DOES HIGHER HOSPITAL COST IMPLY HIGHER QUALITY OF CARE?

Gabriel A. Picone, Frank A. Sloan, Shin-Yi Chou, and Donald H. Taylor, Jr.*

Abstract——This study investigates whether higher input use per stay in the hospital (treatment intensity) and longer length of stay improve outcomes of care. We allow for endogeneity of intensity and length of stay by estimating a quasi-maximum-likelihood discrete factor model, where the distribution of the unmeasured variable is modeled using a discrete distribution. Data on elderly persons come from several waves of the National Long-Term Care Survey merged with Medicare claims data for 1984–1995 and the National Death Index. We find that higher intensity improves patient survival and some dimensions of functional status among those who survive.

I. Introduction

A major concern of policymakers in the United States, as in other countries, is the escalating cost of health care. There is a widespread belief that valuable resources are spent on excessive medical care which are not productive at the margin. Supporting this view is the well-known fact that there are large unexplained geographic variations in medical expenditures which are not reflected in appreciable differences in health (e.g., Skinner and Wennberg, 1998).

Among developed countries, the hospital sector in the United States is not particularly high in terms of many indicators, such as bed days, admissions, or hospital days per capita or length of hospital stays (Schieber et al., 1993), but the United States is an outlier in input use per patient day and per stay in the hospital. The additional inputs may be devoted to producing more diagnostic procedures (such as x-rays or lab tests), more therapies (such as surgery or rehabilitation), more individualized supportive care, and/or managerial services that increase the productivity of services that benefit patients. Or the inputs may be devoted to uses that are valued by managers, employees, and physicians on the hospital staff, but are not productive in improved patient care. High input intensity may be interpreted alternatively as slack chosen by hospital administrators (Pope, 1989), an administrative burden imposed on hospitals by insurers and regulatory agencies (Redelmeier & Fuchs, 1993), or as providing higher quality valued by hospital administrators (Newhouse, 1970) and/or consumers (Feldstein, 1977).

Interestingly, there is no empirical evidence on whether higher input intensity has a positive marginal product measured in terms of improved patient health outcomes, irrespective of how such care may be valued by patients and others. Most studies of the marginal benefits of medical care compare the effects of a specific intensive technology with those of a less intensive one on mortality—for example, the benefits of revascularization procedures to treat heart attacks (see, for example, McClellan & Newhouse, 1995, 1997).

This study determines empirically whether higher treatment intensity and longer patient stays following a health shock are productive in terms of improved health outcomes of patients, using a national sample of Medicare beneficiaries. By merging household data with Medicare claims and area data, we are able to control for many more influences than in past studies. We measure treatment intensity as the total cost incurred on behalf of the patient during an index admission. Using a patient-specific measure of services received during a hospital stay as the measure of treatment intensity has the advantage of not being procedure-specific, and patients who received the more intensive technology are more likely to also receive other extra medical care.

Our strategy is to identify a group of persons having similar health conditions and then to compare the effects of different treatment intensity on outcomes. The outcomes we analyzed are subsequent mortality after an index health shock, and the improvement in the quality of life for those who survive. We use Medicare claims data to measure treatment intensity and length of stay during the index admission.

A problem with using nonexperimental data to study the effects of hospital input use on health outcomes is that there are variables unobservable to researchers but observable to the physician and patient (for example, severity of the shock) that affect both health and treatment. This endogeneity problem is avoided in randomized clinical trials (RCTs). However, RCTs are expensive and difficult to implement. They also have limited external validity, and randomization of therapeutic strategies, for which there is some evidence of benefits, raises some ethical issues. In this study, we allow for the endogeneity of treatment intensity by estimating a quasi-maximum-likelihood discrete factor model, where the distribution of the unmeasured variables (such as the severity of the shock) is modeled using a discrete distribution.

We find that higher intensity of hospital care is productive in both improved survival and improved functional status among those who survive. By contrast, longer inpatient stays are not productive in either dimension.

The rest of this study is organized as follows: Section II presents the conceptual framework, section III develops the econometric strategy, section IV describes the data, section V gives the empirical specification, section VI discusses the results, and section VII concludes the paper.

II. Conceptual Framework

The ultimate goal of our analysis is to relate quantity of inputs in the hospital allocated to individual patients, termed intensity, to various health outcomes of these same patients. Three agents affect the final choice of treatment: hospitals, physicians, and patients. Hospitals make decisions about treatment capacity by selecting labor and nonlabor inputs, based in part on anticipated demand responses from patients and their physicians (see, for example, McClellan, 1996). A hospital can affect treatment intensity at its facility by varying amounts of general capacity (such as for beds) as well as capacity for particular types of treatments (such as open-heart surgery units for patients with heart disease). Conditional on a shock having occurred, the physician computes an optimal treatment plan at each hospital in the market and selects the hospital and plan that maximize patient welfare, which in turn reflects the marginal product of medical care, travel time to the facility, and other considerations. Thus, the treatment plan and hospital are selected jointly. If a patient selects a high-intensity hospital, he/she is more likely to receive intensive care.

The treatment plan has two dimensions: treatment intensity and length of stay. Treatment intensity, the quantity of inputs per stay, may be productive to the extent that extra testing results in a quicker and more accurate diagnosis. Following diagnosis, a more aggressive treatment approach may result in a better health outcome. Longer hospital stays have these potential benefits: more time for a more careful diagnosis; more watching and waiting time to look for complications that may arise after the person is admitted; and not discharging the person too early, which might expose the person to an emergency without ready access to specialized personnel. However, longer hospital stays may also reflect complications after the original diagnosis and treatment approach or slack.

The levels of intensity and length of stay that a patient receives depend not only on his or her health but also on characteristics of hospitals in the local market, such as hospital capacity. Although not endogenous to patient and physician at the time of the shock, hospital capacity is plausibly endogenous in our data. One of the determinants of hospital capacity for high-technology services in the area is the cost of capital to the hospital. Although factor markets for hospital investment goods are largely national (or even international), reimbursement for hospital capital expense is local. More specifically, state Blue Cross plans and state Medicaid programs have differed appreciably in their reimbursement of hospital payments for interest and depreciation and for the opportunity cost of equity held by hospitals in the form of plant and equipment (Wedig, Sloan, & Hassan, 1989). Therefore, the hospital cost of capital net of insurers' reimbursements has varied considerably among hospitals. As that net cost decreases, one expects hospitals to undertake more investment. Such investment may be in general hospital capital or be treatment-specific.

Competition among hospitals may affect patterns of hospital care. The underlying mechanisms are complex (Dranove, Shanley, & Simon, 1992; Kessler & McClellan, 2000; Meltzer & Chung, 2001). Since health insurers have switched from cost-based methods of paying hospitals to prospective payment methods, reduced concentration of output among hospitals within local markets has led to decreased hospital cost. Furthermore, competition from health maintenance organizations (HMOs) may force reductions in cost, even for care delivered on a fee-for-service basis. For these reasons, in the equations for intensity and length of stay we included the share of the population enrolled in HMOs and the Herfindahl index of concentration of beds in short-term general hospitals in the area in which the patient lived.

The third equation is for health outcomes. Physicians and patients observe patient health and treatment patterns among hospitals at the time the shock occurs. They select the hospital that yields higher utility net of cost, such as for travel, to the patient. Since Medicare covers hospital care almost in full, choice is not based on hospital prices. If treatment intensity and length of stay are productive, they will have a positive effect on health outcomes, but other factors should affect such outcomes as well, including the person's health before the shock and the severity of the shock. Effects of previous health and the severity of the shock are unclear. For example, the severity of a shock may imply higher treatment intensity because the patient is sicker, but may also cause less treatment intensity because it may be a terminal case; which effect dominates is unclear. Likewise, some longer stays may be productive for reasons given above. On the other hand, adverse events occurring during the hospitalization may prolong stays. Although we want to test the marginal productivity of additional days in the hospital, empirically, the latter source of endogeneity is likely to be important [see for example the evidence in Walter et al. (2001)] to the effect that elderly persons who stay in the hospital longer have higher subsequent rates at mortality at one year following the admission).

Formally, we assume that treatment intensity (TI), length of stay (LS), and health outcomes (H^*) are related through the following equations:

$$TI_i = X'_{1i}\beta_1 + u_{1i},$$

$$LS_i = X'_{2i}\beta_2 + u_{2i},$$

$$H^*_i = X'_{3i}\beta_3 + TI_i\,\delta_1 + LS_i\,\delta_2 + u_{3i}.$$

The main problem when estimating the effects of treatment intensity on health outcomes is that the initial health during the index hospital admission was observable to the patient and his or her physician but not to us (omitted variable bias). We observed certain health characteristics, such as primary and secondary diagnoses, but not the severity of the health shock, which varied appreciably

among patients with the same diagnoses. For example, some stroke victims may have been admitted practically dead and/or in such weakened conditions that they were subject to more shocks during the hospital stay. Others may have had a mild temporary paralysis. We could control for some important aspects of health and functional status before the shock and various demographic characteristics and household income that plausibly affect treatment choices and health outcomes.

III. Econometric Strategy

The models to be estimated have a binary dependent variable and two continuous endogenous explanatory variables. Maximum likelihood and several two-step estimators could have been used to estimate this model. For this analysis, we used a quasi-maximum-likelihood estimator with discrete factor approximations, as well as two-stage least squares instrumental variables (IV) estimators.

Maximum likelihood estimators estimate jointly the likelihood function of observing a given binary outcome for health and continuous outcomes for treatment intensity and length of stay. This method involves calculating a trivariate integral, which is very expensive to calculate. To overcome this problem, economists have relied on numerous two-step estimators that involve estimating ordinary least squares and probit regressions (River & Vuong, 1988). Many economists have criticized this approach. If the error terms are normally distributed, parameter estimates are consistent, but their standard errors may be unduly high, making the estimates very unreliable. Also, there is the issue of identification. In principle, exclusion restrictions are not necessary, but two-step procedures perform poorly without them. Staiger & Stock (1997) showed that IV methods, like two-stage methods, are highly dependent on the quality of instruments.

The quasi-maximum-likelihood estimator with discrete factor approximations used in this study is well suited to correct for endogeneity and selection bias caused by unmeasured explanatory variables in qualitative dependent-variable models [see for example Goldman (1995), Cameron and Heckman (1998), and Cameron and Taber (1998)]. This approach obviates the need to evaluate multivariate normal integrals, and it is less expensive computationally than the regular maximum likelihood estimator. Mroz (1999), in a Monte Carlo study, found that this estimator is more robust to deviations from normality and quality of instruments than two-stage methods.

To apply this method, we assumed that the error terms can be written as $u_1 = \alpha_1 v + \varepsilon_1$, $u_2 = \alpha_2 v + \varepsilon_2$, and $u_3 = \alpha_3 v + \varepsilon_3$, where ε_1, ε_2, and ε_3 are independently normally distributed and v is a univariate random variable with equation-specific factor loading parameters α_1, α_2, and α_3. Thus the joint density function of u_1, u_2, and u_3 conditional on v can be written as

$$f(u_1, u_2, u_3 | v)$$
$$= \frac{1}{\sigma_1} \phi\left(\frac{u_1 - \alpha_1 v}{\sigma_1}\right) \frac{1}{\sigma_2} \phi\left(\frac{u_2 - \alpha_2 v}{\sigma_2}\right) \frac{1}{\sigma_3} \phi\left(\frac{u_3 - \alpha_3 v}{\sigma_3}\right),$$

where σ_1, σ_2, and σ_3 are the standard deviations of ε_1, ε_2, and ε_3.

The discrete factor approximation method assumes that the distribution of v can be approximated by a discrete distribution, where $\Pr(v = \eta_k) = P_k \geq 0$ for $k = 1, \ldots, K$ and $\Sigma_{k=1}^{K} P_k = 1$. The number of steps, K, the location parameters η_k, and the probabilities P_k are called *incidental parameters*; they are estimated jointly with the other parameters describing the model. By letting K become large as the sample becomes large, this method can approximate any smooth distribution function fairly well. Using discrete factor approximations, the likelihood function associated with our model is given by

$$L = \prod_{i=1}^{N} \sum_{k=1}^{K} P_k \left[\frac{1}{\sigma_1} \phi\left(\frac{u_{1i} - \alpha_1 \eta_k}{\sigma_1}\right) \right]$$
$$\times \left[\frac{1}{\sigma_2} \phi\left(\frac{u_{2i} - \alpha_2 \eta_k}{\sigma_2}\right) \right] \Phi(u_{3i} - \alpha_3 \eta_k)^{d_i} [1 - \Phi(u_{3i} - \alpha_3 \eta_k)]^{1-d_i},$$

where d_i is the binary outcome. This likelihood function only contains integrals of univariate normal densities. If the incidental parameters satisfy some regularity conditions, the estimates obtained from this method are consistent (Lindsay, 1983).

Because the number of points K, η_k, and P_k are unknown, the maximization of the above likelihood function is numerically challenging. Heckman and Singer (1984) estimated η_k and P_k directly and used an iterative procedure to maximize the likelihood function with respect to the regular parameters of the model and the incidental parameters η_k and P_k. In this study, we used the method proposed in Mroz (1999). We parameterized η_k and P_k (see appendix), and then we maximized the likelihood function with respect to all the parameters of the model simultaneously (including the parameters describing η_k and P_k). The location and scale of the distribution of v were not identified. Because both the health and the outcome equations contained an intercept and we estimated the factor loading parameters α_1, α_2, and α_3 in our parameterization, η_k was restricted to be between 0 and 1 with $\eta_0 = 0$ and $\eta_K = 1$.

The marginal effects of treatment intensity and length of stay in the health equations were calculated as the derivative of the predicted probabilities with respective to the treatment intensity and length of stay.[1] We computed the marginal effect for each individual and then took the mean of these individual responses.

---

[1] We calculated marginal effects for treatment intensity and length of stay using the following expression:

Ideally, the number of points of support should increase as the sample size increases. Unfortunately, there is no standard theory regarding how to select K in a finite sample. Some studies have fixed the number of points of support in advance (Goldman, 1995); others add points of support until the likelihood fails to improve (Cameron and Taber, 1998; Heckman and Singer, 1984). Using a likelihood function ratio test to add points of support is intuitively appealing, but it has these drawbacks: If the null hypothesis of fewer points of support is correct, the Hessian matrix of the alternative model is singular. Also, the test statistics does not follow an asymptotic *chi*-squared distribution. However, in a Monte Carlo study, Mroz (1999) demonstrated that this test performs fairly well when deciding between small numbers of points of support (one to four). In this study, we used a likelihood ratio test to determine whether additional points of support should be added—up to four, the largest number that we allowed.

Because our likelihood function had multiple local optima, we drew 1,000 bootstrap samples from the data, and used gradient-descent methods on each of the samples to generate 1,000 estimates of the parameters. The estimate with the highest likelihood score was then used as the starting point to estimate the parameters using the whole data set (see Mroz, 1999). The whole process took 54 hours on average per run on a Sun Ultra-1 workstation.

IV. Data

The study sample was drawn from the National Long-Term Care Survey (NLTCS), which is a panel study fielded in 1982, 1984, 1989, and 1994. Overall, 35,845 Medicare beneficiaries were included in the database for at least some time. NLTCS drew its sample from Medicare enrollment records for persons 65 years of age and older. A screener interview was administered to all beneficiaries. Based on responses to the screener, full interviews were conducted with persons who reported having at least one limitation in activities of daily living (ADLs) or in instrumental activities of daily living (IADLs).[2] Respondents lived in the community or in other facilities, most notably in nursing homes. The NLTCS collected detailed information on functional and cognitive status, health conditions, demographic characteristics of the family including potential caregivers, education, race, and income, including sources of income and wealth.

The NLTCS interviews for 1984, 1989, and 1994 were merged with data from other sources. First, data on all Medicare claims, inpatient, outpatient, Part B physician, home health, skilled nursing facility, and hospice from 1982

through 1995 were merged with all individuals screened by NLTCS in any year (Manton, Stallard, & Corder, 1995). Each claim included information on diagnoses and amounts billed and paid by Medicare. Dates of deaths for all NLTCS respondents were verified from Medicare enrollment records and vital records systems for all NLTCS respondents.

For purposes of this analysis, we selected persons who were admitted to hospitals with primary diagnoses of hip fracture, stroke, coronary heart disease, or congestive heart failure for stays of 30 days or less. These are among the most common reasons for hospitalizations of elderly persons in the United States.

We selected the first admission for these conditions that occurred starting in 1984. Because we had Medicare claims data starting in 1982 and the NLTCS asked about conditions during the preceding year, we had a minimum of a three-year lookback period for ascertaining first admissions for a particular condition. A reason for limiting the empirical analysis to first shocks was to limit hospital choice based on health states or unobserved preference differences, because first shocks are more likely to be unanticipated. After the first shock, persons may move to be near facilities offering high-quality care for their conditions.

Persons with very severe shocks died during the index admission. We dropped these cases from the analysis presented here. Thus, our estimates are all conditional on the person having survived the index hospital admission; 9.1% of cases were lost for this reason. The final sample for the mortality analysis consisted of 5,332 persons. The analysis of living arrangements and change in functional status at the next NLTCS interview was based on fewer cases. The difference mostly reflects loss of sample due to deaths between the date of discharge from the hospital and the next NLTCS interview.

V. Empirical Specification

A. Overview

We specified equations for mortality and two additional outcome measures, conditional on the patient's survival to the interview following the shock, and for hospital length of stay and treatment intensity. The dependent variable for the mortality analysis was a binary one, indicating alternatively whether or not the person who was discharged alive survived 6 months, 1 year, or 2 years. Using data from the NLTCS interview following the shock, we specified dependent variables for (1) living arrangement—in the community versus the nursing home, and (2) having the same (or better) functional status at the interview after the shock as at the interview before the shock, versus worse functional status, measured in terms of the same or fewer limitations in ADLs. The ADLs were eating, bathing, dressing, moving around inside, toileting, and getting in and out of bed. The dependent variable for hospital length of stay was the

---

$$\sum_{k=1}^{K} P_k \frac{1}{\sigma_1} \phi\left(\frac{u_1 - \alpha_1 \eta_k}{\sigma_1}\right) \frac{1}{\sigma_2} \phi\left(\frac{u_2 - \alpha_2 \eta_k}{\sigma_2}\right) \phi(u_3 - \alpha_3 \eta_k)\beta_j.$$

[2] ADLs are very personal activities such as eating, dressing, and bathing. IADLs are less personal activities, such as doing laundry.

logarithm of total number of days the person spent in the hospital during the index admission. The dependent variable for treatment intensity was the logarithm of an estimate of the total cost incurred on behalf of the patient during the index admission. Length of stay and treatment intensity were the key explanatory variables in health outcomes analysis (when length of stay and treatment intensity were used as explanatory variables, we did not take logarithms).

B. Measuring Treatment Intensity

Medicare expenditures for inpatient hospital services do not reflect patient-specific treatment, because (1) hospitals are paid by Medicare prospectively on a fixed-price basis, and (2) Medicare adjusts payments for a hospital's being a teaching hospital and for its providing a disproportionate share of care to indigent persons. We used Medicare Cost Reports from the year of the index admissions to calculate the Medicare inpatient cost-to-charge ratio (see for example Lipscomb et al., 1998).[3] This ratio represented a hospital-level measure of Medicare costs-to-charge ratios in the year of an index admission, thus allowing the ratio for a particular hospital to change over time. This scaling factor was applied to actual charges billed by the hospital on each Medicare claim associated with the index inpatient hospitalization during our study period.[4] We added physician payments reimbursed by Medicare Part B during index admission to adjusted charges to obtain total cost of the index admission. All monetarily expressed variables were converted to 1994 dollars using the Consumer Price Index, all items.[5]

C. Equation Specification

In the analysis of mortality, explanatory variables fell into five categories: hospital care (length of stay and treatment intensity); demographic/income; health before the index admission; primary diagnosis at index admission; and other variables.

Demographic variables were age at the date of the index shock, gender, race, and marital status. Marital status and family income were measured at the NLTCS interview prior to the shock.

Included in the health preshock category were a binary variable indicating whether the person had a screener interview only (versus a full interview), two dummy variables for the number of ADL limitations (no ADL limitations, and four or more than four ADL limitations, out of a maximum

possible of six), and whether or not the person was cognitively unaware—all as reported at the NLTCS interview before the shock. To receive a full interview in 1984 and 1989, the person had to have at least one ADL or IADL limitation for at least 90 days before the screener was administered, or have an ADL or IADL limitation expected to last for at least 90 days. In 1994, a relatively small sample of persons who did not have an ADL or IADL limitation were also administered full interviews. Examples of IADL limitations are needing help with laundry, shopping, or balancing one's checkbook. On the whole, persons who were screened but did not have a full interview were healthier than those who received the full interview.

The NLTCS measured cognitive functioning with a ten-question Short Portable Mental Status Questionnaire (Pfeiffer, 1975). We considered a person to be cognitively unaware if the person answered fewer than seven correctly. Questions dealt with orientation in time (what is today's date?) and place (what is the name of this place?) and ability to perform simple calculations (count backward in threes starting with 20).

For primary diagnosis at index admission, we included a risk-adjuster (comorbidity index) (DxCG, 1996; Ellis et al., 1996). The comorbidity index classified patients on the basis of age, sex, and diagnoses contained in the index-admission hospital claims record. Diagnoses other than the primary reason for the index admission were reflected in this comorbidity score, which allowed for comparison of patients with divergent conditions in terms of future expected Medicare-financed resource use. We included binary variables indicating whether the patients were admitted from a nursing home or from the emergency room. The omitted reference category was admission from home.

There was another source-of-admission category: transfer from other hospital. We had no such cases, because we selected cases from the first hospital to which the patient was admitted. If a patient was transferred to a different hospital (117 cases), we added the total cost and length of stay from the second admission to the original admission. We also included binary variables for stroke, coronary heart disease, and congestive heart failure, with hip fractures the omitted disease type.

Hospital characteristics variables were teaching status and number of hospital beds. Market-characteristic variables were HMO market share and percentage of beds in major teaching hospitals. They were defined for the respondent's Primary Sampling Unit (PSU). The PSUs consisted of groups of counties forming metropolitan areas or, for nonmetropolitan PSUs, single counties. A major teaching hospital was defined as a hospital with interns and residents and with a number of interns and residents per bed above the median for such hospitals. We also included time trend and three region dummies for Census areas.

In the equations for living arrangement and changes in functional status, we included lagged dependent variables

---

[3] For index admissions occurring between January 1, 1984, and September 30, 1994, we calculated the Medicare cost-to-charge ratio using Medicare Cost Report data from October 1, 1984 to September 30, 1995 (PPS-II).

[4] Billed charges are distinct from amounts Medicare and many other insurers paid, which were case-based prices or discounted charges.

[5] For physician payments, Medicare claims listed the date of service provision, allowing us to determine if the physician service was provided during the index hospitalization.

measured at the NLTCS interview before the shock as well as the other explanatory variables described above.

In the equations for length of stay and treatment intensity, we included the same explanatory variables for demographic characteristics and income, for health before the index admission, for primary diagnosis at index admission, and for other hospital and market characteristic variables, as in the mortality analysis. In addition, we included explanatory variables for the Herfindahl index based on beds, the Medicare hospital wage index, and a measure of the cost of capital in the state of the index admission in 1984 relative to its cost in New York state in that year (Wedig et al., 1989). The measures of the Herfindahl index and Medicare hospital wage index were for the PSU in which the respondent lived. The cost of capital was measured at the state level. These area variables were excluded from the health equations, for the theoretical reasons discussed in Section II. The NLTCS did not identify the precise location of the person's residence within the PSU, for confidentiality reasons. Thus, patient location could not be used as an IV.

We tested the validity of our exclusion restrictions by using the test statistics NR^2 from a regression of the health outcomes IV regression residuals on the instruments (exclusion restrictions) and exogenous variables, where N denotes the sample size and R^2 the goodness-of-fit statistic (Staiger and Stock, 1997). Test statistics did not suggest the violation of overidentification restrictions. The cost-of-capital, Herfindahl, and wage index variables were excluded from the health outcomes analysis.

VI.   Results

A.   Descriptive Evidence

On average, $5,344 of resources were allocated per patient stay for care of elderly patients admitted for one of the four conditions (1994$), with a standard deviation about equal to this (Table 1). These patients stayed in the hospital almost 10 days on average. Mortality rates were 0.13 at six months, 0.20 at one year, and 0.30 at two years.

The mean age was 78. Most respondents (61%) were women; only 46 percent were married. The mean annual family income was nearly $17,000. Only 14% of sample persons had no ADL limitations at the NLTCS interview before the shock occurred. Nearly 70% of persons had at least four such limitations. About half of the respondents were classified using our criteria as being cognitively unaware at the previous NLTCS interview. Among the four primary diagnoses, coronary heart disease was the most common. Stroke and hip fractures were the least common. Of the hospital stays we studied, 67% occurred between 1984 and 1989; the remainder occurred between 1990 and 1995.

TABLE 1.—MEANS AND STANDARD DEVIATIONS

	Mean	Standard Deviation
Dependent variables:		
6 month mortality	0.126	0.332
1 year mortality	0.198	0.398
2 year mortality	0.297	0.457
Living in the community	0.925	0.264
ADLs improving or staying the same	0.627	0.484
Hospital care		
Treatment intensity (total cost '000)	5.344	5.713
Length of stay	9.937	5.862
Demographic/income:		
Age ('00)	0.784	0.078
Male	0.388	0.487
White	0.895	0.306
Married	0.460	0.498
Income ('0000)	1.682	0.858
Health preindex admission:		
Screener	0.238	0.426
No ADL limitations	0.695	0.461
More than 4 ADL limitations	0.136	0.343
Cognitively unaware	0.511	0.500
Primary diagnosis at index admission:		
Comorbidity index	1.565	1.176
Admitted from nursing home	0.057	0.233
Admitted from emergency room	0.269	0.443
Stroke	0.174	0.379
Coronary heart disease	0.410	0.492
Congestive heart failure	0.250	0.433
Other		
Major teaching hospital	0.165	0.372
No. of beds in hospital ('000)	0.304	0.227
HMO share	0.094	0.147
% of beds in teach. hosp. by PSU	0.085	0.213
Cost-of-capital index	0.139	0.033
Herfindahl index	0.287	0.297
Wage index	1.018	0.200

$N = 5,332$.

B.   Determinants of Treatment Intensity

The younger elderly, males, those who did not get the full NLTCS interview (screener only), those with higher values on the comorbidity index, those who were admitted from a nursing home or from an emergency room as opposed to the reference group (from home), and those who had a primary diagnosis of hip fracture (the omitted reference group) received more intensive hospital care on average (Table 2). Clearly, the relationship between severity of illness and receipt of intensive hospital care is complex. The very old got less care. On the other hand, those with complex diagnoses got more care on average. Intensity was higher at major teaching hospitals and at larger hospitals.

Among our exclusion restrictions, the cost-of-capital index and the wage index had statistically significant negative and positive effects, respectively, on intensity at better than the 5% level. The result for cost of capital suggests that when that cost, and hence the cost of purchasing expensive technology, is higher, hospitals select less intensive treatment methods. The coefficient of the Herfindahl index had a negative effect on intensity, implying higher intensity in areas with greater competition among hospitals; but the coefficient was not statistically significant at conventional

levels. Although this result seems inconsistent with findings that increased competition among hospitals decreases their cost (see for example Meltzer & Chung, 2001), this relationship is more apparent in recent years. Two-third of the admissions in our sample came from the 1980s.

Length of stay largely reflected case mix. A higher cost of capital led to longer stays, suggesting that hospitals compensated in part for lower intensity in higher-capital-cost areas by prolonging stays. Areas with greater bed concentration had longer hospital stays on average. Higher wage rates also were associated with longer stays.

The results for most of the explanatory variables are plausible. Although we only show regressions for treatment intensity and length of stay corresponding to two-year mortality, the results for other health outcomes were similar.

C. Determinants of Mortality

Higher intensity of care improved survival at two years after the health shock occurred (Table 2). Judging by the marginal effects (shown in square brackets), the probability of dying within two years decreased by 0.0043 for each $1,000 increase in the total cost of the stay (treatment intensity). However, longer length of stay led to higher mortality, which suggests that increased stays were unproductive in prolonging survival. The results are for four points of support. Log likelihood ratio tests rejected specifications with one point of support, indicating that our measure of treatment intensity and length of stay were endogenous to mortality.

Many of the other explanatory variables had plausible and statistically significant effects on mortality: age (+); male gender (+); being married (−); no ADL limitations (−); more than four ADL limitations (+); cognitively unaware (+); comorbidity index (+); having been admitted to the hospital from a nursing home (+) or through an emergency room (+); and having a primary diagnosis of stroke, coronary heart disease, or congestive heart failure (+), with elderly admitted to the hospital with hip fracture being the most likely to be alive at two years following the admission date.

Table 3 compares results on treatment intensity by estimator for mortality at six months, one year, and two years following admission. Using probit analysis, the treatment intensity and length of stay have uniformly negative and positive effects on the probability of dying, respectively, with the marginal effects increasing in absolute value as the period for evaluating mortality lengthens. However, considering the sample means (see Table 1), the relative increases in magnitude of effect are considerably less.

Using discrete factor analysis reduces the effects of patterns of care, as judged from the marginal effects. Treatment intensity has a statistically significant effect at better than the 5% level on mortality at one and at two years. The length-of-stay results uniformly indicate that such care is unproductive at the margin. The two-stage IV results essentially suggest that both types of care are unproductive in terms of improving survival.

D. Determinants of Living Arrangements, Cognition, and Functional Status

As with mortality, higher intensity of care during the index hospital stay was productive in making persons more likely to live in the community and have either an improvement or at least no decline in the number of limitations of ADLs (Table 4). The probability of living in the community rather than in a nursing home at the NLTCS interview following the shock increased by 0.003 on average per $1,000 increase in cost. The probability of having the same number or fewer ADL limitations at this interview than at the interview before the shock was raised by 0.002 per $1,000 increase in cost. All comparisons were between the status at the interview after the shock and that at the interview five years earlier and before the shock. The smaller sample sizes than in the mortality analysis reflected loss of sample due to death. Length of stay, however, is unproductive in terms of these outcomes, as it was for survival.

E. Interpreting the Unmeasured Heterogeneity

Table 5 reports the parameters describing the unmeasured heterogeneity: the location and size of the points of support, the standard deviation, and the constants from the binary and continuous equations. We do not report the estimated parameters for the location and size of the points of support (see appendix), but rather the implied locations and probabilities. All the parameters (not reported) describing the location of the points and their weights were significant at the 1% level. A striking result is that the location of the points is very similar among all the specifications and appears to be symmetric.

Based on a likelihood ratio test, we used four points of support in all specifications, which was the maximum number that we allowed. Although the likelihood ratio test suggested using more than four points, we decided to stop at four because additional points created some instability in the other parameters describing the unmeasured heterogeneity. Specifically, many coefficients describing the unmeasured heterogeneity were not significant. This is already observed with the coefficient for the loading factor in the discrete outcome, which is significant with three points but not significant with four, although the coefficient is very similar. It appeared that the method formed spurious points of support beyond some number of points. The coefficients for intensity and length of stay were robust in the range of numbers of points of support that we evaluated.

VII. Discussion

Our results indicate that greater intensity of care during the initial hospitalization for four common health shocks

TABLE 2.—DETERMINANTS OF TREATMENT INTENSITY, LENGTH OF STAY, AND TWO-YEAR MORTALITY

	Treatment Intensity Coeff. (S.E.)	Length of Stay Coeff. (S.E.)	Mortality Coeff. (S.E.) [M.E.]
Hospital care:			
Treatment intensity (total cost '000)	—	—	−0.028[a]
	—	—	(0.006)
	—	—	[−0.0043]
Length of stay	—	—	0.023[a]
	—	—	(0.008)
	—	—	[0.0035]
Demographic/income:			
Age	−0.690[a]	0.173	3.391[a]
	(0.143)	(0.106)	(0.289)
Male	0.019	−0.045[b]	0.331[a]
	(0.024)	(0.017)	(0.046)
White	0.001	−0.043[c]	−0.036
	(0.033)	(0.024)	(0.063)
Married	0.011	−0.012	−0.141[a]
	(0.023)	(0.017)	(0.048)
Income ('0000)	0.025[b]	0.008	0.025
	(0.012)	(0.009)	(0.023)
Health, Preindex admission:			
Screener	0.121[a]	−0.002	0.074
	(0.035)	(0.026)	(0.070)
No ADL limitations	0.043	0.015	−0.305[a]
	(0.030)	(0.022)	(0.056)
More than 4 ADL limitations	0.041	0.000	0.127[c]
	(0.036)	(0.027)	(0.068)
Cognitively unaware	−0.027	−0.066[a]	0.138[a]
	(0.024)	(0.018)	(0.049)
Primary Diagnosis at Index Admission:			
Comorbidity index	0.160[a]	0.121[a]	0.143[a]
	(0.010)	(0.007)	(0.020)
Admitted from nursing home	0.288[a]	0.228[a]	0.399[a]
	(0.048)	(0.036)	(0.088)
Admitted from emergency room	0.145[a]	0.097[a]	0.115[b]
	(0.026)	(0.019)	(0.050)
Stroke	−0.500[a]	−0.181[a]	0.628[a]
	(0.041)	(0.031)	(0.070)
Coronary heart disease	−0.396[a]	−0.412[a]	0.304[a]
	(0.042)	(0.032)	(0.070)
Congestive heart failure	−0.518[a]	−0.322[a]	0.898[a]
	(0.040)	(0.029)	(0.070)
Other:			
Major teaching hospital	0.060[c]	−0.001	−0.090
	(0.032)	(0.024)	0.061
No. of beds in hospital ('000)	0.438[a]	0.064	0.042
	(0.050)	(0.037)	(0.094)
HMO share	−0.129	−0.129[b]	−0.121
	(0.086)	(0.063)	(0.148)
% of beds in teach. hosp.	−0.039	0.009	−0.120
	(0.042)	(0.031)	(0.093)
Cost-of-capital index	−1.321[a]	0.748[a]	—
	(0.342)	(0.253)	—
Herfindahl index	−0.061	−0.073[b]	—
	(0.043)	(0.032)	—
Wage index	0.546[a]	0.114[b]	—
	(0.067)	(0.050)	—
Time trend	0.482[a]	−0.352[a]	−0.250[a]
	(0.046)	(0.034)	(0.089)
Constant	−0.075	1.275[a]	−4.057[a]
	(0.148)	(0.110)	(0.271)
No. of points of support of unmeasured heterogeneity	4		
N	5,332		
No. of deaths	1,583		
Log likelihood	−10652		

We also included three region dummies for Census Areas.
Treatment intensity and length of stay in logs when used as dependent variable.
Marginal effects are for a $1,000 increase in total cost.
[a]Significant at the 1% level, two-tail test.
[b]Significant at the 5% level, two-tail test.
[c]Significant at the 10% level, two-tail test.

TABLE 3.—THE EFFECT OF TREATMENT INTENSITY AND LENGTH OF STAY FOR THE INDEX ADMISSION ON HEALTH OUTCOMES

Statistic	6 Month Coeff. (S.E.) [M.E.]	1 Year Coeff. (S.E.) [M.E.]	2 Year Coeff. (S.E.) [M.E.]
Probit			
Treatment intensity	−0.004	−0.013[b]	−0.025[a]
('000)	(0.006)	(0.006)	(0.005)
	[−0.00065]	[−0.0032]	[−0.0084]
Length of stay	0.017[a]	0.023[a]	0.031[a]
	(0.005)	(0.005)	(0.005)
	[0.0031]	[0.0057]	[0.0101]
Discrete Factor Quasi Maximum Likelihood			
Treatment intensity	−0.006	−0.015[b]	−0.028[a]
('000)	(0.007)	(0.006)	(0.006)
	[−0.00064]	[−0.0020]	[−0.0043]
Length of stay	0.009	0.016[c]	0.023[a]
	(0.009)	(0.008)	(0.008)
	[0.00089]	[0.0020]	[0.0035]
Two-Stage Instrumental Variables			
Treatment intensity	0.026[c]	0.024	0.003
('000)	(0.013)	(0.016)	(0.018)
Length of stay	−0.010	0.008	0.032[c]
	(0.012)	(0.015)	(0.016)

Marginal effects are for a $1,000 increase in total cost.
[a]Significant at the 1% level, two-tail test.
[b]Significant at the 5% level, two-tail test.
[c]Significant at the 10% level, two-tail test.

affecting elderly persons was productive in better health outcomes. By linking the total cost of inpatient hospital care to such outcomes as mortality, living arrangements, and functional status, our study has expanded the scope of the hospital quality-of-care literature.

In contrast to treatment intensity, we did not find that length of stay was productive in any of the outcomes we measured. One interpretation is that longer stays are not productive. Of course longer stays may have other benefits that we did not measure, such as improving the psychological well-being of hospitalized persons and/or their families and convenience for families. An alternative interpretation is that longer lengths of stays are productive but we fail to measure the true relationship. Possibly even in the discrete factor analysis we could not take account of exogenous complications that arose during the hospital stay. To the extent that this is so, including the length of stay improved the precision of the estimated treatment intensity effects.

The effect of treatment intensity was larger and more distinct for more distant outcomes than for mortality at six months. In one study with a very large sample, research on mortality following heart attacks found no diminution of effect of hospital quality at two years (Allison et al., 2000). Another investigation found no attenuation in the effects of the presence of sophisticated hospital technologies at one year after admission versus at 30 days (Peterson et al., 2000). Like us, these authors did not assess precisely which aspect of quality caused improved survival.

Cost is a summary measure of input intensity. Underlying it are many treatment decisions. For hip fracture, most patients received surgery to repair their hip, and variation in medical intensity was most likely to be driven by comorbidity conditions (such as problems in medication management) that require medical support and that, if untreated, may cause death or added disability. For coronary heart disease and stroke, there are multiple therapeutic paths that can be taken, which differ in intensity. Treatment of a heart attack, for example, can range from pharmaceutical treatment with a simple agent, such as aspirin or beta-blocker drugs, to invasive diagnostic and therapeutic cardiac catheterization procedures or angioplasty, to coronary artery bypass graft surgery. There is evidence that the increased use of such invasive procedures increased the quality of life at 1 year but not 30 days in a U.S.–Canadian comparative study (Mark, Naylor, & Hlatky, 1994). Likewise, congestive heart failure may be treated with pharmaceuticals such as ACE inhibitors, but surgery is not commonly used in heart-failure management programs designed to reduce high readmission rates (for example, Fonarow, Stevenson, & Walden, 1997). Hospitalizations, which can be of varying intensity, are mainly to deal with exacerbations of this disease in the elderly. Following stroke, diagnostic imaging is particularly important in diagnosing the type of stroke, which guides treatment options.

We note two final caveats. First, we excluded persons who died during the index admission, because the stays

TABLE 4.—THE EFFECT OF TREATMENT INTENSITY FOR THE INDEX ADMISSION ON LIVING ARRANGEMENTS AND FUNCTIONAL STATUS

	Probability of Living in the Community: Coeff. (S.E.) [M.E.]	Probability of ADLs Improving or Staying the Same: Coeff. (S.E.) [M.E.]
Hospital care:		
Treatment intensity	0.026[b]	0.029[a]
('0000)	(0.011)	(0.009)
	[0.0030]	[0.0017]
Length of stay	−0.016	−0.430[a]
	(0.013)	(0.134)
	[−0.0044]	[−0.011]
Demographic/income:		
Age	−3.787[a]	−3.815[a]
	(0.485)	(0.460)
Male	0.092	0.210[a]
	(0.078)	(0.067)
White	−0.074	0.102
	(0.105)	(0.096)
Married	0.258[a]	0.008
	(0.077)	(0.067)
Income ('0000)	0.045	0.043
	(0.037)	(0.037)
Health, preindex admission:		
Screener	0.264[a]	0.948[a]
	(0.099)	(0.088)
Lagged dependent variable	1.434[a]	0.183[a]
	(0.119)	(0.021)
Cognitively unaware	−0.025	0.226[a]
	(0.072)	(0.068)
Primary diagnosis at index admission:		
Comorbidity index	−0.052	−0.041
	(0.037)	(0.033)
Admitted from nursing home	−0.587[a]	−0.238[c]
	(0.124)	(0.135)
Admitted from emergency room	−0.199[a]	−0.178[b]
	(0.076)	(0.070)
Stroke	−0.153	−0.158
	(0.098)	(0.098)
Coronary heart disease	0.480[a]	0.467[a]
	(0.101)	(0.095)
Congestive heart failure	0.349[a]	0.196[b]
	(0.104)	(0.099)
Other:		
Major teaching hospital	0.191[b]	0.123
	(0.093)	(0.081)
No. of beds in hospital ('000)	−0.015	0.056
	(0.149)	(0.134)
HMO share	−0.446[b]	0.056
	(0.218)	(0.203)
% of beds in teach. hosp.	0.043	−0.060
	(0.144)	(0.127)
Time trend	0.007	−0.047[a]
	(0.014)	(0.013)
Constant	2.437[a]	2.851[a]
	(0.448)	(0.399)
No. of points of support of unmeasured heterogeneity	4	4
N	3,076	2,586
No. in the community	2,845	
No. having better or the same ADLs		1,621
Log likelihood	−5536	−5219

Marginal effects are for a $1,000 increase in total cost.
[a]Significant at the 1% level, two-tail test.
[b]Significant at the 5% level, two-tail test.
[c]Significant at the 10% level, two-tail test.

DOES HIGHER HOSPITAL COST IMPLY HIGHER QUALITY OF CARE? 61

TABLE 5.—UNOBSERVED HETEROGENEITY RESULTS

	Mortality					
	6 Month		1 Year		2 Year	
	Mass Point	Prob. Weight	Mass Point	Prob. Weight	Mass Point	Prob. Weight
1st support	0.000	0.131	0.000	0.131	0.000	0.131
2nd support	0.336	0.387	0.336	0.387	0.336	0.387
3rd support	0.646	0.352	0.647	0.352	0.647	0.352
4th support	1.000	0.129	1.000	0.130	1.000	0.130
	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
Std. Dev. (2)*	0.377[a]	0.006	0.378[a]	0.006	0.378[a]	0.006
Std. dev. (3)	0.267[a]	0.005	0.266[a]	0.005	0.266[a]	0.005
Factor loading (1)	0.236	0.211	0.201	0.192	0.221	0.177
Factor loading (2)	2.234[a]	0.032	2.233[a]	0.032	2.232[a]	0.032
Factor loading (3)	1.681[a]	0.021	1.680[a]	0.021	1.680[a]	0.021
Corr. coeff. (1–2)	0.049		0.042		0.046	
Corr. coeff. (1–3)	0.046		0.040		0.043	
Corr. coeff. (2–3)	0.495		0.495		0.494	

	Other Health Outcomes			
	Probability of Living in the Community		Probability of ADLs Improving or Staying the Same	
	Mass Point	Prob. Weight	Mass Point	Prob. Weight
1st support	0.000	0.121	0.000	0.116
2nd support	0.329	0.387	0.329	0.376
3rd support	0.631	0.365	0.632	0.364
4th support	1.000	0.127	1.000	0.143
	Coeff.	Std. Dev.	Coeff.	Std. Dev.
Std. dev. (2)	0.381[a]	0.008	0.387[a]	0.009
Std. dev. (3)	0.258[a]	0.006	0.259[a]	0.007
Factor loading (1)	0.054	0.313	0.355	0.300
Factor loading (2)	2.244[a]	0.045	2.231[a]	0.054
Factor loading (3)	1.702[a]	0.029	1.688[a]	0.033
Corr. coeff. (1–2)	0.011		0.073	
Corr. coeff. (1–3)	0.010		0.070	
Corr. coeff. (2–3)	0.491		0.495	

[a]Significant at the 1% level, two-tail test.
[b]Significant at the 5% level, two-tail test.
[c]Significant at the 10% level, two-tail test.
*(1) = health outcome, (2) = treatment intensity, and (3) = length of stay.

were censored by death. Such decedents constituted less than 10% of persons admitted for the four primary diagnoses we analyzed. Although their stays were about 2 days shorter on average than those who survived the index admission, the total expenditure on behalf of such persons was about 10% higher. It is plausible that for the majority of these decedents, the higher intensity was not at all productive in improved outcomes. But since these persons represented a small fraction of the admissions and of total resources, even considering such unproductive care would have modified only slightly our conclusion that higher intensity was productive overall.

Second, it is possible that higher input intensity during the index admission is positively correlated with input use after the patient is discharged from the hospital, and the latter contributes in part to the improved outcomes following the major health shocks. To investigate this possibility, we computed simple correlations between treatment intensity and Medicare expenditures on care during the first six months following the health shock other than care during the index admission. Such expenditures included the cost of readmissions to hospitals. The correlation was 0.08. The corresponding correlation between length of hospital stay and these downstream expenditures was 0.04. Though positive, these correlations are low, suggesting that the bias from this source is at most minor.

Clearly, treatment choices following major shocks have implications for patient health. By documenting that greater intensity of treatment results in better outcomes, we have demonstrated that this type of medical care is productive, at least for the Medicare program. This conclusion is of definite relevance to policymakers who make allocative decisions about public medical programs.

REFERENCES

Allison, Jeroan J., Catarina I. Kiefe, Norman Weissman, J. G. Canto, S. D. Person, O. D. Williams, and R. M. Centor, "Relationship of Hospital Teaching Status with Quality of Care and Mortality for Medicare Patients with Acute MI," *Journal of the American Medical Association* 284:10 (2000), 1256–1262.

Cameron, Stephen V., and James J. Heckman, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts," *Journal of Political Economy* 106 (1998), 262–233.

Cameron, Stephen, and Christopher Taber, "Discount Rate Bias in the Returns to Schooling," Northwestern University manuscript (1998).

Dranove, David, Mark Shanley, and Carol Simon, "Is Hospital Competition Wasteful?" *RAND Journal of Economics* 23 (1992), 247–261.

DxCG, Inc., *DxCG Software Version 02e (Medicare) Program Documentation* (Waltham, MA: DxCG, Inc., 1996).

Ellis, Randall P., Gregory C. Pope, Lisa I. Iezzoni, J. Z. Ayanion, D. W. Bates, H. Burstin, and A. S. Ash, "Diagnosis-Based Risk Adjustment for Medicare Capitation Payments," *Health Care Financing Review* 17 (1996), 101–128.

Feldstein, Martin S., "Quality Change and the Demand for Hospital Care," *Econometrica* 45:7 (1977), 1681–1702.

Fonarow, Gregg C., Lynne W. Stevenson, and Julie A. Walden, "Impact of a Comprehensive Heart Failure Management Program on Hospital Readmission and Functional Status of Patients with Advanced Health Failure," *Journal of the American College of Cardiology* 30 (1997), 725–732.

Goldman, Dana, "Managed Care as a Public Cost-Containment Mechanism," *RAND Journal of Economics* 26 (1995), 277–295.

Heckman, James, and Burton Singer, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica* 52 (1984), 271–320.

Kessler, Daniel P., and Mark B. McClellan, "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics* 115:2 (2000), 577–615.

Lindsay, Bruce, "The Geometry of Mixture Likelihoods: A General Theory," *Annals of Statistics* 11 (1983), 86–94.

Lipscomb, Joseph, Marek Ancukieswicz, Giovanni Parmigiani, Vic Hasselblad, Greg Samsa, and David Matchar, "Predicting the Cost of Illness: A Comparison of Alternative Models Applied to Stroke," *Medical Decision Making* 18:Suppl. (1998), S39–S56.

Manton, Kenneth G., Eric Stallard, and Larry Corder, "Changes in Morbidity and Chronic Disability in the U.S. Elderly Populations: Evidence from 1982, 1984, and 1989 National Long Term Care Surveys," *Journal of Gerontology and Social Sciences* 50B (1995), 194–204.

Mark, Daniel, David Naylor, and Mark Hlatky, "Use of Medical Resources and Quality of Life after Acute Myocardial Infarction in Canada and the United States," *New England Journal of Medicine* 328 (1994), 779–784.

McClellan, Mark B., Medicare Reimbursement and Hospital Cost Growth (pp. 149–184), in David A. Wise (Ed.), *Advances in the Economics of Aging* (Chicago: University of Chicago Press, 1996).

McClellan, Mark, and Joseph Newhouse, "The Marginal Benefits of Medical Treatment Intensity: Acute Myocardial Infarction in the Elderly," NBER working paper (1995).

—— "The Marginal Cost-effectiveness of Medical Technology: A Panel Instrument-variables Approach," *Journal of Econometrics* 77 (1997), 39–64.

Meltzer, David, and Jeanette Chung, "Effects of Competition under Prospective Payment on Hospital Costs among High and Low Cost Admissions: Evidence from California, 1983–1993," NBER working paper 8069 (2001).

Mroz, Thomas, "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome," *Journal of Econometrics* 72:2 (2001), 233–274.

Newhouse, Joseph P., "Toward a Theory of Nonprofit Institutions: An Economics Model of a Hospital," *American Economic Review* 60:1 (1970), 64–74.

Petersen, Laura A., Sharon-Lise T. Normand, Jennifer Daley, and Barbara J. McNeil, "Outcome of Myocardial Infarction in Veterans Health Administration Patients as Compared with Medicare Patients," *New England Journal of Medicine* 343:26 (2000), 1934–1941.

Pfeiffer, Eric, "A Short Portable Mental Status Questionnaire for the Assessment of Organic Brain Deficit in Elderly Patients," *Journal of American Geriatric Society* 23 (1975), 433–441.

Pope, Gregory C., "Hospital Nonprice Competition and Medicare Reimbursement Policy," *Journal of Health Economics* 8 (1989), 147–172.

Redelmeier, Donald A., and Victor R. Fuchs, "Hospital Expenditures in the United States and Canada," *New England Journal of Medicine,* 328 (1993), 772–778.

River, Douglas, and Quang Vuong, "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models," *Journal of Econometrics* 39 (1988), 347–366.

Schieber, G. J., Jean-Pierre Poullier, and Leslie M. Greenwald, "Health Spending, Delivery, and Outcomes in OECD Countries," *Health Affairs* (Summer 1993), 120–129.

Skinner, Jonathan, and John E. Wennberg, "How Much Money is Enough? Efficiency and Medicare Spending in the Last Six Months of Life," NBER working paper 6513 (1998).

Staiger, Douglas, and James Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65 (1997), 517–536.

Walter, Louise C., Richard J. Brand, Steven R. Counsell, Robert M. Palmer, C. Seth Landefeld, Richard H. Fortinsky, and Kenneth E. Covinsky, "Development and Validation of a Prognostic Index for 1-Year Mortality in Older Adults after Hospitalization," *Journal of the American Medical Association* 285:23 (2001), 2987–2994.

Wedig, Gerard, Frank Sloan, and Mahmud Hassan, "Hospital Investment Decisions and the Cost of Capital," *Journal of Business* 62 (1989), 517–536.

APPENDIX

Parameterization for the Points of Support

η_k was restricted to be between 0 and 1 with $\eta_0 = 0$ and $\eta_K = 1$. The other $K - 2$ points of support are given by

$$\eta_k = \frac{\exp(\theta_{1k})}{1 + \exp(\theta_{1k})}, \quad k = 2, \ldots, k - 1.$$

Parameterization for the Probabilities

Let $P_k = \tau_k / \sum_{k'=1}^{k} \tau_{k'}$, where

$$\tau_k = \begin{cases} 1 + \sin(3\pi/2 + \theta_{2k}) & \text{for} \quad k = 1, \ldots, K - 1, \\ 1 + \sin(\sum_{k'=1}^{K} \theta_{2k'}) & \text{for} \quad k' = K. \end{cases}$$