

Outcome-based indicators in healthcare policy: Insights from a hospital competition with heterogeneous agents

Calogero Guccio, Domenico Lisi, Marco Martorana*

Department of Economics and Business, University of Catania

Abstract

Outcome-based measures, such as mortality and readmission rates, are frequently interpreted as good indicators of the unobservable hospital quality and, thus, they are increasingly used as a basis for providing financial incentives to hospitals. In this paper, preliminarily we examine risk-adjusted mortality and readmission rates for three specific medical procedures (CHF, COPD, STROKE) in Italian hospitals and, in contrast to the conventional wisdom, we show that the two outcome measures tell two completely different stories in terms of hospitals' performance. Then, moving from this evidence, we study a model of hospitals' behaviour where patients differ in their idiosyncratic risk and, importantly, we consider explicitly the generating process of standard outcome-based indicators. We find that, when adjustment for the idiosyncratic risk is not fully appropriately made, readmission rates are worse indicators of hospital quality than mortality rates, and their use in performance programmes might penalize the best hospitals. Moreover, we find that higher is the probability of mortality for the specific medical intervention, lower is the reliability of readmission rates as indicators of quality. Therefore, our results emphasize that the use of readmission rates as a basis for performance programmes should be set more appropriately.

JEL Classification: I12, I18

Keywords: hospital quality, outcome-based indicators, heterogeneous agents.

* Corresponding author: Department of Economics and Business, University of Catania, Corso Italia 55, 95129 Catania, Italy. E-mail address: domenico.lisi@unict.it. Domenico Lisi gratefully acknowledges financial support by the European Commission within the Seventh Framework Programme (FP7) for the "International Research Project on Financing Quality in Healthcare". The usual disclaimer applies.

1. Introduction

Over the last decades, the introduction of prospective systems, aiming at stimulating cost containment in the provision of healthcare services by hospitals, has raised concerns that such payment system could provide too little incentives for the healthcare quality. For this reason, many governments have tried to combine prospective systems with further regulatory mechanisms to control, and provide incentives for, the quality of services (Busse et al., 2011). In this regard, the main difficulties lie in the fact that the intrinsic quality of services are either unobservable for policy makers, or the collection of the necessary data is too costly. Therefore, the approach in the healthcare policies has been to look at the outcome-based indicators of the unobservable quality, such as mortality and readmission rates, on the premise that hospitals with better quality of services should exhibit better outcomes, after controlling for differences in patients' characteristics.

In the United States, for instance, the Affordable Care Act has introduced explicitly such financial incentives for hospitals. In particular, the Hospital Readmission Reduction Program (HRRP) has introduced a penalty system into the Medicare reimbursement system for hospitals underperforming in terms of selected 30-day risk-adjusted readmission rates (such as, AMI and CHF) for discharges beginning on October 2012 (Foster and Harkness, 2010; Adashi and Kocher, 2011). In the same spirit, the England's National Health Service has introduced in 2011 financial measures linked to readmissions, in response to the extraordinary increase in the level of readmissions, establishing that hospitals do not receive payment for emergency readmission within 30 days of discharge following an elective admission (Department of Health, 2011). Likewise, hospitals in Germany have the same incentive to avoid readmissions as they will not receive additional reimbursement for cases readmitted within 30 days of a discharge (Geissler et al., 2011). In other countries, such as France and Italy, the abovementioned outcome indicators are not integrated in the hospitals payment system, but are constantly monitored and publicly released by the deputed national agencies to gain more insight in quality of care.

Outcome-based measures are also used in the health literature to evaluate and study hospitals' performance. In particular, low mortality and readmission rates for selected diagnosis (such as, heart failures, acute myocardial infarction, strokes, pneumonia, hip fracture) are often used as a proxy measure for good inpatient care quality (Propper et al., 2004, 2008; Cooper et al., 2011; Cavalieri et al., 2013; Gravelle et al., 2014). Furthermore, a considerable amount of empirical studies used outcome measures also to evaluate other

important issues in health research as the impact of new technologies (Xian et al., 2011) and the impact of specific policies (Evans et al., 2008).

The idea behind outcome-based indicators is that if appropriate risk adjustment is made for patients case mix, differences in mortality and readmission rates are likely to be driven by differences in unobservable quality. Therefore, the advantage of outcome-based indicators is that they would allow to evaluate easily the hospitals' performance in terms of the quality of services, without suffering all costs needed to collect the data on the process of care. However, under the abovementioned policies, it is crucial to ensure that such outcome indicators would lead to an accurate inference on the relative hospitals' performance, as an incorrect ranking would lead to a flawed and perverse design of providers' incentives.

In this perspective, a literature review by Fischer et al. (2014) shows that, among the numerous studies using readmission rates as a hospital quality indicator, very few of them considered somehow the validity of the indicator. Nonetheless, few recent empirical papers looked more carefully at the characteristics of the data generating process of outcome indicators (Mohammed et al., 2009; Chua et al., 2010; Schreyogg and Stargardt, 2010; Papanicolas and McGuire, 2016). Among these, Laudicella et al. (2013) argue that, when patient characteristics are not perfectly observable and there is a significant probability of mortality associated to the specific intervention, the best hospitals with low mortality rates might face a larger share of unobservably sicker patients at risk of readmission and, thus, their quality performance might be underestimated by readmission rates.

In the same spirit, the aim of this paper is to explore from a theoretical perspective the data generating process of standard outcome-based indicators and to get insights on their use in healthcare policy. To this purpose, we preliminary examine risk-adjusted mortality and readmission rates for three specific medical procedures (CHF, COPD, STROKE) in Italian hospitals and, in contrast to the conventional wisdom, we show that the two outcome measures tell completely different stories in terms of hospitals' performance. Then, moving from this evidence, we study a model of hospitals' behaviour where patients differ in their idiosyncratic risk and, importantly, we consider explicitly the generating process of standard outcome-based indicators, specifically mortality and readmission rates.

In line with Laudicella et al. (2013), we find that, when adjustment for the idiosyncratic risk is not fully appropriately made, readmission rates are worse indicators of hospital quality than mortality rates, and their use in performance programmes might penalize the best hospitals. Moreover, we find that higher is the probability of mortality for the specific medical intervention, lower is the reliability of readmission rates as indicators of quality. Therefore, in

the last part of the paper we discuss the implications for the use of outcome-based indicators to provide financial incentives for the healthcare quality.

Our study contributes to the abovementioned literature on outcome-based indicators (e.g., Laudicella et al., 2013; Fischer et al., 2014), by exploring in depth the generating process of standard outcome indicators of the unobservable quality in a theoretical model of hospitals' and patients' behaviour. The microfounded nature of our model consents to disentangle the different components of the outcomes' generating process at the hospital level, and this allows us to get a few insights on the use of outcome-based indicators in the healthcare policy. On the other hand, this paper contributes to the theoretical literature on hospitals' behaviour and payment systems (e.g., Ma, 1994; Beitia, 2003; Brekke et al., 2010, 2011; Siciliani et al., 2013; Guccio et al., 2016), by introducing explicitly the generating process of standard outcome-based indicators in a model with heterogeneous agents. Finally, and most importantly, our results might contribute to the public debate on the optimal design of healthcare policy by shedding further light on the use of outcome-based indicators to provide incentives for the quality of healthcare services.

The outline of the paper is as follows. In Section 2 we examine risk-adjusted mortality and readmission rates in Italian hospitals to highlight that, indeed, the two outcome indicators can tell a different story in terms of hospitals' performance. Then, Section 3 lays out the model of hospitals' behaviour, including the generating process of mortality and readmission rates. Section 4 develops the equilibrium results and derives the implications for the use of outcome indicators in the healthcare policy. Finally, Section 5 provides concluding remarks.

2. Outcome-based indicators in Italian hospitals

Several healthcare systems have recently introduced financial incentive schemes based on indicators of hospital quality, such as mortality and readmission rates (see Section 1). Their use would require that they correctly and reliably rank hospitals' performance in so far as they are able to capture the unobservable quality of healthcare services. That is, such measures should provide a correct and accurate measure of hospital quality to avoid distortions in the incentive schemes.

Whereas we study in the next sections the reliability of such indicators from a theoretical perspective, here we examine risk-adjusted mortality and readmission rates for medical procedures related to Congestive Heart Failure (CHF), Chronic Obstructive Pulmonary Disease (COPD) and Stroke in Italian hospitals, in order to verify whether the

rankings coming out from them are consistent, that is, whether they convey the same information on hospital quality. In Italy, outcome-based indicators are increasingly used to monitor hospitals' and health systems' performance (Nutri et al., 2016), though the hospital payment system is not linked to them¹.

To do so, we use data from the Italian National Program for Outcome Assessment (Programma Nazionale Esiti – PNE) carried out by the National Agency for Regional Health services (Agenzia Nazionale per i Servizi Sanitari Regionali – AGENAS) together with the Italian Ministry of Health, including information on all Italian hospitals in 2012 and 2013. We focus on the most adopted indicators in literature (mortality and readmission rates for CHF, COPD and STROKE) and, coherently with our purpose, use the risk-adjusted rates². Table 1 shows average Mortality (MR) and Readmission (RR) rates. It is worth to be noticed that average RR are lower for STROKE, which shows higher MR.

Table 1. Risk-adjusted Mortality and Readmission rates.

<i>MR</i>				<i>RR</i>			
	COPD	CHF	STROKE		COPD	CHF	STROKE
2012	8.79	10.67	12.06	2012	13.58	14.64	10.29
2013	8.78	10.44	11.56	2013	13.86	14.86	7.27

As argued before, MR and RR can be reliably used as proxies for hospitals' performance as long as performance rankings coming out from them are consistent; that is, if the two outcome measures are strongly correlated and generate comparable rankings. In Table 2, we report Pearson correlation, and Spearman and Kendall rank correlation for RR and MR of Italian hospitals. Looking at the Pearson correlation, RR and MR are poorly positive correlated for COPD and CHF (only 2013 is significant) and are not correlated at all for STROKE. Similarly, their rankings are not consistent, being the Spearman and Kendall rank correlation very low and also not significant for STROKE. Basically, MR and RR do not seem to convey the same information and, in particular, do not draw the same picture in terms of hospitals' performance rankings. Moreover, there is a rough evidence that the higher is the mortality rate (STROKE), the lower is the correlation between the two outcome measures, as can also be seen in Figure 1 showing the scatter plot of MR and RR for the three specific

¹ For a detailed description of the reimbursement system in Italy, see Cavalieri et al. (2013, 2014).

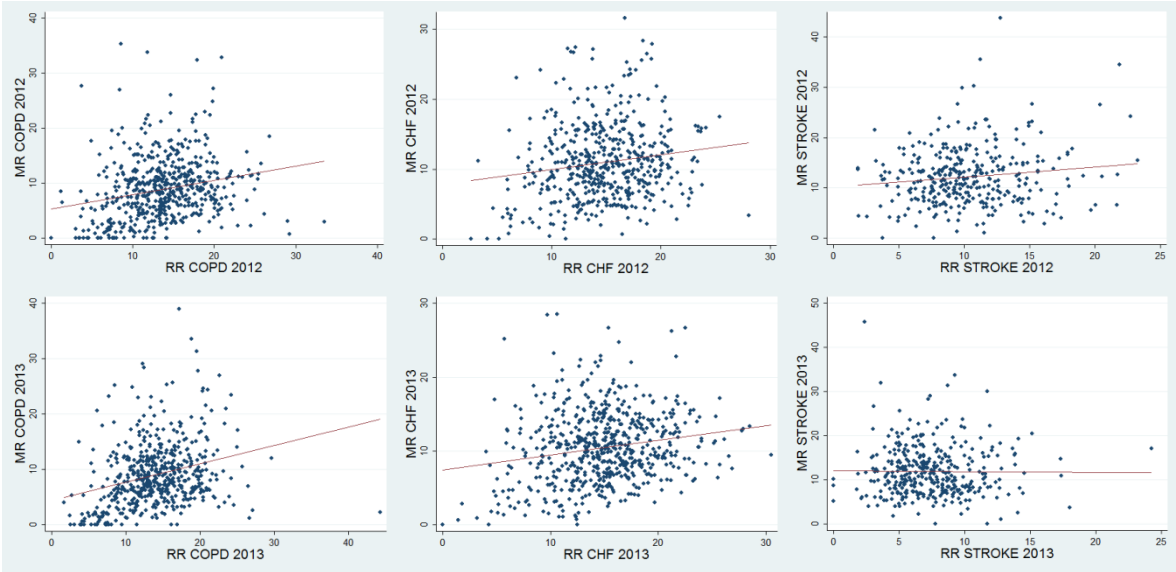
² The risk adjustment procedure, employed through a validated methodology, allows for adjusting mortality and readmission rates with respect to the idiosyncratic risk of patients.

procedures. This evidence might hide a relationship between the mortality risk connected to a specific procedure and the residual distribution of patients' idiosyncratic risk among hospitals. Specifically, hospitals with lower MR might face a larger share of patients at a higher risk of readmission. If such relationship exists, and adjustment for the idiosyncratic risk is not fully appropriately made, RR might be biased and underestimate the quality of best hospitals. This call for further analysis on the data generating process of outcome-based indicators.

Table 2. Risk-adjusted MR and RR correlation

<i>Pearson</i>			
	COPD	CHF	STROKE
2012	0.142*	0.124	0.119
2013	0.278*	0.191*	-0.008
<i>Spearman</i>			
	COPD	CHF	STROKE
2012	0.198*	0.123	0.063
2013	0.296*	0.176*	-0.001
<i>Kendall</i>			
	COPD	CHF	STROKE
2012	0.134*	0.082	0.043
2013	0.205*	0.118*	-0.002

Figure 1. Risk-adjusted MR and RR for COPD, CHF and STROKE



3. The model

In this section, we introduce the main structure of the model. In line with the previous literature on hospital competition (Beitia, 2003; Brekke et al., 2010, 2015), the analysis of providers' behavior is conducted in the framework of the Hotelling spatial competition (Hotelling, 1929), where we study a market for medical treatments with two hospitals located at the either end of the unit line $S = [0, 1]$. Indeed, the oligopolistic framework is well-suited for our research topic, as we are interested to study a setting where few hospitals provide healthcare services with a different treatment quality and this, in turn, generates different mortality and readmission rates.

As standard in the literature, since healthcare providers do not usually choose their location according to the patients distribution, we consider providers location to be exogenously given. On the line segment S , there are K groups of patients, each one uniformly distributed over the line segment, who differ in their idiosyncratic risk h_k and, thus, in the valuation of medical treatment (e.g., Brekke et al., 2008). Furthermore, the total density of patients is normalized to 1 (e.g., Brekke et al., 2010), so that the share of each patients group s_k is the normalized total mass of patients with an idiosyncratic risk h_k – thus, we have that $\sum_{k=1}^K s_k = 1$. Each patient demands only one medical treatment and, as fairly realistic in this market, they do not pay for the treatment they receive; rather, they are only concerned with the quality of their treatment. Assuming full market coverage – that is, patients do not have an attractive outside option – all patients simply choose a hospital for the treatment.

Respect to previous literature, in our model we consider explicitly that each patient treated by hospitals has a positive probability of dying and of being readmitted, which clearly generates a disutility for her/him. Both negative outcomes are governed, on the one hand, by a treatment specific severity index, γ and δ for mortality and readmission, respectively; on the other hand, they depend also on the patient specific idiosyncratic risk h_k . However, moving from the argument that negative outcomes might be “quality dependent”, we also consider that mortality and readmission probabilities can be consistently reduced by a higher quality treatment. In particular, for a patient k treated in hospital i , the probability of dying and the probability of being readmitted are assumed to be given, respectively, by:

$$M(\gamma, h_k, q_i) = h_k(\gamma - \theta q_i) \tag{1}$$

$$R(\delta, h_k, q_i) = h_k(\gamma - \theta q_i)\delta \tag{2}$$

where θ is the parameter governing the impact of quality on the probability of negative outcomes³.

Considering the disutility of negative outcomes, the utility of a patient r , located at $x \in S$, and receiving treatment from hospital i , located at 0 in the unit line, is given by:

$$U(h_r, x) = v + \eta h_r q_i - h_r(\gamma - \theta q_i)(1 + \delta) - \tau x - g\left(\sum_{k=1}^K s_k x_k^*\right) \quad (3)$$

where v is the gross valuation of medical treatment, $q_i \geq \underline{q}$ is the treatment quality at hospital i , ηh_r measures the marginal utility of quality, τ and g are the transportation and congestion cost parameters, respectively, and x_k^* is the amount of patients k demanding treatment from hospital i , as defined hereunder.⁴ The lower bound \underline{q} represents the lowest treatment quality hospitals are allowed to offer – that is, if $q_i < \underline{q}$, then i might lose its license – and without loss of generality (e.g., Brekke et al., 2012) we set $\underline{q} = 0$, implying that $h_k \gamma$ and $h_k \gamma \delta$ represent the upper bound of the probability of dying and of being readmitted, respectively, for patients k .

Form (3), we can see that a higher patient specific idiosyncratic risk h_r , beyond the effect on the probability of negative outcomes, increases the marginal utility of treatment quality ηh_r ; this captures the reasonable idea that patients with a higher illness severity are relatively more careful to get a higher quality of the treatment respect to avoid higher transportation and congestion costs. In this perspective, the presence of congestion costs is important in our model, as it generates a crucial trade-off in the choice of the hospital. On the one hand, a higher quality is attractive for the higher valuation of medical treatment $\eta h_r q_i$; on the other hand, the higher attractiveness implies a larger amount of patients x_k^* , for each group k , demanding treatment from that hospital, which evidently produces higher congestion costs.

Given the assumption of uniform patients distribution, the amount x_r^* of patients r demanding treatment from hospital i , can be characterized by the location of the specific patient r who is indifferent between provider i and j , located at 0 and 1 respectively, that is:

$$v + \eta h_r q_i - h_r(\gamma - \theta q_i)(1 + \delta) - \tau x_r^* - g\left(\sum_{k=1}^K s_k x_k^*\right) = v + \eta h_r q_j - h_r(\gamma - \theta q_j)(1 + \delta) - \tau(1 - x_r^*) - g\left(\sum_{k=1}^K s_k x_k^*\right) \quad (4)$$

which yields

³ Evidently, there is a propulsive role for quality in reducing the probability of negative outcomes as long as $\theta > 0$, which we consider reasonable in the majority of medical branches.

⁴ There is a wide empirical evidence showing that the main predictors of hospital choice by patients are distance to hospitals and quality of treatment (e.g., Tay, 2003; Shen, 2003; Howard, 2006; Varkevisser et al., 2012). Moreover, recent studies have also shown that the congestion, for instance induced by longer waiting lists, affects significantly the demand for hospital services (e.g., Martin et al., 2007).

$$x_r^* = \frac{(\tau+g) + [\eta + \theta(1+\delta)] h_r (q_i - q_j) - 2g \sum_{k \neq r} s_k x_k^*}{2(\tau + s_r g)} \quad (5)$$

Not surprisingly, the amount x_r^* of patients r demanding treatment from hospital i depends also on the same amount x_k^* for the other groups of patients, as a higher x_k^* for each $k = 1, 2, \dots, K$ would increase the congestion costs in hospital i and, in turn, it would marginally reduce its attractiveness. Thus, the amount x_k^* for each $k = 1, 2, \dots, K$ is given by the following system:

$$\begin{cases} x_1^* = \frac{(\tau+g) + [\eta + \theta(1+\delta)] h_1 (q_i - q_j) - 2g \sum_{k \neq 1} s_k x_k^*}{2(\tau + s_1 g)} \\ \vdots \\ x_r^* = \frac{(\tau+g) + [\eta + \theta(1+\delta)] h_r (q_i - q_j) - 2g \sum_{k \neq r} s_k x_k^*}{2(\tau + s_r g)} \\ \vdots \\ x_K^* = \frac{(\tau+g) + [\eta + \theta(1+\delta)] h_K (q_i - q_j) - 2g \sum_{k \neq K} s_k x_k^*}{2(\tau + s_K g)} \end{cases} \quad (6)$$

Solving (6), the amount x_r^* of patients r demanding treatment from hospital i is given by:

$$x_r^* = \frac{1}{2} + \frac{[\eta + \theta(1+\delta)] [\tau h_r + g \sum_{k \neq r} (h_r - h_k)]}{2\tau(\tau+g)} (q_i - q_j) \quad (7)$$

From (7), if hospitals provide medical treatments with the same quality ($q_i = q_j$), they share equally the amount of patients for each group. Furthermore, we can also see that, while for those patients with a high illness severity h_r a higher quality q_i always increases the demand x_r^* for that hospital, for those patients with a low illness severity h_r a higher quality q_i might even reduce the demand x_r^* for that hospital – in fact, notice that when h_r is low then $[\tau h_r + g \sum_{k \neq r} (h_r - h_k)] < 0$. This is not surprising, as the higher treatment quality q_i increases the attractiveness of hospital i especially for those patients at higher risk h_r and, in turn, it produces a higher congestion costs; thus, the abovementioned trade-off in the choice of hospital might even lead patients at lower risk to reduce the demand x_r^* for that hospital.

Finally, the total demand for hospital i is given by the weighted sum of the demands x_r^* of each patients group, that is:

$$x_{D_i}^* = \sum_{k=1}^K s_k x_k^* = \frac{1}{2} + \frac{[\eta + \theta(1+\delta)] (\sum_{k=1}^K s_k x_k^*)}{2(\tau+g)} (q_i - q_j) \quad (8)$$

On the other hand, the demand for hospital j is simply the complement to 1, that is:

$$x_{D_j}^* = (1 - x_{D_i}^*) = \frac{1}{2} + \frac{[\eta + \theta(1+\delta)] (\sum_{k=1}^K s_k x_k^*)}{2(\tau+g)} (q_j - q_i) \quad (9)$$

As can be seen from (8) and (9), if hospitals offer the same treatment quality, they halve the market exactly; otherwise, the hospital with higher treatment quality obtains a larger market share. However, following the above conclusion from (7) on the impact of q_i on the demand of each patients group, we can also say that the larger market share will be represented, for the

most part, by those patients with a higher idiosyncratic risk h_r . Therefore, the hospital with higher treatment quality faces a worse distribution of patients at risk of negative outcomes.

Finally, the role of the other parameters are more standard in the literature (e.g., Brekke et al., 2010, 2012, 2015). In particular, the extent to which the difference in quality affects the hospital market share depends on the marginal utility of quality η and transportation cost τ , which are usually interpreted in the literature as indicators of competition in the healthcare market, as well as on the congestion cost g and the impact of quality on the probability of negative outcomes θ .

Hospitals are financed by a third-party payer through a prospectively regulated price p per treatment and a potential lump-sum transfer T .⁵ The cost of medical treatments is given by the cost function $C(x_{D_i}^*, q_i)$, which is assumed to be increasing and convex both in output and quality, plus some fixed cost F :

$$C(x_{D_i}^*, q_i) = \frac{c}{2} (x_{D_i}^*)^2 + \frac{k}{2} q_i^2 + F \quad (10)$$

where c and k are parameters governing the marginal cost of output and quality, respectively. The convexity of the cost function in output $x_{D_i}^*$ captures a standard feature of the hospital sector, namely that hospitals face some capacity constraints (e.g., Brekke et al., 2008)⁶, and it appears consistent with our specification (3) of patients' utility including congestion costs.

Finally, as extensively discussed in the literature (e.g., Ellis, 1998; Brekke et al., 2011; Makris and Siciliani, 2013), non-financial incentives might also be relevant in the healthcare market. Accordingly, we explicitly include the non-financial motivation of provider management $\alpha_i B(q_i, x_{D_i}^*)$, with the hospital specific parameter $\alpha_i \geq 0$ governing its relative weight in the provider objective function. The main idea behind this non-financial part is that offering high quality medical treatment gives hospital staff a high social and professional status, which, in turn, gives them a higher non-monetary utility. Therefore, we consider the assumptions $B_{q_i} > 0$, $B_{x_{D_i}^*} > 0$, and $B_{q_i, x_{D_i}^*} > 0$ to be reasonable. For simplicity, we assume that the provider's non-financial motivation is given by⁷:

⁵ As standard in this literature, the role of the lump-sum transfer T is simply to ensure that the “non-negative profit” participation constraint are guaranteed (e.g., Brekke et al., 2008, 2010; Ma and Mak, 2015).

⁶ While no scholar disagrees on the presence of capacity constraints and economies of scale in the hospital sector, there is a large discussion in the literature on the optimal scale in the provision of healthcare services. For instance, Aletas (1999) suggests that the optimal scale for hospitals is around 100 – 200 beds.

⁷ The standard way in the literature to introduce the non-financial motivation is to imagine that providers management receive a positive utility from the consumer surplus. However, in that specification even providing a very low quality medical treatment (remember that $\underline{q} = 0$) gives a positive non-financial utility, which is something we consider unreasonable. Differently, in (7) hospitals management receive a positive non-financial utility only providing at least a quality higher than the license standard, emphasizing the interpretation of the

$$\alpha_i B(q_i, x_{D_i}^*) = \alpha_i q_i x_{D_i}^* \quad (11)$$

Hence, the objective function of hospital i is given by:

$$\Omega_i = T + p x_{D_i}^* - \frac{c}{2} (x_{D_i}^*)^2 - \frac{k}{2} q_i^2 - F + \alpha_i q_i x_{D_i}^* \quad (12)$$

In the next section, we consider the simultaneous game in which each hospital chooses its treatment quality q independently to maximize the objective function (12). In our model we do not allow for explicit rationing, implying that hospitals cannot turn down patients seeking treatment. For our purposes, we are interested to study the ‘‘asymmetric’’ Nash equilibrium where the two hospitals provide different treatment qualities which, in turn, generate different mortality and readmission rates. Therefore, in the following we imagine that the two hospitals attach a different weight $\alpha_i \neq \alpha_j$ to the non-financial motivation in the objective function.⁸

4. Equilibrium with heterogeneous agents and outcome-based indicators

The hospital i 's maximization problem is

$$\text{Maximize}_{q_i} \Omega_i = T + p x_{D_i}^* - \frac{c}{2} (x_{D_i}^*)^2 - \frac{k}{2} q_i^2 - F + \alpha_i q_i x_{D_i}^*$$

The optimal quality for hospital i is given by the following first-order condition:

$$q_i^* = \frac{\left(p - \frac{c}{2}\right) + \frac{\alpha_i}{2\Lambda}}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_i} + q_j \frac{c\Lambda - \alpha_i}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_i} \quad (13)$$

where $\Lambda = \frac{\partial x_{D_i}^*}{\partial q_i} = \frac{[\eta + \theta(1 + \delta)] (\sum_{k=1}^K s_k x_k^*)}{2(\tau + g)}$. As standard in this framework, the first-order condition (13) yields the optimal quality for hospital i as the best response respect to the quality of hospital j . Then, the usual way to close the model is to look for the symmetric Nash equilibrium (e.g., Brekke et al., 2008, 2010, 2015), where hospital i is identical to hospital j . Imposing symmetry in our model – that is, $\alpha_i = \alpha_j = \alpha$ and thus $q_i = q_j = q$ – yields the following symmetric Nash equilibrium:

$$q^* = \frac{\left(p - \frac{c}{2}\right) + \frac{\alpha}{2\Lambda}}{\frac{k}{\Lambda} - \alpha} \quad (14)$$

Notice that, the equilibrium quality (14) is consistent with the Nash equilibrium found in previous papers in the literature (e.g., Brekke et al., 2010, 2015), even if it is more general as it considers the general case of K groups of patients who differ in their idiosyncratic risk h_k .

non-financial utility as a higher social and professional status given by a high quality health services. Nonetheless, it can be easily found that the two different interpretations produce the same model implications.

⁸ Clearly, the difference in the weight of non-financial motivation ($\alpha_i \neq \alpha_j$) is not the only way through which we can capture heterogeneity in providers. For instance, we could have also assumed different marginal costs of output and quality between hospitals, that is $c, k_i \neq c, k_j$.

For our purposes, however, we are interested to study the asymmetric Nash equilibrium where the two hospitals provide different treatment qualities which, in turn, generate different mortality and readmission rates. In particular, under the general assumption $\alpha_i \neq \alpha_j$, the asymmetric Nash equilibrium is given by the following⁹:

$$q_i^* = \frac{\left\{ \left(\frac{p-c}{2} \right) + \frac{\alpha_i}{2\Lambda} + \left[\frac{(p-c)}{2} + \frac{\alpha_j}{2\Lambda} \right] \left[\frac{c\Lambda - \alpha_i}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_i} \right] \right\}}{\left\{ 1 - \left[\frac{c\Lambda - \alpha_j}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_j} \right] \left[\frac{c\Lambda - \alpha_i}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_i} \right] \right\}} \quad (15)$$

$$q_j^* = \frac{\left\{ \left(\frac{p-c}{2} \right) + \frac{\alpha_j}{2\Lambda} + \left[\frac{(p-c)}{2} + \frac{\alpha_i}{2\Lambda} \right] \left[\frac{c\Lambda - \alpha_j}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_j} \right] \right\}}{\left\{ 1 - \left[\frac{c\Lambda - \alpha_i}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_i} \right] \left[\frac{c\Lambda - \alpha_j}{c\Lambda + \frac{k}{\Lambda} - 2\alpha_j} \right] \right\}} \quad (16)$$

Then, studying the difference between the equilibrium qualities, not very surprisingly, we have that $(q_i^* - q_j^*)$ is increasing (decreasing) in α_i (α_j) and, in particular, we can conclude that the hospital attaching a higher weight to the non-financial motivation will provide a higher treatment quality – that is, $q_i^* > (<)q_j^*$ whenever $\alpha_i > (<)\alpha_j$.

Clearly, the different equilibrium qualities between hospital i and j generate different outcomes for the two hospitals. In particular, considering the probability of dying specified in (1), we have that the mortality rates for the two hospitals are given by:

$$MR_i = \frac{\sum_{k=1}^K s_k x_k^* h_k}{\sum_{k=1}^K s_k x_k^*} (\gamma - \theta q_i) \quad (17)$$

$$MR_j = \frac{\sum_{k=1}^K s_k (1-x_k^*) h_k}{\sum_{k=1}^K s_k (1-x_k^*)} (\gamma - \theta q_j) \quad (18)$$

Similarly, the different equilibrium qualities generate also different readmission rates for the two hospitals. However, when we compute the readmission rates we have to consider that the population of patients potentially at risk of readmission is not the same as that at the initial admission, because the mortality outcome changes the population of patients that can be potentially readmitted in hospitals. More specifically, we can easily establish that, for each patients group $k = 1, 2, \dots, K$, the survived population of patients that can be potentially readmitted in hospital i is $s_k x_k^* [1 - h_k (\gamma - \theta q_i)]$. Furthermore, we can notice that, as the survival rates are different among patients groups, the mortality outcome changes also the

⁹ It is straightforward to show that, under the symmetric assumption $\alpha_i = \alpha_j$, the equilibrium qualities (15) and (16) are both equal to the symmetric Nash equilibrium (14).

distribution of the idiosyncratic risk h_k among patients that can be potentially readmitted in hospital i . In particular, considering the probability of being readmitted specified in (2) and the population of survived patients, the readmission rates for the two hospitals are given by:

$$RR_i = \frac{\sum_{k=1}^K s_k x_k^* [1 - h_k(\gamma - \theta q_i)] h_k}{\sum_{k=1}^K s_k x_k^* [1 - h_k(\gamma - \theta q_i)]} \delta(\gamma - \theta q_i) \quad (19)$$

$$RR_j = \frac{\sum_{k=1}^K s_k (1 - x_k^*) [1 - h_k(\gamma - \theta q_j)] h_k}{\sum_{k=1}^K s_k (1 - x_k^*) [1 - h_k(\gamma - \theta q_j)]} \delta(\gamma - \theta q_j) \quad (20)$$

In the rest of the paper, by considering the equilibrium results of our model, we attempt to get further insights on the relationship between unobservable quality and outcomes in hospitals and, most importantly, on the use of outcome-based indicators to provide financial incentives in the healthcare policy.

4.1 Relationship between quality and outcomes

In this section, we aim at studying how a change in the equilibrium quality affects the data generating process of mortality and readmission rates, in order to get further insights on how well these outcome indicators reflect differences in the unobservable quality of hospitals. To make the intuition clearer, we consider the simple case with two patient groups H and L , both uniformly distributed over the line segment, characterized by a high h_H and a low h_L idiosyncratic risk (i.e., $h_H > h_L$), respectively. Furthermore, for the sake of simplicity, we imagine that the two patient groups have the same total mass, that is $s_H = s_L = \frac{1}{2}$.

In this case, the outcomes of the two hospitals are fairly straightforward. In particular, the unadjusted mortality rates for the two hospitals are given by:

$$MR_i = \frac{x_H^* h_H + x_L^* h_L}{x_H^* + x_L^*} (\gamma - \theta q_i) \quad (21)$$

$$MR_j = \frac{(1 - x_H^*) h_H + (1 - x_L^*) h_L}{(1 - x_H^*) + (1 - x_L^*)} (\gamma - \theta q_j) \quad (22)$$

from which we can clearly observe the two components of the generating process, that is the distribution of patients idiosyncratic risk and the quality of hospitals. Furthermore, we can easily derive the “theoretical” risk-adjusted mortality rates, defined as the mortality rates generated by hospitals over (as if they would treat) the whole population of patients, and the associated risk-adjustment factor:

$$RA_MR_{i,j} = \frac{1}{2} (h_H + h_L) (\gamma - \theta q_{i,j}) \quad (23)$$

$$RA_F_MR_i = \frac{\frac{1}{2} (h_H + h_L) (x_H^* + x_L^*)}{x_H^* h_H + x_L^* h_L} \quad (24)$$

Similarly, the unadjusted readmission rates for the two hospitals are given by:

$$RR_i = \frac{x_H^* [1 - h_H(\gamma - \theta q_i)] h_H + x_L^* [1 - h_L(\gamma - \theta q_i)] h_L}{x_H^* [1 - h_H(\gamma - \theta q_i)] + x_L^* [1 - h_L(\gamma - \theta q_i)]} \delta(\gamma - \theta q_i) \quad (25)$$

$$RR_j = \frac{(1 - x_H^*) [1 - h_H(\gamma - \theta q_j)] h_H + (1 - x_L^*) [1 - h_L(\gamma - \theta q_j)] h_L}{(1 - x_H^*) [1 - h_H(\gamma - \theta q_j)] + (1 - x_L^*) [1 - h_L(\gamma - \theta q_j)]} \delta(\gamma - \theta q_j) \quad (26)$$

As observed in the general case (19), the survival rates are different between the two patient groups – that is, $[1 - h_H(\gamma - \theta q_i)] < [1 - h_L(\gamma - \theta q_i)]$ – and, thus, the mortality outcome changes the distribution of the idiosyncratic risk of patients that can be potentially readmitted. Finally, as we did in (23) for the mortality rates, we can derive the “theoretical” risk-adjusted readmission rates and the associated risk-adjustment factor:

$$RA_RR_{i,j} = \frac{1}{2} (h_H + h_L) \delta(\gamma - \theta q_{i,j}) \quad (27)$$

$$RA_F_RR_i = \frac{\frac{1}{2} (h_H + h_L) \{x_H^* [1 - h_H(\gamma - \theta q_i)] + x_L^* [1 - h_L(\gamma - \theta q_i)]\}}{x_H^* [1 - h_H(\gamma - \theta q_i)] h_H + x_L^* [1 - h_L(\gamma - \theta q_i)] h_L} \quad (28)$$

Not surprisingly, we can see from (23) and (27) that, when the risk-adjustment is fully appropriately made, both outcome-based indicators perfectly reflect the differences among hospitals quality. On the other hand, we can also argue that, when “empirical” risk-adjusted mortality and readmission rates for the same medical procedure are not consistent – that is, they tend to tell two completely different stories in terms of hospitals’ performance – this could be interpreted as a warning signal of a flawed risk-adjustment. In this perspective, the previous evidence in the literature (e.g., Laudicella et al., 2013; Gravelle et al., 2014), as well as the evidence showed in Section 2 for the Italian outcome-based indicators, would seem to highlight a remarkable inconsistency among “empirical” risk-adjusted mortality and readmission rates and, thus, a presumably faulty risk-adjustment (e.g., Mohammed et al., 2009; Fisher et al., 2012; Papanicolas and McGuire, 2016), which instead is the main premise underpinning the use of outcome-based indicators in healthcare policy.

When risk-adjustment is not fully appropriately made, the unconditional use of such outcome indicators to provide financial incentives for quality appears rather unsatisfactory, if not damaging, and we need to improve our understanding on the relationship between unobservable quality and outcomes in hospitals. In this respect, further investigation of our microfounded model could help to get further insights on how these outcome indicators reflect differences in the unobservable quality and, in turn, on the use of them in healthcare policy. From this perspective, we investigate hereunder how a change in hospital quality affects the data generating process of mortality and readmission rates.

Considering the mortality outcome, from (29) we can see that a change in the hospital quality q_i generates two effects in the mortality rate of hospital i : “*Demand effect*” and “*Quality effect*”. The first effect is the change in the distribution of patients idiosyncratic risk due to the change in patients demand and, in particular, we can see that the *Demand effect* unambiguously increases the mortality rate. This is not surprising, however, as the higher treatment quality q_i increases the attractiveness of hospital i especially for those patients H at higher risk (i.e., $x'_H > 0$), whereas the higher congestion induced by more patients H might even lead patients L at lower risk to reduce the demand for that hospital (i.e., $x'_L \leq 0$ or, at least, $x'_L \leq x'_H$); therefore, the higher quality q_i makes the distribution of patients idiosyncratic risk worse for hospital i and, in turn, increases the mortality rate. On the contrary, the second effect is the pure *Quality effect* of an increase in quality in the probability of negative outcome (1) for each patient and, thus, it unambiguously reduces the mortality rate. Clearly, the overall effect will depend on the relative strength of these two effects and, in particular, when the demand is not extremely quality elastic there might be a scope to use (even not fully appropriately) risk-adjusted mortality rates as outcome indicators of the unobservable quality.

$$\frac{\partial MR_i[x_{H,L}^*(q_i)]}{\partial q_i} = \frac{(h_H - h_L)(x'_H x_L^* - x'_L x_H^*)(\gamma - \theta q_i)}{(x_H^* + x_L^*)^2} - \theta \frac{x_H^* h_H + x_L^* h_L}{x_H^* + x_L^*} \quad (29)$$

that is

$$\frac{\partial MR_i[x_{H,L}^*(q_i)]}{\partial q_i} = \underbrace{\text{Demand effect}}_{> 0} + \underbrace{\text{Quality effect}}_{< 0}$$

Looking at the readmission outcome, instead, we can notice from (30) that a change in quality q_i generates three effects in the readmission rate of hospital i : “*Demand effect*”, “*Quality effect*” and “*Selection effect*”. Similarly to the mortality outcome, a higher quality q_i , on the one hand, makes the distribution of patients idiosyncratic risk worse for hospital i and, in turn, increases the readmission rate (“*Demand effect*”); on the other hand, it reduces the probability of negative outcome (2) for each patient and, thus, reduces the readmission rate (“*Quality effect*”). Respect to (29), however, we can notice that an increase in quality q_i generates a further effect that increases unambiguously the readmission rate of hospital i . In particular, the *Selection effect* is the worsening of the idiosyncratic risk distribution of survived patients, induced by the improvement of hospital i ’ performance in the mortality outcome. As largely discussed above, in fact, the survival rates are different between the two patient groups and, thus, the mortality outcome changes the residual distribution of the idiosyncratic risk of patients that can be potentially readmitted. More specifically, an

improvement in the mortality outcome due to a higher quality q_i will presumably impact more especially those patients H at higher risk of mortality, implying a worsening of the idiosyncratic risk distribution of survived patients. Therefore, in the case of a higher quality q_i the *Selection effect* unambiguously increases the readmission rate of hospital i .

$$\begin{aligned} \frac{\partial RR_i[x_{H,L}^*(q_i)]}{\partial q_i} = & \frac{(h_H - h_L)(x'_H x_L^* - x'_L x_H^*) [1 - h_H(\gamma - \theta q_i)] [1 - h_L(\gamma - \theta q_i)]}{\{x_H^* [1 - h_H(\gamma - \theta q_i)] + x_L^* [1 - h_L(\gamma - \theta q_i)]\}^2} + \\ & + \frac{\theta x_H^* x_L^* (h_H - h_L)^2 \delta(\gamma - \theta q_i)}{\{x_H^* [1 - h_H(\gamma - \theta q_i)] + x_L^* [1 - h_L(\gamma - \theta q_i)]\}^2} + \\ & - \delta \theta \frac{x_H^* [1 - h_H(\gamma - \theta q_i)] h_H + x_L^* [1 - h_L(\gamma - \theta q_i)] h_L}{x_H^* [1 - h_H(\gamma - \theta q_i)] + x_L^* [1 - h_L(\gamma - \theta q_i)]} \end{aligned} \quad (30)$$

that is

$$\frac{\partial RR_i[x_{H,L}^*(q_i)]}{\partial q_i} = \begin{array}{ccc} \text{Demand effect} & + & \text{Selection effect} & + & \text{Quality effect} \\ > 0 & & > 0 & & < 0 \end{array}$$

Clearly, also for the readmission rate the overall effect of a higher hospital quality will depend on the relative strength of these three effects. Respect to the mortality rate, however, we have seen that the *Selection effect* could represent an additional source of bias in the readmission rate. Therefore, the readmission rates might not directly reveal differences in the unobservable quality of hospitals, even when the *Demand effect* is small. Moreover, we can notice that the *Selection effect* is more significant exactly in those hospitals with the best performance in the mortality outcome. From this perspective, when the risk-adjustment is not fully appropriately made, the use of readmission rates in performance programmes aiming to provide incentives for quality appears to be improper, as it might penalize exactly the best hospitals and, thus, might generate perverse incentives to hospitals' behaviour.

To clarify the intuition, we can imagine that, as a result of the incentives for increasing quality given by a performance programme, the best hospitals in the market increase their treatment quality, whereas the worse ones do not change. Then, the higher quality of the best hospitals should allow them to reduce the probability of dying for each patient and, thus, improve their performance in the mortality outcome (as long as the *Demand effect* is not too large); however, such improvement would lead to a worsening of the idiosyncratic risk distribution of survived patients potentially at risk of readmission (*Selection effect*) and, eventually, to an increase in the readmission rates of the best hospitals. Therefore, the result of such performance programme potentially would be higher readmission rates exactly for the best hospitals, which instead could be able to get an improvement in the mortality rates. These results are summarized in the following proposition:

Proposition 1. *When adjustment for the idiosyncratic risk is not fully appropriately made, readmission rates are worse indicators of hospital quality than mortality rates, and their use in performance programmes might penalize the best hospitals.*

As repeatedly underlined above, the source of the additional bias in the readmission rates as indicators of the hospital quality lies in the *Selection effect* generated by the mortality outcome. From this perspective, it seems intuitive that, when the specific medical treatment has by its nature a low risk of death, the extent of the *Selection effect* will be fairly small and, thus, the readmission rates will be more in line with the mortality rates.¹⁰

$$\frac{\partial RA_{F_MR_i}}{\partial \gamma} = \frac{\frac{1}{2}(h_H + h_L) x_H^* x_L^* (h_H - h_L)^2}{\{x_H^* [1 - h_H(\gamma - \theta q_i)] h_H + x_L^* [1 - h_L(\gamma - \theta q_i)] h_L\}^2} > 0 \quad (31)$$

This point is clearly illustrated in (31), showing that the “theoretical” risk-adjustment factor (28) for getting the risk-adjusted readmission rate has to be higher when the treatment specific severity index for mortality γ is higher. This simply says that, when a specific medical treatment has by its nature a higher risk of death, the risk-adjustment procedure has to do a greater work for cleaning the unadjusted readmission rates, because of the significant *Selection effect* in the mortality outcome. Therefore, the burden of a faulty risk-adjustment of readmission rates will be greater when the specific medical treatment has by its nature a higher risk of death. This further insight is summarized in the following proposition:

Proposition 2. *When adjustment for the idiosyncratic risk is not fully appropriately made, higher is the probability of mortality for the specific medical intervention, lower is the reliability of readmission rates as indicators of quality.*

4.2 Model simulation

5. Concluding remarks

In this paper we theoretically study hospitals’ and patients’ behaviour when patients differ in their idiosyncratic risk. We preliminary analyse standard risk-adjusted readmission and mortality rates for Italian hospitals and show that they are poorly correlated and yield completely different rankings of hospital performance, which can be viewed as a signal of

¹⁰ Notice that this statement appears perfectly in line with the evidence on “empirical” outcomes in Section 2.

imperfect risk-adjustment. Moving from this descriptive evidence, we present our model of hospitals' and patients' behavior that explicitly considers the generating process of the above indicators and allows for disentangling their relevant components. In a nutshell, we argue that hospital quality affects the distribution of the idiosyncratic risk of patient among hospitals, which in turn affects outcome indicators. More specifically, high quality hospitals will attract a higher share of patients with higher risk, which will affect their readmission and mortality rates. Then, we largely discuss the related implications and show that, when risk adjustment is not appropriately made, quality of the best hospitals will be undervalued by readmission rates, which are thus worse indicators of the hospital quality than mortality rates. Moreover, the higher is the probability of mortality for the specific medical intervention, the lower the reliability of the readmission rate as an indicator of quality.

Our results provide relevant implications for the optimal design of healthcare policies and for the evaluation of hospitals' performance. Our findings prove that, if patients' characteristics are not properly taken into account, the use of readmission rates as indicators of hospital quality might penalize exactly the best hospitals, yielding incorrect rankings in terms of hospital quality and performance. Consequently, their use as a basis for financial schemes may lead to perverse incentives for hospital behaviour, in contrast with the objective of stimulating the quality of the healthcare services.

References

Adashi, E.Y., Kocher, R.P. (2011), "Hospital Readmissions and the Affordable Care Act. Paying for Coordinated Quality Care", *Journal of American Medical Association* 306: 1794–1795.

Aletras, V.H. (1999), "A comparison of hospital scale effects in short-run and long-run cost functions", *Health Economics*, 8: 521-530.

Beitia, A. (2003), "Hospital quality choice and market structure in a regulated duopoly", *Journal of Health Economics* 22: 1011–1036.

Brekke, K.R., Siciliani, L., Straume, O.R. (2008), "Competition and waiting times in hospital markets", *Journal of Public Economics* 92: 1607–1628.

Brekke, K.R., Cellini, R., Siciliani, L., Straume, O.R. (2010), “Competition and quality in health care markets: A differential–game approach”, *Journal of Health Economics* 29: 508–523.

Brekke, K.R., Siciliani, L., Straume, O.R. (2011), “Hospital Competition and Quality with Regulated Prices”, *Scandinavian Journal of Economics* 113: 444–469.

Brekke, K.R., Siciliani, L., Straume, O.R. (2012), “Quality Competition with Profit Constraints”, *Journal of Economic Behavior & Organization* 84: 642–659.

Brekke, K.R., Siciliani, L., Straume, O.R. (2015), “Hospital competition with soft budget”, *The Scandinavian Journal of Economics* 117: 1019-1048.

Busse, R., Geissler, A., Quentin, W., Wiley, M. (2011), *Diagnosis-Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*. New York: Open University Press.

Cavalieri, M., Gitto, L., Guccio, C. (2013), “Reimbursement systems and quality of hospital care: An empirical analysis for Italy”, *Health Policy* 111: 273-289.

Cavalieri, M., Guccio, C., Lisi, D., Pignataro, G. (2014), “Financial Incentives and Inappropriateness in Healthcare: Evidence from Italian Cesarean Sections”, *FinanzArchiv/Public Finance Analysis* 70(3): 1–28.

Chua et al. (2010), “A two-stage estimation of hospital quality using mortality outcome measures: an application using hospital administrative data”.

Department of Health (2011), *Payment by Results Guidance for 2011-12*. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_124522.pdf.

Ellis, R.P (1998), “Creaming, skimping and dumping: provider competition on the intensive and extensive margins”, *Journal of Health Economics* 17: 537–555.

Evans, W.N., Garthwaite, C., Wei, H. (2008), “The impact of early discharge laws on the health of newborns”, *Journal of Health Economics* 27: 843-870.

Fischer et al. (2014), “Is the readmission rate a valid quality indicator? A review of the evidence”, *PLoS One* 9.

Fischer, C., Anema, H.A., Klazinga, N.S. (2012), “The validity of indicators for assessing quality of care: a review of the European literature on hospital readmission rate”, *European Journal of Public Health* 22: 484–491.

Foster, D., Harkness, G. (2010), *Healthcare reform: pending changes to reimbursement for 30-day readmissions*. Thomson Reuters.

Geissler, A., Scheller-Kreinsen, D., Busse, R. (2011), “Germany: Understanding G-DRGs”. In: Busse R., Geissler A., Quentin W., Wiley M. (Eds), *Diagnosis Related Groups in Europe. Moving Towards Transparency, Efficiency and Quality in Hospitals*. Maidenhead: Open University Press.

Gravelle, H., Santos, R., Siciliani, L. (2014), “Does a hospital’s quality depend on the quality of other hospitals? A spatial econometrics approach”, *Regional Science and Urban Economics* 49: 203–216.

Guccio, C., Lisi, D., Pignataro, G. (2016), “Readmission and Hospital Quality under Different Payment Regimes”, *FinanzArchiv/Public Finance Analysis*, forthcoming.

Hotelling, H. (1929), “Stability in competition”, *Economic Journal* 39: 41–57.

Howard, D.H. (2006), “Quality and consumer choice in healthcare: evidence from kidney transplant”, *Topics in Economic Analysis & Policy* 5: 1–20.

Laudicella, M., Li Donni, P., Smith, P.C. (2013), “Hospital readmission rates: Signal of failure or success?”, *Journal of Health Economics* 32: 909–921.

Ma, C-T. A. (1994), “Health Care Payment Systems: Cost and Quality Incentives”, *Journal of Economics & Management Strategy* 3: 93–112.

Ma, C-T. A., Mak, H. Y. (2015), “Information disclosure and the equivalence of prospective payment and cost reimbursement”, *Journal of Economic Behavior & Organization* 117: 439–452.

Makris, M., Siciliani, L. (2013), “Optimal Incentive Schemes for Altruistic Providers”, *Journal of Public Economic Theory* 15: 675–699.

Martin, S., Rice, N., Jacobs, R., Smith, P. (2007), “The market for elective surgery: Joint estimation of supply and demand”, *Journal of Health Economics* 26: 263–285.

Mohammed, M. A., J. J. Deeks, A. Girling, G. Rudge, M. Carmalt, A. J. Stevens, R. J. Lilford (2009), “Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals”, *British Medical Journal* 338.

Nuti, S., Vola, F., Amat, P. (2016), *Evaluating the network healthcare system performance*, MeS Laboratory, Pisa.

Papanicolas, I., McGuire, A. (2016), “Measuring and forecasting quality in English hospitals”, *Journal of the Royal Statistical Society: Series A*, Doi:10.1111/rssa.12203.

Propper, C., Burgess, S., Gossage, D. (2008), “Competition and Quality: Evidence from the NHS Internal Market”, *The Economic Journal* 118 (525): 138–170.

Propper, C., Burgess, S., Green, K. (2004), “Does competition between hospitals improve the quality of care?: Hospital death rates and the NHS internal market”, *Journal of Public Economics* 88 (7-8): 1247–1272.

Schreyogg, J., Stargardt, T. (2010), “The trade-off between costs and outcomes: the case of acute myocardial infarction”, *Health Services Research* 45:1585–1601.

Shen, Y. S. (2003), “The effect of financial pressure on the quality of care in hospitals”, *Journal of Health Economics* 22: 243–269.

Siciliani, L., Straume, O.R., Cellini, R. (2013), “Quality competition with motivated providers and sluggish demand”, *Journal of Economic Dynamics and Control* 37: 2041–2061.

Tay, A. (2003), “Assessing competition in hospital care markets: the importance of accounting for quality differentiation”, *RAND Journal of Economics* 34: 786–814.

Varkevisser, M., Van Der Geest, S.A., Schut, F.T. (2012), “Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands”, *Journal of Health Economics* 31: 371–378.

Xian, Y., Holloway, R.G., Chan, P.S., Noyes, K., Shah, M.N., Ting, H.H. (2011),
“Association between stroke center hospitalization for acute ischemic stroke and mortality”,
Journal of the American Medical Association 305: 373-380.