Health Care Manag Sci (2017) 20:565–576
DOI 10.1007/s10729-016-9372-4

Using nonparametric conditional approach to integrate quality into efficiency analysis: empirical evidence from cardiology departments

Yauheniya Varabyova [1] · **Carl Rudolf Blankart** [1,2] · **Jonas Schreyögg** [1]

Received: 8 April 2016 / Accepted: 5 July 2016 / Published online: 13 July 2016

Abstract Health care providers are under pressure to improve both efficiency and quality. The two objectives are not always mutually consistent, because achieving higher levels of quality may require additional resources. The aim of this study is to demonstrate how the nonparametric conditional approach can be used to integrate quality into the analysis of efficiency and to investigate the mechanisms through which quality enters the production process. Additionally, we explain how the conditional approach relates to other nonparametric methods that allow integrating quality into efficiency analysis and provide guidance on the selection of an appropriate methodology. We use data from 178 departments of interventional cardiology and consider three different measures of quality: patient satisfaction, standardized mortality ratio, and patient radiation exposure. Our results refute the existence of a clear trade-off between efficiency and quality. In fact, the impact of quality on the production process differs according to the utilized quality measure. Patient satisfaction does not affect the attainable frontier but does have an inverted *U*-shaped effect on the distribution of inefficiencies; mortality ratio negatively impacts the attainable frontier when the observed mortality more than doubles the predicted mortality; and patient radiation exposure is not associated with the production process.

✉ Yauheniya Varabyova
yauheniya.varabyova@uni-hamburg.de

1 Hamburg Center for Health Economics, Universität Hamburg, Esplanade 36, 20354 Hamburg, Germany

2 Center for Gerontology and Health Care Research, School of Public Health, Brown University, Providence, RI, USA

Keywords Quality · Technical efficiency · Cardiology department · Conditional approach · Data envelopment analysis (DEA)

1 Introduction

Rapidly growing health expenditures over the recent decades have raised concerns about the affordability of hospital care and have put pressure on hospitals to increase the efficiency of resource allocation. One way to achieve higher levels of technical efficiency is to produce higher quantities of output with the same quantities of input, or in other words, to treat more patients with the same amount of personnel. However, health care providers argue that lowering the ratios of personnel to patient could lead to a deterioration of the quality of health services. To encourage quality improvement, most health care systems introduced various quality assurance programs and some health systems even rely on programs relating the remuneration of providers to the achieved results on quality indicators, known as pay-for-performance programs [1].

While health care policy emphasizes the importance of both efficiency and quality, only a small proportion of research analyzing efficiency in the health care sector has considered quality so far. Hollingsworth [2] identified more than 317 publications up to mid-2006 that relied on nonparametric, such as Data Envelopment Analysis (DEA) or Free Disposal Hull (FDH), and parametric, such as Stochastic Frontier Analysis (SFA), methods to estimate and compare the efficiency of health care providers. However, only 9 % of these publications integrated measures of quality into the analysis. The paucity of studies accounting for quality is in part caused by the lack of methodological guidance on the integration of quality into the efficiency analysis. We are particularly interested in the nonparametric methods to estimate efficiency and will focus on such throughout this study.

Previous studies have mostly relied on three distinct methods to integrate quality into the efficiency analysis. The first method is to treat an indicator of quality as an additional freely (or strongly) disposable output in the efficiency model [3–11]. This method is referred to as the *one-stage approach* [12]. To represent the idea that more output quality is better for production, the studies relying on the one-stage approach often transformed negative indicators of quality into positive indicators (e.g., by transforming mortality into inverse mortality). The second method to incorporate quality is based on the idea of congestion, which imposes an opportunity cost on the disposal of bad outputs [13, 14]. In the health care context, congestion could mean that reducing mortality requires sacrificing the treatment of further patients. Studies relying on the *congestion analysis* add a negative indicator of quality (i.e., lack of quality) as a weakly disposable output to the efficiency model [15–19]. Finally, the *two-stage approach* involves estimating the values of provider efficiency in the first stage without considering quality and then regressing the obtained efficiency estimates on quality in the second stage of analysis [20–26]. Table 1 provides a summary of the key features of the models integrating quality into efficiency analysis. The conditional approach will be explained in more detail in the next section.

The three widely used methods to integrate quality rely on different assumptions about the channel through which quality influences the production process. The shared element of the one-stage approach and the congestion analysis is that quality indicators are used to augment the input-output production set. Augmenting the efficiency model by the measures of quality suggests that quality has an effect on the attainable set of inputs and outputs. However, this does not need to be the case, as quality may have an effect on the distribution of the inefficiencies inside the production set without affecting the efficient boundary [27]. Moreover, the selected efficiency model, such as DEA or FDH, imposes further assumptions on the augmented production set, such as disposability (or monotonicity), convexity and returns to scale, which may not be appropriate for the measures of quality [12]. In contrast, the two-stage approach treats quality as an external variable, which is not part of the production set, but is helpful in explaining the differences in efficiency across health care providers. The two-stage approach requires that quality does not have an effect on the attainable set but instead has an effect only on the distribution of the inefficiencies inside the production set [28]. Simar & Wilson [29] described the situation in which an external variable has no effect on the attainable set as a separability condition. This condition may or may not be supported by the data, which necessitates a formal test to avoid a bias in empirical results [30]. Benchmarking decision-making units and examining the underlying relationship between efficiency and quality using the above methods may become problematic when the underlying assumptions are not verified in the empirical settings.

Another important aspect in the examination of the trade-off between efficiency and quality is the variety of measures utilized to capture the quality of provided services. Previous studies have included the indicators of *outcome quality* (e.g.,

Table 1 Methods to integrate quality into efficiency analysis and their assumptions on quality

Method	Main assumptions on quality	Estimation model	Interpretation
One-stage approach	o Strong disposability (or monotonicity) o Convexity (DEA), non-convexity (FDH) o Returns to scale	INPUTS → OUTPUTS INPUTS → QUALITY	Efficiency score considers the simultaneous production of output quantity and quality
Congestion analysis	o Weak disposability o Convexity (DEA), non-convexity (FDH) o Returns to scale	INPUTS → OUTPUTS INPUTS → LACK OF QUALITY	Efficiency score considers a potential quality congestion due to the expansion of output quantity
Two-stage approach	o Separability condition	I INPUTS → OUTPUTS II UNCOND → QUALITY	I) Unconditional efficiency score does not consider quality II) Regress unconditional efficiency on quality to examine whether and how the two are related
Conditional approach	o No *a priori* assumptions	I INPUTS → OUTPUTS (QUALITY ↑) II COND / UNCOND → QUALITY III COND → QUALITY	I) Conditional efficiency score incorporates a potential effect of quality on the attainable frontier II) Regress the ratio of conditional to unconditional efficiency on quality to examine the effect on the shift of the attainable frontier III) Regress conditional efficiency on quality to examine the effect on the distribution of inefficiencies

COND = conditional efficiency estimate, DEA = Data Envelopment Analysis, FDH = Free Disposal Hull, UNCOND = unconditional efficiency estimate mortality [3, 7, 8, 10, 21], hospital-acquired infections [16, 18], and readmissions [19]), *process quality* (e.g., acute myocardial infarction (AMI) patients who received aspirin within 24 h of arrival [11]), *structural quality* (e.g., extra nursing hours [4]), and *patient experience* (e.g., patient satisfaction [5]) as well as various combinations of multiple quality measures. However, different measures of quality may have a different relationship to efficiency. For example, if a reduction in physicians would lead to a reduction in the time spent talking to patients without compromising clinical care, this would most likely result in the negative relationship between efficiency and the measure of quality captured in patient satisfaction, but there may be no relationship between efficiency and clinical measures of quality [31]. Moreover, some utilized measures, for instance, mortality rate for AMI, capture only a part of hospital quality and may thereby not be representative of the total hospital quality but rather reflect the quality of particular hospital departments. In fact, previous research has shown that hospitals performing well on one condition (e.g., congestive heart failure) may not perform as well on other conditions (e.g., pneumonia) [32].

In the context of the above, this study aims to demonstrate the application of an advanced nonparametric method – the conditional approach – which allows exploring the relationship between efficiency and quality while avoiding the limitations of the previous studies. The conditional approach provides a flexible way to integrate quality into the efficiency model without the need to transform the measures of quality or impose additional assumptions, such as disposability, convexity, and returns to scale. The conditional approach allows differentiating between the two types of the effect of quality on the production process: the effect on attainable frontier and the effect on distribution of inefficiencies [27]. Furthermore, the conditional approach is based on the probabilistic formulation of the production process and as such is easily extended to a partial frontier analysis [33]. Estimates based on the partial frontier are no longer deterministic and are thus less affected by extreme values than full-frontier measures, such as DEA or FDH, and have better rates of convergence [33].

We take advantage of the hospital data at the department level, namely interventional cardiology departments, which ensures that the compared decision-making units rely on similar production technology and provide consistent quality indicators. Additionally, the analysis at the department level is congruent with hospital organization and planning and thereby increases the value of the results from the managerial and clinical perspectives. We examine three different measures to account for the potential differences between quality dimensions. Thus, we examine two measures of clinical quality: standardized mortality ratio to depict the outcome dimension and patient radiation exposure to depict the process dimension of quality. Moreover, patient satisfaction is used to account for patient experience. This study, therefore, contributes to the existing literature by providing the first empirical application of the conditional approach to the integration of quality into efficiency analysis and analyzing the relationship between efficiency and different measures of quality. Additionally, we explain how the conditional approach relates to the traditional methods to integrate quality and provide guidance on the selection of the appropriate methodology.

2 Methodology

The methods for nonparametrical efficiency analysis have been extensively described in Ozcan [34], Simar & Wilson [12] and elsewhere. The conditional approach was formally described in Bădin et al. [35] and references therein. In this chapter, we will provide an intuitive explanation of the main concepts of the conditional approach to enhance the understanding of this advanced method.

2.1 Conditional approach

Let the production technology be described by the vector of inputs $X \in R_+^p$ and the vector of outputs $Y \in R_+^q$. The production set Ψ includes all technically feasible combinations of inputs and outputs: $\Psi = (x, y) \in R_+^{p+q}$, where x can produce y. In their innovative study, Cazals et al. [33] proposed a probabilistic formulation to describe the production process, in which efficiency is described by the joint probability measure of (X, Y). The resulting estimator coincides with the Free Disposal Hull (FDH) estimator. Daraio & Simar [36] extended the probabilistic approach to allow for convexity and enabled the derivation of an estimator equivalent to the DEA estimator.

Full-frontier nonparametric estimators envelop all data points and are very sensitive to outliers and extreme values. To obtain robust nonparametric estimates, Cazals et al. [33] suggested estimating the partial efficiency measure of order-m. The estimator based on a partial frontier compares a unit (x, y) to m randomly selected peers. The order-m output-oriented efficiency measure is given by the following integral:

$$\lambda(x, y) = \int_0^\infty \left(1 - \left(1 - S_{Y|X}(uy|x)\right)^m\right) \mathrm{d}u, \tag{2.1}$$

where $S_{Y|X}(y|x) = \frac{\sum_{i=1}^{n} \mathrm{I}(x_i \le x, y_i \ge y)}{\sum_{i=1}^{n} \mathrm{I}(x_i \le x)}$ and $I(\cdot)$ is an indicator function, which equals 1 if the condition is true and 0 otherwise. The parameter m represents the number of units used to benchmark performance and determines the degree of robustness of the obtained estimate. At small values of m, there will be many observations beyond the efficient frontier; however, as the value of m increases, fewer observations will be left beyond the frontier. For large value of m, all observations are enveloped by the frontier leading to the full frontier estimator (FDH estimator).

Cazals et al. [33] and Daraio & Simar [37] demonstrated how to incorporate the set of environmental variables $Z \in R^r$ and obtain the conditional measures of efficiency. The attainable conditional production set can be expressed by: $\Psi^Z = (x, y)|Z = z$, where x can produce y conditional on external factors Z, such as quality. The conditional measure of output-oriented order-m efficiency is obtained by solving the following integral:

$$\lambda(x, y|z) = \int_0^\infty \left(1 - \left(1 - S_{Y|X,Z}(uy|x, z)\right)^m\right) \mathrm{d}u \quad (2.2)$$

where $S_{Y|X,Z}(y|x, z) = \sum_{i=1}^{n} \mathrm{I}(x_i \le x, y_i \ge y)\ K((z - z_i) \frac{/h_n)}{\sum_{i=1}^{n} \mathrm{I}(x_i \le x)} K((z - z_i)\ /h_n)$, $K(\cdot)$ is some kernel function with a compact support and h_n is the observation-specific bandwidth. Bădin et al. [38] showed how to derive the optimal value of the bandwidth.

Bădin et al. [27] explained how the conditional approach can be used to disentangle the channels through which an external factor Z enters the production process. In fact, Z may either affect the range of attainable values (X, Y), causing a shift in the attainable frontier, or it may affect the distribution of the inefficiencies inside the production set with the boundary not affected by Z, or it may affect both. Nonparametrically regressing the ratio of the conditional to unconditional efficiency estimates $R(x, y|z) = \lambda(x, y|z)/\lambda(x, y)$ on Z is informative about the potential shift of the attainable frontier due to the influence of Z. In contrast, regressing the conditional efficiency estimates $\lambda(x, y|z)$ on Z allows observing the effect of Z on the distribution of the inefficiencies.

2.2 Illustration using simulation

To illustrate the main concepts of the conditional approach, we simulate two datasets inspired from Bădin et al. [27, 35]. To keep the graphical presentation simple, the input is standardized to one ($X \equiv 1$). Therefore, decision-making units compete on the basis of maximal output Y. The inefficiency term is half-normally distributed $U \sim \mathcal{N}^+ \left(0, \sigma_U^2\right)$ with $\sigma_U^2 = 3$. The external variable Z is uniformly distributed, $Z \sim unif(0, 10)$.

The observations ($n = 200$) are simulated according to the following two data generating processes (DGP):

$$Y_1 = 40 - |Z - 5|^{1.5} - 2U \quad (2.3)$$

$$Y_2 = 40 - U|Z - 5| \quad (2.4)$$

In the first DGP, Z enters the production process by affecting the attainable frontier; the maximal attainable production set increases with Z for $Z < 5$ and decreases with Z for $Z > 5$. In the second DGP, Z affects the distribution of inefficiencies but not the boundary of the attainable set; whereby the firms have a decreasing probability of being inefficient with Z for $Z < 5$ and an increasing probability of being inefficient with Z for $Z > 5$.

The left panel of Fig. 1 illustrates the two datasets resulting from equation (2.3) and equation (2.4). In the upper scatterplot, the effect of Z on the shift of the true attainable frontier is observed in the solid black line, which is different from the dashed line representing maximal output in the absence Z. Therefore, the conditional measure $\lambda(x, y|z)$, which compares units facing similar level of Z, is different from the unconditional measure $\lambda(x, y)$, which does not take Z into account. In the lower scatterplot, Z does not have an effect on the attainable frontier (the solid and dashed lines coincide), leading to the equality of conditional $\lambda(x, y|z)$ and unconditional measures $\lambda(x, y)$. However, the cloud of data points indicates that the distribution of the inefficiencies is affected by Z, because the data points tend to be more dispersed at high and low values of Z and less dispersed when Z is close to 5.

The right panel of Fig. 1 illustrates how two different nonparametric regressions can be used to explain the effect of Z on the production process. In the first nonparametric regression, the ratios of conditional to unconditional efficiency estimates $R(x, y|z) = \lambda(x, y|z)/\lambda(x, y)$ are regressed on Z to investigate the impact on the attainable frontier. In the second nonparametric regression, the conditional efficiency estimates $\lambda(x, y|z)$ are regressed on Z to examine the effect on the distribution of inefficiencies. In the first DGP (the upper panel), Z affects the attainable frontier, therefore, the fitted regression line of $R(x, y|z)$ on Z is increasing for $Z < 5$ and decreasing for $Z > 5$, which exactly describes the shape of the true frontier (solid line) in the simulated dataset. In contrast, the regression line of $\lambda(x, y|z)$ on Z is flat, because, the distribution of the inefficiencies is not affected by Z. In the second DGP (the lower panel), Z enters the production process by affecting the distribution of the inefficiencies, but it does not affect the attainable frontier. Therefore, the fitted regression line of $R(x, y|z)$ on Z is flat, whereas the fitted line of $\lambda(x, y|z)$ on Z is U-shaped. Because higher values of $\lambda(x, y|z)$ represent higher inefficiency, a decreasing regression line represents favorable influence and an increasing regression line represents unfavorable influence of Z on the distribution of inefficiencies. Again, the fitted regression line corresponds to the relationship between Z and the production process in the simulated dataset.

3 Data

We combined data from three sources to obtain structural data on hospital cardiology departments and the corresponding

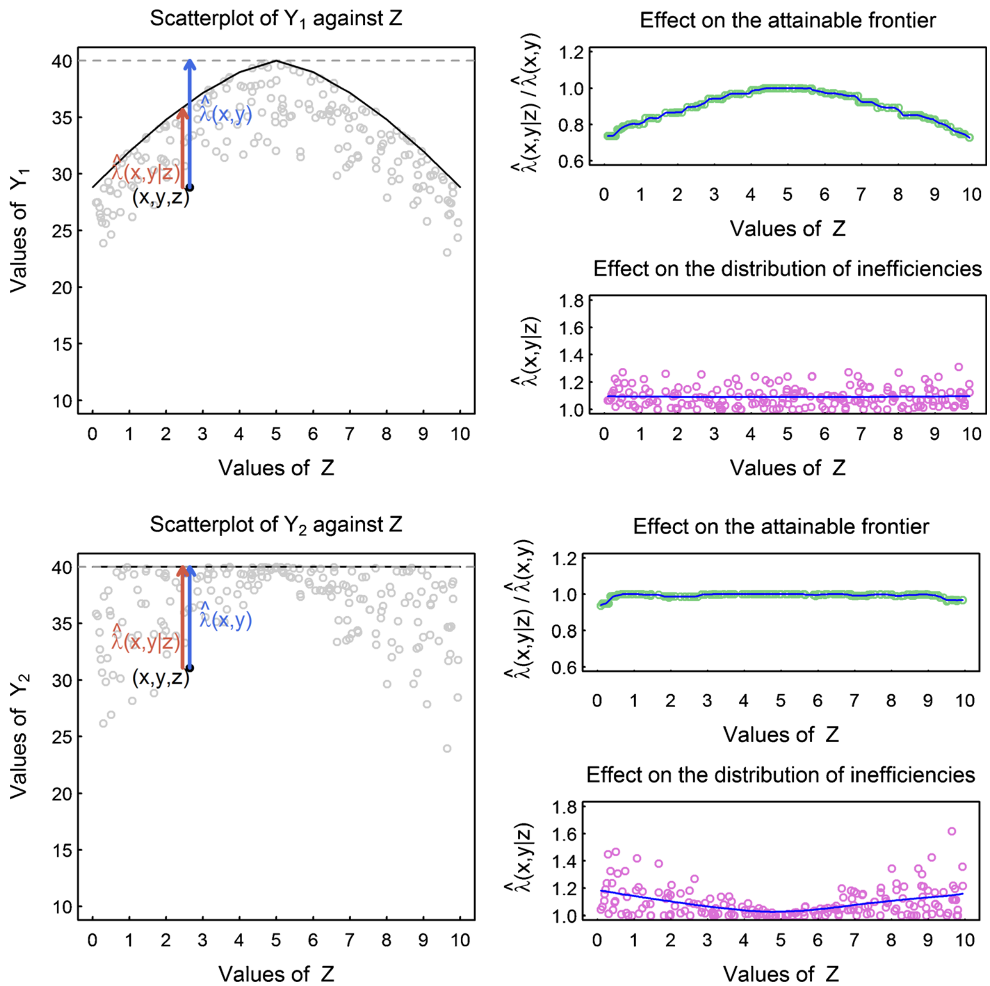

Fig. 1 Two mechanisms of the influence of Z on the production process. In the upper panel, Z has an effect on the attainable frontier by influencing the output Y_1 directly. In the lower panel, Z affects the distribution of inefficiencies but does not affect the attainable level of output Y_2

quality measures from calendar year 2012. Structural data on inputs and outputs were retrieved from Structured Quality Reports, which are released annually by all acute care hospitals in Germany [39]. The quality measure of patient experience was obtained from the independent non-profit institute "Weisse Liste", which conducts the largest nationwide survey of patient satisfaction with roughly a million returned surveys for 2012 [40]. The dataset was supplemented with nationally validated measures of inpatient clinical quality that by law have to be documented by German hospitals [41].

Two inputs included the number of full-time equivalent (FTE) physicians and the number of FTE nurses. The output was measured by annual inpatient discharges adjusted for case-mix. To adjust the number of outputs for case-mix, we used the procedure based on the relative length of stay in different diagnostic categories, which was developed by Herr [42] and subsequently applied in empirical applications in the absence of information on Diagnosis-Related Groups (DRG) [43].

We used three quality indicators: patient satisfaction, standardized mortality ratio, and patient radiation exposure. Patient satisfaction is increasingly accepted as one of the benchmarks of quality in health care and has been shown to be consistenly related to clinical effectiveness and patient safety [44]. In our analysis, patient satisfaction was measured as the patient's willingness to recommend the hospital to a best friend. The responses on the Likert scale range from 1 (very likely to recommend the hospital) to 6 (not at all likely to recommend the hospital). We used the mean value across all obtained responses in a cardiology department with a minimum of 30 responses. Higher values of patient satisfaction represent worse department quality.

Mortality is one of the most frequently used indicators of quality. However, its theoretical relation with efficiency is ambiguous. Thus, an inverse relationship can arise if higher mortality necessitates the provision of intensive care (requiring high personnel ratios) due to a more complex case-mix, whereas a direct relation will be observed if high mortality reflects lower levels of provided care due to low personnel ratios [21]. We used a standardized mortality ratio estimated as the ratio of observed to expected mortality rate during isolated coronary angiography [45]. Higher levels of the mortality ratio represent worse quality.

Finally, patient radiation exposure is an indicator of the process quality. It has been argued that process indicators should not be included into the efficiency analysis because it is not an output of production process [46]; however, several previous studies included process indicators either to augment the production set [6, 11, 47] or to explore the relationship with the distribution of inefficiencies in the two-stage analysis [23, 24, 26]. To shed new light on this discussion, we explored whether and how a process indicator enters the production process. In our study, radiation exposure was measured as the proportion of patients exposed during coronary angiography to a radiation dose over 3.500 cGy*cm^2 [45]. Again, higher values of patient radiation exposure represent worse quality.

4 Results

4.1 Descriptive statistics

Table 2 provides summary statistics for the utilized sample. In total, 178 cardiology departments provided complete data on the clinical measures of quality and at least 30 surveys of patient satisfaction. Thus, our sample represents approximately 25 % of interventional cardiology departments in Germany. On average, a cardiology department in our sample employed 25 full-time physicians and 86 full-time nurses to produce 3950 inpatient discharges adjusted for case-mix.

Considering quality measures, the mean patient satisfaction equaled 2.01, which indicated that most patients were rather satisfied with their stay at the cardiology departments (because 1 is the highest possible value and 6 is the lowest possible value of satisfaction). The department average value of satisfaction varied between 1.30 and 3.16. The mean value of mortality ratio equaled 1.10 (range: 0 to 3.94), meaning that, on average, the observed values of mortality only slightly exceeded the predicted values of mortality. Finally, the mean value of patient radiation exposure was 0.27 (range: 0.01 to 0.64), indicating that, on average, less than a third of patients was exposed to a dangerously high radiation dose during coronary angiography.

Table 2 Descriptive statistics

	N	Mean	SD	Min	Max
Inputs					
Physicians [FTE]	178	25	17	6	125
Nurses [FTE]	178	86	124	9	865
Outputs					
Inpatient discharges	178	3950	1568	874	13,076
Quality					
Patient satisfaction	178	2.01	0.31	1.30	3.16
Standardized mortality ratio	178	1.10	0.76	0.00	3.94
Patient radiation exposure	176	0.27	0.15	0.01	0.64

Notes: Data for calendar year 2012. FTE = full time equivalents, N = number of departments, SD = standard deviation

4.2 Efficiency estimates

Table 3 summarizes the obtained efficiency estimates. In the output-oriented framework, efficiency estimates equal to 1 represent efficient departments and efficiency estimates greater than 1 represent inefficient departments. Because we rely on the partial frontier analysis of order-m ($m = 80$), some efficiency estimates are smaller than 1. These estimates represent departments that are more efficient than the average 80 benchmark departments. The mean value of unconditional efficiency estimates (i.e., not considering quality differences) equals 1.26. This means that expanding the amount of output could lead to a reduction of inefficiency by 26 %.

Next, we condition the efficiency analysis on quality and obtain the mean values of conditional efficiency estimates equal to 1.26, 1.18, and 1.22 for patient satisfaction, mortality ratio, and patient radiation exposure, respectively. The mean values of conditional efficiency estimates controlling for mortality ratio and patient radiation exposure are smaller from the mean value of unconditional efficiency estimates, because we compare units at the similar levels of quality. However, only in case of the mortality ratio, the difference in the mean values is substantial, which provides some indicative evidence that only mortality has an effect on the shift in the attainable frontier.

4.3 Effect of quality on the production process

Using nonparametric regression analysis, we investigate the mechanisms how the measures of quality affect the production process. The left panel of Fig. 2 provides the results of the nonparametric regression of the ratios of conditional to unconditional efficiency estimates $R(x, y|z) = \lambda(x, y|z)/\lambda(x, y)$ on Z. These results are informative about the potential shift of the attainable frontier due to Z. The right panel of Fig. 2 illustrates the results of the regression of conditional efficiency estimates $\lambda(x, y|z)$ as a function of Z that are indicative of the effect of Z on the distribution of inefficiencies within the production set.

In case of patient satisfaction (Z_1), the results suggest that there is no effect of patient satisfaction on the attainable frontier because the mean values of the nonparametric regression of $R(x, y|z)$ on Z_1 form a flat line (the left panel). In contrast, there is a visible inverted U-shaped effect of patient satisfaction on the distribution of inefficiencies (the right panel). Near the center, the distribution of the inefficiencies is largest; however, at both high and low patient satisfaction, the departments are more efficient. This result characterizes the situation in

Table 3 Efficiency estimates

	N	Mean	SD	Min	Max
Unconditional efficiency estimates $\lambda(x, y)$					
	178	1.26	0.30	0.78	2.38
Conditional efficiency estimates $\lambda(x, y\|z)$, *where…*					
$z = Z_1$ (patient satisfaction)	178	1.26	0.29	0.85	2.39
$z = Z_2$ (standardized mortality ratio)	178	1.18	0.25	0.90	2.26
$z = Z_3$ (patient radiation exposure)	176	1.22	0.28	0.82	2.41

Notes: Higher values of efficiency estimates represent higher inefficiency. N = number of departments, SD = standard deviation

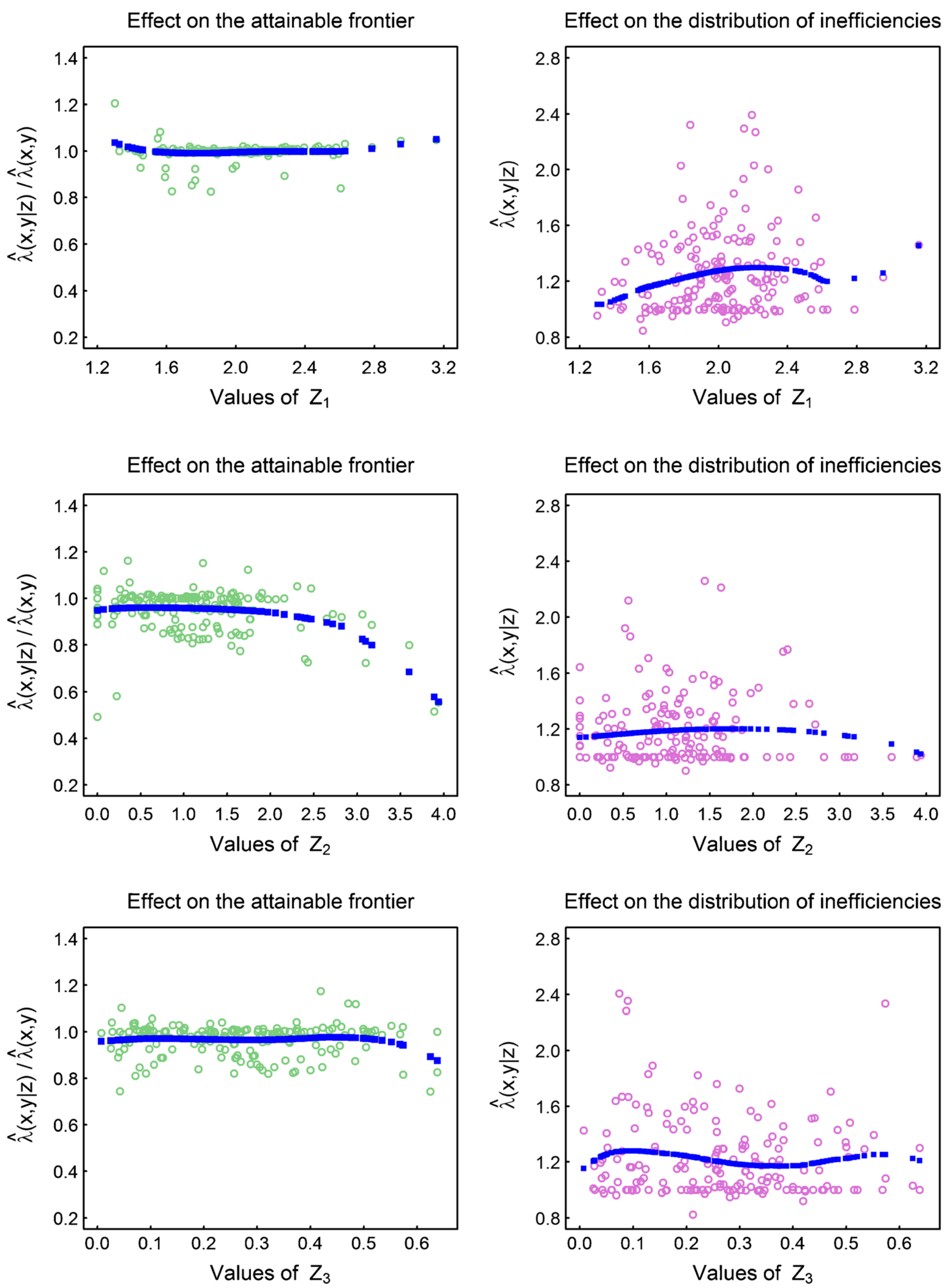

Fig. 2 The relationship of quality measures with the production process. Scatterplot of the ratios $R = \lambda(x, y| z)/\lambda(x, y)$ against Z (*left panel*); scatterplot of the conditional efficiencies $\lambda(x, y| z)$ against Z (*right panel*). The mean values of the corresponding nonparametric regressions are depicted with dark blue squares. Z_1= patient satisfaction, Z_2= standardized mortality ratio, Z_3= patient radiation exposure

which there are some excellent cardiology departments that achieve both high efficiency and high patient satisfaction, some trade-off cardiology departments that achieve high efficiency but low patient satisfaction, and some departments are "stuck-in-the-middle" with low values of efficiency and average values of patient satisfaction.

In case of standardized mortality ratio (Z_2), the results of the nonparametric regression of $R(x, y|z)$ on Z_2 show that there is a shift of the attainable set associated with mortality (the left panel). This supports the indicative evidence of the frontier shift identified by the difference between the average unconditional and conditional estimates. Because we are using the output-oriented model, a decreasing regression line indicates that the production process is adversely affected by the mortality ratio. The results of the nonparametric regression of the $\lambda(x, y|z)$ on Z_2 show that standardized mortality ratio does not affect the distribution of inefficiencies, because the average regression values form almost a flat line (the right panel).

Patient radiation exposure (Z_3) does not seem to have an effect on efficiency. There is almost no observable effect of patient radiation exposure either on the shift in the boundary of the attainable set or on the distribution of inefficiencies. The lines formed by the mean regression values are roughly flat in both regressions.

5 Discussion

In this study, we applied the conditional approach to analyze the relationship between technical efficiency and three different measures of quality: patient experience and outcome and process quality indicators. We used data on 178 departments of interventional cardiology, which ensured a good comparability of the analyzed units. Two different nonparametric regressions were used to investigate the channel through which quality affected the production process. The regression of the ratio of conditional to unconditional efficiency estimates on quality provided evidence about the effect of quality on the attainable frontier, whereas the regression of the conditional efficiency estimates on quality revealed the effect of quality on the distribution of inefficiencies. Our results refute the existence of a trade-off between efficiency and quality. In our study, the relationship between efficiency and quality seems much more complex and turns out to be highly dependent on the type of the utilized measure of quality.

The measure of patient satisfaction does not have an effect on the attainable frontier; however, there is an inverted *U*-shaped effect of patient satisfaction on the distribution of inefficiencies. Cardiology departments with both the highest and lowest values of patient satisfaction are relatively efficient, whereas departments with median values of patient satisfaction are characterized by the highest dispersion in the inefficiencies and are also, on average, the least efficient. The fact that some providers manage to have both high values of efficiency and patient satisfaction indicates that high efficiency may be achieved without a significant sacrifice of service quality, which corresponds to the philosophy of total quality management (TQM) [18]. TQM is an integrated approach that emphasizes the maintenance and continuous improvement of processes to reach high values of quality [48]. TQM supports the idea that improvements in quality may reinforce improvements in efficiency. On the other hand, the departments that are characterized by the trade-off between efficiency and patient satisfaction may be following a different path to improve efficiency. Indeed, patient dissatisfaction may arise when technical efficiency is achieved through a reduced investment in work environment, causing disrupted routines or increased workloads [31, 49].

The effect of mortality on the production process is quite different. The standardized mortality ratio has an unfavorable effect on the attainable frontier whereas the effect on the distribution of inefficiencies is rather small. The effect on the attainable frontier is more pronounced at high than at low values of the mortality ratio. In fact, there is a negative effect on the shift of the attainable frontier for cardiology departments, in which the observed mortality more than doubles the predicted mortality. Departments with a high mortality ratio require more input resources per patient. In other words, a high mortality ratio (or low quality) is associated with low efficiency contradicting the existence of quality-efficiency trade-off. This empirical finding is consistent with the hypothesis that an inverse relation between the mortality and efficiency arises when higher mortality reflects the need for intensive care [21]. For instance, patients with several severe comorbidities and a high chance of dying may require a large amount of labor resources.

The measure of patient radiation exposure represents process quality. In our empirical application, this quality measure does not have an effect either on the attainable frontier or on the distribution of the inefficiencies. The measures of process quality are frequently used in clinical studies as a proxy for operational learning in interventional cardiology. Accumulated experience usually leads to a reduced time of the procedure, which is generally a good sign, as it reduces the time that the patient is exposed to radiation and decreases the chance of infection [50]. However, the reduction of procedure time is probably too small to be translated into the more efficient use of capital and personnel resources. Nevertheless, this empirical result is important because it highlights the difference between process and outcome indicators. It lends some support to a proposition that process measures should not be directly included in efficiency models [46]. Instead, the researchers should concentrate on the quality measures that are justified by a more direct theoretical link to the production process.

5.1 Limitations

Our empirical analysis has some limitations. First, the analysis is based on the departments of interventional cardiology and may not be generalizable to other medical specialties. However, the shift of research focus from the hospital to department level is imperative in making research findings useful in managerial and clinical practice. Several studies suggest that organization and management of hospital activities is more effective at the department level than at the hospital level [51, 52].

Another limitation of our analysis is the fact our input measure contains only labor but not capital resources, which are usually represented by the number of beds [53]. We believe that the focus on one medical specialty mitigates this limitation, because departments of the same medical specialty have similar structures and use similar technology [54]. Moreover, the number of hospital beds and the requirements on technical equipment for interventional cardiology are set by the German system of regional hospital planning [55] and are not fully under control of individual hospitals. Therefore, labor resources play a paramount role in hospital management.

Our quality measures also have some limitations. The empirical analysis is complicated by little variation in patient satisfaction because most patients are, on average, rather satisfied with their hospital stay [56]. The measure of patient satisfaction may also include some response bias, because dying patients and patients with severe post-acute complications are less likely to take part in the survey. However, both efficient and inefficient departments are similarly affected by this bias and, therefore, conclusions can still be drawn from our results. The indicators of clinical quality are limited to standardized mortality ratio and patient radiation exposure. This selection is due to data availability, because many other national indicators of clinical quality for interventional cardiology are contested for their imprecise documentation and risk adjustment. We selected two quality indicators that were available for a sufficient number of cardiology departments, were adjusted for case-mix, and were rated by experts as having good theoretical and empirical explanatory power [57].

We also did not account for hospital characteristics, such as ownership type or university status. Some of the unexplained differences in efficiency may be related to the institutional characteristics of the analyzed departments.

5.2 Methodological and policy implications

This study applies the nonparametric conditional approach to investigate the role of quality in the efficiency performance of hospital cardiology departments. The advantage of our approach is that quality is introduced in a non-restrictive way. Given the empirical findings of our study, different measures of quality can have an effect either on the attainable frontier or on the distribution of the inefficiencies. Furthermore, the direction of the effect of quality on the production process and the linearity of this relationship are not always known a priori. Therefore, a model that allows for a differential effect of quality on the production process is the most appropriate to integrate quality into the analysis of health care efficiency. The conditional approach has the advantage of not requiring a priori assumptions about the relationship between quality and efficiency. Moreover, the conditional approach can accommodate good and bad outputs without the need to transform a quality indicator constructed in a particular way to have meaning.

In contrast, the traditional methods to incorporate quality in the analysis of efficiency require making quite restrictive assumptions on quality. This requirement may pose a challenge because most studies start with the premise that the underlying relationship between quality and the production process is unknown. In fact, quality can either influence the attainable set of inputs and outputs or it can influence the distribution of inefficiencies, or it influence affect both. The test of the separability condition, which has been described at length in Daraio et al. [30], can be used to understand which particular influence mechanism is at place.

The separability condition was first mentioned in Simar & Wilson [29] and it requires that an external variable (such as quality) has no influence on the location of the attainable frontier. When quality shifts the location of the attainable frontier (i.e., the separability condition is violated), then using the two-stage approach is inappropriate, because the efficiency estimates obtained in the first stage are not meaningful [28, 30]. A shift of the attainable frontier indicates that firms at different levels of quality are characterized by different production possibilities. Augmenting the production process by quality may alleviate this problem. The one-stage approach requires that an external variable that is favorable to efficiency is added as an input and a variable that is unfavorable to efficiency is added as an output [12]. In this case, a researcher must know a priori whether the influence of quality is favorable or unfavorable to efficiency. In our empirical analysis, only the measure of mortality had an effect on the attainable frontier and this effect was unfavorable. If we were to transform this measure to obtain the inverse of the mortality ratio (indicating a desirable outcome) then the effect of quality on the attainable frontier would be favorable. Therefore, it would be incorrect to add the transformed measure of mortality as another output variable. Moreover, the transformations of the measures of quality may be another source of bias in the efficiency model [58, 59].

The congestion analysis also augments the production process by quality and because it relaxes the assumption of strong disposability, situations with both favorable and unfavorable influence on the frontier can be accommodated. Therefore, the congestion analysis could be applied in our dataset to the measure of mortality. However, both the one-stage approach and the congestion analysis impose other restrictive assumptions on the dataset augmented by the measures of quality, such as convexity and returns to scale. It is conceivable that not all these assumptions will be supported by empirical datasets. These assumptions need to be tested using the appropriate tests [60, 61].

When quality does not shift the location of the attainable frontier but instead only affects the distribution of inefficiencies inside the production set (i.e., the separability condition holds), then the two-stage approach described in Simar & Wilson [29] can be used. In our empirical analysis, patient satisfaction and patient radiation exposure did not affect the attainable frontier and thereby the two-stage approach could be performed with these quality indicators. On the other hand, the one-stage approach or the congestion analysis would fail to identify the inverted *U*-shaped relationship between patient satisfaction and the distribution of the inefficiencies or the lack of relationship between patient radiation exposure and production. Adding an indicator Z that only affects the distribution of inefficiency to the production set is inappropriate because the new constraint would be binding only for observations with high values of Z in relation to inputs [62]. Particularly in cases when higher values of Z indicate the lack of quality (such as patient satisfaction and patient radiation exposure in our analysis), the managerial implications will be misleading, because decision-making units with very bad quality will be considered efficient.

Therefore, the distinction of the channels through which quality enters the production process is crucial to the selection of the methodology to examine the trade-off between efficiency and quality. A misapplication of the traditionally used methods can potentially lead to incorrect managerial and policy implications. The inference about the relationship between efficiency and quality is only possible when the analyst has a clear understanding of the underlying assumptions.

Future research would benefit from the insights regarding the theoretical foundation that underlies the mechanisms through which quality impacts the production process of health care institutions. This would necessitate exploring different types of quality measures in conjunction with different health care institutions. Another useful research extension would be to analyze the role of institutional and environmental characteristics in the relationship between efficiency and quality.

6 Conclusions

Contemporary health care policy is concerned with increasing the efficiency of the hospital sector while improving the quality of provided care. Policy makers in different countries are interested in reforming reimbursement systems to reward superior quality through pay-for-performance programs [1]. Therefore, understanding the potential trade-off between efficiency and quality is paramount for decision makers to allocate constrained resources between and within hospitals. The literature, however, provides scant and ambiguous empirical evidence on this trade-off, which is to some extent due to the use of methods that are based on different assumptions about the role of quality in the production process. Therefore, we add to the literature by shedding light on the channels through which different measures of quality impact the efficiency of health care providers. Additionally, we provide methodological guidance on the selection of the appropriate methods to integrate quality into the analysis of health care efficiency.

This is the first study to apply the conditional approach to integrate quality into the analysis of efficiency using health care data. The conditional approach allows benchmarking units at similar levels of quality and enables differentiating between the effect of quality on the shift of the attainable frontier and on the distribution of inefficiencies. In our empirical analysis of the data from 178 cardiology departments, each quality measure deserves an individual examination, because the relationship between efficiency and quality varies according to the type of measure. Thus, patient satisfaction does not have an effect on the attainable frontier, but it affects the distribution of the inefficiencies within the production set. Cardiology departments with the highest and the lowest values of patient satisfaction achieve the best efficiency, whereas departments with the median values of patient satisfaction have rather low values of efficiency. The standardized mortality ratio has a negative effect on the attainable frontier, suggesting that departments with the highest mortality ratio are characterized by the highest resource intensity (and technical inefficiency). Therefore, in this case, we observe a positive association between efficiency and quality instead of a trade-off between efficiency and quality. Finally, the measure of patient radiation exposure, which represents the process dimension of quality, has neither an effect on the attainable frontier nor an effect on the distribution of inefficiencies. Our results confirm that, because different measures of quality may have differential effects on the production process, policy makers and researchers should be careful when selecting methods and interpreting the influence of various quality indicators on efficiency.

References

1. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W (2010) Systematic review: effects, design choices, and context of pay-for-performance in health care. BMC Health Serv Res 10(1):247
2. Hollingsworth B (2008) The measurement of efficiency and productivity of health care delivery. Health Econ 17(10):1107–1128

3. Chang S-J, Hsiao H-C, Huang L-H, Chang H (2011) Taiwan quality indicator project and hospital productivity growth. Omega 39(1): 14–22
4. Garavaglia G, Lettieri E, Agasisti T, Lopez S (2011) Efficiency and quality of care in nursing homes: an Italian case study. Health Care Manag Sci 14(1):22–35
5. Mark BA, Jones CB, Lindley L, Ozcan YA (2009) An examination of technical efficiency, quality, and patient safety in acute care nursing units. Policy Politics Nur Pract 10(3):180–186
6. Nayar P, Ozcan YA (2008) Data envelopment analysis comparison of hospital efficiency and quality. J Med Syst 32(3):193–199
7. Tiemann O, Schreyögg J (2009) Effects of ownership on hospital efficiency in Germany. Bus Res 2(2):115–145
8. Varabyova Y, Schreyögg J (2013) International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches. Health Policy 112(1):70–79
9. Yang J, Zeng W (2014) The trade-offs between efficiency and quality in the hospital production: some evidence from Shenzhen, China. China Econ Rev 31:166–184
10. Du J, Wang J, Chen Y, Chou S-Y, Zhu J (2014) Incorporating health outcomes in Pennsylvania hospital efficiency: an additive super-efficiency DEA approach. Ann Oper Res 221(1):161–172
11. Ferrier GD, Trivitt JS (2013) Incorporating quality into the measurement of hospital efficiency: a double DEA approach. J Prod Anal 40(3):337–355
12. Simar L, Wilson PW (2015) Statistical approaches for non-parametric frontier models: a guided tour. Int Stat Rev 83(1):77–110
13. Färe R, Grosskopf S (1983) Measuring output efficiency. Eur J Oper Res 13(2):173–179
14. Fare R, Grosskopf S, Lovell CK (1994) Production frontiers. Cambridge University Press
15. Clement JP, Valdmanis VG, Bazzoli GJ, Zhao M, Chukmaitov A (2008) Is more better? An analysis of hospital outcomes and efficiency with a DEA model of output congestion. Health Care Manage Sci 11(1):67–77
16. Mutter R, Valdmanis V, Rosko M (2010) High versus lower quality hospitals: a comparison of environmental characteristics and technical efficiency. Health Serv Outcome Res Methodol 10(3–4):134–153
17. Valdmanis VG, Rosko MD, Mutter RL (2008) Hospital Quality, Efficiency, and Input Slack Differentials. Health Serv Res 43(5p2):1830–1848. doi:10.1111/j.1475-6773.2008.00893.x
18. Prior D (2006) Efficiency and total quality management in health care organizations: a dynamic frontier approach. Ann Oper Res 145(1):281–299
19. Wu C-H, Chang C-C, Chen P-C, Kuo K-N (2013) Efficiency and productivity change in Taiwan's hospitals: a non-radial quality-adjusted measurement. CEJOR 21(2):431–453
20. Anderson R, Weeks H, Hobbs B, Webb J (2003) Nursing home quality, chain affiliation, profit status and performance. J Real Estate Res 25(1):43–60
21. Bates LJ, Mukherjee K, Santerre RE (2006) Market structure and technical efficiency in the hospital services industry: a DEA approach. Med Care Res Rev 63(4):499–524
22. Gok MS, Sezen B (2013) Analyzing the ambiguous relationship between efficiency, quality and patient satisfaction in healthcare services: the case of public hospitals in Turkey. Health Pol 111(3): 290–300
23. Kooreman P (1994) Nursing home care in the Netherlands: a non-parametric efficiency analysis. J Health Econ 13(3):301–316
24. Laine J, Finne-Soveri UH, Björkgren M, Linna M, Noro A, Häkkinen U (2005) The association between quality of care and technical efficiency in long-term care. Int J Qual Health Care 17(3): 259–267
25. Lee RH, Bott MJ, Gajewski B, Taunton RL (2009) Modeling efficiency at the process level: an examination of the care planning process in nursing homes. Health Serv Res 44(1):15–32
26. Nyman JA, Bricker DL (1989) Profit incentives and technical efficiency in the production of nursing home care. Rev Econ Stat 71(4): 586–594
27. Bădin L, Daraio C, Simar L (2012) How to measure the impact of environmental factors in a nonparametric production model. Eur J Oper Res 223(3):818–833
28. Simar L, Wilson PW (2011) Two-stage DEA: caveat emptor. J Prod Anal 36(2):205–218
29. Simar L, Wilson PW (2007) Estimation and inference in two-stage, semi-parametric models of production processes. J Econ 136(1): 31–64
30. Daraio C, Simar L, Wilson PW (2015) Testing the" separability" condition in two-stage nonparametric models of production. LEM Working Paper Series,
31. Jones J, Sanderson C, Black N (1992) What will happen to the quality of care with fewer junior doctors? A Delphi study of consultant physicians' views. J R Coll Physicians Lond 26(1):36–40
32. Jha AK, Li Z, Orav EJ, Epstein AM (2005) Care in U.S. Hospitals — The Hospital Quality Alliance Program. N Engl J Med 353(3): 265–274. doi:10.1056/NEJMsa051249
33. Cazals C, Florens J-P, Simar L (2002) Nonparametric frontier estimation: a robust approach. J Econ 106(1):1–25
34. Ozcan YA (2014) Health care benchmarking and performance evaluation an assessment using data envelopment analysis (DEA), 2nd edn. Springer, Newton, MA
35. Bădin L, Daraio C, Simar L (2014) Explaining inefficiency in non-parametric production models: the state of the art. Ann Oper Res 214(1):5–30
36. Daraio C, Simar L (2007) Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. J Prod Anal 28(1–2):13–32
37. Daraio C, Simar L (2005) Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J Prod Anal 24(1):93–121
38. Bădin L, Daraio C, Simar L (2010) Optimal bandwidth selection for conditional efficiency measures: a data-driven approach. Eur J Oper Res 201(2):633–640
39. G-BA (2016) Qualitätsbericht der Krankenhäuser.
40. Gehrlach C, Altenhöner T, Schwappach D (2009) Der patients' experience questionnaire. Patientenerfahrungen vergleichbar machen Bertelsmann Stiftung, Gütersloh
41. Bramesfeld A, Willms G Gesetzliche Qualitätssicherung in der medizinischen Versorgung: Stand und Weiterentwicklung–Sektorenübergreifende Qualitätssicherung nach § 137a SGB V. In: Public Health Forum, 2014. vol 2. Elsevier, pp 14. e11–14. e13
42. Herr A (2008) Cost and technical efficiency of German hospitals: does ownership matter? Health Econ 17(9):1057–1071
43. Varabyova Y, Blankart CR, Torbica A, Schreyögg J (2016) Comparing the Efficiency of Hospitals in Italy and Germany: Nonparametric Conditional Approach Based on Partial Frontier. Health Care Manag Sci:1–16
44. Doyle C, Lennox L, Bell D (2013) A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ open 3(1)
45. AQUA (2013) Beschreibung der Qualitätsindikatoren für das Erfassungsjahr 2012: Koronarangiographie und Perkutane Koronarintervention (PCI). Indikatoren 2012
46. van Ineveld M, van Oostrum J, Vermeulen R, Steenhoek A, van de Klundert J (2015) Productivity and quality of Dutch hospitals during system reform. Health Care Manage Sci:1–12

47. Navarro-Espigares JL, Torres EH (2011) Efficiency and quality in health services: a crucial link. Serv Ind J 31(3):385–403
48. Martínez-Costa M, Martínez-Lorente AR, Choi TY (2008) Simultaneous consideration of TQM and ISO 9000 on performance and motivation: an empirical study of Spanish companies. Int J Prod Econ 113(1):23–39
49. Zhang NJ, Unruh L, Wan TT (2008) Has the Medicare prospective payment system led to increased nursing home efficiency? Health Serv Res 43(3):1043–1061
50. Khan N, Abboudi H, Khan MS, Dasgupta P, Ahmed K (2014) Measuring the surgical 'learning curve': methods, variables and competency. BJU Int 113(3):504–508
51. Kristensen SR, Bech M, Lauridsen JT (2015) Who to pay for performance? The choice of organisational level for hospital performance incentives. Eur J Health Econ:1–8
52. Suñol R, Arah O, Wagner C, Groene O (2013) Relationships between the implementation of quality management strategies and clinical outcomes in European hospitals. The International Society for Quality in Health Care
53. Worthington AC (2004) Frontier efficiency measurement in health care: a review of empirical techniques and selected applications. Med Care Res Rev 61(2):135–170
54. Ludwig M, Van Merode F, Groot W (2010) Principal agent relationships and the efficiency of hospitals. Eur J Health Econ 11(3): 291–304
55. Busse R, Blümel M (2014) Germany. Health system review Health Systems in Transition (2)
56. Hannöver W, Dogs CP, Kordy H (2000) Patientenzufriedenheit–ein Maß für Behandlungserfolg? Psychotherapeut 45(5):292–300
57. AQUA (2014) Indikatoren in den Strukturierten Qualitätsberichten. Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH.
58. Ali AI, Seiford LM (1990) Translation invariance in data envelopment analysis. Oper Res Lett 9(6):403–405
59. Blank R, Krusenbaum S, Steven M (2015) Effizienzanalyse mittels data envelopment analysis unter Berücksichtigung von Emissionen. WiSt-Wirtschaftswissenschaftliches Studium 44(11):613–622
60. Simar L, Wilson PW (2002) Non-parametric tests of returns to scale. Eur J Oper Res 139(1):115–132
61. Banker RD (1996) Hypothesis tests using data envelopment analysis. J Prod Anal 7(2–3):139–159
62. Smith P (1997) Model misspecification in data envelopment analysis. Ann Oper Res 73:233–252