HC
7/26/07
10:05 AM

The Cost of Quality in Hospitals: A Revealed-Preference Approach*

John A. Romley
RAND Corporation

Dana Goldman
RAND Corporation and NBER

June 30, 2007

Abstract

The relationship between quality and costs among hospitals is important yet not well understood. Quality is difficult to measure and potentially confounded with productivity. We infer quality at hospitals in the Los Angeles area from the revealed preference of coronary-care patients. We then investigate the joint contribution of quality and productivity to hospital costs. Appealing to a model of hospital behavior and our evidence concerning hospital demand, we decompose this contribution. A one standard-deviation increase in a hospital's quality leads to an increase in its costs on the order of thirty percent. We also find that quality is positively correlated with productivity. An analysis that ignores this relationship would substantially understate the cost of quality.

1 Introduction

In recent years there have been prominent calls for quality improvement within the health-care sector. The Institute of Medicine's report *Crossing*

*Corresponding author: romley@rand.org. We wish to thank Jack Clift and Ying Liu for research assistance, Dan Ackerberg, Anil Bamezai, Melinda Beeuwkes Buntin, Steve Garber, Geoff Joyce, Darius Lakdawalla, Raphael Thomadsen, and Bill Vogt for helpful discussions, and Glenn Melnick and Bob Reddick for sharing their data. The financial support of the Bing Center for Health Economics is gratefully acknowledged.

the Quality Chasm garnered national attention and spawned large research efforts devoted to measuring and improving quality (of Medicine (2001)). Productivity is also a significant concern. Zuckerman et al. (1994)'s stochastic-frontier analysis of the performance of U.S. hospitals in the mid-1980s attributes nearly fourteen percent of total costs to inefficient behavior; such "inefficiency" is observationally equivalent to efficient behavior among firms with heterogeneous productivity (Stigler (1976); Biesebroeck (2004)). Costs may also be influenced by quality, and the stakes are high. Expenditures on hospital care alone totaled $611 billion in 2005 (DummyCite (00))

Despite its importance, the relationship between quality , productivity and costs in the hospital-care industry is not well understood.[1] Carey (2000) find decreasing economies of hospital scale under OLS but increasing returns under fixed-effects and correlated random-effects specifications. In light of strong evidence (described in the next section) that quality influences hospital choice by patients, this pattern is consistent with unobserved but costly quality. The author also remarks that other unobserved factors, such as managerial competence, may be important. By contrast, when Zuckerman et al. (1994) account for various measures of clinical quality, the share of costs attributable to inefficiency is unaffected, suggesting that quality is not costly.

Yet hospital quality is difficult to measure, potentially leading to attenuation bias in estimates of the cost of quality (Newhouse (1994)). A researcher is likely to observe some aspects of the hospital experience (e.g., the attentiveness of nurses to patient comfort) imperfectly at best. Carey and Burgess, Jr. (1999)'s results are consistent with measurement error in hospital quality. These authors find, though, that hospital costs *decrease* in quality.[2]

Gertler and Waldman (1992) deal with unobservable quality in costs by specifying a reduced-form model for a firm's choice of quality. The behavior of consumers is essential to identification: the taste for quality facing each

[1] Ackerberg et al. (2006) consider the relationship between productivity and costs among alcohol-treatment providers but do not directly analyze the cost of quality in this setting.

[2] Carey and Burgess, Jr. (1999) analyze Veterans Administration hospitals over the period 1989-1993. Their notion of quality is clinical as well as managerial, with measures including risk-adjusted mortality rates, readmission rates, and rates of failure to follow up with patients in an ambulatory setting post-discharge. For each measure, a lower value corresponds to higher quality. One-standard-deviation increases in quality (i.e., decreases in the measures) are predicted to decrease hospital costs in the range of three to six percent. The authors explain that mortality and readmission may reflect unmeasured case severity; failure to follow up may be a consequence of inadequate coordination of care and poor management more generally.

firm is assumed to be exogenous, and variation across firms in the strength of these tastes induces variation in the marginal value of quality and thus firms' optimal qualities. The authors estimate their models of cost and quality choice in the nursing-care industry and find that quality is costly, with a 1.3% increase in quality leading to a 10% increase in nursing-home costs.

We would argue that the relationship between a hospital's productivity and its quality may further confound estimates of the cost of quality. Economists have long recognized that a firm's levels of inputs and outputs may depend on productivity (Marschak and Andrews, Jr. (1944); Nerlove (1965)). When higher productivity lowers the marginal cost of quality or quantity, a hospital chooses to supply more quality. High-quality hospitals are then relatively low-cost, and vice versa. If a researcher cannot observe productivity, quality again appears to be less costly than is truly the case.

This paper aims to clarify the relationship among hospital cost, quality and productivity and, in particular, to measure the cost of quality in the hospital-care industry. Our approach supposes that hospital patients are—if not perfectly informed about hospital quality—at least better informed than researchers. We can then infer quality from the revealed preference of patients for hospitals.

We also rely on choice behavior to distinguish the contribution of quality to hospital costs from that of productivity. Their joint contribution can be recovered under certain conditions from longitudinal data on hospital costs. The identification strategy is then similar to Gertler and Waldman (1992)'s. However, we estimate firm demand rather than quality choice.[3] We also treat productivity as heterogeneous across firms.

While our particular application is to coronary care among Medicare patients in the Los Angeles area, we believe that our analytical framework can contribute to a better understanding of the cost structures of firms in differentiated-products industries more generally.

To preview our current results, the cost of hospital quality seems to be substantial. A one-standard deviation increase in quality at an average hospital would increase costs on the order of thirty percent, a result that is robust to alternative specifications. We also find that quality tends to

[3] In placing less structure on quality choice, the current analysis is less likely to be confounded by misspecification. We cannot, however, assess the distribution of costs in a counterfactual world with no productivity differences. In the future, we intend to assess the impact of heterogeneity in productivity on costs.

increase in the responsiveness of hospital demand to quality and that quality and productivity are positively correlated, consistent with our model of quality choice among hospitals. If we ignored this relationship between quality and productivity, a one-standard deviation increase in quality would appear to increase costs by only eight percent.

The remainder of this paper is organized as follows. The next section presents our framework for analyzing hospital quality, productivity and cost. Section 3 describes the empirical analysis. Section 4 reviews our findings. Section 5 concludes, emphasizing the limitations and implications of the current analysis.

2 A Framework for Analyzing Hospital Quality, Productivity and Cost

Our objective is to recover the hospital cost function $C(Y_1, \mathbf{Q}_1, \ldots, Y_n, \mathbf{Q}_n, A, \mathbf{W})$. Hospitals typically produce multiple outputs $Y_1, \ldots, Y_n$. The qualities (or attributes) of an output (denoted by the vector $\mathbf{Q}_i$) differ among hospitals—one hospital may deliver relatively advanced care (cardiac catheterization), while another offers relatively pleasing amenities (such as a high ratio of nurses to beds). Costs depend on the scale, scope and qualities of a hospital's outputs. Costs also depend on the hospital's productivity (denoted A) as well as factor prices $\mathbf{W}$.

Researchers confront two challenges in investigating the cost of quality in this complex industry. First, quality must be measured. We use existing methods for analyzing consumer choices among differentiated products to infer quality. Second, a hospital may exercise its discretion over the scale, scope and qualities of its outputs according to factors that researchers do not observe, particularly productivity. We motivate this concern with a model of hospital behavior and explain an identification strategy that exploits heterogeneity in hospital demand.

2.1 Quality and Choice Behavior

We assume that each patient chooses the hospital that maximizes her utility. Empirical evidence is consistent with the view that patients value a hospital's quality as well its proximity to home and choose accordingly. For a variety of medical diagnoses and procedures, Luft et al. (1990) found that patients in three California hospital markets in 1983 were more likely to receive care at high-quality hospitals, as measured by health outcomes (mortality and complications) and other clinically oriented indicators. The authors suggested that lay referral networks may have been useful sources of information about quality during the period studied, when hospital-level outcomes were not publicly reported. Patients were also more likely to receive care at hospitals relatively close to home. More recently, Tay (2003) has found that quality and distance from home are related to the hospitals at which patients receive care for acute myocardial infarction (AMI), or heart attack.[4] Prompt transport to a hospital is vital for this medical condition. Nevertheless, AMI patients are apparently willing to travel for both "ordinary service" quality (such as nurses per bed) and "high-tech service" quality (such as the availability of a cardiac catheterization laboratory). We also analyze the choices of AMI patients, as well as a group of coronary-care patients who plausibly exercise a freer choice of hospital than do AMI patients. We emphasize, however, that patients need not choose hospitals free of any constraints.[5]

Patient utility U_{ih} is comprised of systematic and idiosyncratic utility, denoted $\overline{U}_{ih}$ and ϵ_{ih}, respectively:

$$U_{ih} = \overline{U}_{ih} + \epsilon_{ih} = \beta_{d,i} D_{ih} + \beta_{p,i} P_{ih} + \boldsymbol{\beta}_{\mathbf{x},i} \mathbf{X}_h + \boldsymbol{\beta}_{q,i} Q_h + \epsilon_{ih}, \tag{1}$$

in which D_{ih} is the distance between patient i's home and hospital h, P_{ih} is the patient's price for the hospital's care, $\mathbf{X}_h$ is a vector of additional hospital attributes that the researcher observes, and Q_h is an unobserved quality index. In this framework, the taste for unobserved quality cannot be disentangled from its level. To appreciate this observation, note that utility is identical for $\boldsymbol{\beta}_{q,i} \cdot Q_h$ and $\left(\boldsymbol{\beta}_{q,i} / \rho\right) \cdot (\rho Q_h)$. Hence the taste parameter $\boldsymbol{\beta}_{q,i}$ must be normalized for some patient.

The potential variability of tastes across patients is important for several reasons. First, the variability of tastes can help clarify the identification

[4] Town and Vistnes (2001), Gaynor and Vogt (2003) and Geweke et al. (2003) present further evidence that patients value quality and proximity to home.

[5] For example, Burns and Wholey (1992) find that the distance between a hospital and the office of the patient's physician is as influential as the distance between the hospital and the patient's residence, again for a variety of diagnoses (AMI included) and procedures. As others have recognized (Tay (2003)), a patient may communicate her preferences to her physician and/or choose her physician on the basis of admitting privileges or physician quality (which may be correlated with hospital quality.)

of hospital demand. Suppose that tastes for the jth hospital attribute depend on observed and unobserved patient characteristics, denoted $\mathbf{K}_i$ and ν_{ij}, respectively, as follows: $\beta_{j,i} = \mathbf{\Pi}_j \mathbf{K}_i + \nu_{j,i}$ For concreteness, consider the plausible assumption that the taste for unobserved quality depends on observed comorbidities and unobserved health status. Then equation 1 becomes:

$$U_{ih} = \beta_{d,i} D_{ih} + \beta_{p,i} P_{ih} + \boldsymbol{\beta}_{\mathbf{x},i} \mathbf{X}_h + \mathbf{\Pi}_q \mathbf{K}_i Q_h + \left\{ \nu_{q,i} Q_h + \epsilon_{ih} \right\}, \tag{2}$$

in which the model's composite disturbance appears in braces. In order to recover demand, all observables in equation 2 must be distributed independently of the interacted taste for and level of unobserved quality. This would not be the case if unobservably healthy or sick patients chose where to live based on the quality of hospitals. An assumption that unobserved tastes are distributed with mean zero and independently of all observables is sufficient for identification of the demand parameters. Ideally, this assumption could be empirically assessed, yet we are not aware of a means of doing so. In our opinion, this assumption, which is typically maintained in studies of hospital choice, is not unreasonable.

Second, variability in the taste for unobserved quality is one source of heterogeneity in the responsiveness of hospital demand to quality. We assume in our empirical analysis in section **??** that all potential patients elect to receive care at some hospital; we also assume that idiosyncratic tastes are distributed type-1 extreme-valued. Then the probability that the ith patient chooses the jth hospital takes the following form:

$$\pi_{ih} \equiv \Pr\left(U_{ih} > U_{ih'} \forall h'\right) = e^{\overline{U}_{ih}} \Big/ \sum\nolimits_{h'} e^{\overline{U}_{ih'}} \tag{3}$$

In expectation, hospital demand, denoted Y_h, is the sum of patients' demands:

$$Y_h = \sum\nolimits_i \pi_{ih} \tag{4}$$

The derivative of demand with respect to unobserved quality is then:

$$\frac{\partial Y_h}{\partial Q_h} = \sum\nolimits_i \boldsymbol{\beta}_{q,i} \pi_{ih} \left(1 - \pi_{ih}\right) > 0 \tag{5}$$

Because patients prefer hospitals that are located near their homes, a hospital's demand will increase more with unobserved quality when patients near the hospital have a stronger taste for quality. This variation in the responsiveness of hospital demand to unobserved quality is central to our strategy, discussed in the next section, for distinguishing the contribution of unobserved quality to costs from that of productivity. Variation in the geographic distribution of patients (i.e., in the number of patients near particular hospitals) affects the average level of π_{ih} at a hospital and can therefore create additional variation in "quality responsiveness."

2.2 Productivity and Quality

In the context of production, it has long been recognized that firms may choose the levels of factor inputs based on heterogeneous productivity, inducing a correlation between the independent variables and the disturbance in an empirical model of the production function (Marschak and Andrews, Jr. (1944)). The related problem for the dual to production, the cost function, has also long been recognized (Nerlove (1965)). We argue here that differences in unobserved productivity across hospitals can lead to differences in quality and then present an instrumental-variable strategy for recovering the true cost of quality.

We assume that a hospital produces a single output whose quality may be represented by an index; moreover, the price of hospital output is fixed at a level that exceeds marginal cost. The hospital's problem is thus to choose quality to maximize utility. Utility depends directly on profits Π and, perhaps, quality Q:

$$U = \Pi + \beta Q \tag{6}$$

Because patients benefit from quality, the interpretation of $\beta > 0$ is that the hospital is altruistic, as in Newhouse (1970) and Lakdawalla and Philipson (2006). Even if $\beta = 0$, utility depends indirectly on quality through its impact on profits:

$$\Pi = PY\left(Q, \mathbf{Q}^{-}, \mathbf{X}\right) - C\left(Y, Q, A, \mathbf{W}\right) \tag{7}$$

in which $Y(\cdot)$ is hospital output/demand, $\mathbf{Q}^{-}$ is a vector of the qualities of competing hospitals, $\mathbf{X}$ is a vector of demand shifters, $C(\cdot)$ is cost, A is productivity, and $\mathbf{W}$ is a vector of factor prices.

Given equations 6 and 7 and some regularity assumptions, the following condition must hold at an optimum:

$$\frac{dU}{dQ} = (PY_Q + \beta) - (C_Y Y_Q + C_Q) = 0 \tag{8}$$

The hospital increases quality until the benefit from additional revenues and "good deeds" just equals the additional costs. Equation 8 implicitly defines functions for a hospital's optimal quality and its realized level of demand, namely, $Q^*(Q^-, \mathbf{X}, A, \mathbf{W}, \beta)$ and $Y^*(Q^-, \mathbf{X}, A, \mathbf{W}, \beta) \equiv Y\left(Q^*(Q^-, \mathbf{X}, A, \mathbf{W}, \beta), \mathbf{Q}^-, \mathbf{X}\right)$.

The relationships between optimal quality and realized demand, on the one hand, and productivity, on the other, can be characterized as follows:

$$Q_A^* = \frac{C_{YA} Y_Q + C_{QA}}{U_{QQ}} \geq 0, \tag{9a}$$

$$Y_A^* = Y_Q Q_A^* \geq 0 \tag{9b}$$

In this model, higher productivity can raise the marginal utility of quality, so that quality increases either directly or indirectly. Quality increases directly when the marginal cost of quality decreases with productivity ($C_{QA} < 0$). It decreases indirectly when the marginal cost of output decreases with productivity ($C_{YA} < 0$), and demand increases with quality (consistent with the evidence reviewed in the preceding section). By equation 9b, if quality increases, so must output.

Figure 1 illustrates an important implication of this model for the empirical analysis of the quality-cost relationship. The true quality-cost curves of a more and less productive hospital appear in the figure; these curves condition on all other determinants of costs, including output (which is typically included in cost functions). Under the model, the more productive hospital chooses to supply higher quality. Hence, when the observed data is fit by least squares, the true cost of quality is understated. The reason is that hospitals that tend to choose high quality also tend to be low-cost. Indeed, quality does not contribute to observed differences in hospital costs here; rather, quality-setting behavior decreases the variability of hospital costs relative to a world in which quality and productivity are unrelated.

If this model is correct, the cost of quality can be identified with variation in quality that is unrelated to productivity. We would then expect that the estimated cost of quality is higher than under least squares. Our approach appeals to the model for additional determinants of hospital quality. Features of hospital demand are promising candidates, as these features can be inferred from the model of hospital choice in the preceding section. Kessler and McClellan (2000) take a similar approach in analyzing the impact of competition on hospital expenditures and outcomes; Gaynor and Vogt (2003) use demand to instrument for price in their analysis of hospital mergers and competitive outcomes.

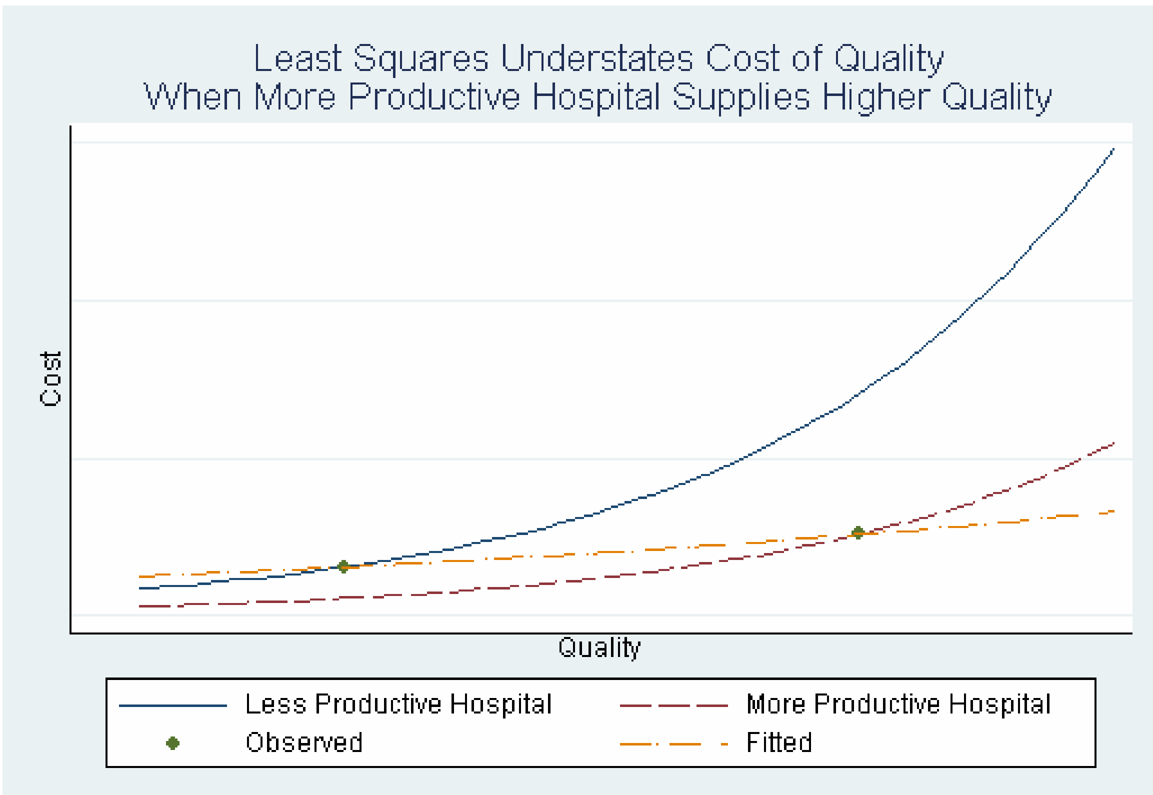

Least Squares Understates Cost of Quality
When More Productive Hospital Supplies Higher Quality
Cost
Quality
Less Productive Hospital
More Productive Hospital
Observed
Fitted

Figure 1:

The marginal utility of quality in equation 8 turns on the responsiveness of hospital demand to quality, i.e., to Y_Q. This aspect of hospital demand, which we term "quality responsiveness," reflects hospitals' "market potential." Within a large market such as the Los Angeles area, quality responsiveness may vary across hospitals as the number of patients and the taste for quality varies within localized submarkets (see equation 5). Figure 9 illustrates the marginal utility schedules for two hospitals in submarkets in which the responsiveness of demand to quality differs and all else (including productivity) is equal. For any level of quality, Y_Q is higher for the first hospital than the second; thus the marginal utility of quality is higher for the first hospital; and so this hospital chooses to supply higher quality. Hence the responsiveness of demand to quality is correlated with quality.

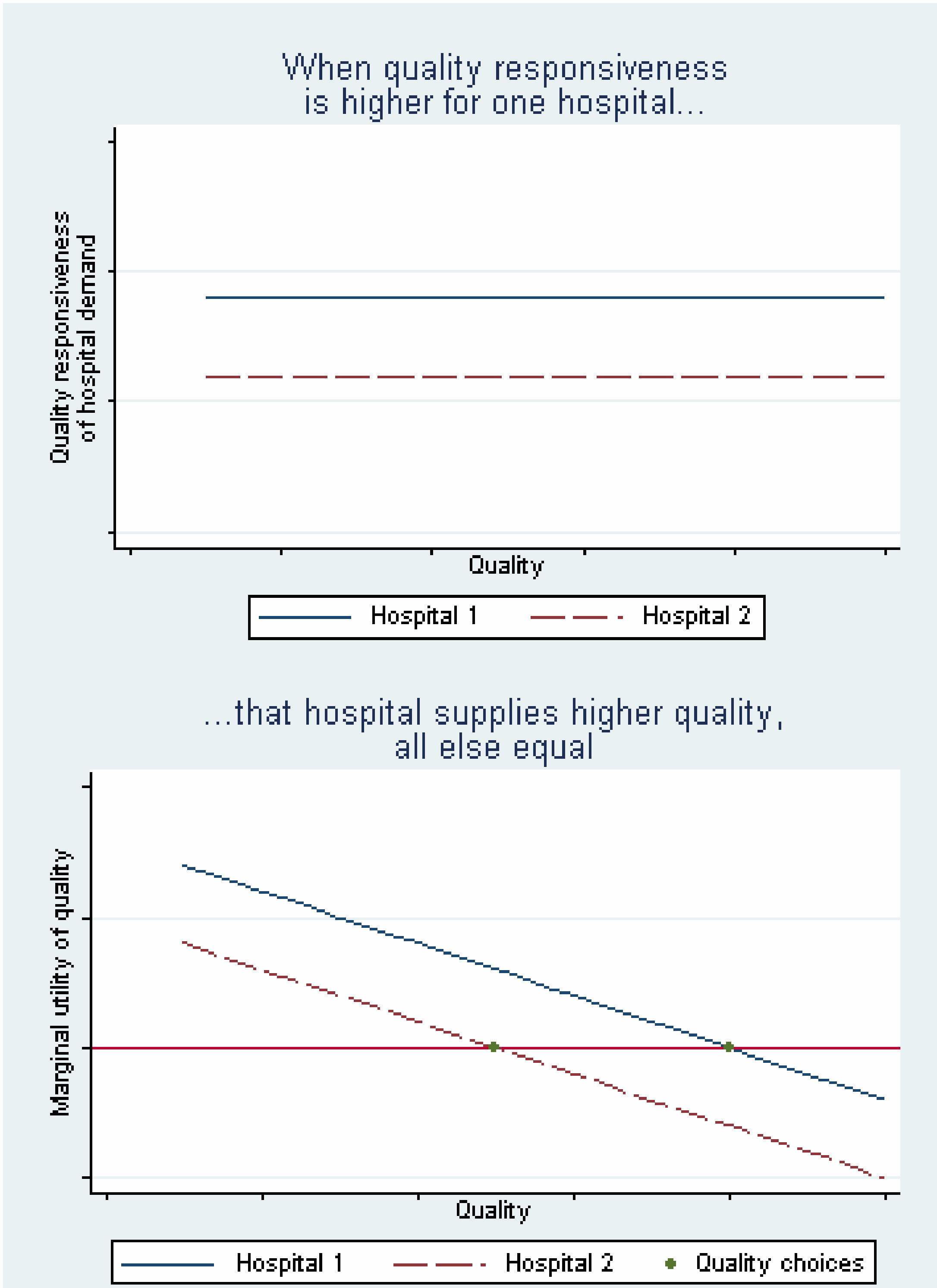
The validity of quality responsiveness as an instrument requires that it be uncorrelated with productivity. Drawing on the results from the model of hospital choice, the response to quality within a market depends on actual quality, recognizing that $Y_Q = Y_Q\left(Q, \mathbf{Q}^{-}, \mathbf{X}\right)$. Variation in this responsiveness is likely to be indirectly correlated with productivity through its impact on actual quality. We therefore fix quality throughout the market at some $\overline{Q}$. $Y_Q\left(\overline{Q}, \overline{\mathbf{Q}}, \mathbf{X}\right)$ does not vary with $\overline{Q}$ under the choice model embodied in equation 3.[6] We therefore fix quality at zero, defining $Y_Q^0 \equiv Y_Q\left(0, \mathbf{0}, \mathbf{X}\right)$.

Quality responsiveness and productivity could be correlated for other reasons. For example, patients with a strong taste for quality could reside near hospitals that are well-managed. Sicker patients may choose to live near high-quality hospitals, which tend to be more productive under the model. We cannot rule out such hypotheses without information on hospital productivity. The general lack of such information is precisely the problem.

3 Empirical Approach

In this section we explain our empirical approach, first to inferring hospital quality from the choice behavior of coronary-care patients, then to assessing the joint contribution of quality and productivity to hospital costs, and finally to distinguishing the contribution of quality to costs from that of productivity.

[6] For any patient, the contribution of $\overline{Q}$ to $\overline{U}_{ih}$ in section 2.1 is identical across hospitals. Thus, assuming that the ε_{ih} is distributed Type-1 extreme-valued and that patients always receive care, π_{ih} is invariant to the level of $\overline{Q}$ for every hospital.


When quality responsiveness
is higher for one hospital...
Quality responsiveness
of hospital demand
Quality
Hospital 1
Hospital 2
...that hospital supplies higher quality,
all else equal
Marginal utility of quality
Quality
Hospital 1
Hospital 2
Quality choices

3.1 Hospital choice

We analyze hospital choice among Medicare fee-for-service patients receiving coronary care in the Los Angeles area in 2002. The behavior and characteristics of these patients are disclosed in a data set on hospital discharges compiled annually by the California Office of Statewide Health Planning and Development. In the remainder of this section, we motivate our focus on this context while describing our empirical analysis of hospital choice.

Coronary care is sensible for our purposes. As we noted in section 2.1, there is persuasive evidence that patients view hospitals as differentiated in the quality of coronary care. Indeed, hospitals vary in their clinical competence, with some providing basic services (including, for example, the administration of beta blockers), others providing sophisticated diagnostics such as catheterization, and still others providing advanced surgical procedures such as coronary artery bypass graft surgery (CABG) or percutaneous transluminal coronary angioplasty (PTCA). The existing evidence on choice of hospital for coronary care also serves as a useful point of reference.

We consider two groups of coronary patients. The first group, following Tay (2003), includes patients with a primary diagnosis of acute myocardial infarction (AMI), i.e., a heart attack.[7] We refer to these patients (and their chosen hospitals) as the "AMI" sample. Such emergent patients are often transported to the hospital by ambulance. Even so, they appear to exercise some discretion over hospital.[8] These patients are typically admitted to the hospital through an emergency room.

The second group includes patients who were admitted for cardiac surgery (CABG or PTCA) through a route other than the emergency room.[9] We refer to these patients and their hospitals as the "planned-surgery" sample. These patients arguably scheduled their surgery and thus exercised greater choice over hospitals than did AMI patients.[10] We exclude coronary patients who do not receive surgery (e.g., those experiencing unstable angina) in an effort to create a relatively homogeneous sample.[11] The trade-off in

[7] These patients are identified in the discharge data by an ICD-9 classification that begins with "410."

[8] Iin Los Angeles county, for example, patient health is the most important factor in determining the hospital to which a patient will be transported, yet the patient's preference is a factor that must also be weighed (DummyCite (00)).

[9] These patients are identfied by the DRG codes 106, 107, 109, and 516-518.

[10] We thank Bill Vogt for this insight.

[11] Coronary-care patients are identified by a major diagnostic category of "diseases and

doing so is that hospitals at which advanced cardiac procedures could not be performed are excluded from the analysis, potentially masking important differences in hospital quality and costs.

Both samples include patients admitted either from home or an inpatient setting to a hospital within fifty miles of home. We retain patients admitted from an inpatient setting because some patients are initially treated at one hospital and then transferred to another once stabilized. The initial and subsequent hospitals may tend to have different qualities, and we wish to exploit this information. The share of patients admitted from an inpatient setting is modest (e.g., 11.7% of patients in the AMI sample). Patients admitted to hospitals more than fifty miles from home are excluded because of uncertainty regarding their circumstances, e.g., the identities of hospitals in their choice set and their locations relative to these hospitals. The number of such patients is quite small—the size of our AMI sample would increase by only 0.4% if these observations were retained. Finally, we exclude patients whose age, sex or race is censored for privacy reasons. The number of such patients is not negligible—the size of our AMI sample would increase by nearly 17.5% if these observations were retained. Yet these patient characteristics are plausibly related to the taste for quality (Tay (2003)).

These samples of patients are characterized in Table **??**. There are 7,014 AMI patients and 5,030 planned-surgery patients. Slightly more than 500 patients appear in both specifications; these may include AMI patients whose doctors arranged for a non-ER admission. Nearly forty percent of AMI patients are eighty years old or older, versus slightly more than twenty percent of planned-surgery patients. AMI patients are also more likely to be female and black. Finally, these patients have poorer health in general, as measured by the Charlson-Deyo index of comorbidities present in the discharge record (Quan et al. (2005)).

We consider the choice of these patients among hospitals in the Los Angeles area. Our definition of this hospital market includes all general acute-care hospitals in the five counties comprising the Los Angeles/Riverside/Orange County consolidated metropolitan statistical area, with the exception of two remote facilities.[12] Figure 2 shows the 160 hospitals chosen by at least one of our AMI patients in 2002; the 62 planned-surgery hospitals are a subset of

disorders of the circulatory system."

[12] These facilities are located in Blythe and Needles. These communities abut the Colorado River and the Arizona-California border, approximately two hundred miles from downtown Los Angeles.

Table 1:

Summary Statistics for Hospital Patients								
Variable	*Specification*							
	AMI				Planned surgery			
	Mean	SD	Min	Max	Mean	SD	Min	Max
80+ years old	41.4%	—	—	—	21.4%	—	—	—
Female	49.8%	—	—	—	36.7%	—	—	—
Black	6.4%	—	—	—	4.6%	—	—	—
Charlson-Deyo index (CDI)	2.9	1.8	1	15	1.4	1.5	0	12
Patients in sample	7014				5030			
Hospitals in sample	160				62			

the AMI hospitals. In the future we will consider an alternative definition that includes hospitals belonging to the Los Angeles hospital referral region, as defined in the Dartmouth Atlas of Health Care (DummyCite (00)). This definition excludes some relatively remote as well as rural hospitals (e.g., in Ventura County or the "high desert," or near Palm Springs) for which the impact of distance on patient choice may be different, potentially confounding inferences about quality.[13]

Our empirical model of the utility that a patient expects to receive from each of these hospitals is as follows:

$$U_{ih} = \beta_{di} D_{ih} + \beta_{qi} \sum\nolimits_{h'} Q_{h'} I\left(h' = h\right) + \epsilon_{ih}, \tag{10}$$

in which D_{ih} is the distance between the patient's home and the hospital, Q_h is the quality that patients observe but the econometrician does not, and $I\left(\cdot\right)$ is an indicator function equaling one when a statement is true and zero otherwise.

In contrast with equation 1, this specification does not include the price of a hospital's care. Like Town and Vistnes (2001) and Tay (2003), we focus on Medicare fee-for-service patients, for whom out-of-pocket costs are essentially identical across hospitals. Under our assumptions (described below) about behavior, these costs do not influence the hospital choice of our patients. This feature of the model is convenient, because price is difficult to measure.

[13] Hospital referral regions are reported in the American Hospital Association's Annual Survey of Hospitals. Fifteen hospitals in the discharge data set could not be linked to the AHA data. The figure reflects the unreported region of these hospitals.

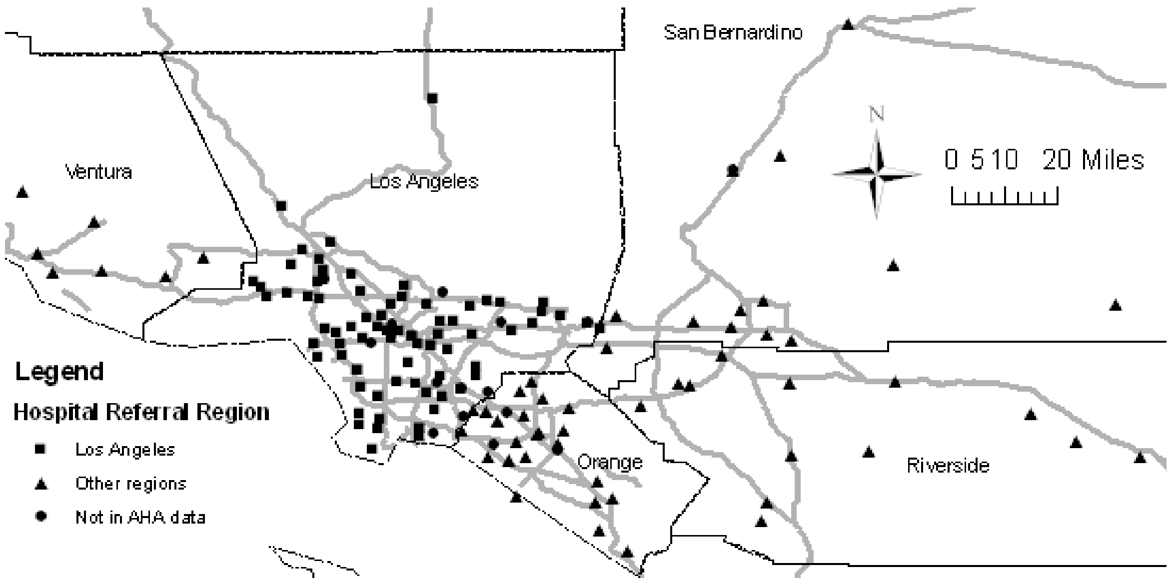


Figure 2: Hospitals in the Los Angeles area chosen by at least 1 patient in AMI sample in 2002

Moreover, in this context, the relationship between price and unobserved quality is likely to be complex and susceptible to misspecification.

Intuitively, unobserved quality is revealed by the preferences of patients over hospitals that differ in their distances from home. As we noted in section 2.1, patients prefer hospitals close to home, all else equal. If patients in an area favor a particular hospital that is farther away over a hospital that is closer, utility can be higher at the former hospital only because Q_h is greater at the former than at the latter. The observed behavior of patients in our samples suggest that this approach to inferring quality can succeed. Whereas the mean distance to the nearest hospital is 1.5 miles among AMI patients, the mean distance to the chosen hospital is 3.7 miles. Only 44.7% of these patients chose the nearest hospital, while 42.4% chose hospitals that are the third nearest to, or even farther from, home. The pattern among planned-surgery patients is similar. The mean distance to the nearest hospital is greater for these patients than for AMI patients, reflecting the fact that advanced cardiac surgery could not be performed at many hospitals.

In the empirical model of utility, we allow the tastes for distance and

Table 2:

Distance and Hospital Choice

Statistic	*Specification*	
	AMI	Planned surgery
Mean distance to nearest hospital (miles)	1.5	3.6
Mean distance to chosen hospital	3.7	6.0
Nearest hospital chosen	44.6%	44.7%
2nd nearest hospital chosen	13.0%	18.9%
3rd nearest or farther hospital chosen	42.4%	36.4%

quality to differ across patients. Tastes are assumed to take the following form:

$$\beta_{x,i} = \beta_{x,0} + \beta_{x,80+\text{ years}} I\left(Age_i \geq 80\right) + \beta_{x,Female} I\left(Sex_i = Female\right) + \beta_{x,Black} I\left(Race_i = Black\right) + \beta_{x,CDI} CDI_i,\ x = d, q,$$

in which CDI_i is the Charlson-Deyo index for the ith patient. Tay (2003) finds that older, female and, in some specifications, black AMI patients have different tastes for hospital attributes. Health status also plausibly affects a patient's valuation of quality and distance. Because quality is latent (see section 2.1), the taste for quality must be normalized relative to the taste for distance; we fix the intercept at one ($\beta_{Q,0} = 1$). This normalization implies that quality as inferred from patients cannot be compared in levels across specifications and, in particular, the AMI and planned-surgery samples.[14]

If tastes indeed differ across patients, we have argued in section 2.2 that such variation can induce variation in hospitals' chosen qualities that is unrelated to productivity. Table 3 demonstrates that hospitals' local submarkets differ in demographic characteristics potentially related to the taste for quality. For example, the coefficient of variation across hospitals in the proportion of AMI patients eighty years old or older residing within 2.5 miles is 0.32. Within both samples and at different distance thresholds, there is significant variation in all of the patient characteristics considered. We have also argued that the number of patients living near a hospital may influence

[14] Similarly, quality could not be compared in levels across markets (e.g., Los Angeles versus Chicago) without imposing additional structure, such as a fully specified model of quality choice by hospitals and a common technology of production (up to differences in productivity and factor prices).

Table 3:

Coefficient of Variation in Mean Characteristics of Patients Surrounding Hospitals

	AMI		Planned surgery	
	2.5 miles	*10 miles*	*2.5 miles*	*10 miles*
80+ years old	0.32	0.17	0.39	0.17
Female	0.21	0.08	0.30	0.13
Black	1.99	1.02	1.99	1.03
Charlson-Deyo index	0.14	0.06	0.25	0.12
Number of patients	0.66	0.60	0.68	0.54

its choice of quality. There is considerable variation in this dimension as well (e.g., a coefficient of variation of 0.66 for AMI patients within 2.5 miles).

In order to derive the likelihood that a patient is observed to choose a hospital, some additional assumptions are necessary. First, we assume that patients in need of coronary care always elect to receive care. This assumption is highly plausible for AMI patients and may be a reasonable approximation for planned-surgery patients due to the importance of coronary treatment and the generosity of Medicare benefits. Second, ϵ_{ih} is assumed to be distributed type-I extreme-valued. Then the probability that patient i chooses hospital h is:

$$\pi_{ih} = \frac{\exp\left[\beta_{di} D_{ih} + \beta_{\kappa i} \sum_{h'} Q_{h'} I\left(h' = h\right)\right]}{\sum_{h''} \exp\left[\beta_{di} D_{jh''} + \beta_{\kappa i} \sum_{h'} Q_{h'} I\left(h' = h''\right)\right]} \tag{11}$$

A hospital's expected demand follows immediately, namely, $Y_h = \sum_i \pi_{ih}$.

We estimate the model's parameters, including each hospital's quality Q_h, by maximizing the joint likelihood of patients' observed choices. As is typical in these models, unobserved quality must be normalized for some hospital (Berry et al. (1995); Town and Vistnes (2001)).[15] The loglikelihood is maximized with a quasi-Newton routine using the BFGS approximation to the Hessian.

3.2 Hospital costs

We recover hospital costs by specifying a translog cost function and estimating this model by fixed-effects regression over the period 2000-2004. The

[15] The quality of UCLA Harbor Medical Center is normalized to zero.

translog is a flexible functional form consistent with cost-minimizing behavior (Christensen et al. (1973)); this specification has been used to study the cost structures of a wide variety of firms, including hospitals.[16]

Our empirical approach to hospital costs rests on three fundamental assumptions. First, hospital productivity is assumed to be fixed in the short run. Ideally, some evidence would speak to this assumption; unfortunately, we are not aware of such evidence. While this assumption may be a reasonable approximation over the five years studied, we also consider the sensitivity of our results to a period as short as two years. In any case, this assumption implies that the fixed effect in each hospital's costs embodies its unobserved productivity, eliminating a potential source of correlation between the model disturbance and observables. We also assume that hospital quality is fixed. This assumption is not essential but, rather, permits the analysis of hospital choice to focus on a single year.[17] We hope in the future to explore the evolution of hospital quality as revealed by patients. Our belief is that quality will be fairly stable in the short run but may change significantly over long periods.

Second, productivity is assumed to shift the hospital's production possibility frontier proportionally, i.e., productivity is Hicks-neutral (Biesebroeck (2004)). Under this assumption, about which we do not have a strong prior, unobserved productivity influences costs only through the hospital fixed effect.[18]

Third, hospital quality for Medicare coronary-care patients is assumed to be positively correlated with quality for other kinds of patients. This assumption implies that quality as inferred from the choice behavior of the former is also a measure of quality for the latter. While this assumption strikes us as reasonable, error in the quality measure for non-coronary patients may lead to attenuation bias in the cost of quality. We plan to explore

[16] Cowing and Holtmann (1983), Zuckerman et al. (1994), and Dor and Farley (1996) are just a few applications of the translog to hospital costs.

[17] There are several reasons why the assumption that quality is fixed does not necessarily follow from the assumption that productivity is fixed. First, recalling section 2.2, demand conditions may change over time. Second, achieving a precise level of quality may be difficult, with hospitals undershooting or overshooting and then reverting toward the optimum. Third, a hospital may implement its optimal level of quality over time, due, for example, to convex adjustment costs.

[18] This assumption also implies, by Shepherd's lemma, that factor shares do not vary with productivity. We do not analyze factor shares here.

this issue in the future.[19]

Given these assumptions, hospitals produce quantity and quality according to the transformation function:

$$F\left(Y_{ht}, Q_h, \mathbf{Z}_{ht}, A_h\right) = 0 \tag{12}$$

in which Y_{ht} is quantity at hospital h in year t, Q_h is quality, $\mathbf{Z}_{ht}$ is a vector of cost shifters (e.g., for patient severity), and A_h is productivity. Quantity/output is one-dimensional here. In the future we will also relax the implied homogeneity of hospital care, distinguishing, for example, between coronary and non-coronary stays.

The solution to the hospital's problem of minimizing costs subject to quality and quantity constraints and the vector of factor prices $\mathbf{W}_{ht}$ gives the cost function:

$$C_{ht} = C\left(Y_{ht}, Q_h, \mathbf{Z}_{ht}, A_h, \mathbf{W}_{ht}\right)$$

This function is assumed to be translog:

$$\begin{aligned} \ln C_{ht} = \; & \alpha_0 + \alpha_Y \ln Y_{ht} + \alpha_Q Q_h + \sum\nolimits_j \alpha_{W_j} \ln W_{htj} + \sum\nolimits_j \alpha_{Z_j} \ln Z_{htj} \\ & + \frac{1}{2}\alpha_{Y,Y}\left(\ln Y_{ht}\right)^2 + \alpha_{Y,Q} \ln Y_{ht} \cdot Q_h \\ & + \sum\nolimits_j \alpha_{Y,W_j} \ln Y_{ht} \ln W_{htj} + \sum\nolimits_j \alpha_{Y,Z_j} \ln Y_{ht} \ln Z_{htj} \\ & + \frac{1}{2}\alpha_{Q^2} Q_h^2 + \sum\nolimits_j \alpha_{Q,W_j} Q_h \ln W_{htj} + \sum\nolimits_j \alpha_{Q,Z_j} Q_h \ln Z_{htj} \\ & + \frac{1}{2}\sum\nolimits_j \sum\nolimits_k \alpha_{W_i,W_k} \ln W_{htj} \ln W_{htk} \\ & - A_h + \varepsilon_{ht} \end{aligned} \tag{13}$$

Several observations are worth making about equation 13. First, the negative sign that precedes A_h reflects the fact that higher productivity entails

[19] At a minimum, we believe that measures of clinical quality for different kinds of care at hospitals may be found in the health-services literature. We may also infer quality from the choice of non-coronary patients. Geweke et al. (2003), for example, have analyzed hospital choice among Medicare patients with pneumonia; Town and Vistnes (2001) consider Medicare patients admitted to hospitals for a variety of conditions. Gaynor and Vogt (2003) also consider hospital choice for a variety of conditions. Given their focus on privately insured patients, these authors construct a measure of a patient's expected cost of hospital care. For our purposes, we would be concerned that this approach might confound price and unobserved quality.

lower costs. Recalling section 2.2, it is easy to show that higher productivity must decrease the marginal costs of both quantity and quality. Second, the appearance of quality and productivity in levels rather than logs is without loss of generality, as productivity and our measure of quality are latent (see section 2.1).[20] Third, the lack of interactions between productivity and other factors follows from the assumption that productivity is Hicks-neutral. Quality and the cost shifters are fully interacted with all factors (except productivity).

We are able to study the costs of hospitals whose quality is revealed by the choices of patients in the AMI and planned-surgery samples. Table 4 summarizes the hospitals analyzed. We consider the years 2000-2004 because this period is a moderately lengthy window around 2002, the year in which hospital choice is studied. Like McClellan and Newhouse (1997) and Picone et al. (2003), we measure hospital costs by summing charges (as reported in the discharge data) for all of a hospital's patients in a particular year and then applying the hospital's cost-to-charge ratio from the hospital's annual cost report filed with the Centers for Medicare and Medicaid Services.[21] This measure includes both capital and operating costs. We have been unable to obtain cost-to-charge ratios for 18 hospitals in the AMI sample. Hence the numbers of hospitals whose costs are analyzed is smaller than the numbers of hospitals whose quality is inferred from choice behavior.

The dependent variable in the cost model is the log of hospital costs. For the hospitals corresponding to our AMI patients, annual charges averaged \$340 million, while costs averaged \$110 million. The difference between charges and costs does not represent profit; charges, as list prices from which substantial discounts are frequently negotiated, overstate actual revenues.

We measure hospital quantity by (log) discharges. The number of discharges averaged approximately 12,000 per year for the AMI hospitals; approximately 1,700 of these were coronary cases. These hospitals are a diverse group, with discharges ranging from 181 to 53,060. Turning to quality, Q_h embodies all attributes that patients value, given our treatment of quality

[20] Q_{ht} could be rescaled as $e^{Q_{ht}}$ in hospital demand. Then, under the translog, $\ln e^{Q_{ht}} = Q_{ht}$.

[21] The discharge data is reported on a calendar-year basis. The CMS data is reported for a hospital's fiscal year, which typically coincides with the federal fiscal year beginning October 1. The synchronization of charges and cost-to-charge ratios is therefore "off," often by three months. We view this problem as insignificant and do not attempt to correct for it.

Table 4:

Summary Statistics for Hospital Samples

Variables	*Specification*			
	AMI			
	Mean	*SD*	*Min*	*Max*
Annual charges (million $)	340.4	349.2	2.8	3208.6
Costs (million $)	110	114.9	1.6	962.6
Annual discharges	12062	8829	181	53060
Mean Charlson-Deyo index, all discharges	1.1	0.3	0.4	2.4
2000	21.0%	—	—	—
2001	22.0%	—	—	—
2002	20.0%	—	—	—
2003	21.0%	—	—	—
2004	16.0%	—	—	—
Hospital-years in sample	644			
Hospitals in sample	142			
	Planned surgery			
	Mean	*SD*	*Min*	*Max*
Annual charges (million $)	569.6	402.9	62.4	3208.6
Costs (million $)	185.9	132.9	20.3	962.6
Annual discharges	18571	8535	2724	53060
Mean Charlson-Deyo index, all discharges	1.1	0.3	0.4	2.4
2000	20.0%	—	—	—
2001	21.0%	—	—	—
2002	20.0%	—	—	—
2003	21.0%	—	—	—
2004	18.0%	—	—	—
Hospital-years in sample	280			
Hospitals in sample	60			

in the analysis of hospital choice. We use quality as inferred from hospital choice. For cost shifters, we include the Charlson-Deyo comorbidity index as a measure of case severity; henceforward, $\mathbf{Z}_{ht} = CDI_{ht}$. The typical coronary patient was more comorbid than the typical patient overall. Statistics also appear in the table for hospitals corresponding to planned-surgery patients. The smaller number of hospitals at which advanced cardiac surgery is available tend to be larger in scale. For both specifications, observations are drawn quite evenly across years.

Factor prices also appear in the cost model. The hospitals studied operate in a geographically limited and fairly homogeneous environment. It therefore seems reasonable to assume, as a first approximation, that hospitals face identical factor prices.[22] We further assume that prices grow at an identical rate during the period studied. Factor prices then equal:

$$\ln W_{htj} = \beta_j + \beta_t t \tag{14}$$

Under these assumptions, factor prices need not be measured. Rather, equation 14 can be substituted into equation 13, which becomes:

$$\begin{aligned} \ln C_{ht} = {} & \widetilde{\alpha}_0 + \alpha_Y \ln Y_{ht} + \alpha_Q Q_h + \alpha_{CDI} \ln CDI_{ht} + \frac{1}{2}\alpha_{Y^2} \left(\ln Y_{ht}\right)^2 \\ & + \alpha_{Y,Q} \ln Y_{ht} \cdot Q_h + \alpha_{Y,CDI} \ln Y_{ht} \ln CDI_{ht} \\ & + \frac{1}{2}\alpha_{Q^2} Q_h^2 + \alpha_{Q,CDI} Q_h \ln CDI_{ht} + \frac{1}{2}\alpha_{CDI^2} \left(\ln CDI_{ht}\right)^2 \\ & + \alpha_t t + \alpha_{Y,t} \ln Y_{ht} \cdot t + \alpha_{Q,t} Q_h \cdot t + \alpha_{CDI} \ln CDI_{ht} \cdot t + \alpha_{t^2} \cdot t^2 - A_h + \varepsilon_{ht} \end{aligned} \tag{15}$$

Costs now change over time according to the rate of input-price inflation. Time is measured relative to the year 2000 (i.e., year 2000 is time 0). This equation can be rewritten as:

$$\begin{aligned} \ln C_{ht} = {} & (\widetilde{\alpha}_0 - \phi_0) + \alpha_Y \ln Y_{ht} + \alpha_{CDI} \ln CDI_{ht} + \frac{1}{2}\alpha_{Y^2} \left(\ln Y_{ht}\right)^2 \\ & + \alpha_{Y,Q} \ln Y_{ht} \cdot Q_h + \alpha_{Y,CDI} \ln Y_{ht} \ln CDI_{ht} \\ & + \alpha_{Q,CDI} Q_h \ln CDI_{ht} + \frac{1}{2}\alpha_{CDI^2} \left(\ln CDI_{ht}\right)^2 \\ & + \alpha_t t + \alpha_{Y,t} \ln Y_{ht} \cdot t + \alpha_{Q,t} Q_h \cdot t + \alpha_{CDI} \ln CDI_{ht} \cdot t + \alpha_{t^2} \cdot t^2 + \phi_h + \varepsilon_{ht}, \end{aligned} \tag{16}$$

[22] If hospitals enjoy a degree of monoposony power in factor markets, larger hospitals may pay lower prices. If this is the case, estimated scale economies will reflect this cost advantage for larger hospitals. Our principal focus is on the cost of quality, not quantity.

in which the hospital's "cost effect" ϕ_h is:

$$\phi_h = \phi_0 + \alpha_Q Q_h + \frac{1}{2}\alpha_{Q^2} Q_h^2 - A_h \tag{17}$$

This effect reflects contributions from both quality and productivity to cost. The intercept in equation 17 is not identified; Stata's *xtreg* routine normalizes ϕ_0 so that ϕ_h is mean zero.

We analyze the cost model 16. The independent variables are demeaned, so that the uninteracted parameters reflect phenomena of interest at a hospital with mean characteristics. For example, the mean hospital experiences scale economies (diseconomies) if and only if $\alpha_Y < 1$ ($\alpha_Y > 1$). We consider the two sets of hospitals corresponding to the two groups of coronary-care patients, using for each the quality as inferred from the corresponding group of patients.

The model's parameters are estimated with random- and fixed-effects regressions. Our model in section 2.2 implies that terms involving quantity and quality will generally be correlated with the hospital cost effect. Random-effects regression is then generally inconsistent, while fixed-effects regression remains consistent. We use a Hausman (1978) test to compare the results from these two methods; a rejection of the null hypothesis that the results are identical would be consistent with our model of hospital behavior.

Finally, quality in equation 16 is measured with sampling error from the analysis of hospital choice. Conventional standard errors will not reflect this sampling error (Murphy and Topel (1986)). Valid standard errors can be computed by reestimating the choice model on a set of resampled patients and then reestimating the cost model using the revised estimates of hospital quality. Standard errors for phenomena of interest—the cost of quality, in particular—can also be boostrapped.

3.3 The Cost of Quality

Understanding the cost of quality is the aim of this paper. Equations 16 and 17 imply that the cost of quality is:

$$\begin{aligned} \frac{\partial \ln C_{ht}}{\partial Q_h} &= \frac{\partial \phi_h}{\partial Q_h} + \alpha_{Y,Q} \ln Y_{ht} + \alpha_{Q,CDI} \ln CDI_{ht} + \alpha_{Q,t} t \qquad (18) \\ &= \alpha_Q + \alpha_{Q^2} Q_h + \alpha_{Y,Q} \ln Y_{ht} + \alpha_{Q,CDI} \ln CDI_{ht} + \alpha_{Q,t} t \end{aligned}$$

The final three terms in this equation make clear that quality may be more or less costly as scale increases, as patients are more comorbid, and as time passes. The results for the cost model 16 include estimates of these parameters. The remaining parameters in equation 18 appear in the cost effect ϕ_h.

We wish to decompose the hospital cost effect:

$$\phi_h = \phi_0 + \alpha_Q Q_h + \frac{1}{2}\alpha_{Q^2} Q_h^2 - A_h \tag{19}$$

Estimates of the dependent variable ϕ_h are available from fixed-effects regression of the cost model, while estimates of the independent variables Q_h and Q_h^2 are available from the analysis of hospital choice. We use both sets of estimates in our analysis.

Two challenges remain. First, we have argued that observed quality is plausibly related to unobserved productivity. A valid instrument is correlated with observed quality but uncorrelated with productivity. As we explained in section 2.2, the derivative of a hospital's demand with respect to its own quality (i.e., "quality responsiveness") affects the marginal utility of quality and thus its optimal level. Quality responsiveness depends on observed quality and is therefore generally correlated with productivity. We therefore evaluate the derivative fixing quality at zero quality for every hospital, i.e., at $Y_Q^0 \equiv Y_Q(0, \mathbf{0}, \mathbf{X})$. The relationship between observed quality and this measure of quality responsiveness can then be assessed as follows:

$$Q_h = \theta_0 + \theta_q Y_{Q,h}^0 + \nu_h \tag{20}$$

For the case of one endogenous variable, Stock and Yogo (2005) suggest that a set of instruments has sufficient power if this first-stage regression's F statistic exceeds ten. Our main analysis restricts the cost effect to be linear in quality, and we instrument for quality with Y_Q^0. We relax this restriction in the sensitivity analysis in section 4.3, instrumenting for quadratic quality with $(Y_Q^0)^2$.

The second challenge is that the dependent and independent variables in equation 19 include sampling error. Moreover, the IV estimator for the dependent variable is unbiased but inconsistent in a finite panel, implying that the finite-sample performance of the estimator is paramount for purposes of inference (DummyCite (00)). Hospital choice can again be boostrapped, applying the resulting estimates first in the cost model and then in the decomposition of quality and productivity.

4 Findings

We now describe our findings from the analyses of hospital choice and costs. We then present our evidence on the cost of quality in hospitals.

4.1 Hospital Choice

The results of our analyses of hospital choice appear in Table 5. For both samples of patients, observed and predicted hospital-level demand are perfectly correlated, due to the inclusion of hospital indicator variables whose parameters measure quality. Our treatment of quality contributes substantially to the model's fit. Focusing exclusively on distance (by evaluating hospital-level demands at zero quality for every hospital, denoted $\widehat{Y}(0, \mathbf{0}, \mathbf{X})$), these correlations decrease to 0.505 and 0.550 for AMI and cardiac-surgery patients, respectively.

To aid in interpretation and assessment, Table 6 characterizes tastes for distance and quality among different groups of patients. For the AMI sample as a whole, the marginal rate of substitution between quality and distance averages 1.67. Our reference group consists of white males under 80 with mean comorbidity; this group is indifferent between a one-unit increase in hospital quality and a 1.87-mile increase in distance. Older men value quality less (relative to distance) than the reference group ($MRS_{Q,D} = 1.51$); women value quality slightly less ($MRS_{Q,D} = 1.81$), and blacks value quality much less ($MRS_{Q,D} = 1.30$).[23] White men under 80 in the top quartile of the comorbidity distribution also value quality slightly less than the reference group ($MRS_{Q,D} = 1.81$). Among planned-surgery patients, the reference group again values quality more than do others. Tastes are now more homogeneous among the elderly, female, and comorbid. Comparing AMI and planned-surgery patients, the latter value hospital quality more highly (again relative to distance) than do AMI patients. This finding is not surprising, as travel time to the hospital (and hence distance) is itself an important determinant of heart-attack outcomes, and many AMIs occur close to home.

The distributions of hospital quality are described in Table 7. While the means and standard deviations for the AMI and planned-surgery samples cannot be compared, the two sets of quality estimates are strongly positively

[23] Racial disparities in the quality of treatment could bias our inferences about hospital quality if the magnitude of these disparities vary with hospital quality.

Table 5: Hospital Choice

Variable	*Specification*	
	AMI	Planned surgery
Distance	-5.232 (0.124)***	-3.667 (0.086)***
Distance*80+ years old	-0.462 (0.110)***	-0.859 (0.145)***
Distance*Female	-0.200 (0.105)*	-0.186 (0.107)*
Distance*Black	-0.337 (-0.240)	-1.705 (0.367)***
Distance*Charlson-Deyo index (CDI)	0.110 (0.028)***	0.134 (0.032)***
Quality	1.000 (—)	1.000 (—)
Quality*80+ years old	-0.104 (0.025)***	0.102 (0.052)*
Quality*Female	0.011 (-0.026)	-0.043 (-0.040)
Quality*Black	-0.230 (0.054)***	-0.169 (0.094)*
Quality*CDI	-0.028 (0.006)***	-0.074 (0.011)***
Hospital indicators	Included	Included
Other Statistics		
Log likelihood	-12620.68	-7620.31
$Corr\left(Y, \widehat{Y}\right)$	1.000	0.999
$Corr\left(Y, \widehat{Y}\left(0, \mathbf{0}, \mathbf{X}\right)\right)$	0.505	0.550

Notes: Standard errors appear in parentheses. * denotes significance at the 10% level, ** at 5%, and *** at 1%.

Table 6:

Mean Tastes and Marginal Rate of Substitution			
Reference Group	*Specification*		
	AMI		
	Quality	*Distance*	*MRS*
All	0.87	-0.52	1.67
	0.03	0.01	0.06
White males under age 80, mean Charlson-Deyo index (CRI)	0.92	-0.49	1.87
	0.02	0.01	0.05
White males over age 80, mean CRI	0.81	-0.54	1.51
	0.03	0.01	0.06
White females under age 80, mean CRI	0.93	-0.51	1.81
	0.03	0.01	0.07
Black males under age 80, mean CRI	0.68	-0.52	1.3
	0.06	0.02	0.11
White males under age 80, top quartile of CRI	0.82	-0.45	1.81
	0.04	0.01	0.09
	Planned surgery		
	Quality	*Distance*	*MRS*
All	0.90	-0.38	2.38
	0.03	0.01	0.11
White index (CRI) males under age 80, mean Charlson-Deyo	0.90	-0.35	2.58
	0.01	0.01	0.08
White males over age 80, mean CRI	1.01	-0.44	2.31
	0.06	0.02	0.13
White females under age 80, mean CRI	0.85	-0.37	2.32
	0.04	0.01	0.12
Black males under age 80, mean CRI	0.68	-0.51	1.33
	0.09	0.04	0.21
White males under age 80, top quartile of CRI	0.71	-0.31	2.25
	0.04	0.01	0.14

Note: Standard errors appear in parentheses.

Table 7:

Summary Statistics for Hospital Quality				
Specification	*Mean*	*SD*	*Min*	*Max*
AMI	-1.93	4.31	-23.89	8.38
Planned surgery	1.87	2.90	-7.92	8.57

Correlation in Quality Estimates		
	AMI	Planned surgery
AMI	1.000	—
Planned surgery	0.752	1.000

correlated ($\rho = 0.752$). These apparent differences in quality and tastes within each sample suggest that our strategy of using demand to instrument for quality can succeed. We assess the relationship between quality and quality responsiveness in the subsequent section that decomposes the contributions of quality and productivity to costs.

4.2 Hospital Costs

The results of the random- and fixed-effects panel regressions appear in Table 8. The regressions explain nearly eighty-five percent of the "within" variation in log costs.

In the AMI sample of hospitals, the elasticity of costs with respect to scale $\eta_{C,Y}$ equals the parameter on $\ln Y_{ht}$ for a hospital with average characteristics, due to the demeaning of the covariates in the interaction terms relating to scale. Constant returns to scale cannot be rejected for an "average" hospital under the random-effects model. (The statistical tests in this section do not presently account for sampling error in the quality estimates but will do so in the future.) Under fixed effects, an average hospital experiences significant returns to scale $\left(\eta_{C,Y} = 0.76\right)$. As we argued in section 2.2, the apparent understatement of scale economies in the random-effects results is consistent with greater demand due to costly quality. The null hypothesis that the random- and fixed-effects estimates are identical is strongly rejected, as the cost effect ϕ_h is positively correlated with observably higher costs ($\rho = 0.426$). For the planned-surgery hospitals, constant returns cannot be rejected under either model. Differences in unobserved factors (in particular, quality) apparently do not lead to widely divergent estimates. This smaller

Table 8: Hospital Costs

Variable	*Specification*			
	AMI		Cardiac Surgery	
	Random Effects	*Fixed effects*	*Random Effects*	*Fixed effects*
Constant	18.010***	18.065***	18.807***	18.842***
	(0.041)	(0.019)	(0.066)	(0.022)
$\ln Y_{ht}$	0.943***	0.742***	0.989***	0.964***
	(0.035)	(0.049)	(0.083)	(0.106)
$\ln CDI_{ht}$	0.382***	0.266***	0.466***	0.413**
	(0.070)	(0.079)	(0.135)	(0.159)
$\frac{1}{2}(\ln Y_{ht})^2$	0.056*	-0.078**	0.117	0.037
	(0.030)	(0.037)	(0.142)	(0.161)
$\ln \mathrm{Y}_{ht} \cdot Q_h$	-0.009	0.014	0.035*	0.047**
	(0.007)	(0.014)	(0.020)	(0.022)
$\ln Y_{ht} \ln CDI_{ht}$	0.097	-0.034	0.037	-0.207
	(0.070)	(0.076)	(0.278)	(0.314)
$Q_h \ln CDI_{ht}$	-0.020	-0.022	-0.039	-0.016
	(0.014)	(0.017)	(0.049)	(0.059)
$\frac{1}{2}(\ln CDI_{ht})^2$	0.110	-0.190	-0.479	-0.889**
	(0.219)	(0.230)	(0.392)	(0.432)
t	0.125***	0.130***	0.126***	0.128***
	(0.004)	(0.004)	(0.005)	(0.006)
$\ln Y_{ht} \cdot t$	0.010**	0.012***	0.017	0.018
	(0.005)	(0.005)	(0.012)	(0.013)
$Q_h \cdot t$	0.000	-0.001	-0.002	-0.002
	(0.001)	(0.001)	(0.002)	(0.003)
$\ln CDI_{ht} \cdot t$	0.015	0.020*	0.040**	0.041**
	(0.012)	(0.012)	(0.019)	(0.019)
t^2	-0.015**	-0.016***	-0.011	-0.010
	(0.006)	(0.006)	(0.008)	(0.008)
Other Statistics				
R^2 (within)	0.834	0.840	0.837	0.838
p value, model Wald stat.	0.000	—	0.000	—
p value, Model F stat.	—	0.000	—	0.000
Observations	672	672	291	291
p value, F stat.on fixed effects	—	0.000	—	0.000
$Corr(\mathbf{X}\boldsymbol{\beta}, \varphi_h)$	—	0.426	—	0.000
p value, Hausman test	0.000		0.799	

Notes: * denotes significance at the 10% level, ** at 5%, and *** at 1%.
Standard errors are not currently adjusted for variability of quality estimates from the analysis of hospital choice.

sample of hospitals is more homogeneous (at least with respect to scale) than the AMI sample.

The results for the two samples are similar in two important respects. First, the comorbid are much more costly to treat. At the mean hospital, a one-standard-deviation increase in comorbidity would increase costs by roughly ten to fifteen percent. Second, cost was increasing in excess of ten percent per year (though the rate of increase diminished in the AMI sample). There is also some evidence that patient comorbidity came to be more costly over time in the AMI sample, while hospital scale came to be more costly in the planned-surgery sample.

For the subsequent analysis of the cost of quality, the interaction terms relating to hospital quality are retained. These interactions are only marginally jointly significant for the AMI hospitals ($p = 0.15$). The cost of quality was increasing in hospital scale for the planned-surgery hospitals.

4.3 The Cost of Quality

Turning to the cost of quality, we first consider the relationship between hospital quality and the responsiveness of demand to quality. The results of first-stage regressions appear in Table 9. For the AMI sample, quality is increasing in quality responsiveness, consistent with our model of hospital behavior in section 2.2. This relationship is illustrated in Figure 3. The power of this instrument exceeds the usual threshold.

For the planned-surgery sample, hospital quality is unrelated to the quality responsiveness of planned-surgery patients. These patients, however, represent only one segment of the demand for a hospital's coronary care. We therefore considered the responsiveness of *AMI* patients to these alternative estimates of hospital quality. The table's final column describes the relationship between quality as inferred from planned-surgery patients and the quality responsiveness of demand among AMI patients. Quality again increases in responsiveness, though the instrument's power is somewhat weak.

We use the quality responsiveness of hospital demand to decompose the contributions of quality and productivity to hospital costs. For the planned-surgery sample, the quality responsiveness of AMI patients, as just discussed, is our instrument. Table 10 reports the results of regressions that restrict the parameter on quadratic quality (α_{Q^2}) to be zero. In the instrumental-variables regression, the parameter estimate on quality is positive for both samples. While the levels are similar, the estimate is statistically insignif-

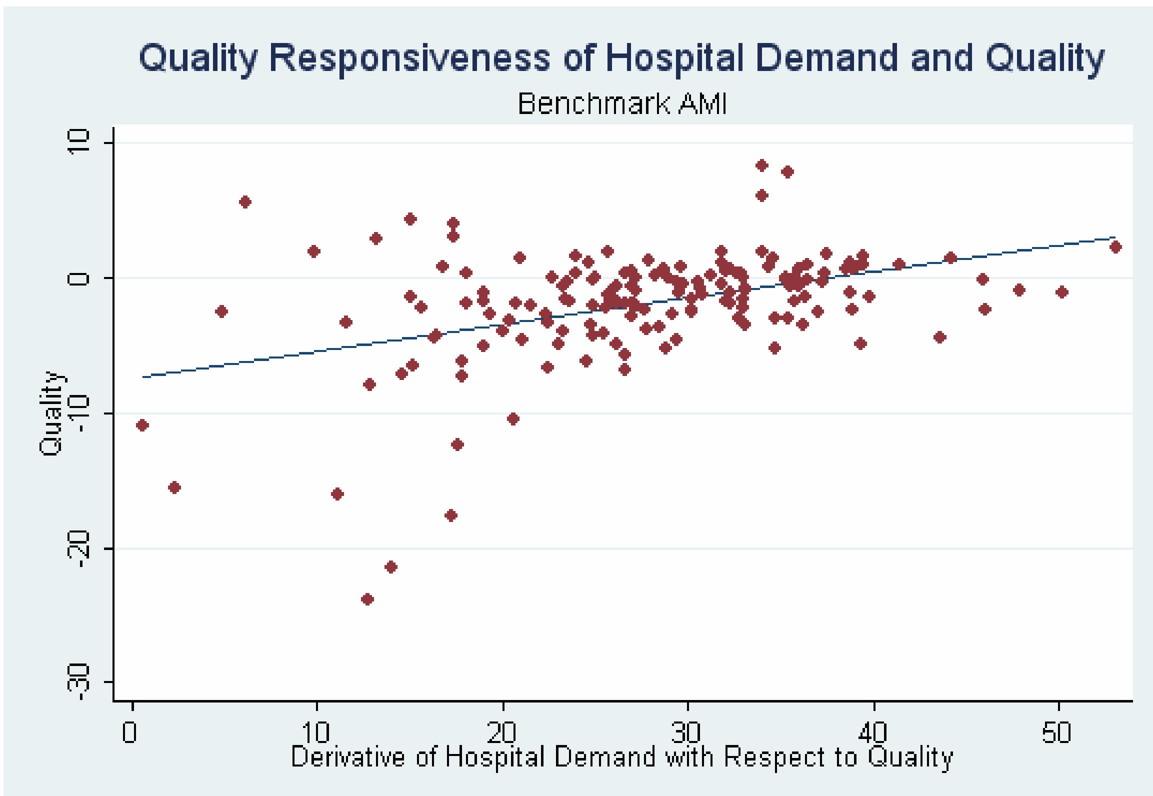

Quality Responsiveness of Hospital Demand and Quality
Benchmark AMI
Quality
10
0
-10
-20
-30
0
10
20
30
40
50
Derivative of Hospital Demand with Respect to Quality

Figure 3:

Table 9:

Estimates (Standard Errors) from First-Stage Quality Regressions

Variable	*Specification*		
	AMI	Planned surgery	
	AMI patients	*Surgery patients*	*AMI patients*
Constant	-7.405***	2.980**	-0.858
	-1.087	-1.326	-1.421
Y_Q^0	0.203***	-0.022	0.093**
	-0.037	-0.026	-0.046
Other Statistics			
R^2	0.176	0.012	0.065
F statistic	29.92	0.72	4.04
N	142	60	60
$Corr\,(Q_h, A_h)$	0.367	0.156	0.486

Notes: * denotes significance at the 10% level, ** at 5%, and *** at 1%. Standard errors are not currently adjusted for variability of quality estimates from the analysis of hospital choice.

icant for the planned-surgery sample, consistent with the weakness of the instrument.

The results are consistent with the model's prediction that more productive hospitals supply higher quality. For both samples the estimated parameter is biased downward under OLS; in the case of the AMI sample, the null hypothesis that the OLS and IV estimates are identical can be rejected. Figure 4 shows the relationship between quality and productivity in the AMI sample. This relationship suggests that, if quality is indeed costly, then quality differentation among firms may compress the distribution of costs per patient relative to a world of homogeneous quality. Quality may also affect this distribution through its impact on hospital demand and thus any scale economies.

Having decomposed the contributions of hospital quality and productivity, the cost of quality can be discerned. We consider one-standard-deviation increases in latent quality. For the AMI sample of hospitals, Table 11 indicates that increased quality raises costs by twenty-eight percent at the mean hospital, with a bootstrapped standard error of five percent, under the preferred IV results. The cost of quality is only eight percent under the OLS

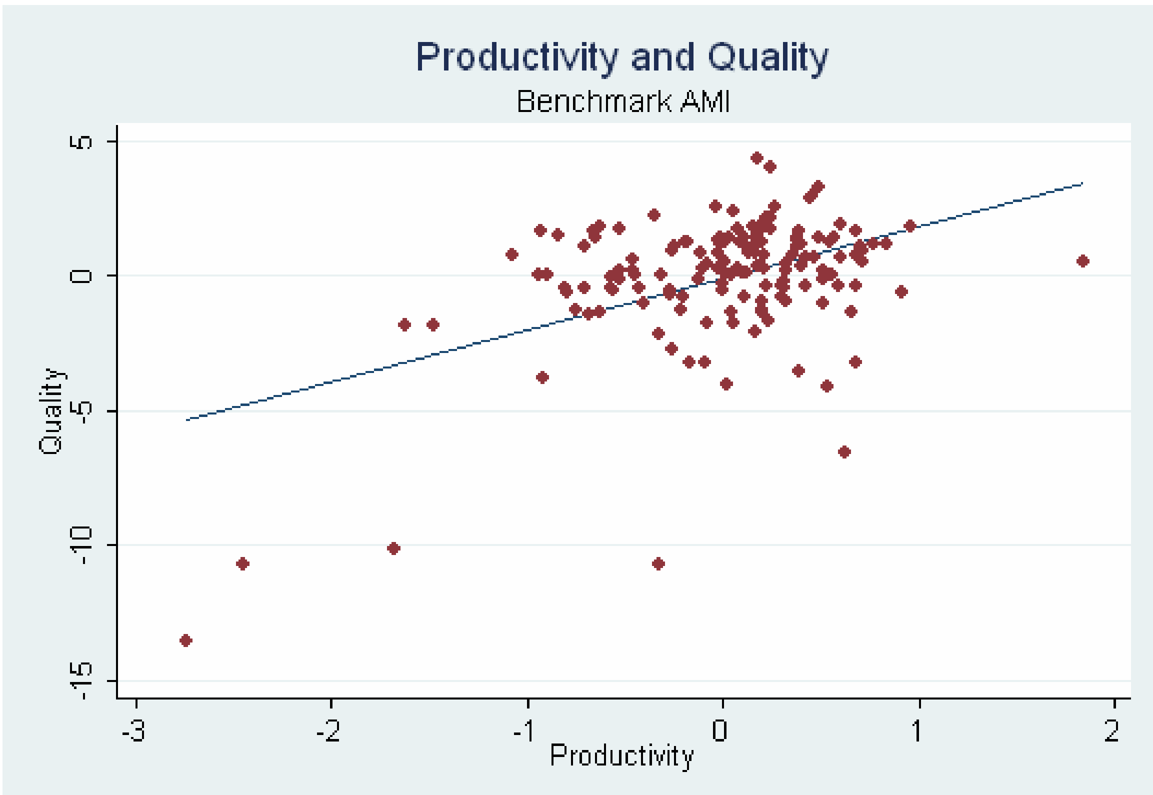

Productivity and Quality
Benchmark AMI
Quality
5
0
-5
-10
-15
-3
-2
-1
0
1
2
Productivity

Figure 4:

Table 10:

Estimates (Standard Errors) from Decomposition of Quality and Productivity

Variable	*Specification*			
	AMI		Cardiac surgery, AMI instrument	
	OLS	*IV*	*OLS*	*IV*
Constant	-0.015	-0.009	0.005	0.002
	(0.046)	(0.049)	(0.056)	(0.064)
Q_h	0.020* (0.010)	0.068** (0.026)	-0.005 (0.019)	0.076 (0.086)
Other Statistics				
R^2	0.424	—	0.904	—
N	142		60	
Hausman test p value	0.048		0.333	

Notes: * denotes significance at the 10% level, ** at 5%, and *** at 1%. Standard errors are not currently adjusted for variability of quality estimates from the analysis of hospital choice.

results. Defining small and large hospitals as those whose annual discharges lie a standard deviation below and above the mean, the cost of quality appears to be higher in large hospitals than in small ones, namely, 33.0% vs. 23.0%.

The robustness of these results to alternative specifications can be assessed. For the statistically imprecise planned-surgery results, increased quality raises costs by 26.8% at the mean hospital. The standard deviation of quality for this sample of hospitals is lower, however (see Table 7). We therefore scaled these estimates by the ratio of the standard deviation of quality, as inferred from AMI patients, to the standard deviation of quality, as inferred from planned-surgery patients, both for the planned-surgery hospitals. The results, which appear under the column header *AMI norm*, are quite similar to the earlier estimates.

In addition, we consider alternative treatments of the AMI sample. First we reestimated costs for 2001-2002, because productivity may be unstable over the five years studied. Increased quality now raise costs by 31.7% at the mean hospital.[24] Second, outliers in the quality distribution may

[24] The results of these cost regressions are available from the authors on request.

Table 11: Estimated Cost Impact (Standard Error) of a One-Standard-Deviation Increase in Quality

AMI, AMI instrument		
Mean hospital	28.0%	
	(5.1%)	
Small hospital	23.0%	
Large hospital	33.0%	
Planned surgery, AMI instrument		
	Surgery norm	*AMI norm*
Mean hospital	21.0%	29.5%
Small hospital	15.3%	21.4%
Large hospital	26.8%	37.5%
Cost for mean hospital, alternative specifications		
AMI, OLS	8.3%	
2001-2002 only	31.7%	
Hospitals within 10th-90th percentile of quality	48.7%	
Quadratic quality	64.2%	

Note: Standard errors appear in parentheses. Standard errors derived from bootstrap sample with 25 replications.

influence our findings. We therefore exclude hospitals whose quality lies in the bottom or top decile from the cost analysis. Increased quality now raises costs by 48.7%. Third, we explore the possibility that the hospital cost effect is quadratic in quality. Instrumenting with the squared quality responsiveness of hospital demand, increased quality raises costs by 64.2% at the mean hospital; while we still need to compute valid standard errors, this estimate appears (on the basis of the regression of hospital cost effects on quality) to be significantly less precise than the restricted estimate.

5 Conclusion

In clarifying the relationship among hospital quality, productivity and costs, we confronted two significant challenges. Hospital quality is difficult for a researcher to observe. In addition, quality is potentially confounded with unobserved differences in hospital productivity, in the sense that hospitals that are able to produce quality at lower cost have incentives to supply higher quality. Each of these problems can make quality appear to be less costly than is truly the case.

Our strategy for dealing with these problems exploits the observed behavior of consumers in this industry. Patients (or their agents) plausibly know more about hospital quality than a researcher does, and so we inferred quality at hospitals in metropolitan Los Angeles from the revealed preference of Medicare fee-for-service patients receiving coronary care. We then recovered the joint contribution of quality and productivity to costs from longitudinal data on hospital costs. In order to distinguish the contribution of quality from that of productivity, we again appealed to consumer behavior, arguing that differences in patients' tastes for quality within localized submarkets for hospital care lead to exogenous variation in hospitals' chosen qualities.

Our empirical findings suggest that the cost of quality is indeed substantial. A one-standard deviation increase in quality at an average hospital would increase costs on the order of thirty percent, a result that appears to be at least somewhat robust to alternative specifications. When, however, we did not instrument for quality, a one-standard deviation increase in quality appeared to increase costs by only eight percent.

These findings rest on a foundation of related results. Quality as perceived by patients was an important determinant of hospital choice, consistent with existing evidence. As a result, a random-effects regression significantly understated scale economies within the broader sample of AMI hospitals. We also found that hospital quality increases in the responsiveness of hospital demand to quality, and that quality and productivity are positively correlated, consistent with our model of quality choice among hospitals. These results suggest that quality differentiation may compress the distribution of cost per patient in the hospital market studied.

The present analysis suggests worthwhile directions for future research. For example, by specifying and estimating an empirical model of quality choice, we could assess the distribution of costs in a counterfactual world with no productivity differences. We could also investigate the role of altruism in the hospital-care industry and, in particular, about the relationship between altruism and hospital ownership. More generally, we believe that our approach holds promise for understanding the cost structure of firms in other differentiated-products industries.

References

Ackerberg, D. A., M. P. Machado, and M. H. Riordan (2006). "Benchmarking for Productivity Improvement: A Health-Care Application." *International Economic Review* 47(1):161–201.

Berry, S., J. Levinsohn, and A. Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica* 63(4):841–890.

Biesebroeck, J. V. (2004). "Robustness of Productivity Estimates." *Working Paper 10303*. Cambridge, MA: National Bureau of Economic Research.

Burns, L. R., and D. R. Wholey (1992). "The impact of physician characteristics in conditional choice models for hospital care." *Journal of Health Economics* 11(1):43–62.

Carey, K. (2000). "A Panel Data Design for Estimation of Hospital Cost Functions." *Review of Economics and Statistics* 79(3):443–453.

Carey, K., and J. F. Burgess, Jr. (1999). "On Measuring the Hospital Cost/Quality Trade-Off." *Health Economics* 8(6):509–520.

Christensen, L. T., D. W. Jorgenson, and L. J. Lau (1973). "Transcendental Logarithmic Production Frontiers." *Review of Economics and Statistics* 55(1):28–45.

Cowing, T. G., and A. G. Holtmann (1983). "Multiproduct Short-Run Hospital Cost Functions: Empirical Evidence and Policy Implications from Cross-Section Data." *Southern Economic Journal* 49(3):637–653.

Dor, A., and D. E. Farley (1996). "Payment Source and the Cost of Hospital Care: Evidence from a Multiproduct Cost Function with Multiple Payers." *Journal of Health Economics* 15(1):1–21.

DummyCite (00) .

Gaynor, M., and W. B. Vogt (2003). "Competition among Hospitals." *RAND Journal of Economics* 34(4):764–785.

Gertler, P. J., and D. M. Waldman (1992). "Quality-Adjusted Cost Functions and Policy Evaluation in the Nursing Home Industry." *Journal of Political Economy* 10(6):1232–1256.

Geweke, J., G. Gowrisankaran, and R. J. Town (2003). "Bayesian Inference for Hospital Quality in a Selection Model." *Econometrica* 71(4):1215–1238.

Hausman, J. (1978). "Specification Tests in Econometrics." *Econometrica* 46(6):1251–1271.

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

Lakdawalla, D., and T. Philipson (2006). "The Nonprofit Sector and Industry Performance." *Journal of Public Economics* 90(8-9):1681–1698.

Luft, H. S., et al. (1990). "Does Quality Influence Choice of Hospital?" *Journal of the American Medical Association* 263(21):2899–2906.

Marschak, J., and W. H. Andrews, Jr. (1944). "Random Simultaneous Equations and the Theory of Production." *Econometrica* 12(3/4):143–205.

McClellan, M., and J. P. Newhouse (1997). "The Marginal Cost-Effectiveness of Medical Technology: A Panel Instrumental-Variables Approach." *Journal of Econometrics* 77(1):39–64.

Murphy, K. M., and R. H. Topel (1986). "Estimation and Inference in Two-Step Econometric Models." *Journal of Business and Economic Statistics* 3(4):370–379.

Nerlove, M. (1965). *Estimation and identification of Cobb-Douglas production functions*. Chicago, IL: Rand McNally.

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

Newhouse, J. P. (1994). "Frontier Estimation: How Useful a Tool for Health Economics?" *Journal of Health Economics* 13(3):317–322.

of Medicine, I. (2001). *Crossing the Quality Chasm*. Washington, DC: National Academies Press.

Picone, C. A., et al. (2003). "Does Higher Hospital Cost Imply Higher Quality of Care?" *Review of Economics and Statistics* 85(1):51–62.

Quan, H., et al. (2005). "Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data." *Medical Care* 43(11):1130–1139.

Stigler, G. J. (1976). "The Xistence of X-Efficiency." *American Economic Review* 66(1):213–216.

Stock, J. H., and M. Yogo (2005). "Testing for Weak Instruments in Linear IV Regression." *Identification and Inference in Econometric Models: Essays in Honor of Thomas Rothenberg*. D. W. Andrews and J. H. Stock, eds. New York: Cambridge University Press. 80–108.

Tay, A. (2003). "Assessing Competition in Hospital-Care Markets: The Importance of Accounting for Quality Differentiation." *RAND Journal of Economics* 34(4):786–814.

Town, R., and G. Vistnes (2001). "Hospital Competition in HMO Networks." *Journal of Health Economics* 20(5):733–753.

Zuckerman, S., J. Hadley, and L. Iezzoni (1994). "Measuring Hospital Efficiency with Frontier Cost Functions." *Journal of Health Economics* 13(3):255–280.