observed. Only the appearance of the discrete hospital choice in the mortality probit equation

---

[8] Examples include the effect of school performance based on graduation rates, of prison rehabilitation programs based on recidivism rates, of job training programs based on the incidence of harassment complaints, and many medical outcome evaluations.

would depart from the classical specification that gives rise to instrumental variable methods. The model developed in this paper handles the unobserved nature of the latent variables through the use of Bayesian Markov chain Monte Carlo (MCMC). These methods iteratively simulate latent variable values conditional on data and parameters, and parameters conditional on data and latent variables. The discrete hospital choice in the mortality probit equation does not pose a problem, and hence the second step is computationally similar to classical instrumental variables. In this way, the simulation methods simultaneously recover the joint posterior distribution of parameters and latent variables.[9] By transforming the integration problem posed by the latent variables to an MCMC problem, the methodology developed here can be used to compute estimates orders of magnitude faster than the method of maximum likelihood.[10] This makes inference feasible in this type of simultaneous equations model. Albert and Chib (1993) used this approach in the binary probit model and Geweke, Keane and Runkle (1997) extended them to the multinomial probit model. The methods developed here extend this approach to a new class of models.

In addition to handling the latent variable problem, Bayesian inference makes it possible to address the motivating policy questions directly, by providing marginal posterior distributions for any functions of parameters. These functions include the probability of mortality in the conceptual random assignment experiment, and the posterior probability that this mortality rate is lower for one hospital than for another. Lastly, Bayesian inference allows us to specify hierarchical priors for hospital quality. This approach, which combines some characteristics of classical fixed- and random-effects models, efficiently extracts the signal from the data.

The data used in this study are taken from the hospital discharge records of 74,848 Medicare patients admitted to 114 hospitals in Los Angeles County during 1989 through 1992 with a diagnosis of pneumonia. The discharge records contain demographic information, including patient addresses, and summary measures of severity of illness at the time of admission. The address data is used to construct the distance of each patient's home from each hospital. Functions of this distance variable, alone and in combination with demographic

---

[9] Surveys that discuss convergence include Chib and Greenberg (1996), Geweke (1997) and Geweke (1999).

[10] Maximum likelihood evaluation for one parameter vector for one individual would require evaluating the joint density of the mortality and hospital choice outcome for that individual. Given that the mortality error and hospital choice error are correlated, this would take several minutes on a fast supercomputer. Multiplied by a data set of roughly 80,000 patients (necessary because of the small signal to noise ratio), it would take months to evaluate the likelihood for a single parameter vector.

characteristics, play a role in the model analogous to that played by the instrumental variables in coping with endogeneity in linear models. The large size of the data set and hierarchical priors are essential because of the low signal-to-noise problem: the ratio of patients to hospitals is roughly 660, but many hospitals treated fewer than 300 patients, and the overall mortality rate is .096. The number of latent variables is roughly the product of the number of patients and number of hospitals, on the order of 10^7, making this one of the largest models of its kind ever applied. This places a premium on issues of computational efficiency, addressed in this study.

Conditional on this data set, the posterior distribution for the parameters of the model has a number of interesting substantive implications. There is substantial variance in the posterior distribution of quality of most individual hospitals: for about 60 percent of the hospitals, there is a posterior probability of at least 10 percent that the hospital is in one of each of three quartiles of the quality distribution. Nonetheless, there appear to be two key relations between hospital characteristics and quality. The smallest and largest hospitals – with 150 or fewer beds or more than 300 beds – have higher quality than medium-sized hospitals, and public hospitals have lower quality than other hospitals.

Turning to the process of hospital admission, there is strong evidence that the level of unobserved severity of illness differs across hospitals. Unobserved severity of illness is found to be positively correlated with hospital quality. This variation in selectivity is quantitatively much more important than variations in hospital quality in explaining the observed variation in mortality rates across hospitals conditional on observed patient characteristics. The simple probit model attributes all variation in hospital mortality rates, conditional on observed characteristics, to hospital quality differences, and therefore leads to strikingly different conclusions about comparative quality. In particular, the simple model does not reveal any sharp relations between hospital characteristics and quality.

Section 2 provides the specification of the model and methods for inference, with some details relegated to an appendix. The database is described in Section 3. Section 4 presents findings on hospital quality and the role of nonrandom admission to hospitals. Section 5 concludes. Six appendices are available on the web.[11] Appendix A1 details the construction of the prior, Appendix A2 details the likelihood function and computation, Appendix A3 provides the posterior distribution of hyperparameters, Appendix A4 gives evidence on the numerical

---

[11] See http://www.econ.umn.edu/~gautam/pdf_papers/bayesquality_appendices.pdf.

accuracy of our Markov chain Monte Carlo (MCMC)) algorithm, Appendix A5 provides posterior rankings for all the hospitals in our data set, and Appendix A6 provides robustness results with alternative priors.

2. The Model

The central component of the model is a structural probit equation, in which the probability of mortality is a function of the hospital to which a patient is admitted, the observed severity of the patient's illness, and the observed demographic characteristics of the patient. The objective is to learn about the way the hospital to which the patient is admitted influences the probability of mortality in this equation. A multinomial probit model of hospital admission supplements the mortality model, to permit non-random assignment of patients to hospitals. This section describes, in turn, the specification of the model, the prior distribution of the model parameters, and methods to recover the posterior distribution of these parameters.

2.1 Model specification

Let $i = 1, \ldots, n$ index the patients in the sample, and let $j = 1, \ldots, J$ index hospitals in the sample. There are two groups of exogenous variables in the model. The $k \times 1$ vector x_i consists of individual characteristics of patient i that may affect mortality, including indicators for age, race, sex, and disease stage, and measures of income. The specifics of these variables are presented in Section 3. The $q \times 1$ vector z_{ij} consists of characteristics specific to the combination of individual i and hospital j. The variables in z_{ij} are the distance between the home of patient i and hospital j, the square of this distance, and the products of distance with age, disease stage, and income, respectively; $q = 5$.

There are two sets of endogenous variables in the model. The mortality indicator m_i is 1 if the patient dies in the hospital within ten days of admission and is 0 otherwise. The $J \times 1$ indicator vector c_i has j'th entry 1 if patient i is admitted to hospital j, and 0 otherwise.

To present the structural mortality equation, let ε_i $(i=1,\ldots,n)$ be independent $N\left(0,\sigma^2\right)$ random variables conditional on the exogenous variables. The mortality probit m_i^* is a latent random variable,

(1) $\qquad m_i^* = c_i'\beta + x_i'\gamma + \varepsilon_i$.

The mortality indicator $m_i = 1$ if $m_i^* > 0$ and $m_i = 0$ if $m_i^* \le 0$. The structural interpretation of (1) is that if patient i were randomly assigned to hospital j, then $m_i^* = \beta_j + x_i'\gamma + \varepsilon_i$ and consequently $P(m_i = 1) = \Phi\left(\left(\beta_j + x_i'\gamma\right)/\sigma\right)$. Note that the parameters β and σ are jointly unidentified in (1) because they can be scaled by the same arbitrary positive constant without changing the behavior of m_i. In the conventional probit model this problem is avoided by setting $\sigma = 1$. We return to this matter in the context of the complete model below.

If c_i were in fact independent of ε_i – as it would be if patients were randomly assigned to hospitals, for example – then c_i would be exogenous in (1). After resolution of the above identification issues this model would conform with the conventional textbook specification of the binary probit model. However, it is likely that in observed data, c_i depends in part on ε_i: the admission of patient i to hospital j takes into account information that is correlated with ε_i. The conventional probit model is then misspecified.

To develop a more plausible model of hospital choice, we assume that patients become infected with one of the many bacterial or viral agents that can cause pneumonia and it has been determined that they are sufficiently ill to benefit from inpatient treatment. At that point the patient (or the patient's agent) selects from the set of J hospitals the hospital to which the patient will be admitted. The actual choice decision will be a complex function of many factors, such as severity of illness, characteristics of the hospital, the patient's primary care physician, etc. One important observable influence on choice is distance: previous research has shown that the farther a patient is from a hospital, the less likely is the patient to be admitted to that hospital, other observables constant.[12]

To present the reduced form model of hospital choice define the $J \times q$ matrix $\tilde{Z}_i$, $\tilde{Z}_i = \left[z_{i1}, z_{i2}, \ldots, z_{iJ}\right]'$. Let the $J \times 1$ vectors $\tilde{\eta}_i \sim N\left(0, \tilde{\Sigma}\right)$ $(i = 1, \ldots, n)$ be mutually independent

conditional on the exogenous variables, and let $\tilde{\rho}_j$, $j = 1, \ldots, J$ denote the correlation between ε_i

and $\tilde{\eta}_{ij}$. Define the $J \times 1$ hospital choice latent vector multinomial probit $\tilde{c}_i^* = \left(\tilde{c}_{i1}^*, \ldots, \tilde{c}_{iJ}^* \right)'$ as

(2) $\qquad \tilde{c}_i^* = \tilde{Z}_i \alpha + \tilde{\eta}_i$.

The choice indicator vector $c_i = \left(c_{i1}, \ldots, c_{iJ} \right)'$ has entry $c_{ij} = 1$ if $\tilde{c}_{ij}^* \geq \tilde{c}_{ik}^*$ $(k = 1, \ldots, J)$ and

$c_{ij} = 0$ otherwise. As above with (1), the parameters α and $\tilde{\Sigma}$ are jointly unidentified since

scaling α by any positive constant and $\tilde{\Sigma}$ by the square of that constant leaves the distribution of

c_i conditional on Z_i unaffected. We return to this matter in the context of the prior distribution

in Section 2.2.

As is customary in models with J choices, it is easier to work with $J - 1$ latent utilities,

and normalize the Jth utility to 0. Accordingly, we define the $(J-1) \times q$ matrix

$Z_i = \left[\tilde{z}_{i1} - \tilde{z}_{iJ}, \tilde{z}_{i2} - \tilde{z}_{iJ}, \ldots, \tilde{z}_{i,J-1} - \tilde{z}_{iJ} \right]'$, the $(J-1) \times 1$ vectors $\eta_i = \left[\tilde{\eta}_{i1} - \tilde{\eta}_{iJ}, \ldots, \tilde{\eta}_{i,J-1} - \tilde{\eta}_{iJ} \right]'$ and

$c_i^* = \left[\tilde{c}_{i1}^* - \tilde{c}_{iJ}^*, \ldots, \tilde{c}_{i,J-1}^* - \tilde{c}_{iJ}^* \right]'$, and the $(J-1) \times (J-1)$ matrix $\Sigma = \text{var}(\eta_i)$. Note that

(3) $\qquad c_i^* = Z_i \alpha + \eta_i$.

If the unobserved severity of illness affects hospital choice, the mortality and choice error

terms will be correlated. Let ρ_j denote the correlation between ε_i and η_{ij} $(j = 1, \ldots, J-1)$. The

larger is ρ_j, the more likely is a patient with a high unobserved severity of illness (ε_i) to be

admitted to hospital j. Thus we shall refer to ρ_j as the *hospital j severity correlation*. The

hospital severity correlations are a useful way to characterize severity of illness by hospital since

they are independent of the scale of ε_i which we know from (1) is unidentified.

Now, we can write the variance of the joint error terms as:

(4) $\qquad \text{var}(\varepsilon_i, \eta_i') = \begin{bmatrix} \sigma^2 & \pi' \\ \pi & \Sigma \end{bmatrix}$

where π is a $(J-1) \times 1$ vector with $\pi_j = \rho_j \sigma \Sigma_{j,j}^{1/2}$.

---

[12] See Luft *et al*. (1990) and Burns and Wholey (1992).

9

To permit unobserved severity of illness to affect hospital choice in any way consistent with the model, the only restriction we place on π is that $\mathrm{var}\left(\varepsilon_i, \eta_i'\right)$ be positive definite. Since this implies complicated restrictions on π, a more graceful treatment is to work with the population regression of the shock ε_i in (1) on the shock vector η_i in (3),

(5) $\qquad \varepsilon_i = \eta_i'\delta + \zeta_i; \ \mathrm{cov}\left(\eta_i, \zeta_i\right) = 0$.

In this regression δ is a $(J-1)\times 1$ parameter vector and the scale of ε_i is normalized by $\mathrm{var}\left(\zeta_i\right) = 1$. This specification simultaneously resolves the identification problem due to the scaling in (1) and incorporates all permissible values of $\pi = \Sigma\delta$ in (4).

With this reparametrization, the variance of the shock in the mortality probit equation is $\sigma^2 = \delta'\Sigma\delta + 1$, and the correlation between ε_i and η_{ij} is

(6) $\qquad \rho_j = \left(\sum_{k=1}^{J-1}\delta_k\Sigma_{kj}\right)\bigg/\left[\Sigma_{jj}\left(\delta'\Sigma\delta + 1\right)\right]^{1/2}$.

In the hypothetical experiment in which patient i is admitted to hospital j by means of a random assignment c_i, $\mathrm{P}\left(m_i = 1 \big| x_i\right) = \Phi\left[\left(c_i'\beta + \mathbf{x}_i'\gamma\right)\big/\left(\delta'\Sigma\delta + 1\right)^{1/2}\right]$. We shall refer to

(7) $\qquad q_j = -\beta_j\big/\left(\delta'\Sigma\delta + 1\right)^{1/2}$,

as the *hospital j quality probit*. Differences in these probits across hospitals may be used to address quality comparisons for individual hospitals. To compare groups of hospitals, we shall make use of the quantities $q_G = \sum_{j \in G}\omega_j q_j$, where G is the group of interest and the weight ω_j is proportional to the number of patients admitted to hospital j. In the conventional probit model with normalization $\sigma = 1$, the hospital j quality probit is $q_j^* = -\beta_j$.

2.2 Prior distributions

In order to solve for the posterior of the model, we need to define priors for the parameters that are used in the estimation process, which are Σ, α, δ, γ, and β. Given the complexity of the model and the low signal-to-noise ratio in the data, the prior distribution must be chosen carefully to reflect reasonable beliefs about hospital choice and mortality.

The number of parameters in the variance matrix Σ in the reduced form multinomial probit model for hospital choice is $J(J-1)/2$, that is, 6,441 in our sample with J=114 hospitals.

The identifying scale normalization reduces the number by only one. Because of the large number of parameters, and because the purpose of this study is to model mortality while permitting non-random hospital admission rather than to study the hospital admission process *per se*, we fix $\tilde{\Sigma} = I_J$, the variance matrix that would result if the random components of utility associated with hospital admission were independently and identically distributed across hospitals prior to normalization. (We introduce some evidence on the plausibility of this assumption in Section 4.4.) Fixing Σ avoids the most common source of instability in inference for multinomial probit models (Keane, 1992). After differencing, $\Sigma = I_{J-1} + e_{J-1}e'_{J-1}$, where e_n denotes an $n \times 1$ vector of units.

The prior distributions of α, δ, γ, and β are independent of one another. The first three distributions are Gaussian, and the last is Gaussian within a hierarchical prior. All have mean zero. The prior variances are chosen to bring reasonable values of the parameter vectors well within two standard deviations of their means. Appendix A1 contains detailed descriptions of all the priors.[13]

The coefficients in α, the coefficient vector in the multinomial probit hospital choice model, are independent in the prior. Each prior standard deviation is chosen so that if the corresponding coefficient is one standard deviation below the mean and all other coefficients are zero, then the probability of choice for the closest hospital is between 0.1 and 0.2, in the sample as a whole, whereas the probability that any one of the 27 hospitals farthest away is chosen is about 0.003. The prior standard deviations are 6, 12, 4, 5 and 35 for the coefficients on the five variables, respectively, listed in the bottom panel of Table 4.

The prior variance of δ was derived by first considering the linear projection of the mortality probit equation shock, ε_i, on the disturbance $\tilde{\eta}_i$ of the multinomial probit model (2) before differencing utilities. We consider prior distributions on the coefficients $\tilde{\delta}_i$ of this projection that are Gaussian, independent, and exchangeable – i.e., that all have the same prior variance $\sigma_{\tilde{\delta}}^2$. It can be shown that the implied prior distribution for δ in (5) is not Gaussian, but is well approximated by the distribution $N\left(0, c \cdot \Sigma^{-1}\right)$ with $c = \sigma_{\tilde{\delta}}^2 / \left(1 + \sigma_{\tilde{\delta}}^2\right)$. (Appendix A1 provides the derivation.)

To choose the value of $\sigma_{\tilde{\delta}}^2$, we examined the implications of alternative values of this hyperparameter for $\tilde{\rho}_j$ (the sample correlation between ε_i and $\tilde{\eta}_{ij}$ of (2)) and $\tilde{R}^2$, the squared multiple correlation between ε_i and $\tilde{\eta}_i$. From (6) the $\tilde{\rho}_j$ all have the same prior distribution with mean 0, because the $\tilde{\delta}_j$ are independent with mean 0. As $\sigma_{\tilde{\delta}}$ increases from 0 to 0.2, $E\left(\left|\tilde{\rho}_j\right|\right)$ increases from 0 to 0.067, and $\mathrm{corr}\left(\left|\tilde{\rho}_i\right|,\left|\tilde{\rho}_j\right|\right)$ increases from 0.001 to 0.005, given $J = 114$ choices (as is the case). For $\sigma_{\tilde{\delta}} > 0.2$, $E\left(\left|\tilde{\rho}_j\right|\right)$ does not increase much beyond 0.067. However $\tilde{R}^2$ rises from 0 at $\sigma_{\tilde{\delta}} = 0$ to 0.27 at $\sigma_{\tilde{\delta}} = 0.2$ to 0.68 at $\sigma_{\tilde{\delta}} = 0.5$, and $\tilde{R}^2$ approaches unity as $\sigma_{\tilde{\delta}} \to \infty$. We chose $\sigma_{\tilde{\delta}} = 0.2$ for the final specification. At this value, the prior mean and standard deviation of the $\left|\tilde{\rho}_j\right|$ are 0.067 and 0.060, respectively, $\mathrm{corr}\left(\left|\tilde{\rho}_i\right|,\left|\tilde{\rho}_j\right|\right) = 0.005$, and the prior mean and standard deviation of $\tilde{R}^2$ are 0.27 and 0.09, respectively.

Of course the prior distribution of the correlation coefficients ρ_j between ε_i in (1) and η_{ij} in the differenced model (3) are quite different. The prior mean and standard deviation of $\left|\rho_j\right|$ remain unchanged (0.067 and 0.05), but $\mathrm{corr}\left(\left|\rho_i\right|,\left|\rho_j\right|\right) = 0.233$, reflecting the differencing used to create (3). The squared multiple correlation coefficient R^2 between ε_i in (1) and η_i in (3) has prior mean 0.82 and prior standard deviation 0.02.[14]

The demographic covariates x_i are of two types: dichotomous variables, and two continuous variables (income and its square). The coefficients γ_i of the dichotomous variables are independent in the prior, all with standard deviation 0.5 in the conventional probit model with $\sigma = 1$. The coefficients on income and its square (call them γ_1 and γ_2) are derived from the independent priors

$$
\gamma_1 \bar{y} + \gamma_2 \bar{y}^2 \sim N\left(0, 0.25^2\right),
$$
$$
\gamma_1 \left(2\bar{y}\right)^2 + \gamma_2 \left(2\bar{y}\right)^2 \sim N\left(0, 0.25^2\right).
$$

---

[13] The appendices of this paper are available at http://www.econ.umn.edu/~gautam/pdf_papers/ bayesquality_appendices.pdf.
[14] Full details on these effects are provided in Table A1 in the Appendix.

Substituting for average income $\bar{y}$, scaling y by 10^{-5} and y^2 by 10^{-9} (as was done in variable construction) yields prior variances of 3.603 and 0.1437 for γ_1 and γ_2, respectively, and a prior covariance of –0.7026.

The prior distribution for β in the conventional probit model is chosen by considering a variance components structure. Each hospital belongs to one of four ownership categories and one of four size categories. If hospital i is of ownership category j and size category k, then decompose $\beta_i = \beta_1 + p_j + s_k + u_i$. To construct the prior let the components $\beta_1, p_1, \ldots, p_4, s_1, \cdots, s_4$ and $u_1, \ldots, u_{114}$ be jointly Gaussian, all with mean zero, and mutually independent. The common term β_1 has standard deviation 3 (essentially a flat prior). The other components have variances τ_p^2, τ_s^2 and τ_u^2, respectively, which we group together in the vector $\tau' = \left(\tau_p^2, \tau_s^2, \tau_u^2 \right)$. This induces a variance matrix for the entire vector β: for example, $\mathrm{cov}\left(\beta_i, \beta_j \right) = 9 + \sigma_p^2$ if hospitals i and j are in the same ownership category but different size categories.

The variance terms τ_p^2, τ_s^2 and τ_u^2 are treated as unknown rather than fixed parameters, with the independent prior distributions $1.25 / \tau_j^2 \sim \chi^2 (5) \, (j = p, s, u)$.[15] Thus the prior distribution of β is hierarchical. The centered 99% prior credible interval for the τ_j^2 is (.22, 1.7). The effect is to permit anything from modest to drastic variation in quality across hospitals, after controlling for demographic covariates. Compared with a normal prior, this hierarchical prior assigns a more important role to the data in the posterior distributions of the τ_j^2, which are in turn summary measures of variation in quality within and across classes of hospitals. Appendix A3 provides details of the posterior distribution of these hyperparameters.

From (5), $\mathrm{E}\left(\varepsilon_i^2 \right) = \delta \Sigma \delta + 1$ in the selection model, whereas $\mathrm{var}\left(\varepsilon_i \right) = 1$ in the probit model. Since (to a good approximation) $\delta \sim N\left(0, c\Sigma^{-1} \right)$, with $c = 0.2^2 / \left(1 + 0.2^2 \right) = 0.0385$, it follows that $\delta \Sigma \delta + 1 \sim c \cdot \chi^2 (J - 1) + 1$ and $\mathrm{E}(\delta \Sigma \delta + 1) = c(J - 1) + 1$ (all to a good approximation). This suggests that the prior standard deviations for β_1 and the γ_j in the probit

---

13

model be scaled by $\left[c(J-1)+1\right]^{1/2} = 2.38$ in the selection model; thus for instance $\beta_1 \sim N\left(0, 2.38^2 \times 9\right)$. Similarly, the scale parameter in the priors for the τ_j^2 are increased by the factor 2.38^2, so that $7.08/\tau_j^2 \sim \chi^2(5)\,(j = p, s, u)$.

2.3 Inference

The observed data are $\left(x_i, Z_i, c_i, m_i, i = 1, \ldots, n\right)$, which can be abbreviated as y. The model contains latent variables $\left(m_i^*, c_i^*, i = 1, \ldots, n\right)$, which can be abbreviated y^*. The parameter vectors are α, β, γ and δ, which can be collected in the vector θ. The model specified in Section 2.1 provides $p\left(y, y^* | \theta\right)$ and the prior distributions in Section 2.2 provide $p(\theta)$. Explicit expressions for these densities are given in Appendix A2. From Bayes rule, the distribution of the unobservables y^* and θ conditional on the data and model specification is

(8) $\qquad p\left(y^*, \theta | y\right) = \dfrac{p(\theta)p\left(y, y^* | \theta\right)}{p(y)} \propto p(\theta)p\left(y, y^* | \theta\right).$

The objective is to obtain the posterior distribution of functions like the hospital quality probits q_j, and $\Phi\left(-q_j + x_i'\gamma\right)$, the probability of mortality under random hospital admission of a patient with observed characteristics x_i to hospital j. A closely related quantity of interest is $P\left(q_j > q_k \mid y\right)$, the posterior probability that hospital j ranks above hospital k in the conceptual motivating experiment. This objective requires integrating a highly nonlinear function over millions of dimensions, most of which correspond to latent variables. It cannot be achieved through analytical means.

Instead, we take advantage of the fact that the parameter vector and latent variables can be partitioned into groups, such that the posterior distribution of any one group conditional on all the others is of a single, easily recognized form that is easy to simulate. Details of the partition are given in Appendix A2. The problem is thus well suited to attack by execution of a Gibbs sampling algorithm (Gelfand and Smith, 1990; Geweke, 1999). In this approach, each group of parameters and latent variables is simulated conditional on all the others. Following each pass

14

through the entire vector of latent variables and parameters, all parameter values are recorded in a file.

As detailed in Appendix A2, the Gibbs sampling algorithm is ergodic and its unique limiting distribution is the posterior distribution. Therefore, dependent draws from the posterior distribution of any function of the parameters $g(\theta)$ can be made by computing the value of g corresponding to the recorded parameter values, after discarding initial iterations of the Gibbs sampling algorithm to allow for convergence. We used parallel computing methods and a supercomputer, exploiting the fact that in each iteration of the Gibbs sampling algorithm the latent variables $\left(m_i^*, c_i^*, i = 1, \ldots, n \right)$ are conditionally independent across individuals. The iterations themselves are executed serially. The results reported in Section 4 are based on every 10th draw from 9,000 successive iterations (a total of 900 draws), after discarding 1,000 burn-in iterations based on convergence diagnostics. For comparison purposes, the same procedures were applied to a conventional probit model for mortality, with the variance of ε_i in (1) fixed at $\sigma = 1$. This computation, which is much simpler, is based on every 10th draw from the Gibbs sampling algorithm described in Albert and Chib (1993). Appendix A4 provides details on the numerical accuracy of our Gibbs sampling algorithm.

3. The Data

The primary source of data for this study is the Version B Discharge Data from the State of California Office of Statewide Health Planning and Development. These data provide records for all patients discharged from any California acute-care hospital during the years 1989 through 1992. We chose to analyze four years because the number of patients per hospital was then large enough to obtain meaningful inference but small enough to be computable. We did not choose more recent data, because increased managed care penetration among Medicare enrollees during the 1990s changed the environment in which hospital admission decisions were made. We confine our attention to Los Angeles County. A large metropolitan area is best suited to our purposes, because it has a large base of patients and contains multiple hospitals in every size and ownership class. We limit our study to a single disease, because there is evidence that the

relation between mortality and covariates is disease specific.[16] We choose pneumonia in particular for three reasons. First, it is a common disease[17] that provides the large sample needed to draw inferences about hospital quality. Second, in-hospital death is a relatively frequent outcome for pneumonia patients, which makes it an attractive disease to examine through the medium of hospital discharge records. Third, there is independent evidence that an appropriately adjusted in-hospital mortality rate for pneumonia is correlated with the quality of in-hospital care.[18] We further confine our attention to patients who were over 65 at the time of admission. Medicare is the common primary source of medical insurance for this group, and in the case of pneumonia limiting the study to patients over 65 leaves a large patient base.

The secondary source of data is the Annual Survey of Hospitals Database published by the American Hospital Association (AHA). Among other information, the AHA data contain the addresses, ownership status, and size of each hospital in our sample.

3.1 Sample construction

The sample was selected through a process of eliminating patients from the 1989-1992 Version B Discharge Data. The first qualification for selection is that the patient live in a Los Angeles County zip code, be admitted to a Los Angeles County hospital and be over 65 at the time of admission.

The second qualification is that one of the five ICD-9-CM disease codes specified in the discharge data be 48.1, 48.2, 48.5, 48.6, or 48.38. This procedure is suggested by Iezzoni *et al*. (1996) to define pneumonia. There is substantial non-random variation across hospitals in the sequence of ICD-9 diagnoses listed. Thus, choosing the first listed ICD-9 code may induce biases. (Iezzoni (1997), Chapter 3).

The third qualification is that the source of admission must be either routine, or from the emergency room. This eliminates patients transferred into the hospital from another medical facility, or admitted from an intermediate care or skilled nursing facility. To the extent that placement in these facilities is correlated with unobserved disease severity, and to the extent that such facilities may be systematically located near higher quality hospitals, the key assumption

---

[16] See Wray *et al.* (1997)

[17] Pneumonia and influenza alone constitute the sixth leading cause of death in the US, and the fourth leading cause of death for those over 65 (National Center for Health Statistics, 1996). Pneumonia is also the leading cause of death among patients with nosocomial (hospital acquired) infections (Pennington, 1994).

[18] See Keeler et al. (1990) and McGarvey and Harper (1993).

that distance from the hospital is exogenous in our model would be violated. This step eliminates approximately 23 percent of the patients from the sample.

The fourth qualification for inclusion in the sample is that the patient be admitted to a hospital with at least 80 admissions for pneumonia in our data set. This qualification was imposed for two reasons. First, the fewer admissions to a hospital in our data the less we can learn about the quality of that hospital. The hospitals eliminated through this qualification would have had a very low signal-to-noise ratio. Second, computation time in the Gibbs sampling algorithm is largely driven by the number of latent variables. To have included the 16 hospitals eliminated through this qualification would have markedly increased computation costs while providing little additional information about the unknown parameters. In principle, this qualification introduces a problem of choice based sampling, but because only 606 patients were thereby eliminated we believe that this is a negligible difficulty.

3.2 Variable construction

The covariate vector x_i in the mortality probit equations contains a constant term, demographic variables and indicators of disease severity. Most of the demographic variables are constructed from the discharge records. These are four age indicators (70-74, 75-79, 80-84, and 85 or older), an indicator for female, and indicators for black, Hispanic, Native American and Asian respectively. The discharge records contain no information on socioeconomic status. As a proxy for the patient's household income, we use the mean 1990 census household income for households with the same zip code, race, and age class as the patient.[19]

Indicators of disease severity in x_i are constructed from the admission disease staging information contained in the discharge records. Disease staging has been shown to be as good as some risk adjustment data based on chart review of medical records.[20] Since some of the 13 stages have very few patients, we aggregated stages into five groups: stage 1.1, stages 1.3

---

[19] The census provides only two relevant age categories, 65 - 74 and 75+, instead of four. Thus, we aggregated the discharge data age categories to this level. Additionally, the census provides income only within cells. To find the mean income, we took the mean value for each cell as the income for each household in that cell. For the highest cell, $100,000 or more, we assumed a mean income of $140,000. Income is measured in units of $100,000 and income squared in units of billions of dollars squared.

[20] See Thomas and Ashcroft (1991). Iezzoni *et al.* (1996) showed excellent agreement of disease stage with the ratings of other systems.

through 2.3, stages 3.1 through 3.6, stage 3.7, and stage 3.8. Indicator variables for all but stage 1.1 are included in x_i.

The indicator for mortality, m_i, is set to 1 if the patient died in the hospital within ten days of admission to the hospital; otherwise it is set to 0. The horizon for mortality is limited to ten days, because beyond this point hospitals sometimes transfer terminally ill patients to other facilities, and this decision appears to vary considerably by hospital. To control for differential patient transfer, Gowrisankaran and Town (1999) used a hazard model as an alternative to the 10-day inpatient mortality, but found little difference between the two specifications. In two separate studies of heart disease patients, McClellan, McNeil and Newhouse (1994) and McClellan and Staiger (1999b), find that there is a very strong correlation between 7-day mortality and 30-day mortality rates across hospitals.[21]

Table 1 provides a summary of the distribution of demographic characteristics and disease severity in the sample, together with mortality rates. For each age group the breakdown of the sample by race and sex closely reflects the demographics of Los Angeles County. Older individuals enter the sample in greater proportion to their numbers in the population than do younger ones. In each age group three-quarters of the sample is classified in the least severe disease stage. Mortality rates increase gradually with age, increase sharply with disease stage, are a little higher for men than for women, and are lower for Asians and Hispanics than for whites or blacks.

The covariate matrix Z_i contains variables specific to the combination of patient i and each hospital. The additional information in Z_i not contained in x_i is the distance of the patient's home from each hospital. We obtained patient zip codes from the discharge data and the hospital zip codes from the AHA data. We then used the Census TIGER database to find the latitude and longitude of the centroid of each zip code. Given the latitudes and longitudes, we computed the distance between each patient home and hospital using standard great circle trigonometric formulas.[22] We then constructed the five variables in Z_i: distance (always measured in hundreds of kilometers); distance-squared; the product of distance and an age indicator (1 for 65-69, 2 for

---

[21] As caveats, note that heart disease is very different from pneumonia and that these studies examine mortality, not inpatient mortality.

[22] For zip codes that contain more than one hospital, we used address-level latitude and longitude data from the Census TIGER database, which stores the geographic location of every block corner and will interpolate from that to find the latitude and longitude of any address.

70-74, 3 for 75-79, 4 for 80-84, 5 for 85+); the product of distance and disease stage (1.1, …, 3.8); and the product of distance and income (in units of $100,000).

The size and ownership status of different hospitals is important in our subsequent analysis. This information was obtained from the AHA survey, and is summarized in Table 2. Note that we have grouped private teaching hospitals as a separate ownership category from private not-for-profit hospitals.[23] Most hospitals are private, split about evenly between for-profit and not-for-profit. Only ten of the 114 hospitals in the sample are teaching or public, but on average they are larger than private hospitals and together admitted over 12% of the patients in the sample. Slightly less than one-quarter of the hospitals are classified in the largest size group (at least 300 beds) but they account for over 40% of the admissions in our sample.

While mortality rates differ slightly by ownership category none of the differences are significant at conventional levels. The same is true by size category. Contrasts in mortality rates are stronger between cross-classified cells in Table 2. Some of these comparisons, for example private for-profit with 201-300 beds (10.54%) and teaching with more than 300 beds (9.17%) are significant at or below the 5% level.

Table 3 summarizes the distribution of severity of illness, as measured by disease stage, across the different categories of hospitals. Patients in larger hospitals tend to be at a more advanced disease stage. The differences in the distribution are small, but because of the large sample size they are highly significant: the test statistic for categorical independence is $\chi^2(12) = 96.9$ $(p = 0.000)$. The distribution of patients by disease stage over hospitals of different ownership type is yet more uneven: almost 79% of the patients in teaching hospitals are in the earliest disease stage, whereas at private for-profit hospitals only a little over 74% are at this stage. The test statistic for categorical independence is $\chi^2(12) = 261.0$ $(p = 0.000)$. Thus it is the case in this data set, as in similar data sets, that observed severity is not randomly distributed across hospitals. This underscores the importance of examining and controlling for nonrandom assignment by unobserved severity, as well.

The summaries of the data provide no simple interpretation of mortality rates. They indicate systematic differences in measured disease severity across hospitals by size and ownership classes. They hint at the possibility of important differences between individual

---

[23] Los Angeles County operated teaching hospitals are assigned to the public hospital category.

hospitals within size and ownership classes. Thus we now turn to the application of the model developed in Section 2 to inform our understanding of the relationship between choice of hospital admission and mortality.

4. Findings

The model set forth in Section 2 applied to the data described in Section 3 yields evidence on systematic differences in quality across hospitals, provides insight into the interaction between hospital choice and hospital quality, and suggests quality orderings among hospitals. This section summarizes and examines the robustness of these findings. Recall that the posterior functions that we are primarily interested in are the hospital quality probits q_j and q_j^*, groupings of these quality indicators q_G and q_G^*, the hospital severity correlations ρ_j, and rankings of hospitals by quality probits.

4.1 Patient mortality and hospital choice

Table 4 presents the posterior means and standard deviations of some parameters and functions of parameters in the selection and standard probit models. In the case of the selection model Table 4 presents the posterior means and standard deviations of the normalized mortality covariates $\gamma_j/(\delta'\Sigma\delta+1)^{1/2}$, the group hospital quality probit q_G, and the coefficients α on hospital choice. The normalization of the γ_j facilitates comparison between models and interpretation of the functions of interest as probits.

The mortality equation has three groups of covariates: demographics, disease severity, and hospital indicators. In the case of the demographic and disease severity covariates, coefficient posterior means in the selection and probit models are similar to each other, and closely reflect the mortality rates presented in Table 1. Posterior standard deviations indicate substantial information about differences in mortality probabilities across demographic groups. This, too, is not surprising in view of the summary statistics in Table 1.

In the case of the group hospital quality probits, there are greater and more interesting differences between the selection model, the probit model, and the raw data. Both the probit model and the raw data (Table 2) do not draw any sharp distinctions in hospital quality by size or

ownership class. However, the selection model finds sharp distinctions by size and ownership type. This suggests that controls for both observed and unobserved severity of illness are important.

Posterior means and standard deviations of the choice covariate coefficient vector α are presented in the bottom panel of Table 4. As expected, distance is an important factor in describing the hospital of admission. To interpret the posterior mean of –13.61, recall that the disturbances in (3) all have variance 2, and that all disturbance covariances are 1, and distance is measured in hundreds of kilometers. The difference of any two disturbances also has variance 2. Hence the posterior mean of –13.61 implies that a hospital that is 20 kilometers farther from a patient's home than another has a normalized probit that is $13.61 \times 0.2 / \sqrt{2} \approx 2$ units lower. The quadratic term in the equation is highly significant, but since distances are at most 100 kilometers within Los Angeles County, its substantive effect is not great.

Other covariates have important impacts on hospital choice as well. Interactions of distance with age, and with severity, all have negative coefficients with posterior standard deviations small relative to their posterior means. Since the age variable is at least 1 and the severity variable is at least 1.1, this lowers the posterior mean of the distance coefficient downward from –13.61, at least as far as –14.42 for everyone in the sample, and to –17.11 for the oldest (85+, coded 5) and most severely ill (disease stage 3.8) patients in the sample. Older and more severely ill patients clearly have the greatest propensity to be admitted to hospitals closer to where they live. The reason for this is likely due to the increased cost and difficulty of transport for severely ill patients. Patients in zip codes with higher average income are more likely to be admitted to nearby hospitals.

Table 5 provides explicit posterior probabilities for hospital group quality comparisons using the selection model. There are sharp differences based on hospital size (Panel A). The posterior probability that the group hospital quality probit for the largest-hospital group exceeds that of the smallest-hospital group is 0.80, and the posterior probability that it exceeds that of the other two size groups exceeds 0.99. The posterior probability that the smallest-hospital group quality probit exceeds that of the two middle-sized groups similarly exceeds 0.99.

These findings are in rough agreement with the literature. A study by Keeler *et al.* (1992), which examined the relationship between hospital quality and size using a very detailed and expensive data set that included pneumonia patients along with patients with other, more

complex diagnoses, found that hospital quality increases with bed size. However, in their study they did not allow for a nonlinear relationship between hospital size and morality rates, thus they could not uncover the U-shaped relationship between hospital quality and size that we do. Successful pneumonia treatments are linked to identifying the pathogen responsible for the infection and administering the appropriate antibacterial agent early in the progression of the disease, and subsequently monitoring and adjusting the dosage of the drug (Rello and Valles (1998), Pennington (1994), McGarvey and Harper (1993)). There is evidence that smaller hospitals may be better at the timely administration of antibiotics (Fine *et al.* (1998)) which may explain why we observe that they have better outcomes. Furthermore, since small hospitals are likely to treat a disproportionate number of pneumonia patients relative to more technically challenging illnesses[24] they may also develop expertise in this disease. That, in turn, may overcome advantages that medium-sized hospitals may have in other dimensions, such as laboratory facilities.

There are less sharp differences in the selection model based on ownership (Panel B). However, the selection model does draw some clear distinctions: the group quality probit for private teaching hospitals appears to be better than that of public hospitals with a posterior probability 0.99, better than other private not-for-profit hospitals with posterior probability 0.95, and better than that of for-profit hospitals with posterior probability 0.93. Overall, the model predicts that private teaching hospitals are the best, other non-public hospitals are in the middle, and public hospitals have the worst group quality probit. For-profit and non-teaching private not-for-profit hospitals appear to be similar, with posterior probability 0.68 that the not-for-profit hospitals are of higher quality.

There is debate in health policy circles regarding the role that for-profit hospitals should play in the U.S. health system (Gray (1991)). Some have argued that private, not-for-profit hospitals may better serve the public interest because they are more likely to provide better care. Our results indicate that for the treatment of pneumonia in older patients and the hospitals in our sample, there is no evidence of this. Keeler *et al.* (1992) also found public hospitals in large cities to be of lower quality, while the difference in quality between for-profit and not-for-profit

---

[24] Performing a simple multinomial logit regression of Southern California patients, we found that pneumonia patients were more likely to be admitted to smaller hospitals than were hospital patients generally. In contrast, acute myocardial infarction (heart attack) patients were more likely to be admitted to larger hospitals than the average hospital patient. Unlike pneumonia treatments, acute myocardial infarction treatments often include high-technology surgery such as cardiac catheterization, angioplasty or bypass.

hospitals is less pronounced. McClellan and Staiger (1999a) conclude that the quality difference in for-profit and not-for-profit hospitals is small and if anything for-profits likely provide better care in the treatment of heart attacks. Private teaching hospitals, which are generally viewed as providing superior care (Keeler *et al.* (1992)) , do appear to offer significantly higher quality according to the selection model.

Table 6 converts the group hospital quality probits from Tables 4 and 5 to mortality differences, beginning from a base mortality probability of 0.10, which is typical for pneumonia in admitted patients over the age of 65. In the conceptual experiment underlying these tables, a patient is assumed to be admitted to one type of hospital with mortality probability 0.10, and the mortality probability for that patient in other types of hospitals is then inferred from the data and the selection model. The cell entries are posterior means of the latter probabilities, and the corresponding posterior standard deviations are indicated parenthetically. For changes in hospital size (Panel A), the range of probability differences is greatest in the comparison of the largest hospitals with 151-200 bed hospitals, where the posterior distribution indicates an increase of 29% in moving from the former to the latter. Other changes are somewhat smaller, but the direction of change is often clearly indicated by the data and the selection model. In the case of classification by ownership (Panel B), changes in probability are again somewhat less pronounced. Consistent with Tables 4 and 5, contrasts are most marked in movements between teaching and public hospitals. Note the expected increase in mortality probability of 53% associated with a move from the former to the latter.

4.2 Selection and selection bias

We present some statistics on the relationship between the posterior means of q_j, q_j^* and ρ_j across the 114 hospitals in Table 7. These statistics allow us to uncover the importance of selection and the relationship between selection and quality.

We start by analyzing the quantitative importance of selection in influencing patient mortality. In the simple probit model (1) the variance in unobserved disease severity ε_i is normalized to be 1. From the posterior means of the coefficients on observed disease severity in the model (Table 4) and the distribution of observed disease severity in the population (Table 1), one may approximate the variance in the contribution of observed disease severity to the

mortality probit: it is about 0.45. The posterior means of the quality probits q_j^* in the model provide the variance in the mortality probit due to variation in hospital quality: it is about 0.013 (Table 7 Panel A). Thus about two-thirds of variation in the mortality probit is due to unobserved disease severity, almost one-third to observed severity, and about one percent to hospital quality. This decomposition of variance is about the same in the selection model – variation in hospital quality is slightly higher (Table 7) but it is still quite small relative to disease severity.

In the selection model the variation in unobserved disease severity is decomposed into a component that is independent of the hospital assignment process (ζ_i from (5)) with variance 1, and a component that is a function of the hospital assignment probits, $\eta_i'\delta$ (also from (5)). The variance of the latter term, $\delta'\Sigma\delta$, has a posterior mean of 4.4, which is much larger than the independent component. This constitutes strong evidence against random assignment of patients, and suggests that the simple probit model provides misleading inferences about hospital quality.

Since patient selection is important, we are interested in understanding the relationship between selection and quality. Table 7 Panel A reveals a positive relationship between the posterior means of q_j and ρ_j: the correlation between posterior means is 0.517 (Panel A) and a simple least squares regression of the posterior means of the ρ_j on the posterior means of the q_j shows a slope coefficient of 0.183 that is significantly positive (*t* of over 6).[25]

To interpret the relation between q_j and ρ_j, consider patient i with high unobserved severity of illness and therefore a large value of ε_i in the mortality equation (6). As ε_i increases, this patient is increasingly likely to be admitted to hospitals with larger values of ρ_j rather than hospitals with smaller values of ρ_j because the shocks η_{ij} in the choice model (8) are likely to be higher for these hospitals. Hospitals with high severity correlations ρ_j also tend to be

---

[25] Since results in Table 7 are based on posterior means, they do not take into account dispersion in the posterior, either. Within the selection model this can be accommodated by regarding the sample relation between q_j and ρ_j as a function of the unknown parameters in the model, and then considering the posterior uncertainty associated with this relationship. To make this approach operational, one produces panels A and B of Table 7 for each draw from the posterior simulator, using the values of q_j and ρ_j for that simulation instead of the posterior means. This yields a posterior mean of .414 and posterior standard deviation of .059 for the sample correlation between q_j and ρ_j over the 117 hospitals. In the regression equation of panel B the posterior mean of the slope coefficient is .198 and its posterior standard deviation is .028. The OLS slope coefficient standard error has a posterior mean of .041 and posterior standard deviation of .003; "*t*" has a posterior mean of 4.92 and posterior standard deviation of 0.85. Relationships among hospitals are thus not much affected by uncertainty about individual hospitals. As these results

hospitals with higher quality q_j. Thus, other things equal, patients with greater unobserved severity of illness are more likely to be admitted to higher quality hospitals.

The hospital-by-hospital impact of the adjustment for nonrandom assignment in the selection model is portrayed in Figure 1. Each axis measures the posterior mean of the hospital-specific mean probability of mortality across patients in our sample, with the conventional probit model on the horizontal axis and the selection model on the vertical. The difference in the two mortality rates exceeds 0.02 in absolute value for 26 of the 114 hospitals and exceeds 0.04 for three. The correlation coefficient of 0.741 between q_j and q_j^* noted in Panel A of Table 7 is clearly reflected in Figure 1.

In any selection model, conditional on observed characteristics (including observed severity), the observed mortality rate for each hospital will be decomposed into a hospital quality component and an unobserved severity component. Panel C of Table 7 shows that in this relationship hospital quality q_j^* in the probit model is well described as a linear function of hospital quality q_j and severity correlation ρ_j in the selection model. From the regression relation reported in panel C of Table 7, it is clear that variation in hospital severity correlation substantially drives variation in inferred hospital quality q_j^* in the probit model. From the regressions in panels B and C, one can infer the slope coefficient of .585 $\left(=.888\text{-}1.651\times.183 \right)$ in panel D. Thus, variation in hospital severity correlation accounts for a substantial portion of the variation in hospital mortality rates in the selection model, whereas in the simple probit model this variation must be attributed to quality differences.

4.3 Ordering by quality

The model and approach to inference described in Section 2 provide the complete posterior distribution of all the parameters in the model, and any functions of these parameters. In particular, corresponding to the parameter values in any iteration of the Gibbs sampling algorithm, it is a simple matter to compute the corresponding hospital quality probits q_j. The 900 draws used to obtain the posterior moments reported in this section therefore also provide 900 draws from the joint distribution of the hospital quality probits q_j. Pairwise comparisons

---

indicate, hospital-specific parameters are roughly independent in the posterior, and consequently relations between them are well summarized by relations between their posterior means.

between hospitals are then straightforward. For example, for two hospitals j and k, the numerical approximation to the posterior probability that $q_j > q_k$ is the fraction of iterations in which $q_j > q_k$, and the joint distribution of q_j and q_k could easily be plotted.

Comparing all 114 hospitals simultaneously is more challenging. A formal approach to ordering hospitals by quality would begin with a loss function for orderings. Suppose the 114-element vector of quality ranks is $\mathbf{r}$, and the estimated quality rank vector is $\hat{\mathbf{r}}$. If the loss function is $(\hat{\mathbf{r}} - \mathbf{r})' \mathbf{A} (\mathbf{r} - \hat{\mathbf{r}})$, where $\mathbf{A}$ is a positive definite matrix, then $\hat{\mathbf{r}}$ should be the posterior mean of $\mathbf{r}$.[26] This estimate may, in turn, be approximated numerically by sorting hospital qualities q_j in each iteration of the Gibbs sampler, finding the corresponding rank for each hospital, and then averaging the ranks across all iterations. The resulting estimated ranks $\hat{r}_j$ are generally not integers.

Table 8 provides the results of this approach for the selection model for a subset of the hospitals. The hospitals listed in the table are sorted by the values of the $\hat{r}_j$ and then selected at roughly evenly spaced points based on this sorting. The number to the left of each hospital name indicates its order in this sorting. The $\hat{r}_j$ are shown in the second column following the hospital names. Posterior mean quality is shown in the first column following the names. Ordering by hospital quality posterior mean does not lead to the same ordering as ordering by $\hat{r}_j$, but it is very close. If the loss function were $\sum_{j=1}^{117} a_j |\hat{r}_j - r_j|$, where all $a_j > 0$, then $\hat{r}_j$ should be the median of the posterior distribution of r_j, which in turn is an integer (with probability one). Medians are shown in the third column following the hospital names. They provide yet another ordering, but it too is similar to ordering by mean quality and mean rank. We conclude that choice of loss function is not likely to affect orderings of point estimates of relative quality very much.

Of much greater significance is posterior uncertainty about comparative quality. Table 8 conveys this in several ways. The last four columns provide the probability of being in each quality quartile for each hospital. Placement within a quartile is most certain for hospitals of very high or very low quality. On the other hand for 10 of the 17 hospitals in Table 8, all far from these extremes, the posterior probability is at least .10 that the hospital is in one of each of *three*

---

[26] See, for example, Bernardo and Smith (1994, Section 5.1.5), for this standard result, as well as the one on medians used in the next paragraph.

quartiles. The uncertainty conveyed by the posterior distribution is also reflected in the mean and median ranks. If there were no posterior uncertainty about ranks, the mean and median ranks would be identical to each other, and to the ordering number to the left of each hospital name. At the other extreme, if hospital qualities were completely exchangeable in the posterior distribution, the mean and median ranks would be 57.5 Note that the situation in Table 8 is intermediate between these two extremes, but closer to the former than the latter.

Table 9 provides an alternative expression of the uncertainty about ranks conveyed by the posterior distribution, for hospitals chosen from alternating lines of Table 8 Then, pairwise posterior probabilities of orderings were computed from the iterations of the Gibbs sampler. For the first and last hospitals, fairly confident conclusions can be draw in comparisons with the other eight, but note that the posterior probability that the quality of Sherman Oaks (15) exceeds St. Johns (1), and that Harbor/UCLA (114) exceeds Charter Community (99), are each above .10. For hospitals ordered 15, 29, 43, 57, 71, 85, and 99, only two orderings that can be made with .90 posterior probability: the quality of Sherman Oaks (15) exceeds that of Charter Community (99) and Robert F. Kennedy (71). Table 9 suggests that for half the hospitals in the sample, one cannot order quality pairwise with posterior probability that exceeds .85.

As suggested by Figure 1 the results of the same ranking exercise for the probit model is substantially different. For example, Sherman Oaks, ranked 15th in the selection model, is ranked 74th in the probit model; Charter Suburban Hospital is ranked 65th in the selection model but 18th in the probit model. Detail of rankings for both models are provided in Appendix A5, Tables A5–A8.

4.4 Specification and robustness

A key assumption in the selection model is that the distances between the patients' homes and the 114 hospitals in the sample constitute variables that may be used to control for the non-random assignment patients to hospitals. Because of the nonlinear relationship between the endogenous variables (hospital choice) in the mortality equation and the instruments, this relationship was modeled explicitly. The posterior mean of the coefficient on distance exceeds zero by more than 100 posterior standard deviations and that on distance squared differs from zero by more than 30 posterior standard deviations (Table 4, bottom panel). The variance of the shock to each probit in the differenced hospital choice equations is 2.0. The variance due to

27

distance and its square, in each equation, is 0.84, evaluated at the posterior means and using the moment matrix of instruments. Thus distance and its square explain about 30% of the variance of the probits. The instruments are indisputably strong, a finding in accord with the literature.[27]

The further assumption that distances from hospitals to patients are uncorrelated with unobserved disease severity cannot be examined so directly. One plausible alternative is that there remain geographic variations in unobserved disease severity after accounting for the observed covariates listed in the first two panels of Table 4. We examined this possibility from three angles. First, in a conventional probit model for mortality using the observed covariates, hospital choice dummies and patient zip code dummies, the zip code dummies are insignificant. Second, the same is true if dummies for nearest hospital replace zip code dummies. In both equations, the coefficients on the hospital choice dummies are jointly significant in the presence of the zip code dummies. Finally, we conducted a more direct examination by retrieving the unobserved disease severity component from the mortality probit equation in each iteration of the MCMC algorithm. In the regression of this component on zip code dummies and the other regressors, the dummies were jointly insignificant in every iteration. All these findings are consistent with the absence of any unobserved geographic component of disease severity.

Given the large number of endogenous variables in the selection model, quite a few assumptions about functional form were required. The dimensionality of the problem is perhaps most evident in the 6,440 potentially independent free parameters in Σ, the prior variance matrix in the multinomial hospital assignment model. The selection model takes the extreme step of assuming that shocks to the probits in this model are iid normal before differencing (Section 2.2, second paragraph). If this assumption is reasonable, then the 113×1 vectors of posterior shocks $\eta_i \left(i=1,\ldots,n\right)$, which may be retrieved in each iteration of the MCMC algorithm, should be consistent with the specification $\Sigma = \mathbf{I}_{J-1} + \mathbf{e}_{J-1}\mathbf{e}'_{J-1}$. If it is not – for example, if patients with certain characteristics all choose from one small group of hospitals – then this will be evidenced by a constructed covariance matrix $\mathbf{S} = \left(n-1\right)^{-1}\sum_{i=1}^{n}\left(\eta_i - \bar{\eta}\right)\left(\eta_i - \bar{\eta}\right)'$ being substantially different from Σ. A conventional goodness of fit test, carried out at the 5% level, rejects the null hypothesis in slightly over half the iterations of the MCMC algorithm. We conclude that there may well be misspecification of the covariance structure in the multinomial hospital assignment

---

[27] See Luft *et al.* (1990) and Gowrisankaran and Town (1999).

covariance matrix, but it is probably not severe. Due to the large number of parameters in Σ, information about the covariance structure beyond the data would be required to deal constructively with this potential misspecification.

The sensitivity of findings to the specification of the prior distribution can be examined in a number of ways. To convey the nature of the sensitivity we set up three further variants of the selection model. Variant A effectively eliminates the instruments, by scaling the prior standard deviations of the coefficient vector α in the multinomial hospital assignment model by the factor 10^{-4}. This variant effectively eliminates the instruments from the entire model, leaving only the functional form to identify the hospital-specific parameters in the mortality equation. Variant B scales the prior standard deviations of α in the original selection model downward by a factor of 5. It does the same for the hospital coefficients β in the mortality probit equation, by taking $0.283/\tau_j^2 \sim \chi^2(5)$ rather than $7.08/\tau_j^2 \sim \chi^2(5)\,(j=p,s,u)$. Variant C is like Variant B except that prior standard deviations are increased by a factor of 5 relative to the base model. Thus, variants B and C provide alternative priors that are plausibly reasonable from the perspective of the subjective prior in the base selection model.

Appendix A6 provides a detailed set of results for each of these prior distributions. As one might expect, coefficients on covariates in the mortality probit equation show very little sensitivity to the choice from among the four prior distributions. The same is true in the hospital choice multinomial probit model, with the obvious exception of prior A. The means square difference between the hospital group quality probit posterior means in the base model and those in variants B and C is about 0.5, whereas it is 0.77 between the base model and variant A. (The same distance measure between the base selection model and the probit model is 0.92.) The probabilities reported in Table 5, for orderings of hospital group quality probits, are similar across the base model and priors B and C, whereas prior variant A (elimination of instruments) produces very different probabilities. We conclude that reasonable variants on the prior produce distinct but insubstantial differences, whereas elimination of the instruments from the model has strong and substantial effects.

5. Conclusion

This study has extended existing econometric methods in order to measure hospital quality using the experience of patients admitted to hospitals in nonrandom fashion. Using discharge records for almost 75,000 older pneumonia patients from 114 hospitals in Los Angeles County, we find strong evidence of differences in quality between hospitals of different size and ownership classifications. The smallest and largest hospitals, as well as private teaching hospitals exhibit higher quality, while public hospitals exhibit lower quality. We also detect substantial differences in quality for a sizable minority of individual hospitals.

As an important by-product, our methods produce information about the hospital admissions process. Patients with greater unobserved severity of illness tend, overall, to be admitted to hospitals of higher quality. Consequently more conventional methods that ignore nonrandom admission, when applied to this data set, tend to lower the inferred quality of good hospitals and raise that of poor ones, relative to our findings. We find that variation across individual hospitals in the unobserved severity of illness is at least as great as variation in quality, and that this variation accounts for most of the large discrepancy between inference about hospital quality in our model and with more conventional methods.

The procedures used here are at the current frontier of intensive computational methods in econometrics. A supercomputer and several days of computing were required to obtain the results reported here. Recent and imminent innovations in numerical methods and computing technology should sharply reduce the real costs of these procedures in the near term. Given the policy importance of assessing quality of care in hospitals, we believe there is a significant return to further investment in these methods and their application to similar questions in health policy and related fields.

Table 1[a]
Frequency and Mortality rates by Age,
Disease Stage, Racial and Sex Categories

Severity and Demographic Categories		Age Categories					Row Totals
		65-69 years	70-74 years	75-79 years	80-84 years	Over 84 years	
Disease Stage	Disease Stage 1.1	8,409 5.01	10,254 5.09	11,524 5.83	11,168 5.82	14,864 10.18	56,217 6.94
	Disease Stage 1.3-2.3	846 5.91	1,021 5.97	1,017 6.88	912 10.09	1,069 10.20	4,865 7.85
	Disease Stage 3.1-3.6	670 12.69	769 12.87	1,018 14.83	973 16.07	1,478 21.99	4,908 16.70
	Disease Stage 3.7	1,350 15.33	1,598 14.77	1,707 16.81	1,381 22.13	1,664 28.18	7,700 19.56
	Disease Stage 3.8	156 45.51	228 42.10	218 44.03	239 56.49	317 53.94	1,158 49.14
Race	White	7,100 7.20	9,301 7.68	10,796 8.75	10,542 10.44	14,256 13.89	51,995 10.10
	Black	1,498 9.74	1,405 8.61	1,295 7.80	1,207 10.60	1,433 13.32	6,919 10.04
	Hispanic	2,013 6.31	2,032 5.41	1,941 6.85	1,978 7.79	2,709 11.04	10,830 7.70
	Asian	794 6.17	1,106 6.06	1,129 6.38	930 8.27	971 11.33	4,990 7.59
	Native American	24 4.17	26 7.69	25 8.00	16 37.50	23 26.09	114 14.91
Sex	Female	5,335 6.61	7,010 6.22	8,116 7.34	7,955 9.25	12,092 13.24	40,899 9.14
	Male	5,703 8.12	6,860 8.42	7,368 9.23	6,718 10.87	7,300 13.51	33,949 10.12
Column Totals		11,429 7.30	13,387 7.31	15,484 8.24	14,673 9.99	19,392 13.34	74,848 9.59

[a]The cell frequency is the top number and the mortality rate is the second number in each cell.

Table 2
Hospital Frequency, Patients Treated, and Mortality
By Hospital Classification

	150 Beds or Less	151-200 Beds	201-300 Beds	Over 300 Beds	Row Totals
Private, Not-for-Profit	9 4,741 9.17	4 2,369 11.11	18 15,526 9.42	19 21,545 9.71	50 44,181 9.62
Private, For-profit	31 9,523 9.32	15 6,627 9.57	7 4,412 10.54	1 973 10.48	54 21,535 9.70
Private Teaching	1 269 6.32	0	0	5 6,802 9.17	6 7,071 9.07
Public	0	0	1 232 8.62	3 1,829 9.57	4 2,061 9.46
Column Totals	41 14,533 9.22	19 8,996 9.97	26 20,170 9.65	28 31,149 9.61	114 74,848 9.59

Note: The first number in each cell is the number of hospitals in that category, the second number is the total number of pneumonia patients discharged from hospitals in that cell, and the third number is the mortality rate (patient weighted) for patients who were discharged from hospitals in that cell.

Table 3
Proportion of Patients in Severity Category
By Hospital Classification

	150 Beds or Less	151-200 Beds	201-300 Beds	Over 300 Beds	Private, Not-for-Profit	Private, For-profit	Teaching	Public	Row Means
Disease Stage 1.1	0.758*	0.751	0.753	0.747	0.751	0.739*	0.789*	0.756	0.751
Disease Stage 1.3-2.3	0.064	0.069	0.063	0.065	0.066	0.069*	0.053*	0.052*	0.065
Disease Stage 3.1-3.6	0.059*	0.064	0.063	0.071*	0.067	0.064	0.058*	0.071	0.066
Disease Stage 3.7	0.104	0.102	0.106	0.100	0.101	0.113*	0.086*	0.094*	0.103
Disease Stage 3.8	0.015	0.014	0.016	0.017	0.015	0.015	0.014	0.026*	0.015

*Significantly different than category mean at the 5% level.

Table 4[a]

Mortality equation parameter posterior means and standard deviations

	Coefficient	Selection model		Probit model	
		$\gamma_j \big/ \left(\delta'\Sigma\delta+1\right)^{1/2}$		γ_j	
Demographic covariates	Age 70-74	-0.008	(0.025)	-0.008	(0.025)
	Age 75-79	0.068	(0.024)	0.069	(0.025)
	Age 80-84	0.187	(0.024)	0.189	(0.025)
	Age > 84	0.372	(0.022)	0.374	(0.023)
	Female	-0.086	(0.013)	-0.087	(0.013)
	Black	-0.022	(0.028)	-0.025	(0.028)
	Hispanic	-0.12	(0.023)	-0.126	(0.023)
	Native	0.158	(0.13)	0.173	(0.131)
	Asian	-0.089	(0.032)	-0.091	(0.031)
	Income	0.260	(0.203)	0.251	(0.192)
	Income^2	-0.032	(0.024)	-0.028	(0.023)
		$\gamma_j \big/ \left(\delta'\Sigma\delta+1\right)^{1/2}$		γ_j	
Disease severity covariates	Emergency admit	0.180	(0.016)	0.180	(0.015)
	Disease stages 1.3-2.3	0.087	(0.027)	0.089	(0.028)
	Disease stages 3.1-3.6	0.492	(0.023)	0.496	(0.024)
	Disease stage 3.7	0.636	(0.019)	0.641	(0.019)
	Disease stage 3.8	1.398	(0.038)	1.411	(0.037)
		q_G		q_G^*	
Negative of hospital group quality probits	150 beds or less	0.032	(0.017)	0.007	(0.012)
	Between 151 and 200 beds	-0.095	(0.030)	-0.030	(0.018)
	Between 201 and 300 beds	-0.030	(0.019)	-0.003	(0.013)
	Over 300 beds	0.051	(0.015)	0.004	(0.012)
	Private, not for profit	0.010	(0.011)	-0.002	(0.009)
	Private, for profit	-0.001	(0.018)	-0.008	(0.009)
	Private Teaching	0.078	(0.041)	0.024	(0.023)
	Public	-0.175	(0.098)	-0.017	(0.043)
		α			
Hospital choice covariates	Distance	-13.61	(0.134)	---	
	Distance^2	12.42	(0.076)	---	
	Distance×Age	-0.45	(0.022)	---	
	Distance×Severity	-0.33	(0.035)	---	
	10^{-5}×Distance×Income	-0.946	(0.280)	---	

[a]Posterior means are accompanied by posterior standard deviations in parentheses. In the case of the hospital group quality probits, the weights are proportional to the number of patients in the sample for each hospital. Distance is measured in hundreds of kilometers. The age variable takes on the value 1 for ages 65-69, 2 for 70-74, 3 for 75-79, 4 for 80-84, and 5 for 85 and above. The severity variable is disease stage (1.1., …, 3.8). Income is in tens of thousands of dollars per year.

Table 5[a]
Posterior Probability Comparisons of Group Hospital Quality Probits, Selection Model

	A. Hospitals grouped by size		
	≤ 150 beds	151-200 beds	201-300 beds
151-200 beds	1.00		
201-300 beds	0.99	0.08	
> 300 beds	0.20	0.00	0.01
	B. Hospitals grouped by ownership classification		
	Private Not-for -profit	Private For-profit	Public
Private For-profit	0.68		
Public	0.96	0.91	
Private Teaching	0.05	0.07	0.01

[a] Entries indicate the posterior probability that the group quality probit q_G in the column classification exceeds the group quality probit q_G in the row classification. Figures are population weighted.

Table 6[a]
Expected Mortality Comparisons, from a base of .10, Selection Model

	A. Hospitals grouped by size			
	≤ 150 beds	151-200 beds	201-300 beds	> 300 beds
≤ 150 beds	0.10	0.125 (0.008)	0.111 (0.005)	0.097 (0.004)
151-200 beds	0.080 (0.006)	0.10	0.090 (0.005)	0.077 (0.006)
201-300 beds	0.090 (0.004)	0.112 (0.008)	0.10	0.087 (0.004)
> 300 beds	0.104 (0.004)	0.129 (0.008)	0.115 (0.005)	0.10
	B. Hospitals grouped by ownership classification			
	Private Not-for-profit	Private For-profit	Public	Private Teaching
Private Not-for-profit	0.10	0.099 (0.004)	0.136 (0.023)	0.087 (0.008)
Private For-profit	0.102 (0.004)	0.10	0.138 (0.022)	0.089 (0.006)
Public	0.074 (0.016)	0.073 (0.014)	0.10	0.064 (0.015)
Private Teaching	0.115 (0.009)	0.113 (0.007)	0.153 (0.028)	0.10

[a] Cells contain expected mortality at hospitals in the column classification, given a patient with an expected mortality of 0.100 at hospitals in the row classification.

Table 7

Relations between hospital quality probits and severity correlations in the sample

A. Variances and correlations of posterior means of q_j, q_j^*, and ρ_j			
q_j	.0207	.741	.517
q_j^*	.0121	.0129	-.160
ρ_j	.0038	-.0009	.0026

(Covariances shown below main diagonal; correlations shown above the main diagonal)

B. OLS regression of ρ_j (posterior means) on q_j (posterior means)

$$\rho_j = .183\ q_j;\ R^2 = .267, s = .043$$
$$(.029)$$

C. OLS regression of q_j^* (posterior means) on q_j and ρ_j (posterior means)

$$q_j^* = .888\ q_j - 1.651\ \rho_j\ ;\ R^2 = .952, s = .025$$
$$(.019)\quad (.054)$$

D. OLS regression of q_j^* (posterior means) on q_j (posterior means)

$$q_j^* = .585\ q_j;\ R^2 = .549, s = .077$$
$$(.050)$$

Table 8

Posterior Distribution of Hospital Quality Probits, Selection Model

|	Hospital name	q_j	Rank		Quartile probabilities			
		Mean	Mean	Median	1st	2nd	3rd	4th
1	ST. JOHNS HOSPITAL AND HEALTH CENTER	0.329	6.3	5	0.997	0.003	0.000	0.000
8	CEDARS SINAI MEDICAL CENTER	0.200	17.8	15	0.861	0.133	0.006	0.000
15	SHERMAN OAKS COMMUNITY HOSPITAL	0.190	21.8	18	0.747	0.201	0.048	0.004
22	KAISER FOUNDATION HOSPITAL - BELLFLOWER	0.151	30.3	22	0.596	0.238	0.133	0.033
29	NU MED REGIONAL MED CENTER WEST VALLEY - ROSCOE	0.091	36.9	36	0.343	0.542	0.109	0.006
36	PRESBYTERIAN INTERCOMMUNITY HOSPITAL	0.064	43.0	42	0.214	0.574	0.203	0.008
43	MIDWAY HOSPITAL MEDICAL CENTER	0.046	47.5	46	0.223	0.451	0.290	0.036
50	NORWALK COMMUNITY HOSPITAL	0.010	55.7	57	0.178	0.327	0.383	0.112
57	LONG BEACH DOCTORS HOSPITAL	-0.005	58.9	61	0.156	0.316	0.370	0.159
64	PACIFIC ALLIANCE MEDICAL CENTER	-0.023	62.5	66	0.106	0.308	0.407	0.180
71	ROBERT F. KENNEDY MEDICAL CENTER	-0.051	68.1	70	0.057	0.263	0.441	0.239
78	WHITTIER HOSPITAL MEDICAL CENTER	-0.071	72.5	76	0.047	0.200	0.432	0.321
85	UCLA MEDICAL CENTER	-0.100	76.9	79	0.024	0.228	0.329	0.419
92	COMMUNITY HOSPITAL OF GARDENA	-0.128	81.7	87	0.034	0.148	0.312	0.506
99	CHARTER COMMUNITY HOSPITAL	-0.168	89.5	93	0.000	0.049	0.330	0.621
107	COAST PLAZA MEDICAL CENTER	-0.201	94.5	100	0.001	0.046	0.208	0.746
114	LOS ANGELES CO. HARBOR/UCLA MEDICAL CENTER	-0.301	104.8	109	0.000	0.006	0.077	0.918

Table 9[a]

Comparison of Selected Hospital Quality Probits, Selection Model

Posterior probability that hospital with rank in row ranks below hospital with rank in column								
	1	15	29	43	57	71	85	99
15	0.885							
29	0.978	0.675						
43	1.000	0.801	0.623					
57	0.993	0.822	0.695	0.557				
71	0.999	0.903	0.791	0.714	0.623			
85	0.997	0.894	0.807	0.696	0.722	0.566		
99	1.000	0.950	0.896	0.862	0.842	0.799	0.814	
114	1.000	0.999	1.000	1.000	0.997	0.986	0.990	0.881

[a]The integer in the first row and column correspond to the first column of Table 8, which indicates the corresponding hospital names.

Figure 1[a]

Scatterplot of selection model hospital mortality means,
versus probit model hospital mortality means
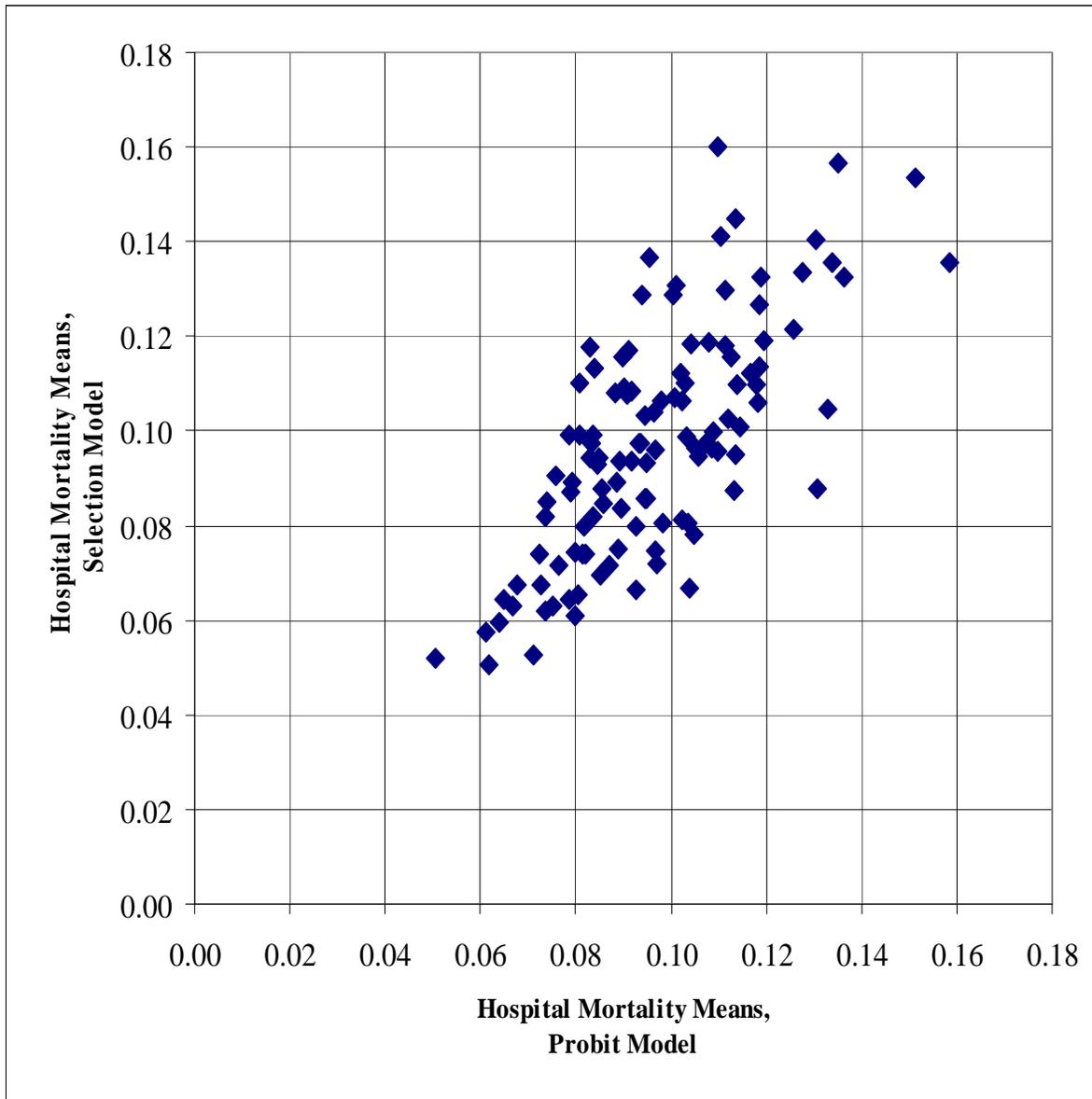
[a]Figure indicates the mean posterior mortality by hospital. The number for each hospital is calculated by integrating over the population distribution of observable characteristics.

References

Albert and Chib, 1993, "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 422, 669-79.

Bernardo and Smith (1994), *Bayesian Theory*. Chichester: John Wiley and Sons.

Burns L, Wholey D., 1992, "The Impact of Physician Characteristics in Conditional Choice Models for Hospital Care," *Journal of Health Economics* 11, 43-62.

Chib, S., and Greenberg, 1996, "Markov Chain Monte Carlo Simulation Methods in Econometrics," *Econometric Theory*, 12, 409-31.

Cutler, D. M., 1995, "The Incidence of Adverse Medical Outcomes Under Prospective Payment," *Econometrica*, 63(1): 29-50.

Fine, J.M., J.D. Scinto, D.H. Galusha, M.K. Petrillo, T.P. Meehan, 1998, "Patient and Hospital Characteristics Associated with Timely Care of Elderly Patients Hospitalized with Pneumonia: Results from the Medicare Quality Indication System Pneumonia Module," *Abstract Book*, Association of Health Services Research, 15.

Gelfand, A.E., and A.F.M. Smith, 1990, "Sampling Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association* 85: 398-409.

Geweke, J., 1997, "Posterior Simulators in Econometrics," in D. Kreps and K.F Wallis (eds.), *Advances in Economics and Econometrics: Theory and Applications*, vol. III. Cambridge: Cambridge University Press, 128-165.

Geweke, J., 1999, "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication," (with discussion and rejoinder) *Econometric Reviews* 18: 1-126

Geweke, J., M. Keane and D. Runkle, 1997, "Statistical Inference in the Multinomial, Multiperiod Probit Model," *Journal of Econometrics*, 80, 125-65.

Gowrisankaran, G., and R.J. Town, 1999, "Estimating the Quality of Care in Hospitals Using Instrumental Variables," *Journal of Health Economics* 18, 747-67.

Gray, B., 1991, *The Profit Motive and Patient Care*, Cambridge, MA: Harvard University Press.

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

Iezzoni, L.I. *et al.*, 1996, "Severity Measurement Methods and Judging Hospital Death Rates for Pneumonia," *Medical Care*, 34(1): 11-28.

Keane, M.P., 1992, "A Note on Identification in the Multinomial Probit Model," *Journal of Business and Economic Statistics*, 10, 193-200.

Keeler, E. B., K.L. Kahn, D. Draper, M.J. Sherwood, L.V. Rubenstein, E.J. Reinisch, J. Kosecoff, and R.H. Brook, 1990, "Changes in Sickness at Admission Following the Introduction of the Prospective Payment System." *Journal of the American Medical Association*, 264, 1962-68.

Keeler, E. B., et al., 1992, "Hospital Characteristics and Quality of Care." *Journal of the American Medical Association*, 268 (13): 1709-14.

Kessler, D. and McClellan, 2000, "Is Hospital Competition Socially Wasteful*?*" *Quarterly Journal of Economics*, forthcoming.

Lohr, K.N., ed., 1990, *Medicare: A Strategy for Quality Assurance*. Volume I. Washington DC: National Academy Press.

Luft, H., *et al.*, 1990, "Does Quality Influence Choice of Hospital?" *Journal of the American Medical Association* 263, 2899-2906.

McClellan M, B. McNeil, and J. Newhouse, 1994, "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?" *Journal of the American Medical Association*. 272, 859-866.

McClellan, M. and H. Noguchi, 1998, "Technological Change in Heart-Disease Treatment: Does High Tech Mean Low Value," *American Economic Review (Papers and Proceedings)*, 88 (2): 90-96.

McClellan M., and D. Staiger, 1999a, "Comparing Hospital Quality at For-Profit and Not-For-Profit Hospitals," NBER Working Paper 7324.

McClellan M., and D. Staiger, 1999b, "The Quality of Health Care Providers," NBER Working Paper 7327.

McGarvey, R. and J. Harper, 1993, "Pneumonia Mortality Reduction and Quality Improvement in a Community Hospital," *Quality Review Bulletin*, 19, 124-30.

National Center for Health Statistics (1996). Available on the Internet at http://www.cdc.gov/nchswww/nchshome.htm.

Pennington, J., 1994, *Respiratory Infections: Diagnosis and Management*, New York: Raven Press, 3rd edition.

Rello, J. and J. Valles, 1998, "Mortality as an Outcome in Hospital-Acquired Pneumonia," Infection Control and Hospital Epidemiology, 19(10):795-7.

Thomas, J.W. and M.L.F. Ashcroft, 1991, "Measuring severity of illness: Six severity systems and their ability to explain cost-variations," *Inquiry*, 28, 39-55.

United States General Accounting Office, 1994, *"Report Cards" Are Useful But Significant Issues Need to be Addressed.* (GAO/HEHS-94-219) Washington, DC: United States General Accounting Office.

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, 326-356.