The role of safe practices in hospitals’ total factor productivity

Abstract: The dual aims of improving safety and productivity are a major part of the health care reform movement hospital leaders must manage. Studies exploring the two phenomena conjointly and over time are critical to understanding how change in one dimension influences the other over time. A Malmquist approach is used to assess hospitals’ relative productivity levels over time. Analysis of variance (ANOVA) algorithms were executed to assess whether or not the Malmquist Indices (MIs) correlate with the safe practices measure. The American Hospital Association’s annual survey and the Centers for Medicare and Medicaid Services’ Case Mix Index for fiscal years 2002–2006, along with Leapfrog Group’s annual survey for 2006 were used for this study. Leapfrog Group respondents have significantly higher technological change (TC) and total factor productivity (TFP) than nonrespondents without sacrificing technical efficiency changes. Of the three MIs, TC ($P < 0.10$) and TFP ($P < 0.05$) had significant relationships with the National Quality Forum’s Safe Practices score. The ANOVA also indicates that the mean differences of TFP measures progressed in a monotonic fashion up the Safe Practices scale. Adherence to the National Quality Forum’s Safe Practices recommendations had a major impact on hospitals’ operating processes and productivity. Specifically, there is evidence that hospitals reporting higher Safe Practices scores had above average levels of TC and TFP gains over the period assessed. Leaders should strive for increased transparency to promote both quality improvement and increased productivity.

Keywords: safety, productivity, quality, safe practice, cost, operations

Introduction

Health care inflation remains a major policy issue in the United States and other countries despite decades of attempts to control the problem. Increased expenditure of hospitals represent a significant portion of the overall spending and is a frequent target for policymakers attempting to ‘bend’ the inflation curve. In the best of cases, reforms provided only a temporary or one-time reduction in inflation rates as hospitals found new ways to profit without necessarily changing their underlying care processes. As a result, policymakers have been pressuring the Department of Health and Human Services to institute hospital reimbursement systems designed to improve efficiency, accelerate TC, and increase productivity through value-based purchasing (VBP) programs. However, such efforts to control health care costs have been shown to come at the expense of quality in some settings.

The interest in VBP has led to numerous studies focusing on hospital inefficiency and the environmental and cultural conditions that make change difficult. In particular, researchers of health services have applied frontier estimation techniques...
such as Data Envelopment Analysis (DEA)\textsuperscript{12,13} and Stochastic Frontier Analysis\textsuperscript{14} to measure hospitals’ X-inefficiency levels.\textsuperscript{15} Given the pressure to be more productive, not just more efficient, it is critical that the entire productivity function be analyzed if health system leaders are going to make meaningful changes to their organizations.\textsuperscript{16}

The purpose of this study is to examine the 5-year period following the release of the Institute of Medicine\textsuperscript{17–20} reports that describe process failures in US hospitals and the efforts to address such failures.\textsuperscript{21} This study’s plan of work involves four steps. First, the Malmquist Indices (MIs) for measuring Technical Efficiency Change (EFFCH), Technological Change (TC), and Total Factor Productivity (TFP) over time are introduced and defined. Next, an empirical analysis of US hospitals from 2002 to 2006 (5 years) is conducted to measure changes in efficiency, technological process change, and total productivity. Third, the MIs are correlated with a measure of quality drawn from the Leapfrog Group’s annual survey. Finally, the dynamic interactions between efficiency change and TC and their impact on overall productivity are discussed in terms of the value proposition\textsuperscript{10} and the productivity paradox.\textsuperscript{22}

TFP analyses can be linked to the use of discreet strategies for improving quality and their potential impact on operating efficiencies through TC. Further, the analyses detail individual hospital’s performance over time in a manner that managers can use to assess how their efforts to promote quality and patient safety impact efficiency. For policymakers and purchasers, having indices that measure both Safe Practices and TFP is ideal for developing VBP programs designed to simultaneously address two of the most intractable problems in health care: cost and quality.

**Conceptual framework**

Calculating a productivity index is relatively simple when a single output is produced from a single input. In such cases, productivity is generally defined as the output created per unit input. Hospitals, however, use many inputs that can be configured in numerous ways to produce multiple outputs. Further, the value of a given hospital’s services is measured relative to the quality and price of other facilities’ services delivered in the same market. Therefore, measures of value and performance among facilities are inherently comparative in nature.

Hospital performance is a relative concept that can be measured in three ways. The performance of a hospital, or any organization for that matter, can be measured relative to its own prior performances over time. Alternatively, hospitals’ performances can be measured relative to other facilities in the same period. Finally, the two strategies can be combined and facilities can be compared to one another over multiple time periods. The latter strategy is desirable because sustained performance is a necessary goal for any organizational leader, and that performance is inherently linked to its ability to succeed in the marketplace. Therefore, having a longitudinal assessment of multiple organizations’ performances increases analyses rigor and explanatory power. Frontier analyses provide a means for examining these phenomena.

Frontier analyses attempt to find organizations that transform inputs into outputs in an optimal fashion. Having identified the ‘frontier’ at varying levels of input, the analyses then assess other firms’ distances from the frontier. The distance from the frontier is a measure of relative efficiency, more specifically relative inefficiency among competing firms.

DEA can provide a comprehensive picture of the operation at the hospital level. This approach can accommodate multiple inputs and outputs simultaneously without requiring a priori knowledge on their relative importance, and the inputs and outputs can be in different units of measurement. Such attributes increase the applicability and practicality of this method. DEA provides a single efficiency score for each unit and, in doing so, identifies a frontier comprised of the best-performing organizations with the other units being at some measured distance from the frontier, making comparisons to other units easy. Finally, the MIs provide information on efficiency, technology, and productivity that can be linked to other performance metrics over time and among competing facilities.

This article extends methodologies of earlier studies for evaluating hospitals’ productivity performance using frontier analysis. It does so by calculating the MIs to assess hospitals’ relative productivity levels. The MIs are part of a family of indices that calculate productivity against a summative input and output data using a nonparametric frontier approach.\textsuperscript{23} The MIs were introduced by Caves et al\textsuperscript{24} as a method for measuring productivity when output pricing information is not available. We know this to be the case in health care, as contracting arrangements with insurance companies and Medicare fee structures routinely mask true reimbursement rates relative to the largely fictional ‘charges’ the consumer usually sees. Malmquist analyses are also useful when the objectives of the firms are not known, have not been achieved, or differ from firm to firm. In the health care sector, the mix of for-profit, nonprofit (often religiously affiliated),
and community ownership arrangements leads to differing strategies and goals for the organizations. Uri25 argued that the Malmquist approach was clearly superior for market- or industry-level assessment under such conditions.

The radial distance feature is a powerful advantage of the Malmquist model, in that it combines the geometric means of two other productivity indices. The DEA produced three MIs: EFFCH, TC, and TFP. The TFP can, therefore, be ‘decomposed’ to the geometric means of EFFCH and TC as described in Eq. 1. Further, the function is constructed so that degradation in performance manifests itself as an index of <1, whereas an improvement in performance is presented by a TFP score greater than unity (ie, one).a

\[
TFP = \sqrt{EFFCH \times TC} .
\]

The TFP is a variable that accounts for changes in total output not caused by varying the amount of inputs. Over time, firms are generally able to increase productivity in one of three ways: 1) they can increase outputs while holding inputs constant; 2) they can produce the same level of outputs while reducing the amount of inputs; or 3) they can simultaneously increase outputs and decrease inputs. Such changes often arise in the form of automation where the human resource input is decreased while the outputs stay constant or increase. TFP is often seen as the real driver of growth within an economy or an organization. Some comparisons of national economies indicate that TFP may account for up to 60% of growth over time.27 Therefore, understanding TFP, its parts and its impact, is an important aspect of changing the way hospitals do business.

The EFFCH (ie, X-inefficiency) component, on the other hand, is an indicator of the ‘management effect’ on organizational performance. Such inefficiency occurs when more of each input is used than required for a given level of output, which is often due to insufficient competitive pressures that allow management to ‘get away’ with subpar productivity or, as Fare et al28 describe, catching-up with the productivity frontier.

The TC component represents organizational innovation due to improvements in the ‘technology’ of organizational processes. The term ‘technology’ is general in nature and does not just refer to information or physical forms of innovations.29 The introduction of the M-form corporate structure is a technological innovation.30 In addition, behavioral modifications are often required to yield significant TCs in an organization and represent a form of TC in and of themselves. For example, the implementation of safe practices and an organization’s culture may also affect productivity.31 Collectively evaluating changes in hospitals’ efficiency and technological acumen are thus critical to evaluating improvements in the productivity of such facilities.

While improvements in productivity yield a TFP index greater than unity, similar improvements in any of the component parts (EFFCH and TC indices) are also associated with values >1. It is important to note that the component parts can move in opposite directions. For example, a hospital may have a TFP index greater than unity signaling a gain, but could have a decline in innovation (indicated by TC value <1) and a gain in EFFCH indicated by index >1.

This decomposition provides insights into the impact of safe practices to spur increases in efficiency, innovation, and productivity.

Methods
Three data sources were employed for the analyses: American Hospital Association’s (AHA’s) annual survey (N = 5500 hospitals3) and Centers for Medicare and Medicaid Services’ Case Mix Index (CMI) for fiscal years 2002–2006, along with Leapfrog Group’s annual survey for 2006 provided the National Quality Forum’s (NQF’s) Safe Practices measure. The identification numbers of AHA were used for merging cases across the various datasets. In order to be included in the MIs analyses, it was necessary to limit to facilities with complete data for the entire panel span. Altogether, 2978 hospitals had complete responses for the 5 years included in the study using variables drawn from the AHA survey in stage one of the analysis. Among the facilities analyzed in the first stage, 1470 responded to the Leapfrog Group’s survey.

Two additional features of the Leapfrog survey further reduced the number of matched cases on some variables. First, responding to the survey is voluntary, and some facilities elected not to answer some items. Second, not every facility was ‘targeted’ for every intervention for a variety of reasons. The stringent casewise deletion criterion significantly reduced the number of observations in the second

---

aThe number of hospitals in the AHA Survey varies from year to year and typically averages ∼5500 facilities.

bThe Leapfrog Group targets larger facilities that perform a high volume of surgical procedures, thus creating a smaller subsample for comparison.
set of analyses to 860 facilities common to both the Frontier Analyses dataset and Leapfrog survey respondents.

This study draws upon Carey’s\(^\text{12}\) input/output DEA model. In particular, the average annual salary and number of beds for each hospital were used as inputs, and the adjusted admissions and adjusted patient days were used as outputs. However, a slight modification to the model’s specification was made by making CMI an input variable similar to the approach used in other recent studies.\(^3\) This perspective views hospitals as taking patients into the system and possessing them until discharge. As a result, we define the inputs of hospitals by the money invested into the people and facilities (total facility expenditures), the number of beds available at the facility (as a measure of capacity), and a metric reflecting the clinical complexity of the patient population admitted – measured using the CMI. In terms of outputs, we assume that adjusted patient days (‘adjusted’ patient days of care taken into account the outpatient care provided by the hospital because staffing level data does not distinguish between inpatient and outpatient staffing) and number of admissions are appropriate output measures. Both measures are indicative of organizational throughput. The former establishes a metric related to revenue for the organization, while the latter not only measures input, but also measures output under the expectation that individuals admitted are discharged as well.

The respective 5-year average, standard deviation, and value range for each of the three MIs (ie, EFFCH, TC, and TFP) are presented in Table 1. These variables were then compared to the Leapfrog Group’s measure of patient safety assessments. In order to assess whether or not the MIs means differed significantly from the NQF’s Safe Practices Composite Score (SP) measure, analysis of variance (ANOVA) algorithms were executed using SPSS software (v. 16.0 for Mac; SPSS Inc., Chicago, IL). Hospitals were placed into five equal-sized groups based on their productivity index score in order to make the analyses. Similar approaches to evaluating productivity are not new;\(^3\) however, this is the first such effort to link longitudinal productivity using MI with other measures of quality. Table 2 provides a comparison of the Leapfrog Group sample and the balance of the facilities assessed in stage one (ie, the DEA). The SP measure is added to the Malmquist information in the second stage of the analysis. This is a composite measure of 27 activities that create a culture of safety and infection control practices. The scores ranged from a low of 44 to 1000 with a mean of 869.62. Because the scores were skewed toward reporting high compliance levels, the variable was recoded into quintiles with value ‘1’ being the lowest attaining group of hospitals and ‘5’ being the highest classification. Post hoc multiple comparisons of means were conducted to ensure that higher levels of attainment were positively correlated with the MIs. The results are presented in the next section.

### Findings

Table 1 summarizes the three MIs for the 2002–2006 period for the 2984 hospitals that provided complete responses to the AHA. While hospitals averaged a 3.3% annual increase in EFFCH over the 5-year period, the TC component largely offsets the EFFCH factor gains yielding an increase in TFP of about three-tenths of 1% each year (mean = 1.0027).

To gain a better understanding of using MIs for studying hospitals, we segmented the respondents from the Leapfrog Group’s annual survey. Table 2 compares the nonrespondents with the Leapfrog respondents’ scores on the MIs factors. While Leapfrog respondents differed significantly on TC and TFP factors relative to the rest of the sample, the main purpose of analyzing this group was to compare their MIs scores to the SP measure.

Table 3 analyzes the relationship between the SP measure and the MIs. Of the three indices, both TC (\(P < 0.10\)) and TFP (\(P < 0.05\)) had significant relationships with SP. Post hoc analyses of the ANOVA measures were also performed to identify where significant differences exist in the Leapfrog survey’s variable response levels vis-à-vis MI. The mean differences of TFP measures progressed in a positive linear fashion moving up the SP scale. Further, the highest SP attainment level differed significantly from the next highest level using the Tukey test for pairwise comparisons. Therefore, the inference is that higher attainment levels on the SP scale are positively correlated with increased levels of TFP.

### Discussion

Over the study period (2002–2006), we find that hospital administrators were able to increase the output-to-input ratio and increase EFFCH. However, they were not able to
improve the underlying care processes (ie, TC) to the extent that they made a positive contribution to TFP. Our results indicate that there has been a trade-off between EFFCH and TC resulting in small gains in TFP.

This trade-off between EFFCH and TC may shed more light on the hospital ‘value proposition’.10 The value proposition is a key feature of every quality improvement philosophy and system (eg, total quality management, Lean manufacturing, and six sigma),31 and in WE Demming’s words that define the value proposition, ‘when we improve quality we also improve productivity’.34 However, the value proposition’s main tenet – that improvements in TC lead to increases in TFP – cannot be supported. The Leapfrog respondents score higher in both TC and TFP relative to the other hospitals meaning that technological improvements to organizational processes can be made without negatively affecting the EFFCH measure. The higher TC scores for Leapfrog respondents still do not exceed one.

The competing concept which is antithetical to the value proposition is the so-called ‘productivity paradox’. When productivity measures fail to improve despite large investments in new technologies, it is referred to as the productivity paradox.35 A significant amount of research into the productivity paradox has revolved around the use of information technology. However, the productivity paradox concept also addresses other major changes to organizations such as modifications to new policies such as adherence to safe practices. Redesigning management structures, replacing staff, changing fundamental clinical practices through the use of incentives (eg, pay-for-performance), and undertaking a redesign of core business technologies are not easy. In many cases related to patient safety practices, TC involves abandoning behaviors that have been the norm for decades in favor of new work systems with which many health care organizations have little experience.36 Further, many such innovations create abrupt, radical, and discontinuous changes in organizational structures and cultures. There are numerous examples of hospitals that have experienced difficulties and failures related to major TCs that have resulted in lost productivity.37–39 The results indicate that the productivity paradox explanation cannot be completely discounted.

Our second stage analyses of the Leapfrog respondents’ progress on making TCs are used to further explore the hospital value proposition versus the productivity paradox. The NQF’s SP measure was significantly related to the TC index. This stands to reason as the 27 items that comprise the SP scale have the most far-reaching organizational impact. In particular, the scale addresses the need for creating a culture of safety, aligning organizational capabilities to care services, improving care processes, and coordinating information flows throughout the hospital.40 Such wide-scale efforts are strategic in nature, take extended periods of time

### Table 2: Student’s t-test for difference between Leapfrog Group survey respondents and nonrespondents

<table>
<thead>
<tr>
<th></th>
<th>Leapfrog respondent (yes = 1)</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
<th>t-test</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFFCH</td>
<td>0</td>
<td>1508</td>
<td>1.0333</td>
<td>0.0746</td>
<td>0.0019</td>
<td>0.368</td>
<td>0.713</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1470</td>
<td>1.0323</td>
<td>0.0804</td>
<td>0.0020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>0</td>
<td>1508</td>
<td>0.9699</td>
<td>0.0584</td>
<td>0.0015</td>
<td>-3.368</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1470</td>
<td>0.9770</td>
<td>0.0556</td>
<td>0.0014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFP</td>
<td>0</td>
<td>1508</td>
<td>0.9996</td>
<td>0.0390</td>
<td>0.0015</td>
<td>-2.407</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1470</td>
<td>1.0060</td>
<td>0.0666</td>
<td>0.0017</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** EFFCH, Technical Efficiency change; SD, standard deviation; SEM, standard error mean; TC, Technological change; TFP, Total Factor Productivity.

### Table 3: ANOVA for Malmquist Indices and the NQF safe practices metric by hospitals

<table>
<thead>
<tr>
<th>Name</th>
<th>Sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFFCH</td>
<td>Between groups</td>
<td>0.002</td>
<td>4</td>
<td>0.001</td>
<td>0.171</td>
</tr>
<tr>
<td></td>
<td>Within groups</td>
<td>2.538</td>
<td>828</td>
<td>0.003</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2.540</td>
<td>832</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TC</td>
<td>Between groups</td>
<td>0.053</td>
<td>4</td>
<td>0.013</td>
<td>1.973</td>
</tr>
<tr>
<td></td>
<td>Within groups</td>
<td>5.569</td>
<td>828</td>
<td>0.007</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>5.622</td>
<td>832</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TFP</td>
<td>Between groups</td>
<td>0.050</td>
<td>4</td>
<td>0.012</td>
<td>2.387</td>
</tr>
<tr>
<td></td>
<td>Within groups</td>
<td>4.308</td>
<td>828</td>
<td>0.005</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>4.357</td>
<td>832</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**Notes:** The ANOVA only includes the Leapfrog Group survey respondents; Post hoc analyses indicated that the mean differences between response levels progressed in a manner indicating a positive correlation between TFP and the NQF SP scores. Further, the top-performing quintile’s mean differed significantly from the third and fourth quintiles’ means.

**Abbreviations:** EFFCH, Technical Efficiency change; F, F-test statistic; TC, Technological change; TFP, Total Factor Productivity; NQF, National Quality Form
for complete implementation, and often encounter significant resistance from physicians.41

Given their broad scope and the fundamental changes to hospital activities they require, it stands to reason that the NQF’s recommendations would have a major impact on an organization’s processes as measured by the TC index (see Table 3). Both the post hoc analyses and the comparison with the rest of the sample (see Table 2) indicate that hospitals reporting higher NQF SP scores through the Leapfrog survey have above average TC performance compared with other facilities. Therefore, wide-scale efforts to change the hospitals’ cultures and safety-related processes do appear to be having a positive impact on TC and TFP levels, although their influence on EFFCH remains unclear.

Practice implications

For hospital administrators and physicians, the goal of continuous improvement in quality care may be challenging in the face of hospital performance pressures. Therefore, we examine over a 5-year window the effect of improving quality (via NQF’s safe practices measure) on EFFCH, TC, and TFP. While our results demonstrate a trade-off between efficiency change and TC resulting in small gains in TFP, there appears to be evidence among the Leapfrog respondents of improvements to organizational processes without negatively affecting EFFCH. Significant gains in hospital productivity may only take place after some extended period of time when improvements in safety-related practices and other organizational changes are fully implemented. As such, practitioners or managers can use this information to better understand the role of safe practices on hospital productivity. In particular, hospital leaders should feel confident that they can implement safety and quality improvement programs without compromising their relative efficiency.

In addition to the results, hospital administrators can use this methodology to compare productivity performance relative to other hospitals or units. The DEA is scalable to the units smaller than the hospital. For example, hospital leaders can compare nursing units over time and test new programs and assess their relative impact on efficiency.

Conclusions

For the sample of US hospitals studied, total productivity levels (TFP) increased from 2002 to 2006, but only to a small degree. Further, the productivity gains are the result of increased EFFCH rather than changes in the underlying technological processes (TC) used in facilities. Nevertheless, there are positive trends occurring in parts of the hospital sector. For example, hospitals that are implementing evidence-based patient safety practices publicly available through the Leapfrog Group’s annual survey are experiencing above average gains in TC without significantly sacrificing EFFCH. Another important finding is that public reporting or transparency may play a critical role in the way hospitals execute their activities.42

While there is evidence that keeps us from ruling out the productivity paradox, these findings are important for promoting and realizing the value proposition. For example, researchers have shown that hospitals engaged in public reporting deliver above average quality.43 However, TCs may be occurring and difficult to rationalize in terms of short-run return on investments. Therefore, policymakers should promote productivity gains, quality improvement, and increased reporting transparency in their initiatives and not just payment reforms. Such payment reform policies, without any connection to TC, may lead to an ineffective long-term strategy.44

Disclosure

The authors report no conflicts of interest in this work.

References


