

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law. Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Translating Frontiers Into Practice: Taking the Next Steps Toward Improving Hospital Efficiency

Medical Care Research and Review
Supplement to 68(1) 3S–19S
© The Author(s) 2011

Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>
DOI: 10.1177/1077558710384878
<http://mcr.sagepub.com>



**Ryan L. Mutter¹, Michael D. Rosko²,
William H. Greene³, and Paul W. Wilson⁴**

Abstract

Frontier techniques, including data envelopment analysis (DEA) and stochastic frontier analysis (SFA), have been used to measure health care provider efficiency in hundreds of published studies. Although these methods have the potential to be useful to decision makers, their utility is limited by both methodological questions concerning their application, as well as some disconnect between the information they provide and the insight sought by decision makers. The articles in this special issue focus on the application of DEA and SFA to hospitals with the hope of making these techniques more accurate and accessible to end users. This introduction to the special issue highlights the importance of measuring the efficiency of health care providers, provides a background on frontier techniques, contains an overview of the articles in the special issue, and suggests a research agenda for DEA and SFA.

Keywords

efficiency, hospital, frontier approaches

Cost, access, and quality continue to be important problems confronting health care delivery in the United States. Total national health expenditures were \$1.35 trillion in 2000 (Hartman, Martin, Nuccio, Catlin, & the National Health Expenditure Accounts

¹Agency for Healthcare Research and Quality, Rockville, MD, USA

²Widener University, Chester, PA, USA

³New York University, New York, NY, USA

⁴Clemson University, Clemson, SC, USA

Corresponding Author:

Ryan L. Mutter, Center for Delivery, Organization and Markets, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, USA
Email: Ryan.Mutter@ahrq.hhs.gov

Team, 2010) and were estimated to be \$2.57 trillion in 2010, an increase of over 90% (Truffer et al., 2010). Hospital care mirrored this pattern, starting the period at \$416.9 billion and ending at \$788.9 billion, which was a growth of more than 80%. Health care expenditures as a percentage of gross domestic product (GDP) are expected to be 17.3 in 2010 (Truffer et al., 2010). This is much more than any other industrialized country. While the United States devotes a greater share of its economy to health care than other developed countries (Anderson, Reinhardt, Hussey, & Petrosyan, 2003), broad health indicators suggest that it may not all be money well spent. For example, among the 30 countries that participate in the Organization for Economic Cooperation and Development (OECD), the United States ranks 23rd in life expectancy at birth and 28th in infant mortality rate (Peterson & Burton, 2007). Anderson et al. (2003) indicate that the higher share in GDP allocated to health care in the United States is mostly due to much higher prices and found that most of the broad health services utilization figures in the United States were below the OECD median.

In 2008, there were more than 43 million Americans without health insurance coverage (Heyman, Barnes, & Schiller, 2009). Many more were without adequate health insurance. The high cost of premiums or premium contributions at work has created a financial barrier to both health insurance and health care. According to the 2008 Kaiser Survey, about 29% of the uninsured postponed health care because of cost considerations (Kaiser Commission on Medicaid and the Uninsured, 2009). In contrast, 7% of the insured postponed health care. Delaying health care can lead to complications requiring more expensive treatment, as well as premature death. Thus, there is evidence in health care of a vicious cycle of high costs, leading to poor access, which in turn leads to undesirable outcomes. While the Patient Protection and Affordable Care Act is estimated to help about 28 million Americans gain new health insurance coverage by 2016 (McGlynn, Cordova, Wasserman, & Girosi, 2010), the impact of the health reform legislation on costs and quality is less certain.

While it might be relatively easy to deal with one element of the trifecta of cost, access, and quality, dealing with all three simultaneously is a much more difficult challenge. Conceptually, one of the best approaches to addressing all three issues is to improve the efficiency of health care providers. For example, an increase in efficiency could reduce health care costs. These savings could be passed on to consumers in the form of reduced health insurance premiums or health care prices, which could then potentially increase access. Finally, more efficient processes could lead to better quality. (In the hospital literature, there is some evidence that more efficient institutions have better outcomes. See, e.g., McKay & Deily, 2008.) In the general literature, Deming (1982) called this the “value proposition” (i.e., simultaneously decreasing costs and increasing quality). It is a major tenet of total quality management.

Hospital Inefficiency Measurement With Frontier Techniques

The above discussion suggests the importance of improving efficiency. But to do so, it is essential to accurately measure it. Indeed, the following aphorism about the importance of measurement is frequently attributed to management thought leader

Peter Drucker: “If you cannot measure something you cannot control it. And if you cannot control it you cannot manage it.” The development of frontier methods, a set of techniques that measure inefficiency as the distance between a best practice frontier (BPF) and actual performance, has advanced the practice of efficiency measurement in health care.

While a variety of frontier techniques exist, data envelopment analysis (DEA) and stochastic frontier analysis (SFA) are the most frequently applied approaches. Farrell (1957) was the first to estimate productive efficiency as a distance from a BPF using linear programming methods. A significant breakthrough occurred when Charnes, Cooper, and Rhodes (1978) generalized Farrell’s single input/output measure to a multiple-input/multiple-output technique, which they termed *DEA*. Charnes, Cooper, Lewin, and Seiford (1994) viewed the development of two-stage analysis as a significant advance in DEA-based research. Combining nonparametric and parametric methods, researchers in the early 1990s began to explore the factors that determine inefficiency.

The first published health care application of DEA was by Wilson and Jadlow (1982). Nunamaker (1983) published the first DEA study of U.S. hospitals. The most recent literature review article reported that, as of mid-2006, 317 articles on frontier efficiency of health care organizations had been published. Of these, more than 200 used DEA and 57 used SFA. About 25 studies were classified as Malmquist-based (an extension of DEA) productivity studies (Hollingsworth, 2008).

The history of SFA begins with the near simultaneous publication of articles by Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977). A major breakthrough was accomplished by Jondrow, Lovell, Materov, and Schmidt (1982) who derived an estimation of one-sided residuals, interpreted as inefficiency scores, in a cross-sectional setting. This advance permitted the estimation of inefficiency for individual units. The ability to obtain producer-specific estimates of efficiency greatly enhanced the appeal of SFA (Kumbhakar & Lovell, 2000). In the 1990s, panel models were developed that relaxed the assumption of time-invariant efficiency (Battese & Coelli, 1992; Cornwell, Schmidt, & Sickles, 1990; Kumbhakar, 1990). Paralleling DEA, SFA studies initially used a two-stage approach to examine the exogenous factors associated with efficiency. This became outmoded with the development of more efficient (in a statistical sense) single-stage approaches that incorporate explanatory variables into the efficiency error component (Battese & Coelli, 1995; Kumbhakar, Gosk, & McGuckin, 1991).

The first SFA study of a health care organization was published by Wagstaff (1989), who examined 49 Spanish hospitals. Zuckerman, Hadley, and Iezzoni (1994) published the first SFA study of U. S. hospitals. Since then at least 27 other U.S studies of hospitals have been conducted. Their results are synthesized by Rosko and Mutter (2011) later in this Special Issue.

Agency for Healthcare Research and Quality (AHRQ) Conference and This Special Issue

In April 2008, the AHRQ released a commissioned report prepared by the RAND Corporation titled “Identifying, Categorizing, and Measuring Health Care Efficiency Measures” (McGlynn, 2008). This report noted that efficiency measurement techniques

developed in the academic literature, namely frontier techniques, have not been applied in the policy setting. McGlynn (2008) identified these measures, including SFA and DEA, as among the most promising approaches for measuring provider efficiency and for suggesting strategies for improving the delivery of health care services. Consequently, AHRQ hosted an invitational meeting, "Translating Frontiers Into Practice: Taking the Next Steps Towards Improving Efficiency," on August 27-28, 2008. This meeting brought together policy makers, stakeholders, and leading technical experts to discuss how frontier techniques can be used most effectively to address the problems confronting the health care system and to identify how the needs of end users should shape the agenda of the research community.

AHRQ followed the conference by sponsoring this special issue of *Medical Care Research and Review*. The aim of this special issue is to assess the state of the science of using frontier techniques to measure provider efficiency and to identify steps that can be taken to increase its accessibility to end users. The articles contained in it review the contributions made using these techniques, identify where gaps remain, and offer examples of the use of frontier approaches that hold promise for the future. An overview of the articles in this special issue follows.

Overview of Special Issue Articles

Ozcan and Luke

Ozcan and Luke (2011) used an extension of DEA, called the Malmquist technique, to examine the impact of a major restructuring that the Veterans Health Administration (VHA) implemented in 1995 on productivity change. The VHA decentralized decision making from VHA headquarters to regions and integrated services and assets regionally to create 21 regional providers called Veterans Integrated Service Networks (VISNs). The VHA's restructuring mirrors efforts in other countries to reorganize health care along regional lines.

The authors calculated a Malmquist Index, a measure of productivity change over time, which can be decomposed into technical change (i.e., change in productivity associated with an intertemporal shift in the best practice production frontier) and efficiency change (i.e., each unit's change or "catching up" in measured efficiency relative to each year's respective efficiency frontier). Another notable feature of their article was a demonstration of the use of the Malmquist technique for benchmarking. For example, they found that in 2004 12 VISNs were off the best-practice production frontier. Ozcan and Luke indicated how much the VISNs would need to reduce each of the four inputs in their model to reach the frontier.

Their article provides valuable insights into a methodology for analyzing innovations that are similar to the restructuring that occurred in the VHA. Looking to the future, the Patient Protection and Affordable Care Act is funding demonstration projects for the creation of Accountable Care Organizations (ACO; "Health Policy Brief," 2010). Ozcan and Luke's article provides a description of a methodology that might be useful for understanding the impact of the ACOs.

Bernet, Moises, and Valdmanis

Bernet, Moises, and Valdmanis (2011) used DEA to measure the social efficiency of hospital care by incorporating costs from the consumers' perspective. They did this by adding patient travel time to the DEA model. This is an important innovation because most evaluations of hospital cost containment programs consider only hospital expenditures and neglect to consider the opportunity costs of patients and their families (Folland, Goodman, & Stano, 2007). An example will highlight the effect of this omission. The Medicare Prospective Payment System (PPS) created strong incentives for hospitals to discharge their patients earlier. Consequently, they were less clinically stable when they left the hospital (Rosko & Broyles, 1988). Recognizing this, many patients were discharged to other post-acute care facilities, such as skilled nursing facilities or rehabilitation hospitals, while many patients who were discharged home were given a coordinated home care plan and received visits from home health nurses. However, it is highly likely that care provided by visiting nurses had to be augmented by informal care givers, such as family and friends. While the Medicare PPS saved billions of dollars in inpatient hospital expenditures (Folland et al., 2007), the "true" social saving are overstated because the increased cost of informal care was not included in the cost calculations. Bernet et al. partially overcame this type of specification error by modeling travel time as an input in the hospital care production process.

Methodologically, they made an interesting adaptation of the Malmquist technique. Rather than comparing changes over time as is done in the Malmquist approach, they compared the frontier differences by urgent versus nonurgent condition in an approach similar to that used by Grosskopf, Margaritis, and Valdmanis (2001). The authors concluded that including patient travel distance in the model yields a more comprehensive view of social welfare and social costs. The nuances identified by their methodology have the potential to yield interesting insights. For example, while their results suggest that public hospitals are less efficient than their private counterparts, when travel costs are included, the efficiency gap narrows. The authors suggest that this indicates that public hospitals provide an increased value-added by being closer to the population in need. In future applications, the methodology proposed by Bernet et al. could be valuable in the assessment of the social efficiency impact of hospital closures, mergers, and relocations.

Murphy, Rosenman, McPherson, and Friesner

Many of the studies in the frontier literature yield a single measure of inefficiency for a decision-making unit. However, Murphy, Rosenman, McPherson, and Friesner (2011) describe an approach that generates efficiency measures within the organization at the level of the hospital cost center. Their technique used a combination of DEA, spreadsheet modeling, and regression, and it took into account the sequential nature of hospital production rather than assuming, as many frontier studies do, that production takes place jointly. They assessed the extent of shared inefficiency across eight cost centers that encompass many of hospitals' activities, and they identified the cost centers within the hospitals that contribute the most to the shared inefficiency.

Murphy et al. applied their approach to a panel of mid- to large-size nonprofit hospitals in the State of Washington using Washington State Department of Health data from 2003 to 2007. They found evidence of considerable shared inefficiency across hospital cost centers. Their results suggest that three cost centers affect the total performance of the firm. They are plant, the emergency department, and an "other" category, which consists of housekeeping, laundry, nutrition, and central services.

Hospital managers and policymakers have sometimes struggled with the results of frontier studies because they do not know what to do in response to them. Understanding the relative performance of an institution can be useful, but decision makers want to know what they need to do to improve efficiency. The approach described by Murphy et al. offers insight into the performance of institutions (i.e., inefficiency is shared) and holds promise to help managers and policymakers design interventions to improve the efficiency of operations (e.g., focus on centers that are the greatest contributors to the overall inefficiency of the firm).

Although the methodology of Murphy et al. has some technically appealing aspects to it, it is also a fairly novel approach. As such, continued work is needed to maximize its potential. The technique does not have explicit statistical foundations, and Murphy et al. operate without a conceptual framework concerning the mechanism through which inefficiency among hospital cost centers is shared. Addressing these limitations through future research could add to this approach, which has the potential to help bridge the gap between frontier analysis and improved practice.

Rosko and Mutter

In the final article in this issue, Rosko and Mutter reviewed 27 studies where SFA has been applied in the analysis of U.S. hospitals. They used X-Inefficiency Theory as the framework for much of their discussion. X-Inefficiency Theory posits that environmental pressures, including internal and external factors, can influence the level of effort (and hence efficiency) given by firm managers. Rosko and Mutter reviewed where SFA has been used to examine the impact of these factors on hospital cost inefficiency.

The internal factors examined in the literature include ownership type, system membership, and specialty hospital status. Although inconsistent results were found for ownership type, the literature provided support that other internal characteristics, including system membership and the provision of general medical and surgical services, were associated with lower cost inefficiency. There were some important nuances related to these findings, which are discussed in the review article.

Public payment policy, critical access hospital (CAH) status, market unemployment rate (a proxy for uncompensated care), health maintenance organization (HMO) penetration, and hospital competition were the external factors analyzed in the literature. In general, higher Medicare and Medicaid share of inpatients, a higher unemployment rate, and greater HMO penetration were associated with lower hospital cost inefficiency. Conversion to CAH status was associated with greater cost inefficiency, and cost inefficiency increased, on average, with the number of years an institution participated in the CAH program. The findings with respect to hospital competition

were mixed. Overall, these results supported the contention of X-Inefficiency Theory that external pressures can shock organizations into better performance.

Rosko and Mutter reviewed other results from the hospital SFA literature of potential interest to policymakers and practitioners, including factors associated with changes in hospital inefficiency over time; the relationship of SFA-derived measures of inefficiency with other measures of hospital performance, including outcome measures of quality; the association between hospital inefficiency and hospital expansion and closure; and the application of SFA to hospital benchmarking and budgeting.

Since the usefulness of a study's findings are determined, in part, by the appropriateness of the methods employed, Rosko and Mutter also reviewed some of the recent advances in controlling for output heterogeneity. Unmeasured variations in output (i.e., type or quality of product) might result in overestimates of cost inefficiency. For example, if it costs more to treat more severely ill patients, and no variables for severity are included in the SFA model, then hospitals that tend to treat these more expensive types of patients will have higher cost-inefficiency estimates. This problem is compounded by the likelihood that certain types of hospitals (e.g., tertiary care hospitals) tend to attract patients with more complicated problems. The controls for hospital quality and patient burden of illness highlighted in the review have the potential to yield more accurate SFA estimates of hospital efficiency.

Future Directions

Much of the research presented in this special issue consists of innovative applications of frontier approaches that are suggestive of future directions that can be taken that could improve the estimates generated by these techniques and that could make their results more actionable by decision makers. To build on the directions indicated in these articles, the following section presents a research agenda for DEA and SFA.

A Research Agenda for DEA

While numerous references to DEA *models* can be found in the academic literature, DEA is but one of several *estimators* one might use in a nonparametric *model* of production. A nonparametric model of production consists of a set of minimal assumptions about the process that generates observed data on input and output quantities. Typically, these assumptions describe a production set consisting of all feasible combinations of input and output quantities; the production set is usually (but not always) assumed to be convex. A nonparametric model must also include assumptions about the probability process that generates observed data; for DEA estimators to be statistically consistent, one must assume that the joint probability density of inputs and outputs is zero everywhere outside the production set and strictly positive over at least some region of the production set. Other assumptions on the smoothness of the boundary (or frontier) of the production set may be made, but in a nonparametric model, no assumptions about the specific form of the frontier or the joint density of inputs and outputs are made. See Simar and Wilson (2000b, 2008) and Kneip, Simar, and Wilson (2008) for specific assumptions and additional details.

The DEA efficiency estimator provides an *estimate* of distance from a fixed point to the boundary of the production set along some path chosen a priori by the researcher. This amounts to measuring distance (typically, in the case of DEA estimators, by solving a linear program) from the fixed point of interest to a DEA estimate of the frontier. DEA estimates of either the frontier or efficiency involve uncertainty, creating a need for statistical inference; both the true frontier and hence the true level of efficiency are unobservable.

Although DEA estimators have been used since Farrell (1957), researchers have only recently derived the estimator's statistical properties. Under certain assumptions, the DEA frontier estimator is a consistent, maximum likelihood estimator (Banker, 1993), with rates of convergence given by Korostelev, Simar, and Tsybakov (1995). Consistency and convergence rates of DEA *efficiency* estimators have been established by Kneip, Park, and Simar (1998). See Simar and Wilson (2000b) for a survey of the statistical properties of DEA estimators.

It is important to note that *consistency*, while perhaps the most fundamental property of an estimator, is at the same time a very weak property. In particular, DEA estimators suffer from the curse of dimensionality in the sense that their convergence rates decrease with increasing numbers of dimensions (i.e., numbers of inputs and outputs). Consequently, for a given sample size, estimation error increases exponentially with dimensionality. In nonparametric regression, various approaches to dimension reduction have been developed to deal with this problem, but to date, it is not clear how such methods might be used with DEA estimators. Given that many applications of DEA estimators involve relatively small sample sizes, more research is needed in this area.

While DEA estimators have been widely used, inference about the underlying model structure or the efficiencies that are estimated is often ignored even today. In the past, inference was problematic since properties of DEA estimators were unknown and consistent bootstrap methods were not available, but this is no longer true today. Simar and Wilson (1998, 2000a) proposed bootstrap methods for inference about efficiency based on DEA estimators in a multivariate framework, and Simar and Wilson (2001a, 2001b) proposed bootstrap methods for testing hypotheses about the structure of the underlying nonparametric model of production, but consistency of these procedures has not been established. Banker (1993, 1996) proposed tests of model structure based on ad hoc distributional assumptions, but simulation results obtained by Kittelsen (1999) and Simar and Wilson (2001a) show that these tests perform poorly in terms of both size and power.

Kneip et al. (2008) and Park, Jeong, and Simar (2010) derived the limiting distributions of DEA efficiency estimators under variable returns to scale and constant returns to scale (respectively), with arbitrary numbers of inputs and outputs. These distributions contain several unknown quantities and are not useful in a practical sense for inference. Kneip et al. (2008) also proposed two bootstrap procedures for inference about efficiency and proved consistency of both methods. The first approach uses subsampling, where bootstrap samples of size $m < n$ are drawn (independently, with replacement) from the empirical distribution of the original n sample observations. Simulation results provided by Kneip et al. (2008) indicate that in finite-sample scenarios, coverage of confidence intervals for efficiency estimated by bootstrap subsampling are quite sensitive to the choice of the subsample size m . In a more recent article, Simar

and Wilson (2009) provide a data-based method for optimizing the choice of m ; Monte Carlo evidence indicates that the method works well in terms of coverage of estimated confidence intervals. Simar and Wilson (2009) also show how subsampling methods can be used for hypothesis testing in the DEA context and give examples including testing of returns to scale (i.e., constant vs. variable returns to scale). This should be especially useful in hospital studies, where there is some debate about whether returns to scale are constant. Moreover, it is easy to adapt the methods described by Simar and Wilson (2009) to other testing situations (e.g., testing whether for-profit (FP) hospitals differ in terms of efficiency from not-for-profit (NFP) hospitals).

The second, full-sample bootstrap procedure described by Kneip et al. (2008) requires for consistency not only smoothing of the distribution of the observations as proposed in Simar and Wilson (1998, 2000a) but also smoothing of the initial DEA estimate of the frontier itself, resulting in a formidable computational burden. Kneip, Simar, and Wilson (2009) developed a computationally efficient full-sample bootstrap method for inference; while the method works well for estimating confidence intervals for efficiency scores, it is not suitable for estimating critical values for test statistics. For the latter, the subsampling methods of Simar and Wilson (2009) appear to be more promising for most applied researchers.

When DEA estimators are used to estimate Malmquist indices, infeasibilities often arise when (estimated) frontiers from two time periods (or two groups) cross each other. Using the hyperbolic efficiency estimator introduced by Färe, Grosskopf, and Lovell (1985) would avoid this problem, but until very recently, this estimator has only been used while imposing constant returns to scale to avoid computational difficulties. Wilson (2010) provides a simple numerical procedure for computing hyperbolic DEA estimates and derives asymptotic properties for the hyperbolic DEA estimator.

DEA estimators are known to be quite sensitive to outliers; apart from the curse of dimensionality, this may be the most serious drawback to using DEA estimators, particularly with hospital data in which outliers are common. One approach is to use outlier detection methods to find the outliers. An alternative approach is to measure efficiency relative to an estimated *partial* frontier instead of an estimate of the full frontier as with DEA.

Cazals, Florens, and Simar (2002) developed the notion of an order- m frontier. In this approach, efficiency is measured in the input direction by comparing a given hospital's input usage to an estimate of the expected minimum input usage among m hospitals producing at least as much of each output as the given hospital of interest. In the output orientation, efficiency is measured by comparing a hospital's output usage to an estimate of the expected maximum output among m hospitals using at least as much of each input as the hospital of interest. This approach has been shown to be robust with respect to outliers; in addition, the usual, parametric rate of convergence is obtained, avoiding the curse of dimensionality, although the variance of the estimator is affected by the number of inputs and outputs. The order- m efficiency estimator is asymptotically normally distributed, provided m is not too large. Wilson (2010) extended the order- m idea to the hyperbolic orientation.

Daouia (2003), Aragon, Daouia, and Thomas-Agnan (2005), and Daouia and Simar (2007) developed input- and output-oriented conditional order- α estimators and derived

their asymptotic properties. These results were subsequently extended to an unconditional, hyperbolic order- α estimator by Wheelock and Wilson (2008). In each case, efficiency is measured relative to a quantile of the joint distribution of inputs and outputs lying close to the full frontier, as opposed to the full frontier itself. This has advantages similar to those of the order- m approach (i.e., the estimators attain the parametric rate of convergence while avoiding the curse of dimensionality, are robust with respect to outliers, and are asymptotically normally distributed).

Only a few applications of the order- m and order- α estimators have been attempted at the time of this writing, and most of these have been in banking (e.g., Wheelock & Wilson, 2008, 2009). Given that hospital data often contain outliers, and the sensitivity of DEA estimators to outliers, these methods seem promising for researchers in health care. Moreover, the methods have been included in the FEAR (Frontier Efficiency Analysis with *R*) package for use with *R* (See Wilson, 2008, for details.) This makes them more accessible for applied researchers.

Regardless of whether DEA, partial frontier, or even stochastic frontier estimators are used, one must decide which variables to include in the model and what their role should be. Nonparametric statistical methods are available for testing whether a particular variable acts as an input or output (Simar & Wilson, 2001b) or how environmental variables might play a role (Daraio, Simar, & Wilson, 2010). In health care research, it seems important to control for “quality,” but it is less clear how quality enters the production process. The available testing methods may be useful here. While most would agree that quality is an important variable, there is less agreement on how quality should be measured. Necessarily, anything that is to be measured must be clearly defined, but unfortunately, discussions of health care quality often proceed without a clear definition of what is being measured. Part of the difficulty surrounding notions of quality relates to data availability; presumably, quality is related to health outcomes, but data on patients’ overall level of health are often sketchy at best. With regard to quality, more research is needed on all sides, including theoretical econometricians and statisticians, data collectors, health care practitioners, and empirical researchers.

A Research Agenda for SFA

The stochastic frontier model builds on a production/cost function view of the process of hospital care. The starting point of the analysis is a decision on the level of aggregation that is of interest, possibly the hospital as a whole or some functional unit within it, such as the emergency department. This is both a matter of appropriate model selection and a question of the interests of the stakeholder engaged in the research.

Although the analysis can theoretically take place with respect to production or costs, in the hospital care setting, it is the latter that is usually of greater interest. Indeed, that this is the primary dimension on which DEA and SFA will differ—DEA will provide results about the process of production while SFA will focus immediately on the cost consequences of production. One of the important avenues of research will be to aggregate and to reconcile these two sources of information.

A natural place to begin the search for best practice is total cost. A simple comparison of “average” costs of hospitals immediately hits several obstacles. Costs differ greatly across hospitals for many reasons having little to do with relative success. Hospitals differ on dimensions that will substantively affect costs, such as the setting—urban/rural, local labor market conditions—and case mix, whether mainly emergency/outpatient or inpatient, and with respect to the latter, which types of care are mostly provided and whether the hospital is involved in teaching as well as treatment or only the latter. This presents two challenges to the analyst: (a) determining what to use as the output (i.e., the activity level of the hospital [patient days, discharges, etc.]) and (b) to distinguish clearly the aspects of the cost consequences of these activities that are the target of the study from the features of the hospital and its activity mix that are not related to efficiency.

The overall purpose of frontier modeling, as discussed at several other points in this issue, is to compare hospitals (their costs), not directly with respect to their costs, but rather on the level of costs compared with what they could be. Thus, there are two degrees of comparison: First, in absolute terms, efficiency as measured with respect to how well a given hospital does compared with how well it could do, and second, in relative terms, with respect to how well hospitals do compared with each other.

The stochastic frontier is one of two commonly used methods that are specifically designed to measure costs from the first viewpoint above, then to compare firms (hospitals) on the second basis as a corollary. The methods differ on specific assumptions about the form of the technology and how deviations from the frontier (best practice) arise and are ultimately measured. The stochastic frontier model assumes that best practice differs across hospitals both systematically and randomly, in ways that are not always measurable. The technique seeks to measure efficiency against a hospital-specific benchmark that accommodates differences such as those noted above (e.g., case mix). A comparison of hospitals then measures them in terms of a frontier that recognizes the differences that substantially affect costs. Superficially, the approach makes sense. A hospital that treats an inherently expensive case mix of patients is not inefficient compared with one that treats a less expensive case mix *because of the case mix per se*, though it may be so (or not) for other reasons. Because there is vast heterogeneity in hospital types and in the settings (markets) in which they operate, an important element of the ongoing research on hospital cost-efficiency relates to the appropriate ways to handle these differences. Where these elements can change over time, such as in local labor markets, repeated observations of hospitals that can reveal dynamic effects becomes yet another feature in the ongoing research.

What follows is a consideration of some specific dimensions of the ongoing research about hospital costs. There are two large issues that impinge on the type of analysis that can be done and which, ultimately, will color what can be expected of a cost study however carefully done: ownership type and the impact of “quality” on hospital costs.

Hospitals differ discretely between three ownership types: government, FP, and NFP. This discussion will focus on differences between FP and NFP institutions. It is distinctly possible (though not inevitable) that cost structures and outcomes will

differ substantively across these two types. In a statistical setting, it might seem appropriate to perform separate analyses of these two broad types of hospitals. However, this approach might turn up what appears to be differences in efficiency aspects of hospital costs that are actually related to differences between FP and NFP types that are not what they appear superficially to be. This is known as a “sample selection” problem. Unfortunately (for the analyst), the differences between FP and NFP may well be real, but the conventional analytical methods are inappropriate for uncovering those differences. A second related, though possibly of lesser consequence, dimension of hospital environment is whether the hospital is a member of network or a multihospital system. Compounding this is the existence of a variety of different types of system structures (Bazzoli, Shortell, Dubbs, Chan, & Kralovec, 1999). Zuckerman et al. (1994) provide some empirical evidence and wrestle with the issue of using pooled versus group-specific frontiers. (A disadvantage of group-specific frontiers is that inefficiency cannot be compared across groups.) This topic could benefit from further consideration by the field.

A significant component of the costs of hospital patient care is “quality,” broadly defined. Fundamentally, quality is not directly measurable—it is impossible to agree on an objective measure—but it is obviously a substantive element of the cost structure. It is an important question at the outset of the discussion whether quality is something that hospitals produce along with other outputs, as measured by inpatient days or outpatient discharges, or whether quality is an input into the process, along with materials, labor (nurses, doctors, technicians, etc.), and capital (beds, etc.). Properly accounting for quality of care might reasonably be viewed as the signature open-end in contemporary research on hospital costs. Regardless of how quality is measured and interpreted, its presence in the cost structure of the hospital implies that cost-efficiency always embodies this hidden feature of production. For example, the most natural outcome will be that unaccounted for investment in quality will show up in a cost model as if it were inefficiency, a distinctly unappealing result if one seeks a cost basis for measuring (in)efficiency. Recent and ongoing research in econometric methodology has considered this problem of how to accommodate unmeasured heterogeneity, such as quality, in hospital costs. One aspect of this study is how best to make use of observable proxies for quality such as investments in technology (i.e., structures), institution of broad procedure regimes (i.e., processes), and in-hospital mortality (i.e., outcomes; Romano & Mutter, 2004).

Finally, there are many fine, but still substantive, technical points about the model building itself that the researcher will want to consider. The search for the right model is ultimately the search for a model specification that will properly isolate inefficiency from the production process itself and both observed (e.g., at least to some extent, case mix) and unobserved (e.g., quality) heterogeneity from the measurement of inefficiency. One such feature of the model is the way that it handles the inevitable variation across hospitals in the technical parameters of the model equations. These are econometric issues including, for example, “heteroscedasticity” and “parameter heterogeneity,” that impinge on the ability of the model to produce appropriate estimates of

hospital-efficiency measures. One of the largely unanswered (satisfactorily) questions in the broad field of efficiency analysis is the extent to which departures from purity in the structural features of the cost model are propagated in the implied estimates of quantities of interest, particularly measures of efficiency. In short, can the imperfect model be trusted to provide useful results in spite of its imperfections? The answer to this question is a matter of degree, not of type.

The preceding suggests an agenda for the researcher interested in formulating a modeling framework for measuring efficiency in hospital care. There remains a detail that, ultimately, is of crucial importance. That is, packaging the results of the stochastic frontier modeling in a way that is useful for the policy maker or manager. Stochastic frontier results come in the form of quantitative efficiency measures, for example, a hospital- (and possibly time-) specific measure of the percentage difference between what is achieved and what appears to be theoretically possible. There is a question of the statistical uncertainty in the estimates. It may be appropriate, for purposes of overall policy analysis, to suggest a credible range of values for this quantity rather than a single number. If so, the methodology remains incomplete on how such a range should be constructed. Finally, in the search for best practice, it may evolve that ranks of hospitals, not hospital-specific estimates are the measures of interest. This implicitly suggests that the best practice is determined in the field, not in the abstract. Turning the analysis in this direction will then call upon the researcher to continue the study to pinpoint more precisely the features of those hospitals that provide the standards for best practice.

Conclusion

The application of DEA and SFA in hundreds of published studies over the past several decades indicates the need for techniques to measure efficiency in the health care sector. It also suggests the usefulness of frontier methods. Nevertheless, these approaches are not without drawbacks. Some important methodological questions remain unanswered, and the accessibility of the insights from these methods to end users needs to be increased. The articles in this special issue are intended to make contributions in both these areas, and the research agendas for DEA and SFA are intended to highlight where particular gaps in understanding remain and to suggest where future work might be particularly useful.

Authors' Note

This article does not represent the policy of either the Agency for Healthcare Research and Quality (AHRQ) or the U.S. Department of Health and Human Services (DHHS). The views expressed in this article are those of the authors and no official endorsement by AHRQ or DHHS is intended or should be inferred.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

Funding

This research was partially funded by contracts from the Agency for Healthcare Research and Quality (AHRQ): AHRQ Contract Number HHSP233200800267P (Michael Rosko) and AHRQ Contract Number HHSN263999900147B, Task Number 2516 (William Greene, Paul Wilson).

References

- Aigner, D., Lovell, C., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production models. *Journal of Econometrics*, 6, 21-37.
- Anderson, G., Reinhardt, U., Hussey, P., & Petrosyan, V. (2003). It's the prices, stupid: Why the United States is so different from other countries. *Health Affairs*, 22, 89-105.
- Aragon, Y., Daouia, A., & Thomas-Agnan, C. (2005). Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory*, 21, 358-389.
- Banker, R. (1993). Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. *Management Science*, 39, 1265-1273.
- Banker, R. (1996). Hypothesis tests using data envelopment analysis. *Journal of Productivity Analysis*, 7, 139-159.
- Battese, G., & Coelli, T. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis*, 3, 153-169.
- Battese, G., & Coelli, T. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20, 325-332.
- Bazzoli, G., Shortell, S., Dubbs, N., Chan, C., & Kralovec, P. (1999). A taxonomy of health networks and systems: Bringing order out of chaos. *Health Services Research*, 33, 1683-1717.
- Bernet, P., Moises, J., & Valdmanis, V. (2011). Social efficiency of health care delivery: Frontier analysis from the consumer's perspective. *Medical Care Research and Review*, 68S, 36S-54S.
- Cazals, C., Florens, J., & Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics*, 106, 1-25.
- Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444.
- Charnes, A., Cooper, W., Lewin, A., & Seiford, L. (1994). *Data envelopment analysis: Theory, methodology, and application*. Boston, MA: Kluwer Academic.
- Cornwell, C., Schmidt, P., & Sickles, R. (1990). Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics*, 46, 185-200.
- Daouia, A. (2003). *Nonparametric analysis of frontier production functions and efficiency measurement using nonstandard conditional quantiles* (Unpublished doctoral dissertation). Groupe de Recherche en Economie Mathématique et Quantitative, Université des Sciences Sociales, Toulouse I, et Laboratoire de Statistique et Probabilités, Université Paul Sabatier, Toulouse III, France.
- Daouia, A., & Simar, L. (2007). Nonparametric efficiency analysis: A multivariate conditional quantile approach. *Journal of Econometrics*, 140, 375-400.

- Daraio, C., Simar, L., & Wilson, P. (2010). *Testing whether two-stage estimation is meaningful in non-parametric models of production* (Discussion Paper No. 1031). Louvain-la-Neuve, Belgium: Institut de Statistique, Université Catholique de Louvain.
- Deming, W. (1982). *Out of the crisis*. Cambridge: MIT Press.
- Färe, R., Grosskopf, S., & Lovell, C. (1985). *The measurement of efficiency of production*. Boston, MA: Kluwer-Nijhoff.
- Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 120, 253-290.
- Folland, S., Goodman, A., & Stano, M. (2007). *The economics of health and health care* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Grosskopf, S., Margaritis, D., & Valdmanis, V. (2001). Comparing teaching and non-teaching hospitals: A frontier approach (teaching vs. non-teaching hospitals). *Health Care Management Science*, 4, 83-90.
- Hartman, M., Martin, A., Nuccio, O., Catlin, A., & the National Health Expenditure Accounts Team. (2010). Health spending growth at a historic low in 2008. *Health Affairs*, 29, 147-155.
- Health policy brief: Accountable care organizations (2010, July 27). *Health Affairs*. Retrieved from <http://www.healthaffairs.org/healthpolicybriefs>
- Heyman, K., Barnes, P., & Schiller, J. (2009). *Early release of selected estimates based on data from the 2008 Health Interview Survey*. Atlanta, GA: National Center for Health Statistics, Centers for Disease Control and Prevention. Retrieved from <http://www.cdc.gov/nchs/nhis.htm>
- Hollingsworth, B. (2008). The measurement of efficiency and productivity of health care delivery. *Health Economics*, 17, 1107-1128.
- Jondrow, J., Lovell, K., Materov, I., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19, 233-238.
- Kaiser Commission on Medicaid and the Uninsured. (2009). *The uninsured and the difference health insurance makes*. Washington, DC: Kaiser Family Foundation.
- Kittelsen, S. (1999). *Monte Carlo simulations of DEA efficiency measures and hypothesis tests* (Unpublished Working Paper, Memorandum No. 09/99). Oslo, Norway: Department of Economics, University of Oslo.
- Kneip, A., Park, B., & Simar, L. (1998). A note on the convergence of nonparametric DEA efficiency measures. *Econometric Theory*, 14, 783-793.
- Kneip, A., Simar, L., & Wilson, P. (2008). Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models. *Econometric Theory*, 24, 1663-1697.
- Kneip, A., Simar, L., & Wilson, P. (2009). A computationally efficient, consistent bootstrap for inference with nonparametric DEA estimators (Discussion Paper No. 0903). Louvain-la-Neuve, Belgium: Institut de Statistique, Université Catholique de Louvain.
- Korostelev, A., Simar, L., & Tsybakov, A. (1995). On estimation of monotone and convex boundaries. *Publications de l'Institut de Statistique de l'Université de Paris*, 39, 3-18.
- Kumbhakar, S. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics*, 46, 201-212.

- Kumbhakar, S., Gosk, S., & McGuckin, J. (1991). A generalized production frontier approach for estimating determinants of inefficiency in US dairy farms. *Journal of Business and Economic Statistics*, 9, 279-286.
- Kumbhakar, S., & Lovell, C. (2000). *Stochastic frontier analysis*. Cambridge, England: Cambridge University Press.
- McGlynn, E. (2008). *Identifying, categorizing, and evaluating health care efficiency measures* (AHRQ Publication No. 08-0030). Washington, DC: Agency for Healthcare Research and Quality.
- McGlynn, E., Cordova, A., Wasserman, J., & Girosi, F. (2010). Could we have covered more people at less cost? Technically, yes; politically, probably not. *Health Affairs*, 29, 1142-1146.
- McKay, N., & Deily, M. (2008). Cost inefficiency and hospital health outcomes. *Health Economics*, 17, 833-848.
- Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18, 435-444.
- Murphy, S., Rosenman, R., McPherson, M., & Friesner, D. (2011). Measuring shared inefficiency between hospital cost centers. *Medical Care Research and Review*, 68S, 55S-74S.
- Nunamaker, T. (1983). Measuring routine nursing services efficiency: A comparison of cost per patient day and data envelopment analysis models. *Health Services Research*, 18, 183-205.
- Ozcan, Y., & Luke, R. (2011). Health care delivery restructuring and productivity change: Assessing the Veterans Integrated Service Networks (VISNs) using the Malmquist approach. *Medical Care Research and Review*, 68S, 20S-35S.
- Park, B., Jeong, S., & Simar, L. (2010). Asymptotic distribution of conical-hull estimators of directional edges. *Annals of Statistics*, 38, 1320-1340.
- Peterson, C., & Burton, R. (2007). *U.S. health care spending: Comparison with other OECD countries* (Order Code RL34175). Washington, DC: Congressional Research Service.
- Romano, P., & Mutter, R. (2004). The evolving science of quality measures for hospitals: Implications for studies of competition and consolidation. *International Journal of Health Care Finance and Economics*, 4, 131-157.
- Rosko, M., & Broyles, R. (1988). *The economics of health care: A reference handbook*. New York, NY: Greenwood Press.
- Rosko, M., & Mutter, R. (2011). What have we learned from the application of stochastic frontier analysis to U.S. hospitals? *Medical Care Research and Review*, 68S, 75S-100S.
- Simar, L., & Wilson, P. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44, 49-61.
- Simar, L., & Wilson, P. (2000a). A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics*, 27, 779-802.
- Simar, L., & Wilson, P. (2000b). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis*, 13, 49-78.
- Simar, L., & Wilson, P. (2001a). Nonparametric tests of returns to scale. *European Journal of Operational Research*, 139, 115-132.
- Simar, L., & Wilson, P. (2001b). Testing restrictions in nonparametric efficiency models. *Communications in Statistics*, 30, 159-184.

- Simar, L., & Wilson, P. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives. In H. Fried, C. Lovell, & S. Schmidt (Eds.), *The measurement of productive efficiency* (2nd ed., pp. 421-521). Oxford, England: Oxford University Press.
- Simar, L., & Wilson, P. (2009). *Inference by subsampling in nonparametric frontier models* (Discussion Paper No. 0933). Louvain-la-Neuve, Belgium: Institut de Statistique, Université Catholique de Louvain.
- Truffer, C., Keehan, S., Smith, S., Cylus, J., Sisko, A., Poisal, J., Lizonitz, J., & Clemens, M. (2010). Health spending projections through 2019: The recession's impact continues. *Health Affairs*, 29, 522-529.
- Wagstaff, A. (1989). Estimating efficiency in the hospital sector: A comparison of three statistical cost frontier models. *Applied Economics*, 21, 659-672.
- Wheelock, D., & Wilson, P. (2008). Non-parametric, unconditional quantile estimation for efficiency analysis with an application to Federal Reserve check processing operations. *Journal of Econometrics*, 145, 209-225.
- Wheelock, D., & Wilson, P. (2009). Robust nonparametric quantile estimation of efficiency and productivity change in U. S. commercial banking, 1985–2004. *Journal of Business and Economic Statistics*, 27, 354-368.
- Wilson, P. (2008). FEAR: A software package for frontier efficiency analysis with R. *Socio-Economic Planning Sciences*, 42, 247-254.
- Wilson, P. (2010). Asymptotic properties of some non-parametric hyperbolic efficiency estimators. In I. van Keilegom & P. Wilson (Eds.), *Exploring research frontiers in contemporary statistics and econometrics*. Berlin, Germany: Physica-Verlag.
- Wilson, P., & Jadow, J. (1982). Competition, profit incentives, and technical efficiency in the provision of nuclear medicine services. *Bell Journal of Economics*, 13, 472-482.
- Zuckerman, S., Hadley, J., & Iezzoni, L. (1994). Measuring hospital efficiency with frontier cost functions. *Journal of Health Economics*, 13, 255-280.