Health Care Manag Sci (2018) 21:348–375
DOI 10.1007/s10729-016-9387-x

Identifying congestion levels, sources and determinants on intensive care units: the Portuguese case

Diogo Ferreira[1] · **Rui Cunha Marques**[1]

Received: 27 April 2016 / Accepted: 10 October 2016 / Published online: 28 December 2016

Abstract Healthcare systems are facing a resources scarcity so they must be efficiently managed. On the other hand, it is commonly accepted that the higher the consumed resources, the higher the hospital production, although this is not true in practice. Congestion on inputs is an economic concept dealing with such situation and it is defined as the decreasing of outputs due to some resources overuse. This scenario gets worse when inpatients' high severity requires a strict and effective resources management, as happens in Intensive Care Units (ICU). The present paper employs a set of nonparametric models to evaluate congestion levels, sources and determinants in Portuguese Intensive Care Units. Nonparametric models based on Data Envelopment Analysis are employed to assess both radial and non-radial (in)efficiency levels and sources. The environment adjustment models and bootstrapping are used to correct possible bias, to remove the deterministic nature of nonparametric models and to get a statistical background on results. Considerable inefficiency and congestion levels were identified, as well as the congestion determinants, including the ICU specialty and complexity, the hospital differentiation degree and population demography. Both the costs associated with staff and the length of stay are the main sources of (weak) congestion in ICUs. ICUs management shall make some efforts towards resource allocation to prevent the congestion effect. Those efforts shall, in general, be focused on costs with staff and hospital days, although these congestion sources may vary across hospitals and ICU services, once several congestion determinants were identified.

Keywords Intensive Care Units · Data Envelopment Analysis · Congestion · Environment adjustment · Bootstrap

Abbreviations

CapC	Capital Costs
CI	Confidence Interval
CGSC	Costs of Goods Sold and Consumed
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
DSE	Degree of Scale Economies
EPE	*Entidade Pública Empresarial*
EU	European Union
GDP	Gross Domestic Product
GSI	Gini's Specialization Index
HC	Hospital Center
HospDays	Hospital Day(s)
ICU	Intensive Care Unit
InpD	Inpatient Discharge(s)
LHU	Local Health Unit
LVP	Law of Variable Proportions
MP	Marginal Product
RTS	Returns to Scale
SA	*Sociedade Anónima*
SDH	Strong Disposability Hull
SEServ	Supplies and External Services
SH	Singular Hospital
SMI	Service-Mix Index

✉ Diogo Ferreira
diogo.cunha.ferreira@tecnico.ulisboa.pt

Rui Cunha Marques
rui.marques@tecnico.ulisboa.pt

[1] CESUR, CERIS, Instituto Superior Técnico, University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

SPA	*Serviço Público Administrativo*
StaffC	Staff Costs
VRS	Variable Returns to Scale
WDH	Weak Disposability Hull

1 Introduction

The congestion effect on healthcare systems assumes a particular importance. By definition, congestion refers to situations where the reduction of one or more inputs generates an increase of at least one output. Cooper et al. [1] state that evidence of congestion is present when reductions in one or more inputs can be associated with increases in one or more outputs or, when increases in one or more inputs can be associated with decreases in one or more outputs, without worsening any other variable. Most of health managers, policy makers and even the public opinion may think that an increase of resources will always return an increase of produced outputs. This is not true in practice as the congestion phenomenon is not as unusual as it sounds. Particularly, within a period featured by both the necessity of resources saving and a growing demand for healthcare, any inefficiency source (as the inputs-related congestion is) must be carefully analyzed.

In 2013, about 9.1 % of the gross domestic product (GDP) was relative to the health sector in Portugal. The healthcare demand predictions are necessary for the hospital management in order to avoid situations such as poor service quality or inefficiency and congestion, being health care delivery systems particularly prone to congestion. According to Nayar et al. [2], hospitals face pressures to maximize performance in terms of production efficiency and quality, especially due to difficult financial environment factors, such as expansion of managed care, changes in public policy, growing market competition for certain services, and growth in the number of uninsured, [3]. In most situations, clinical practices may be characterized by intense health resources consumption, e.g. a large drugs supply, more surgeries and more hospital days than the necessary, which is not always beneficial for patients, eventually leading to the congestion phenomenon. That is why greater expenditures do not imply better quality on health care outputs.

By definition, the Intensive Care Unit (ICU) service is a specific internment ward dealing with critically ill inpatients, i.e. those ones who need for advanced, close and constant life support for 24 h a day due to their life-threatening illnesses/ injuries (e.g. trauma, multiple organ failure, sepsis, preterm birth, congenital disorder, birthing complications, cardiac arrest, acute myocardial infarction, intracranial hemorrhages). Most patients arriving to the ICU are admitted from the emergency department. After their treatment, patients from the ICU service(s) are usually transferred to another medical unit for further care, if their severity of illness is not sufficiently high to be considered for the ICU.

ICUs account for a considerable amount of hospital costs across the world with, historically, up to two-fold variation in risk-adjusted mortality, [4–5]. More recently, Halpern et al. [6] stated that ICU is the most expensive, technologically advanced and resource-intensive area of medical care, consuming about 13 % of hospital costs. As a consequence, ICU should, in general, be earmarked for those patients with severe and complex illness. According to Portuguese health care data, we conclude that, on average, a patient in ICU presents much higher costs than the national average. The ICU services cost, in Portugal, about seven times (per inpatient) more than a standard service, defined as the geometric mean of all Portuguese hospital services' related unitary costs. Additionally, Barrett et al. [7] state that critical care costs have been rising for decades, representing a costly segment of health care spending. Given these facts and the high severity of the inpatients threated in the ICU wards, resources allocated to ICUs must be managed in the most efficient and effective ways.

Several studies are devoted to the efficiency measurement of this service. For instance, Puig-Junoy [8] studied the Spanish ICUs performance. Some other studies about the ICU efficiency assessment can be found in the literature: Tsekouras et al. [9] and Dervaux et al. [10] are some remarkable ones. The latter uses a robust partial non-parametric and non-conditional frontier to assess some French ICU services efficiency. Meanwhile, the former analyses the Greek ICU services performance and the impact of the "*significant amount of financial resources [that] has been devoted by the Greek Government and the European Union*" to the ICU. That study uses a bootstrap-based bias-corrected efficiency measurement and the double bootstrap. Spain, France and Greece have similar demographic and epidemiological patterns to Portugal.

Several authors have dealt with the congestion phenomenon in healthcare provision. Table 5 on Appendix A provides a literature review on congestion measurement and/or on ICU performance evaluation, as well as some comments/ critiques on previous literature. For example, Clement et al. [11], Valdmanis et al. [12], Ferrier et al. [13], Arrieta and Guillén [14] and Matranga and Sapienza [15] employ an output-oriented nonparametric method to assess whether some undesirable outputs (e.g. mortality rates) are congesting hospital performance. In this case, the congestion is investigated over the *bad* outputs and how they could be reduced so as to improve the efficiency and effectiveness of hospitals. In all cases, radial models were employed, still ignoring the environment effect, the existence of other inefficiency sources and the possible existence of data noise. This means that a robustness analysis over their results is lacking, which may jeopardize conclusions that could be drawn from there. On the other hand, they investigated solely the impact of undesirable outputs, ignoring the fact that resources may also have a role on congestion (i.e. congestion over inputs), which is the way we take in the present study, in line with Simões and Marques [16]. This seems to be the most appropriate route within a period of resources scarcity, which in turn may have a considerable impact on ICU management and on the capacity and ability to treat highly critical ill patients. Quotes like "the more, the better" are common among the public opinion and healthcare managers but usually they are wrong precisely due to the congestion effect. This effect may then jeopardize the quality of care, in particular the one in ICUs. Ensuring the best resource allocation at the same time people's lives are saved is a hot topic in ICUs management. Since the congestion is an inefficiency parcel it must be mitigated. Although the congestion of ICUs has been previously studied under queuing theories, e.g. [17–18], so far no study has neither simultaneously employed a bias-corrected environment-based nonparametric congestion model, with both radial and non-radial (in)efficiency measures, concerning the ICU services, and identified the main sources of such phenomenon, over a strong statistical background, nor investigated the impact of the environment on ICU congestion. This paper then tries to overcome those faults, with an empirical application to the Portuguese ICU case.

This study is structured as follows: section 2 presents some different ways to measure congestion and its sources; section 3 presents the sample, the variables and the methods for environment and biasing adjustment; in section 4 we present and discuss the main results; finally, section 5 concludes this study.

2 Measuring the congestion levels and sources

2.1 Measuring the efficiency through non-parametric methods: an overview

The assessment of technical efficiency employing non-parametric linear envelopment of the data dates back to Farrell's work, [19]. Charnes et al. [20] introduced the Data Envelopment Analysis (DEA) estimator of technical efficiency. The technical efficiency of each decision making unit (DMU) is obtained through the comparison by distance with an efficient frontier formed by the best practices, [21], which use the lower level of inputs for a given output level, or produce the higher output level for a given level of resources. A DMU is an entity to be compared with others in similar conditions and produces the same kind of outputs from the same kind of resources (usually, in different proportions).

DEA has become the dominant approach to efficiency measurement in health care, as well as in many other sectors of the economy, such as education or justice, [22]. Ruggiero [23] and De Witte and Marques [24] point out several advantages of nonparametric methods over the parametric ones, such as the possibility of multiple inputs and multiple outputs inclusion and the fact that *a priori* it is not necessary to define the frontier shape. As a matter of fact, unlike the parametric approaches, where the analysis is driven by economic theory, DEA is a data-guided approach, [25]. Regarding health care, the techniques used are mainly based on DEA, [26], which is consistent with the economic theory underlying the optimizing behavior, and the frontier deviations can be interpreted as inefficiency and there are several ways to overshoot the noise. These are the reasons why DEA is considered hereafter.

The following DEA radial model, Eq. (1) [θ−model], can evaluate the (in)efficiency of a specific DMU_0 concerning the production of s different outputs, $y_{rj} \in \mathbb{R}^{s}_{+\cup\{0\}}$, $r = 1 \ldots s$, using m different inputs, $x_{ij} \in \mathbb{R}^{m}_{+}$, $i = 1 \ldots m$, [27−29], and under the potential influence of ϑ exogenous variables, $z_{hj} \in \mathbb{R}_{\vartheta}$, $h = 1 \ldots \vartheta$. In Eq. (1), ε is a non-Archimedean, $\varepsilon \sim 0$, and $p_i^- \in \mathbb{R}^{m}_{+\cup\{0\}}$ and $p_r^+ \in \mathbb{R}^{s}_{+\cup\{0\}}$ are slacks (non-radial inefficiencies) to be optimized by the linear model (1). The unit DMU_0 $(x_{i0}, y_{r0}, z_{h0}) \in \mathbb{R}^{m+s}_{+} \times \mathbb{R}^{\vartheta}$ is evaluated concerning a set of comparable units, Ω_0, which empirically determines a conditional frontier. This topic will be discussed below, see subsection 3.3. This model is output-oriented, thus $\theta_0^* \geq 1$.[1] A DMU_0 is technically efficient if and

[1] Hereinafter, stars * stand for linear programming models' variables optima.

only if $\theta_0^* = 1$, so it cannot increase its outputs without increasing at least one input. However, it is only *strongly efficient* concerning the strong disposability hull (*SDH*) if and only if $\theta_0^* = 1 \cap \forall_{i=1\ldots m} \forall_{r=1\ldots s} p_i^{-*} = p_r^{+*} = 0$.

$$\left\{\theta_0^*, p_1^{-*}\ldots p_m^{-*}, p_1^{+*}\ldots p_s^{+*}\right\}_{SDH} = \max_{\theta_0,\lambda_1\ldots\lambda_n,p_1^{-*}\ldots p_m^{-*},p_1^{+*}\ldots p_s^{+*}} \left\{ \theta_0 + \varepsilon\left(\sum_{i=1}^{m} p_i^- + \sum_{r=1}^{s} p_r^+\right) \middle| \begin{array}{c} \sum_{j=1;\, j\in\Omega_0}^{n} \lambda_j x_{ij} + p_i^- = x_{i0} \\ \sum_{j=1;\, j\in\Omega_0}^{n} \lambda_j y_{rj} - p_r^+ = \theta_0 y_{r0} \\ \sum_{j=1;\, j\in\Omega_0}^{n} \lambda_j = 1 \\ \lambda_j, p_r^+, p_i^- \geq 0 \\ j \in \Omega_0 \\ i = 1\ldots m; r = 1\ldots s \end{array} \right\} \quad (1)$$

2.2 Measuring the congestion

By replacing the objective function of Eq. (1) by $\beta_0^* = \max_{\beta_0,\lambda_j}\left(\beta_0 + \varepsilon \sum_{r=1}^{s} \tilde{p}_r^+\right)$ and the second constraint by $\sum_{j=1;\ j\in\Omega_0}^{n} \lambda_j y_{rj} - \tilde{p}_r^+ = \beta_0 y_{r0}$, we get a weak disposability output-oriented based model, henceforward β–model, cf. Eq. (2), [30]. Therefore, the DMU_0 is *strongly efficient* regarding the weak disposability hull (*WDH*) if and only if $\beta_0^* = 1 \cap \forall_{r=1\ldots s} \tilde{p}_r^+ = 0$. In this model, $\tilde{p}_r^+$ keeps the same meaning as p_r^+ in Eq. (1).

$$\left\{\beta_0^*, \tilde{p}_1^{+*}\ldots \tilde{p}_s^{+*}\right\}_{WDH} = \max_{\beta_0,\tilde{\lambda}_1\ldots\tilde{\lambda}_n,\tilde{p}_1^{+}\ldots \tilde{p}_s^{+}} \left\{ \beta_0 + \varepsilon \sum_{r=1}^{s} \tilde{p}_r^+ \middle| \begin{array}{c} \sum_{j=1;\, j\in\Omega_0}^{n} \tilde{\lambda}_j x_{ij} = x_{i0} \\ \sum_{j=1;\, j\in\Omega_0}^{n} \tilde{\lambda}_j y_{rj} - \tilde{p}_r^+ = \beta_0 y_{r0} \\ \sum_{j=1;\, j\in\Omega_0}^{n} \tilde{\lambda}_j = 1 \\ \tilde{\lambda}_j, p_r^+ \geq 0 \\ j \in \Omega_0 \\ i = 1\ldots m; r = 1\ldots s \end{array} \right\} \quad (2)$$

It is possible to show that, under the output-oriented framework, $\theta_0^* \geq \beta_0^* \geq 1$. In view of that, we can construct the output congestion score, C_0, on its standard way, as in Eq. (3), [31]. $C_0 < 1$ indicates the presence of congestion in the evaluated DMU, [30–35]. The value of C_0 indicates the amounts of outputs that should be increased to reach the non-congestion situation, so the lower C_0, the higher the congestion level. If $C_0 = 1$, the DMU_0 is not congested, i.e., there is absence of congestion inefficiency. However, even an inefficient DMU may be non-congested: it is sufficient that it does not belong to a weak disposability region, i.e., $\theta_0^* = \beta_0^* > 1$.

$$C_0\left(\theta_0^*, \beta_0^*\right) = \frac{\beta_0^*}{\theta_0^*} \leq 1 \quad (3)$$

Let $f_q : \mathbb{R}^q \rightarrow \mathbb{R}$ be an aggregation function of q arguments, e.g. the geometric mean. So, let's define the following inefficiency index, $\tilde{\Lambda}_0$, where the optima set $\left(\theta_0^*, p_i^{-*}, p_r^{+*}, \beta_0^*, \tilde{p}_r^{+*}\right) \in \mathbb{R}_{+\cup\{0\}}^{m+2s+2}$ are obtained by using all ϑ environment variables and Eqs. (1–2):

$$\tilde{\Lambda}_0 = f_m\left(\frac{x_{i0} - s_i^{-*}}{x_{i0}}\right) \cdot f_s\left(\frac{\beta_0^* y_{r0} + \tilde{p}_r^{+*}}{\theta_0^* y_{r0} + p_r^{+*}}\right) \leq 1 \quad (4)$$

If $\forall_{i=1\ldots m} p_i^{-*} = 0 \cap \forall_{r=1\ldots s} p_r^{+*} = \tilde{p}_r^{+*} = 0 \cap \theta_0^* = \beta_0^* = 1$, then the DMU_0 is technically efficient regarding the SDH and $\tilde{\Lambda}_0 = 1$. However, $\tilde{\Lambda}_0 < 1$ does not imply that congestion is present. As a matter of fact, if $\forall_{r=1\ldots s} p_r^{+*} = \tilde{p}_r^{+*} = 0 \cap \theta_0^* = \beta_0^* = 1 \cap \exists_{i=1\ldots m} p_i^{-*} \neq 0 \Rightarrow f_m\left(\frac{x_{i0}-p_i^{-*}}{x_{i0}}\right) < 1 \cap f_s\left(\frac{\beta_0^* y_{r0}+\tilde{p}_r^{+*}}{\theta_0^* y_{r0}+p_r^{+*}}\right) = 1 \Rightarrow \tilde{\Lambda}_0 < 1$ and the DMU_0 is just technically inefficient regarding SDH. Only when $f_s\left(\frac{\beta_0^* y_{r0}+\tilde{p}_r^{+*}}{\theta_0^* y_{r0}+p_r^{+*}}\right) < 1$ congestion can be identified, [36]. This means that the non-radial inputs inefficiency component must be removed from Eq. (4), leading to a congestion composite index:

$$\Lambda_0 = f_s\left(\frac{\beta_0^* y_{r0} + \tilde{p}_r^{+*}}{\theta_0^* y_{r0} + p_r^{+*}}\right) \leq 1 \tag{5}$$

Λ_0 is comparable to $C_0\left(\theta_0^*, \beta_0^*\right)$. Indeed, if $\forall_{r=1\ldots s} p_r^{+*} = \tilde{p}_r^{+*} = 0$, then $\Lambda_0 = C_0\left(\theta_0^*, \beta_0^*\right)$. Still, Λ_0 encompasses both radial and non-radial inefficiency sources, so it is a more robust measure of congestion than $C_0\left(\theta_0^*, \beta_0^*\right)$. Additionally, the ratio $\frac{\tilde{\Lambda}_0}{\Lambda_0}$ measures the extension of congestion in the whole inefficiency. It is easy to show that $f_m\left(\frac{x_{i0}-s_i^{-*}}{x_{i0}}\right) \leq 1$, Eq. (4), which means that $\frac{\tilde{\Lambda}_0}{\Lambda_0} \leq 1$.

2.3 Marginal Products and Scale Elasticities

Dual formulation of Eq. (1) is given by Eq. (6), where u_r and v_i represent, respectively, the virtual weights of the r-th output and the i-th input, μ is a variable that controls for returns to scale (RTS) (CRS – constant returns to scale, VRS – variable returns to scale), and ε (a non-Archimedean) ensures variables' weights are non-zero. Replacing the first constraint of Eq. (1) by $\sum_{j=1;\ j\in\Omega_0}^{n} \lambda_j x_{ij} = x_{i0}$ is equivalent to unrestraint the inequity $v_{i0} \geq \varepsilon$ in Eq. (6), [30]. That is, $\beta_0^* = \min_{\tilde{v}_{10}\ldots\tilde{v}_{m0},\tilde{u}_{10}\ldots\tilde{u}_{s0},\tilde{\mu}_0} \left[\sum_{i=1}^{m} \tilde{v}_{i0} x_{i0} - \tilde{\mu}_0\right]$ if $\tilde{v}_{i0}$ is free in sign, $\forall i = 1 \ldots m$, cf. Eq. (7), [31, 35]. In Eqs. (6) and (7), u_{r0} and $\tilde{u}_{r0}$ share the same meaning; the same applies to v_{i0} and $\tilde{v}_{i0}$. Tiles are utilized to differentiate them. As usually there is no reason to believe that one input has a considerable higher impact on congestion than the remaining resources, no further restrictions to Eq. (7) are required. That is, Eq. (7) allows finding out which input is a congestion source with no further assumptions.

$$\left\{\theta_0^*\right\}_{SDH} = \min_{v_{10}\ldots v_{m0},u_{10}\ldots u_{s0},\mu_0} \left\{ \sum_{i=1}^{m} v_{i0} x_{i0} - \mu_0 \,\middle|\, \begin{array}{c} \sum_{r=1}^{s} u_{r0} y_{rj} - \sum_{i=1}^{m} v_{i0} x_{ij} + \mu_0 \leq 0 \\ \sum_{r=1}^{s} u_{r0} y_{r0} = 1 \\ u_{r0} \geq \varepsilon \\ v_{i0} \geq \varepsilon \\ \mu_0 \text{ free in sign} \Leftrightarrow \mu_0 \geq -\infty \\ i = 1\ldots m; r = 1\ldots s \\ \mathrm{j} \in \Omega_0 \end{array} \right\} \tag{6}$$

$$\left\{\beta_0^*\right\}_{WDH} = \min_{\tilde{v}_{10}\ldots\tilde{v}_{m0},\tilde{u}_{10}\ldots\tilde{u}_{s0},\tilde{\mu}_0} \left\{ \sum\nolimits_{i=1}^{m} \tilde{v}_{i0} x_{i0} - \tilde{\mu}_0 \,\middle|\, \begin{array}{c} \sum_{r=1}^{s} \tilde{u}_{r0} y_{rj} - \sum_{i=1}^{m} \tilde{v}_{i0} x_{ij} + \tilde{\mu}_0 \leq 0 \\ \sum_{r=1}^{s} \tilde{u}_{r0} y_{r0} = 1 \\ \tilde{u}_{r0} \geq \varepsilon \\ \tilde{v}_{i0} \text{ free in sign} \Leftrightarrow \tilde{v}_{i0} \geq -\infty \\ \tilde{\mu}_0 \text{ free in sign} \Leftrightarrow \tilde{\mu}_0 \geq -\infty \\ i = 1\ldots m; r = 1\ldots s \\ \mathrm{j} \in \Omega_0 \end{array} \right\} \tag{7}$$

The Law of Variable Proportions (LVP) states that as the quantity of one factor increases, keeping the other factors fixed, the marginal product (MP), or marginal rate of production, of that factor will eventually decline after a certain stage, [32–34]. When the variable factor becomes abundant, the MP may become negative. Then, the total product (defined as the total of outputs resulting from efforts of all factors of production) decreases *if and only if* the MP value of that factor becomes negative, i.e., when the entity is congested. Additionally, Sueyoshi [35] demonstrated that the MP between the i-th input x_{i0} and the r-th output y_{r0}, of DMU_0, is given by Eq. (8). Such formula was derived from the first restriction of Eq. (7). Under the weak disposability assumption, $\tilde{v}_{i0}^*$ can be negative; if this is the case and since $\tilde{u}_{r0}^* \geq \varepsilon > 0$, by Eq. (7), then MP (x_{i0}, y_{r0}) is negative as well, which means that the i-th input is one source of congestion because an increase of such an input reduces the quantity of the r-th produced output. That is, MP $(x_{i0}, y_{r0}) < 0$, such that *MP* is computed between the r-th output and the i-th input reveals the existence of congestion in the unit (x_{i0}, y_{r0}), and then $\beta_0^* < \theta_0^*$ and $\Lambda_0 < 1$, in other words, the i-th input of DMU_0 is congested. In view of that, congestion sources can be easily identified as those inputs leading to a negative *MP* (which, in turn, is consistent with the economic meaning of congestion).

$$MP(x_{i0}, y_{r0}) = {}^{\partial y_{r0}}\!/\!_{\partial x_{i0}} = {}^{\tilde{v}_{i0}^*}\!\Big/\!_{\tilde{u}_{r0}^*} \tag{8}$$

Consistent with the preceding argument, one can easily compute the corresponding scale elasticity, ρ_0, as in Eq. (9), being ρ_0^{+*} given by Eq. (10), [30, 36][2]:

$$\rho_0 = \left(\rho_0^{+*} + \rho_0^{-*}\right)/2 \tag{9}$$

$$\rho_0^{+*} = 1 + \max_{\tilde{\tilde{v}}_1,\ldots,\tilde{\tilde{v}}_m,\tilde{\tilde{u}}_1,\ldots,\tilde{\tilde{u}}_s,\rho_0^+} \left\{ \rho_0^+ \;\middle|\; \begin{array}{c} \sum_{r=1}^{s} \tilde{\tilde{u}}_r y_{rj} - \sum_{i=1}^{m} \tilde{\tilde{v}}_i x_{ij} + \rho^+ \leq 0 \\ \sum_{r=1}^{s} \tilde{\tilde{u}}_r y_{r0} - \sum_{i=1}^{m} \tilde{\tilde{v}}_i x_{i0} + \rho^+ = 0 \\ \sum_{r=1}^{s} \tilde{\tilde{u}}_r y_{r0} = 1 \\ \tilde{\tilde{u}}_r \geq \varepsilon \\ \tilde{\tilde{v}}_i \text{ free in sign} \Leftrightarrow \tilde{\tilde{v}}_i \geq -\infty \\ \rho_0^+ \text{ free in sign} \Leftrightarrow \rho_0^+ \geq -\infty \\ i = 1, \ldots, m; r = 1, \ldots, s \\ \mathrm{j} \in \Omega_0 \end{array} \right\} \tag{10}$$

Clearly, $\rho_0^{+*} \geq \rho_0^{-*}$, which means that if $\rho_0^{+*} < 0$, then $\rho_0 < 0$ as well, i.e., the Degree of Scale (Dis) Economies (DSE) is negative for the DMU_0 (x_{i0}, y_{r0}, z_{h0}). However, a DMU can exhibit $\rho_0 < 0 \cap \rho_0^{+*} \geq 0$, which happens when $\rho_0^{-*} < -\rho_0^{+*}$. Negative RTS exists if and only if $\rho_0^+ < 0 \Rightarrow \rho_0 < 0$, [30, 36], i.e., $MP(x_{i0}, y_{r0}) < 0 \Leftrightarrow \rho_0^+ < 0 \Rightarrow \rho_0 < 0$.

2.4 Weak congestion

The previous approach assumes that, under the congestion phenomenon, a "*proportional reduction in all inputs warrants an increase in all outputs*", [30], which is a rather restrictive assumption. Tone and Sahoo [36] call it *strong congestion*, so they relax that assumption and introduce the *weak congestion* concept, such that "*an increase in one or more inputs causes a decrease in one or more outputs*", [30]. Strong congestion implies weak congestion, but the reciprocal is not necessarily true, [36]. Because of that, their proposal relies on a semi-radial (and units invariant) approach, as in Eq. (11), where t_r^+ and t_i^- are slacks to be optimized, as before ε is a non-Archimedean number and $(x^{\vee}{}_{i0}, y^{\vee}{}_{r0}) = \left(x_{i0}, \beta_0^* y_{r0} + \tilde{p}_r^{+*}\right)$, i.e., (x_{i0}, y_{r0}) is projected on WDH, cf. Eq. (2), that is $(x^{\vee}{}_{i0}, y^{\vee}{}_{r0})$ is efficient concerning the WDH technology (frontier).

[2] ρ_0^{-*} is achieved by minimizing the linear program in Eq. (10), instead of maximizing.

$$\left\{t_1^{+*}\ldots t_s^{+*},t_1^{-*}\ldots t_m^{-*}\right\}_{SDH} = \max_{t_1^+,\ldots,t_s^+,t_1^-,\ldots,t_m^-,\lambda_1,\ldots,\lambda_n}\left\{\frac{1}{s}\sum_{r=1}^{s}\frac{t_r^+}{y_{r0}}+\frac{\varepsilon}{m}\sum_{i=1}^{m}\frac{t_i^-}{x_{i0}}\;\middle|\;\begin{array}{c}\sum\limits_{j=1;j\in\Omega_0}^{n}\lambda_j x_{ij}+t_i^-=\check{x}_{i0}\\ \sum\limits_{j=1;j\in\Omega_0}^{n}\lambda_j y_{rj}-t_r^+=\check{y}_{r0}\\ \sum\limits_{j=1;j\in\Omega_0}^{n}\lambda_j=1\\ \lambda_j,t_i^-,t_r^+\geq 0\\ j\in\Omega_0\\ i=1,\ldots,m;r=1,\ldots,s\end{array}\right\} \tag{11}$$

From the $\left\{t_1^{+*},\ldots,t_s^{+*},t_1^{-*},\ldots,t_m^{-*}\right\}$ optima obtained in Eq. (11), it is possible to construct a ratio measuring the average improvement in outputs to the average reduction in inputs, [30, 36], as in Eq. (12), where s' and m' respectively represent the numbers of positive t_r^{+*} and positive t_i^{-*}, [30]. In short, such a ratio is a DSE measure for weakly congested units. Still, sources of congestion can be identified by those slacks, $\left\{t_1^{+*},\ldots,t_s^{+*},t_1^{-*},\ldots,t_m^{-*}\right\}\neq\overrightarrow{0}$.

$$DSE_0=-\left(\frac{1}{s'}\sum_{r=1}^{s}\frac{t_r^{+*}}{y_{r0}}\right)\Big/\left(\frac{1}{m'}\sum_{i=1}^{m}\frac{t_i^{-*}}{x_{i0}}\right) \tag{12}$$

As in Eq. (5), we define a composite congestion index for weak congestion:

$$\begin{aligned}\Lambda_{\text{weak},0}&=f_m\left(\frac{x_{i0}-t_i^{-*}}{x_{i0}}\right)\Big/f_s\left(\frac{\beta_0^*\cdot y_{r0}+\tilde{p}_r^{+*}+t_r^{+*}}{y_{r0}}\right)\\&=f_m\left(1-\frac{t_i^{-*}}{x_{i0}}\right)\Big/f_s\left(\frac{\Lambda_0\cdot\left(\theta_0^*\cdot y_{r0}+p_r^{+*}\right)+t_r^{+*}}{y_{r0}}\right)\\&=f_m\left(1-\frac{t_i^{-*}}{x_{i0}}\right)\Big/f_s\left(\Lambda_0\cdot\theta_0^*+\frac{\Lambda_0\cdot p_r^{+*}+t_r^{+*}}{y_{r0}}\right)\leq 1\end{aligned} \tag{13}$$

Where Λ_0 is obtained from Eq. (5), p_r^{+*} from Eq. (1), and $\left\{t_1^{+*}\ldots t_s^{+*},t_1^{-*}\ldots t_m^{-*}\right\}$ from Eq. (11). It is worthy to mention that strong congestion implies weak congestion ($\Lambda_0<1\Rightarrow\Lambda_{weak,0}<1$), but even DMUs with no strong congestion can exhibit weak congestion though, i.e., $\Lambda_0=1\not\Rightarrow\Lambda_{weak,0}=1$. As a matter of fact, $\Lambda_0=1\Rightarrow\Lambda_{\text{weak},0}=f_m\left(1-\frac{t_i^{-*}}{x_{i0}}\right)\Big/f_s\left(\frac{\theta_0^*y_{r0}+p_r^{+*}+t_r^{+*}}{y_{r0}}\right)$ which is unitary if and only if $\forall_{r=1,\ldots,s}p_r^{+*}=t_r^{+*}=0\cap\forall_{i=1,\ldots,m}t_i^{-*}=0\cap\theta_0^*=1$, i.e., if the unit is technically efficient regarding SDH. If there is no weak congestion ($\Lambda_{weak,0}=1$), then by definition $\forall_{r=1\ldots s}t_r^{+*}=0\cap\forall_{i=1\ldots m}t_i^{-*}=0$, which means that the DMU $(\check{x}_{i0},\check{y}_{r0})$ is strongly efficient regarding SDH and $\left\{\theta_0^*=1\cap\forall_{r=1\ldots s}p_r^{+*}=0\right\}\Rightarrow\left\{\beta_0^*=1\cap\forall_{r=1\ldots s}\tilde{p}_r^{+*}=0\right\}$, and finally, $\Lambda_0=1$. That is, $\Lambda_{weak,0}=1\Rightarrow\Lambda_0=1$, in other words, if the DMU_0 has no evidence of weak congestion, then it is not congested. So, $\Lambda_{weak,0}$ defines the link between the weak and the strong congestion measures; the ratio $\Lambda_{weak,0}/\Lambda_0$ gives, then, the extent of weak congestion in the whole congestion. It is easy to show that $0<\frac{\Lambda_{\text{weak},0}}{\Lambda_0}\leq 1$.

2.5 Final considerations regarding the congestion models

Input-congestion models have been defined so far. However, a strategy must be employed so as to check whether DMUs are effectively congested or not, and if so, whether there are either strongly or weakly congested, or not. Such strategy is summarized in Table 1.

3 Data and methodological issues

3.1 Sample

For this paper, each DMU is a different ICU service in a specific hospital. That is, a hospital, which has several different ICU services, has several DMUs. A measure of homogenization between them is required and discussed below. The sample is constituted by 630 DMUs, distributed across 8 years (2002-2009),[3] and four ICU specialties: (a) Polyvalent ICUs, # = 278 DMUs, (b) Cardiology ICUs, # = 143 DMUs, (c) Pediatric, Gynecology, Obstetrics and Neonatology ICUs, # = 105 DMUs, and (d) Surgical ICUs, # = 104 DMUs. Surgical ICUs' classification

[3] We are aware that data can be somehow old. Still, there is no apparent reason to believe that both congestion sources and the environment impact on congestion could significantly change till the present days.

Table 1 Strategy

Step	Equation(s) to be solved	Expected results	Hypothesis being tested	Condition(s) to not reject that hypothesis *	Next step, if that condition(s) holds	Next step, otherwise
1	(9-10)	$\tilde{\rho}_0^{+*}$, $\tilde{\rho}_0^{-*}$ and $\tilde{\rho}_0$ (as a DSE measure) **	DMU_0 is neither strongly nor weakly congested	$\begin{cases} \rho_0 > 0 \\ \rho_0^{+*} \geq 0 \\ \rho_0^{-*} \geq 0 \end{cases}$	2	4
2	(1)	$\{\theta_0^*, p_1^{-*} \ldots p_m^{-*}, p_1^{+*} \ldots p_s^{+*}\}_{SDH}$	DMU_0 is technically and strongly efficient regarding SDH (then, not congested)	$\begin{cases} CI(\theta_0^*) \supset \{1\} \\ \forall_{i=\ldots m} \max\{CI(p_i^{-*})\} \leq \varepsilon \\ \forall_{r=1\ldots s} \max\{CI(p_r^{+*})\} \leq \varepsilon \end{cases}$	Stop	3
3		N.A.	DMU_0 is technically inefficient regarding SDH, but not congested	$\begin{cases} CI(\theta_0^*) \supset \{1\} \\ \exists_{i=1\ldots m} \max\{CI(p_i^{-*})\} >> \varepsilon \text{, or equivalently,} \\ \min\{CI(p_{r=1}^{+*})\} = \varepsilon \end{cases}$ $\begin{cases} \forall_{r=1\ldots s} y_{r0} \in CI(\theta_0^* \cdot y_{r0} + p_r^{+*}) \\ \exists_{i=1,\ldots,m} x_{i0} \notin CI(x_{i0} + p_i^{-*}) \end{cases}$	Stop	
4	(5, 7)	$\{\beta_0^*, \tilde{u}_{10}^* \ldots \tilde{u}_{s0}^*, \tilde{v}_{10}^* \ldots \tilde{v}_{m0}^*\}_{WDH}$ $\Lambda_0 = \frac{\beta_0^* y_{10} + \tilde{p}_1^{+*}}{\theta_0^* y_{10} + p_1^{+*}}$	DMU_0 exhibits strong congestion	$\rho_0^{+*} < 0$ Use Λ_0 to assess the strong congestion level.	5	6
5	(8)	$MP(x_{i0}, y_{10})$	The i-th input is a source of (strong) congestion	$\max\{CI(MP(x_{i0}, y_{10}))\} < 0$	Stop	N.A.
6	(2, 11–13)	$\{t_1^{+*} \ldots t_s^{+*}, t_1^{-*} \ldots t_m^{-*}\}_{SDH}$ DSE $\Lambda_{weak,0}$ ****	DMU_0 exhibits weak congestion; the i-th input is a source of weak congestion if $t_i^{-*} \neq 0$	$\begin{cases} \rho_0^{+*} \geq 0 \\ \rho_0 < 0 \end{cases}$ Use $\Lambda_{weak,0}$ to assess the weak congestion level.		

* CI – confidence interval, 95 %; ε is a Non-Archimedean; **. Under this framework, there is solely one output, which means that $\tilde{\rho}_0^{+*} = \sum_{i=1}^{m} \overset{\approx}{v}_i^* x_{i0}$, from Eq. (10). Identically, $\tilde{\rho}_0^{-*} = 1 + \min_{\bar{\bar{v}}_i, \bar{\bar{u}}_r, \rho^-} \tilde{\rho}^- = \sum_{i=1}^{m} \bar{\bar{v}}_i^* x_{i0}$, which results into $\tilde{\rho}_0 = \frac{1}{2} \sum_{i=1}^{m} \left(\overset{\approx}{v}_i^* - \bar{\bar{v}}_i^* \right) x_{i0}$, from Eq. (9). That is, the scale elasticity is basically the relationship between the optimal virtual weights of inputs obtained from maximization and minimization linear problem. *** N.A. – not applicable; **** The aggregating function f_m is assumed to be the average of its arguments, i.e., $f_m(\vec{g}) = \sum_{i=1}^{m} g_i / m$. By means of the present case, as there is only one output, there is no need of an aggregation function f_s

includes "*General surgical ICUs*" (# = 41), "*Neurosurgical ICUs*" (# = 36), "*Cardiothoracic surgical ICUs*" (# = 9), "*Transplants' ICUs*" (# = 6) and "*Burn Units*" (# = 12). Those 630 DMUs are spread over a range of 25 – 40 hospitals (depending on the year), giving an average of 2 – 3 ICUs per hospital. Still, the DMU definition remains as the ICU specialty service.

3.2 Variables

Given, at least, the theoretical financial unsustainability of the health system and/ or the high financing provided to this department by most governments, [9], an economic outlook is desirable. Therefore, the following inputs were chosen, [27, 28, 9–10][4,5]:

i. X_1 – *Costs of Goods Sold and Consumed* (CGSC) – expenditures with drugs and clinical materials;

[4] We do not include "*work force*" variables, such as number of nurses and doctors, as inputs, once they are multidisciplinary, working in different hospital dimensions, but the information provided by the official sources do not allow to disentangle the staff number working in ICU from other departments.

[5] All required data for this research is available at the official database of the Portuguese Ministry of Health, the Central Administration of Health Systems, cf. http://www.acss.min-saude.pt/, in lawful annual reports of each hospital, and in http://www.pordata.pt/en/Municipalities.

ii. X_2 – *Supplies and External Services* (SEServ) – expenditures with external labor outsourcing;
iii. X_3 – *Staff Costs* (StaffC) – expenditures with staff, including salaries and bonuses to physicians, nurses and other (non-administrative) ancillary staff;
iv. X_4 – *Capital Costs* (CapC) – expenditures with technological asset investments;
v. X_5 – *(ICU) Hospital Days* (HospDays) – total number of days used by all inpatients treated on the ICU service, within 1 year, [27], as time is a resource required to produce the output(s), [10].

Additionally, the following output was chosen, [27, 28, 9, 10]:

i. *Y – Inpatient Discharges* (InpD) – total number of patients treated within a specific ICU service (DMU) in a year, excluding deaths.[6]

Costs (inputs i. up to iv.) were all updated to 2009 by using the GDP deflator. Note that the higher the length of stay (HospDays), the higher the probability of other diseases appearance, such as nosocomial infections and pressure ulcers in bedridden patients, increasing the ICU mortality rate (or, at least, its associated death probability), as stated by Ferreira and Marques [28] and Chan et al. [18]. Given the limited capacity (number of beds) of the service, the higher the number of (ICU) hospital days, the lower the possible discharges. That is, HospDays may contribute to the ICU services congestion. We also assume that the remaining inputs are prone to congestion. Indeed, we have observed a considerable and positive correlation between all costs and HospDays, which is an expected result because more inpatient days require more expenses with staff (nurses), drugs and other clinical material. As a result, since we suspect that HospDays can be a source of congestion, the remaining inputs can be either. As a matter of fact, by definition, if there is an entity producing more InpD than a specific DMU_0, spending fewer resources, say StaffC and keeping the remaining inputs unchanged, then StaffC is obviously congested on DMU_0 because the production could be increased at the same time that StaffC would be decreased. The advantage of using the previously described nonparametric methods is that we do not need to make strong assumptions over the inputs; we only assume that it is somehow possible that they can be eventually congested.

Table 2 contains the descriptive statistics of the main variables utilized in our analysis, by ICU specialty. As we can observe, there is considerable resources consumption in ICUs, but also a huge heterogeneity on both resources consumption and outputs production (standard deviations and averages are quite close). In general, ICUs resource consumption lies essentially on staff costs (StaffC) and costs with drugs and clinical material (CGSC), while capital investments and expenses with outsourcing are generally low compared with the other costs. Surgical ICUs are responsible for the majority of inpatients treated, being followed by Cardiology ICUs. Furthermore, if we assume that the average delay is a measure of the inpatients complexity (because the higher the average delay, the higher the expected inpatient needs as well as their own complexity, [38]), as is the case-mix, then we also observe a significant diversity. As expected, the most complex services (as Burn Units and Transplants ICUs) deal with more complex inpatients, which present *a priori* higher probability of death, and then require a higher length of stay.

3.3 Adjusting for internal and external environment variables

Section 2 has detailed the models to be utilized so as to assess both congestion levels and sources. However, for the sake of comparability issues, units that create the appropriate reference set are addressed to Ω_0, which is computed for each unit $(x_{i0}, y_{r0}, z_{h0}) \in \mathbb{R}_+^m \times \mathbb{R}_+^s \times \mathbb{R}^{\vartheta}$. It is commonly accepted that DMUs efficiency must be assessed taking into account the environment they face, which in turn may jeopardize/ benefit the units' performance. By environment we mean both internal (e.g. legal status) and external environments (e.g. demographic patterns). Adjusting for the environment allows homogenizing the sample.

The question is how to derive such set Ω_0. Let's consider a generic unit $(x_{i0}, y_{r0}) \in \mathbb{R}_+^m \times \mathbb{R}_+^s$, characterized by ϑ different characteristics, $z_{h0} \in \mathbb{R}^{\vartheta}$, $h = 1\ldots\vartheta$. These features (variables) can be either independent, $z_{h_u0} \in \mathbb{R}^{\vartheta_u}$, h_u=1… ϑ_u, or dependent, $z_{h_w0} \in \mathbb{R}^{\vartheta_w}$, h_w=1… ϑ_w, being $\vartheta = \vartheta_w + \vartheta_u$ and $z_{h0} = \{z_{h_u0} \cup z_{h_w0}\}$. Let also $K_{h_u} : \mathbb{N} \to \mathbb{R}_{[0;1]}$ and $K_{h_w} : \mathbb{R} \to \mathbb{R}_{[0;1]}$ be two kernel functions with compact support (e.g. Epanechnikov), and $h_u > 0$ and $h_w > 0$ some appropriate bandwidths for those kernels, triggering the comparability between units for each criterion. In other words, only those DMUs whose operational environment variables are close to DMU $(x_{i0}, y_{r0}, z_{h0}) \in \mathbb{R}_+^m \times \mathbb{R}_+^s \times \mathbb{R}^{\vartheta}$ can be utilized to compose the latter reference set, Ω_0. This reference set is, then, achieved by a global kernel function, as discussed below.

Under a multivariate framework, i.e., ϑ>0, it is common to assume that environmental variables share no dependence between them, so a natural choice for the global kernel, $K : \mathbb{R}^{\vartheta} \to \mathbb{R}_{[0;\,1]}$, would be the product of all univariate kernels, $K_h : \mathbb{R} \to \mathbb{R}_{[0;\,1]}$, [27], as in Eq. (14), being $\vec{u} = \{u_h; h = 1\ldots\vartheta\} = \left\{\frac{z_{h0} - z_{hj}}{H_h}; h = 1\ldots\vartheta\right\} \in \mathbb{R}^{\vartheta}$.

[6] The adjustment for environment (subsection 3.3) shall be enough for inpatients complexity accounting, as claimed by Ferreira and Marques [27].

Table 2 Variables basic statistic, per ICU specialty. Values are presented in the form "average ± std. deviation" and costs (CGSC, SEServ, StaffC and CapC) in thousand €

ICU Specialty	*Y (InpD)*	*X₁ (CGSC)*	*X₂ (SEServ)*	*X₃ (StaffC)*	*X₄ (CapC)*	*X₅ (HospDays)*	AvDelay *
Polyvalent ICUs	360.04 ± 394.47	816.67 ± 746.74	108.87 ± 110.06	1519.15 ± 1355.10	120.39 ± 168.04	2983.47 ± 2760.94	11.48 ± 9.78
Cardiology ICUs	611.45 ± 425.42	545.56 ± 891.16	104.25 ± 134.85	608.74 ± 405.18	74.93 ± 192.80	2075.50 ± 1175.20	4.57 ± 3.57
Pediatric, Gynecology, Obstetrics and Neonatology ICUs	491.48 ± 1157.69	307.91 ± 216.41	98.48 ± 135.78	1249.42 ± 913.96	70.25 ± 69.46	3388.45 ± 2953.04	11.48 ± 9.86
General surgical ICUs	807.15 ± 616.02	504.82 ± 281.97	58.99 ± 39.73	935.31 ± 451.83	226.63 ± 429.94	2330.73 ± 1048.27	5.45 ± 9.33
Neurosurgical ICUs	576.81 ± 386.92	478.26 ± 249.26	36.44 ± 38.29	932.31 ± 546.34	46.39 ± 32.32	3153.78 ± 1575.25	7.76 ± 7.77
Cardiothoracic surgical ICUs	605.33 ± 374.26	529.75 ± 233.19	11.64 ± 14.66	751.97 ± 411.80	173.52 ± 348.15	2256.89 ± 993.61	9.40 ± 17.41
Transplants ICUs	94.33 ± 74.50	146.82 ± 69.67	28.86 ± 38.94	319.93 ± 164.80	32.19 ± 55.15	1083.50 ± 770.01	14.06 ± 5.10
Burn Units	89.42 ± 53.72	464.72 ± 230.93	50.15 ± 27.31	853.94 ± 232.28	71.92 ± 48.22	1916.17 ± 744.25	24.45 ± 5.52

*AvDelay – Average Delay (= HospDays/ InpD)

$$K\left(\overrightarrow{u}\right)=\prod_{h=1}^{\vartheta}K_h(u_h) \tag{14}$$

However, it is not always true that environment variables share no interdependence. Unless their correlation is quite low, we may take advantage of such dependence. To do so, we adapt the approach proposed by Daraio and Simar [38], which can be synthetized as in Eq. (15), where S is the covariance matrix of the ϑ variables u_h, and as usual n is the sample size and $\mathbb{I}$ is the indicator function.

$$K(\overrightarrow{u})=\frac{\mathbb{I}\left(\operatorname{diag}\left(\overrightarrow{u}^T S^{-1}\overrightarrow{u}\leq^{\vartheta+4}\sqrt{\frac{16}{n^2(\vartheta+2)^2}}\right)\right)}{\int_{\mathbb{R}^{\vartheta}} d\overrightarrow{u}\,\mathbb{I}\left(\operatorname{diag}\left(\overrightarrow{u}^T S^{-1}\overrightarrow{u}\leq^{\vartheta+4}\sqrt{\frac{16}{n^2(\vartheta+2)^2}}\right)\right)}\in[0;1] \tag{15}$$

Splitting environment variables into independent (categorical), u, and depedent (either discrete, categorical or continous), w, allows us to create a global multivariate kernel, $K:\mathbb{R}^{\vartheta_w}\times\mathbb{R}^{\vartheta_u}\rightarrow\mathbb{R}_{[0;1]}$, as follows:

$$K\left(\overrightarrow{u},\overrightarrow{w}\right)=\prod_{h_u=1}^{\vartheta_u}K_{h_u}(u_{h_u}) \times\frac{\mathbb{I}\left(\operatorname{diag}\left(\overrightarrow{w}^{\mathrm{T}}(S)^{-1}\overrightarrow{w}\leq^{\vartheta_w+4}\sqrt{\frac{16}{n^2(\vartheta_w+2)^2}}\right)\right)}{\int_{\mathbb{R}^{\vartheta_w}} d\overrightarrow{u}\,\mathbb{I}\left(\operatorname{diag}\left(\overrightarrow{w}^{\mathrm{T}}(S)^{-1}\overrightarrow{w}\leq^{\vartheta_w+4}\sqrt{\frac{16}{n^2(\vartheta_w+2)^2}}\right)\right)}\in[0;1] \tag{16}$$

Where $\overrightarrow{w}=\{w_{h_w};h_w=1\ldots\vartheta_w\}=\{z_{h_w0}-z_{h_wj};h_w=1\ldots\vartheta_w\}\in\mathbb{R}^{\vartheta_w}$, $\overrightarrow{u}=\{u_{h_u};h_u=1\ldots\vartheta_u\}=\left\{\frac{z_{h_u0}-z_{h_uj}}{H_{h_u}};h_u=1\ldots\vartheta_u\right\}\in\mathbb{R}^{\vartheta_u}$, and S only regards the dependent variables. Kernel functions for independent variables shall be triggered by a small bandwidth, such that no other units belonging to different categories can be utilized for comparability issues; in other words, as those categories are usually defined by integer values (typically 1, 2, 3...), the bandwidth H_{h_u} shall be lower than 1, i.e., $H_{h_u}\in\mathbb{R}:0<H_{h_u}<1$; in this work, we impose $H_{h_u}=0.5,\forall_{h_u=1\ldots\vartheta_u}$, a choice that does not impact on final results, [40]. Finally, the comparability set Ω_0, for the DMU$_0$: $(x_{i0},y_{r0},z_{h0})\in\mathbb{R}_+^m\times\mathbb{R}_+^s\times\mathbb{R}^{\vartheta}$, is composed by only those DMUs verifying $K\left(\overrightarrow{u},\overrightarrow{w}\right)>0$.

Hospitals can be classified by different points of view and face a meaningful environment impact on their performance, [27, 28, 37]. Such environment must adjust efficiency scores, which can be either internal or external. Similarly, each ICU faces (a) the same external environment as the whole hospital, and (b) specific ICU-related environment variables, mostly due to the inherent complexity of inpatients that ICUs take care. Table 3 identifies and describes the environment variables, either internal or external, to be utilized in the present study.

The adoption of a single independent categorical variable, Z_1, i.e., $\vartheta_u=1$ under the proposed framework (assuming a

Table 3 Environment variables

Variable	Variable name	Description and comments	Statistics
Z_1	ICU Specialty	• Independent categorical variable • Categories:* 1 (Polyvalent ICUs), 2 (Cardiology ICUs), 3 (Pediatric, Gynecology, Obstetrics and Neonatology ICUs), 4 (Surgical ICUs) • ICU specific variable	Frequencies: 1 – 44 % 2 – 23 % 3 – 17 % 4 – 16 %
Z_2	Year	• Dependent discrete variable • Categories: 2002 – 2009 • The below mentioned hospital reforms have been progressively employed over time, [27], so we consider this variable as discrete so as to take into account the relationship between Z_2 and the remaining environment variables, Z_3 up to Z_{10}.	Frequencies: 2002 – 10 % 2003 – 12 % 2004 – 12 % 2005 – 13 % 2006 – 13 % 2007 – 15 % 2008 – 16 % 2009 – 9 %
Z_3	Legal Status †	• Dependent categorical variable • Categories: 1 (Administrative Public Sector, SPA), 2 (hospital enterprises with limited liabilities, SA), 3 (corporate public entities, EPE)** • Hospital related variable (internal environment) • Portuguese public hospitals can be classified either into (a) hospitals belonging to the Administrative Public Sector, under the public/administrative law [in Portuguese, *Sector Público Administrativo*, SPA], (b) hospital enterprises with limited liabilities and under the commercial/private law [in Portuguese, *Sociedade Anónima*, SA], and (c) corporate public entities [in Portuguese, *Entidade Pública Empresarial*, EPE]; for details about the corporatization of the Portuguese public hospitals, please see [27, 28, 41]; SA hospitals were extinct in 2005, giving rise to EPE hospitals.	Frequencies: 1 – 39 % 2 – 11 % 3 – 50 %
Z_4	Hospital Type	• Dependent categorical variable • Categories: 1 (Undifferentiated hospitals), 2 (Maternities), 3 (Oncology centers) • Hospital related variable (internal environment) • Hospitals can either be undifferentiated or differentiated (maternities, oncology centers and psychiatric hospitals**); accordingly, the former may have different technologies than the latter and explore some economies of scope that differentiated hospitals eventually may not.	Frequencies: 1 – 95 % 2 – 2.5 % 3 –2.5 %
Z_5	Merging Status †	• Dependent categorical variable • Categories: 1 (Singular Hospitals, SH), 2 (Hospital Centers, HC), 3 (Local Health Units, LHU) • Hospital related variable (internal environment) • With the aim of exploiting economies of scale and scope, in the past few years, several SHs have merged either horizontally (with other SHs), creating HCs, or vertically (with primary health centers), creating LHUs.	Frequencies: 1 – 69 % 2 – 29 % 3 – 2 %
Z_6	Service-Mix Index (SMI)***	• Dependent continuous variable • ICU-specific variable (internal environment) • Different services usually have different technologies (capital investments, [42]) and different complexities, so the SMI is employed to homogenize ICUs by such technological differences; therefore, for instance, Pediatric ICUs verify $Z_1 = 3$, but even within such classification there is a considerable technological heterogeneity,	Average: 1.3824 Std. Dev: 1.9260 SMI $\in$ [0.3518; 3.6695]

Table 3 (continued)

Variable	Variable name	Description and comments	Statistics
		so employing a SMI as an (internal) environment variable restricts the comparability-based reference set for Pediatric ICUs to only those ones with similar technologies (eventually, other Pediatric ICUs as well) • The higher the SMI, the higher the expected inherent complexity of the ICU service, because costs with drugs, doctors and technological devices are usually dependent on services complexities *per se*. • The SMI computation details are based on the work of Ferreira and Marques [27] (and others cited over there, e.g. [43, 44]) and are left to the Appendix (see Appendix A); • Worth noting that the SMI proposed by Ferreira and Marques [27], as the well-known Case-Mix Index, encompasses an inefficiency component on its formulation; in the case of the present paper, an initial step was introduced to correct for such biasing (see Appendix A).	
Z_7	Gini's Specialization Index (GSI)***	• Dependent continuous variable • Hospital-specific variable (internal environment) • Based on Daidone and D'Amico [45] and Lindlbauer and Schreyögg [46], we introduce the GSI as a complement to variables Z_4 and Z_6 so as to measure the specialization (case-mix) degree of the hospitals where the ICUs are placed: GSI ranges from 0 (lower specialization degree) to 1 (higher specialization degree), [45], based on Disease Related Groups (DRGs) data for each hospital • GSI computational details are left for the Appendix C. Worth to note that our database differentiates the number of inpatients by DRG and by severity level (1, low severity, to 4, very high severity), so we compute 4 different GSIs (as detailed in Appendix C), say GSI_κ, κ=1,2,3,4, and aggregate them by using a weighted geometric mean, assuming weights equal to the standardized severity category, i.e., $w_k = 0.10, 0.20, 0.30$ and $0.40 \Rightarrow GSI = \prod_{\kappa=1}^{4} GSI_\kappa^{w_\kappa}$. • $\overline{GSI_{\kappa=1}} = 0.7029 \in [0.6381; 0.8095]$; $\overline{GSI_{\kappa=2}} = 0.6911 \in [0.5872; 0.7859]$; $\overline{GSI_{\kappa=3}} = 0.6927 \in [0.4398; 0.7621]$; $\overline{GSI_{\kappa=4}} = 0.5269 \in [0.1667; 0.6451] \Rightarrow$ $\overline{GSI} = 0.6283 \in [0.3605; 0.6975]$ where the overbar represents the 10 % trimmed mean. A considerable heterogeneity on specialization degree can be observed among hospitals.	Average: 0.6437 Std. Dev: 0.0269 GSI ∈ [0.5858; 0.6741]
Z_8	Population density	• Dependent continuous variable**** • Demographic related variable (external environment) • It is measured by the ratio inhabitants per km^2 • This is an important variable to adjust for environment as in urban regions (where the population density is higher) secondary health care access is easier than in rural regions, [27, 28]	Average: 68.19 Std. Dev: 67.29 $Z_8 \in [1.96; 245.57]$
Z_9	Wealth index	• Dependent continuous variable**** • Demographic related variable (external environment) • This variable is commonly known as the purchasing power, so it must be adjusted for inflation • Usually wealthier populations have higher education levels, so they dwell in urban areas, [27, 28]	Average: 104.22 Std. Dev: 21.94 $Z_9 \in [71.40; 134.20]$
Z_{10}	Aging index	• Dependent continuous variable****	Average: 129.43

Table 3 (continued)

Variable	Variable name	Description and comments	Statistics
		• Demographic related variable (external environment) • The aging index is measured by the number of elderly (>65 years old) per 100 youth (0-14 years old) • From the Grossman's model, health care expenses commonly increase with age due to diseases complexity, [27, 28]	Std. Dev: 42.46 $Z_{10} \in [82.10; 228.90]$

*The following ICU specialties were excluded from the sample due to the low number of DMUs: gastroenterology ICUs, psychiatric ICUs and infectious diseases ICUs. Polyvalent ICUs are, by the standard international definition, undifferentiated services receiving and treating severely ill patients from various specialties. Depending on the inpatient severity/ complexity level, he can be treated on Polyvalent or on differentiated ICUs (such as Cardiology ICUs or Surgical ICUs, for inpatients whose illness is more severe/ complex)

**Psychiatric hospitals were discarded from the classification of Z_2 as there are no hospitals of such a kind in our sample

***The adoption of the SMI and the GSI as environment variables follows Simões and Marques [16], the first approach of Ferreira and Marques [27] and Dervaux et al. [10]. In line with Dervaux et al. [10], this adjustment is able to control for patients' heterogeneous severity

****For these environment variables, data is available per municipality (http://www.pordata.pt/en/Municipalities). Usually, hospitals serve the population of more than one municipality. Eventually, they may serve a whole district or a specific region (as in the case of General hospitals, like "*CH Lisboa Norte*"). Therefore, municipal, district and regional data are suitably allocated for ICUs taking into account the type of hospitals where those ICUs are placed

†The inclusion of these classifications intends only to differentiate hospitals (ICUs) because the underlying reforms have significantly changed the management of the entities. Such inclusion follows the general and governmental classification of hospitals as well as Ferreira and Marques [27]. For instance, it is not correct to compare hospital centers to singular hospitals, e.g. due to their different scale sizes. Although these were important reforms, it is not intended to deeply explore their effect on congestion or congestion sources, rather only to check whether they are determinant factors on congestion or not

sufficiently small bandwidth, $H_{h_u} = 0.5$, and a uniform kernel, K_{h_u}) avoids unadvisable comparisons among DMUs from different ICU specialties. For instance, by using the variable Z_1 – ICU Specialty, DMUs from a specific ICU specialty are only compared with those ones from the same very ICU specialty. That is, '*Polyvalent ICUs*' are not compared with '*Cardiology ICUs*', for instance. In other words, the best practice frontier for each DMU is ICU specialty-specific.

On the other hand, variables such as Z_2 – Year, Z_3 – Legal Status, Z_4 – Hospital Type and Z_5 – Merging Status, are defined as dependent categorical variables to be included in the multidimensional kernel function as defined in Eqs. (15)-(16) and to enjoy possible interactions they may have between them and with continuous variables, Z_6 up to Z_{10}, so $\vartheta_w = 9$.[7]

Table 6 (Appendix E) contains the Pearson's correlation coefficients for those dependent discrete and continuous environment variables. There is a considerable correlation among some of them, which justifies the adoption of the multivariate kernel of Eq. (16). Clearly, the SMI, Z_6, is not correlated with the remaining environment variables as their effect (as well as the technical inefficiency of resources consumption) was filtered in the SMI computation, see Appendix A. Finally, the GSI and the demographic variables show significant correlation among them: hospitals located in urban regions, where the population density and the purchasing power are higher and the aging indexes are lower, tend to present more medical specialties.

3.4 Methodological issues

Efficiency under both SDH and WDH is computed by using the models presented in section 2, and following a conditional framework imposed by the method provided in subsection 3.3.[8] However, those models do not provide robust bias-corrected efficiency estimates due to their deterministic nature, neither do they allow achieving statistic-based results (such as confidence intervals) nor doing statistical inference tests over some hypotheses. Accordingly, we employ the bootstrap technique, as introduced by Simar and Wilson [47] and detailed in Appendix D, over a pooled conditional frontier. Such a technique allows obtaining B (a large number, say $B \sim 1000$) pseudo-frontiers, which are close to the true, still unobserved, frontier. Additional model features include: VRS and output-orientation, to be in line with the models described

[7] This results into $\sqrt[\vartheta_w+4]{\frac{16}{n^2(\vartheta_w+2)}}^{\,2} = \sqrt[9+4]{\frac{16}{630^2(9+2)}}^{\,2} \approx 0.3175$, which represents a bandwidth for the multidimensional kernel function.

[8] The authors, using the software Matlab®, developed all computational frameworks.

Table 4 Bias- and environment-corrected efficiency and congestion results (10 % trimmed mean scores and 95 % confidence intervals)

	Category	Global Sample	Congested DMUs				
		$\boldsymbol{\theta}_0^*$	$\boldsymbol{\theta}_0^*$ (congested DMUs)	*Kruskal-Wallis p-value on $\boldsymbol{\theta}_0^*$ (congested vs non-congested DMUs)*	β_0^*	Λ_0	$\Lambda_{\mathbf{weak},0}$
ICU Specialty	Polyvalent ICUs	4.2462, [3.6847; 5.0587]$_{95\%}$	3.5021, [2.9858; 4.1716]$_{95\%}$	*0.2538*	1.3409, [1.2125; 1.4776]$_{95\%}$	0.4610, [0.4191; 0.4912]$_{95\%}$	0.2091, [0.1794; 0.2399]$_{95\%}$
	Cardiology ICUs	2.8728, [2.4181; 3.5580]$_{95\%}$	2.3894, [2.0137; 2.9263]$_{95\%}$	*0.0294***	1.0844, [0.9893; 1.1880]$_{95\%}$	0.5603, [0.5000; 0.5984]$_{95\%}$	0.3688, [0.3078; 0.4769]$_{95\%}$
	Pediatric, Gynecology, Obstetrics and Neonatology ICUs	6.9479, [5.8290; 8.2732]$_{95\%}$	6.6879, [5.6142; 7.9082]$_{95\%}$	$\sim 10^{-4}$*	1.8232, [1.6517; 2.0019]$_{95\%}$	0.3469, [0.3234; 0.3727]$_{95\%}$	0.1082, [0.0940; 0.1242]$_{95\%}$
	Surgical ICUs	6.0029, [5.0365; 7.0070]$_{95\%}$	4.9713, [4.1571; 5.9102]$_{95\%}$	$\sim 10^{-6}$*	1.0828, [0.9822; 1.1901]$_{95\%}$	0.4576, [0.4981; 0.4094]$_{95\%}$	0.2296, [0.1868; 0.2798]$_{95\%}$
	Kruskal-Wallis p-value	$\sim 10^{-18}$*	$\sim 10^{-12}$*	-	$\sim 10^{-3}$*	$\sim 10^{-8}$*	$\sim 10^{-11}$*
Year	2002	3.7923, [3.2459; 4.5003]$_{95\%}$	3.1359, [2.6496; 3.7032]$_{95\%}$	*0.7873*	1.0367, [0.9362; 1.1404]$_{95\%}$	0.4607, [0.4203; 0.4926]$_{95\%}$	0.2167, [0.1905; 0.2595]$_{95\%}$
	2003	3.8895, [3.3168; 4.6686]$_{95\%}$	3.0130, [2.5366; 3.6513]$_{95\%}$	*0.3026*	1.2660, [1.1483; 1.3897]$_{95\%}$	0.5068, [0.4523; 0.5442]$_{95\%}$	0.2543, [0.2188; 0.3158]$_{95\%}$
	2004	3.8780, [3.3145; 4.5752]$_{95\%}$	3.3499, [2.8351; 3.9728]$_{95\%}$	*0.3869*	1.2010, [1.0792; 1.3169]$_{95\%}$	0.4330, [0.3926; 0.4630]$_{95\%}$	0.2043, [0.1792; 0.2480]$_{95\%}$
	2005	3.6197, [3.0807; 4.3337]$_{95\%}$	3.1425, [2.6299; 3.8567]$_{95\%}$	*0.1677*	1.0166, [0.9160; 1.1209]$_{95\%}$	0.4279, [0.3768; 0.4609]$_{95\%}$	0.2695, [0.2024; 0.3464]$_{95\%}$
	2006	5.0763, [4.3064; 6.1084]$_{95\%}$	4.9394, [4.1668; 5.9071]$_{95\%}$	*0.0350***	1.4678, [1.3303; 1.6130]$_{95\%}$	0.4736, [0.4304; 0.5064]$_{95\%}$	0.2161,[0.1732; 0.2528]$_{95\%}$
	2007	6.3660, [5.4212; 7.5945]$_{95\%}$	7.3920, [6.2712; 8.6960]$_{95\%}$	*0.5467*	2.4471, [2.2139; 2.6920]$_{95\%}$	0.4483, [0.4153; 0.4788]$_{95\%}$	0.2365, [0.1902; 0.3161]$_{95\%}$
	2008	4.9306, [4.1763; 5.8949]$_{95\%}$	4.9302, [4.1167; 5.8762]$_{95\%}$	*0.4763*	1.6705, [1.5166; 1.8341]$_{95\%}$	0.4466, [0.4061; 0.4838]$_{95\%}$	0.1776, [0.1511; 0.2106]$_{95\%}$
	2009	5.8748, [5.0215; 6.9347]$_{95\%}$	5.7207, [4.8239; 6.7621]$_{95\%}$	*0.2988*	1.6738, [1.5176; 1.8395]$_{95\%}$	0.4410, [0.4079; 0.4704]$_{95\%}$	0.2643, [0.2207; 0.3232]$_{95\%}$
	Kruskal-Wallis p-value	*0.2078*	*0.2472*	-	*0.0133***	*0.9100*	*0.5573*
Legal Status	SPA hospitals	4.3075, [3.6872; 5.0885]$_{95\%}$	3.7132, [3.1287; 4.4260]$_{95\%}$	*0.2879*	1.2384, [1.1191; 1.3609]$_{95\%}$	0.4596, [0.4192; 0.4936]$_{95\%}$	0.2150, [0.1819; 0.2595]$_{95\%}$
	SA hospitals	4.0688, [3.4675; 4.9225]$_{95\%}$	3.3934, [2.8106; 3.9417]$_{95\%}$	*0.2014*	1.2169, [1.1000; 1.3351]$_{95\%}$	0.4613, [0.4155; 0.4902]$_{95\%}$	0.2407, [0.2105; 0.3042]$_{95\%}$
	EPE hospitals	4.9127, [4.1698; 5.9105]$_{95\%}$	4.9133, [4.1274; 5.8839]$_{95\%}$	*0.0090**	1.5271, [1.3827; 1.6803]$_{95\%}$	0.4474, [0.4048; 0.4812]$_{95\%}$	0.2367, [0.1892; 0.2960]$_{95\%}$
	Kruskal-Wallis p-value	*0.7139*	*0.3597*	-	*0.2502*	*0.9738*	*0.7011*
Hospital Type	Undifferentiated hospitals	4.6372, [3.9407; 5.5640]$_{95\%}$	4.4735, [3.7616; 5.3602]$_{95\%}$	*0.0546*	1.3995, [1.2664; 1.5375]$_{95\%}$	0.4461, [0.4043; 0.4789]$_{95\%}$	0.2249, [0.1871; 0.2766]$_{95\%}$
	Maternities	2.3579, [2.0104; 2.7392]$_{95\%}$	1.9348, [1.6553; 2.2264]$_{95\%}$	*0.0896*	1.0000, [0.9074; 1.0947]$_{95\%}$	N.A.***	0.1889, [0.1477; 0.2499]$_{95\%}$
	Oncology centers	4.0153, [3.6292; 4.6353]$_{95\%}$	2.1463, [1.8710; 2.4687]$_{95\%}$	*0.5557*	1.0132, [0.9207; 1.1171]$_{95\%}$	0.5525, [0.5104; 0.5865]$_{95\%}$	0.3696, [0.2801; 0.5117]$_{95\%}$
	Kruskal-Wallis p-value	*0.0350***	*0.0030**	-	*0.0775*	*0.1288*	*0.0613*
Merging Status	Singular Hospitals	4.6854, [3.9972; 5.6110]$_{95\%}$	4.3006, [3.6268; 5.1532]$_{95\%}$	*0.0463***	1.4530, [1.3150; 1.5970]$_{95\%}$	0.4733, [0.4292; 0.5067]$_{95\%}$	0.2319,[0.1904; 0.2854]$_{95\%}$
	Hospital Centers	4.1618, [3.5144; 4.9850]$_{95\%}$	4.0419, [3.3855; 4.8130]$_{95\%}$	*0.5723*	1.1456, [1.0371; 1.2575]$_{95\%}$	0.4098, [0.3720; 0.4435]$_{95\%}$	0.2276, [0.1893; 0.2835]$_{95\%}$
	Local Health Units	5.5904, [4.8806; 6.4759]$_{95\%}$	5.6314, [4.9271; 6.4076]$_{95\%}$	*0.2363*	1.9863, [1.8049; 2.1816]$_{95\%}$	0.4272, [0.4084; 0.4312]$_{95\%}$	N.A. ****
	Kruskal-Wallis p-value	*0.4657*	*0.4277*	-	*0.5614*	*0.1089*	*0.3875*
Global results		**4.5694, [3.8896; 5.4690]$_{95\%}$**	**4.2314, [3.5617; 5.0613]$_{95\%}$**	$\sim 10^{-4}$*	**1.3445, [1.2171; 1.4771]$_{95\%}$**	**0.4539, [0.4116; 0.4871]$_{95\%}$**	**0.2292, [0.1890; 0.2828]$_{95\%}$**

*Reject the null hypothesis (equal distributions) at both 5 % and 1 % levels ($p < 0.01$);

**Do not reject the null hypothesis at the 1 % level ($0.01 \leq p < 0.05$);

***No strongly congested maternity has been identified;

****No weakly congested LHU has been identified.

in section 2.[9] This study adopts the strategy defined in Table 1 (section 2). We can make use of bootstrap iterations to employ a set of statistical tests and check whether:

(1) Strong congestion levels are not significant across the entire sample, i.e., $H_{0(1)}: p_{r=1}^{+*} = \tilde{p}_{r=1}^{+*} = 0 \cap \beta_0^* = \theta_0^* = 1 \Leftrightarrow \Lambda_0 = 1$ vs $H_{1(1)}: p_{r=1}^{+*} \neq \tilde{p}_{r=1}^{+*} \neq 0 \cup \beta_0^* > \theta_0^* \Leftrightarrow \Lambda_0 < 1$
(2) Both strong and weak congestion levels are similar, i.e., $H_{0(2)}: \Lambda_0 = \Lambda_{\text{weak},0}$ vs $H_{1(2)}: \Lambda_0 > \Lambda_{\text{weak},0}$
(3) Weak congestion levels are not noteworthy, i.e., $H_{0(3)}: f_m\left(1-\frac{t_i^{-*}}{x_{i0}}\right) = \Lambda_0 \cdot \theta_0^* + \frac{\Lambda_0 \cdot p_r^{+*} + t_r^{+*}}{y_{r0}}$ vs $H_{1(3)}: f_m\left(1-\frac{t_i^{-*}}{x_{i0}}\right) < \Lambda_0 \cdot \theta_0^* + \frac{\Lambda_0 \cdot p_r^{+*} + t_r^{+*}}{y_{r0}}$

Let's define the following general tests to evaluate the second hypothesis[10]:

$$T^b = \frac{f_n\left(\Lambda_{\text{weak},j}^b\right)}{f_n\left(\Lambda_j^b\right)}; \quad T^{obs} = \frac{f_n\left(\Lambda_{\text{weak},j}^{obs}\right)}{f_n\left(\Lambda_j^{obs}\right)} \tag{17}$$

Where, as before, f_n is an aggregating function of n arguments (n is the sample size). The p-value utilized is as follows:

$$p \approx \frac{1}{B} \cdot \sum_{b=1}^{B} \mathbb{I}\left(T^b \leq T^{obs}\right) \tag{18}$$

Where $\mathbb{I}$ is the indicator function. A very small p (say, $p < 0.05$) allows us rejecting the null hypothesis at the 5 % level. In this paper, as aggregating function, f_n, we use the arithmetic mean, the geometric mean, the median and the 10 % trimmed mean.

Measuring the impact of a specific environment variable implies running previous models with and without such a variable, $z_h, h = 1 \ldots \vartheta$. As there are 10 environment variables, this means that the aforementioned analysis is conducted 11 times. Furthermore, let $\Lambda_0(z)$ (*resp*. $\Lambda_0(z \backslash z_h)$) be a congestion index computed by using all environment variables (resp. The same index computed when z_h is excluded). We utilize the ratios $\Gamma_0(z_h) = \Lambda_0(z)/\Lambda_0(z \backslash z_h)$, $\forall h = 6 \ldots 10$, and $\Gamma_{\text{weak},0}(z_h) = \Lambda_{\text{weak},0}(z)/\Lambda_{\text{weak},0}(z \backslash z_h)$, $\forall h = 6 \ldots 10$, to test whether the continuous variable z_h impacts on congestion. It is possible to conclude that $\partial \Gamma_0(z_h)/\partial z_h > 0$ means that the higher z_h, the lower the congestion levels. Nonparametric regressions utilize the Nadaraya-Watson nonparametric regression method, using Gaussian kernels and the Silverman's bandwidth.

4 Empirical Results

4.1 Identifying Congestion levels

4.1.1 Global results of technical efficiency

Table 4 provides the main results of both bias- and environment-corrected efficiency and congestion measures, divided by categories. While the 3rd column is devoted to the whole sample, the 4th column onwards shows the results of congested units only. Regarding the technical efficiency, we observe that ICUs are generally highly inefficient and they could increase their outputs (InpD) into about 78 % $\left(= 1-\left[\theta_0^*\right]^{-1} = 1-[4.5694]^{-1} \approx 0.78\right)$, keeping their resources unchanged, i.e., ignoring the congestion effect. This high inefficiency level is prominent on Pediatric, Gynecology, Obstetrics and Neonatology and Surgical ICUs, all specialties with similar levels of (in)-efficiency. Although we can observe that these inefficiency levels have increased over time, those differences are not statistically significant according to the Kruskal-Wallis nonparametric test. Besides, there is no apparent reason to justify these inefficiency levels based on criteria like the hospital legal status, merging status or type, at least at the 1 % significance level; still, these differences remain on the ICUs specialty basis and their treated inpatients' inherent complexity. As a matter of fact, neither the last hospital reforms (merging and legal statuses) nor the possible existence of scope and scale economies have contributed to the improvement of technical efficiency in ICUs on the period 2002-2009. However, if the sample is divided into congested and non-congested DMUs, we verify that the inefficiency is significantly lower on congested ICUs from differentiated hospitals. Furthermore, non-congested ICUs on average operate on the increasing RTS region, as shown in Table 7 (Appendix E).

4.1.2 Global results of congestion

The sample exhibits considerable levels of strong congestion, as shown in Table 4 (7th column) and as proved by the bootstrap-based test over the hypothesis $H_{0(1)}: p_{r=1}^{+*} = \tilde{p}_{r=1}^{+*} = 0 \cap \beta_0^* = \theta_0^* = 1 \Leftrightarrow \Lambda_0 = 1$. Statistics T^b have returned $p \sim 0$ for all employed aggregating functions, which means that the null hypothesis can be rejected at any significance level. As no output slacks,

[9] Multiple optima (solutions) are not problematic in the present case. Indeed, our results are consistent with those ones obtained through the approach proposed by Sueyoshi and Sekitani [31]. However, so as to avoid a too long paper and to keep the analysis as simple as possible, those results are not displayed but can be provided upon request.

[10] *Mutatis mutandis*, it can be easily adapted to the other two hypotheses.

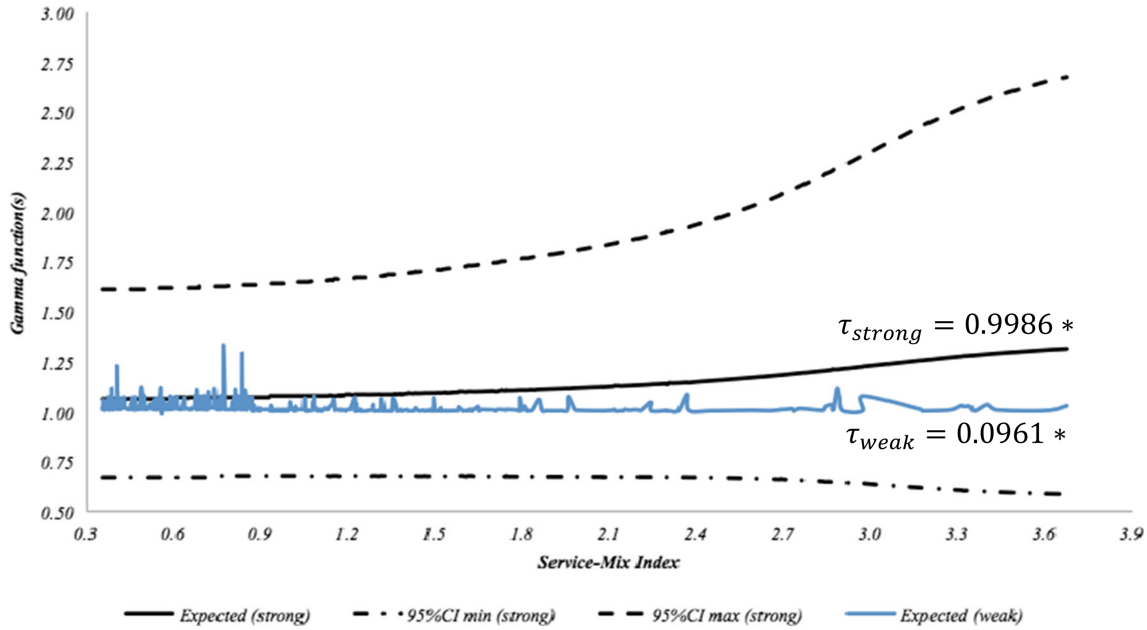

Fig. 1 Nadaraya-Watson regression of Gamma functions, $\Gamma_0(z_6)$ and $\Gamma_{weak,0}(z_6)$, against the Service-Mix Index, z_6. τ is the Kendall's correlation coefficient, * indicates significance at 5 %

$p^{+*}_{r=1}$ and $\tilde{p}^{+*}_{r=1}$, have been identified, then Λ_0, Eq. (5), and $C_0\left(\theta_0^*, \beta_0^*\right)$, Eq. (3), overlap. On average, congested ICUs could have increased their production (number of discharges) into about 120 % ($= \Lambda_0^{-1} - 1 = [0.4539]^{-1} - 1 \approx 1.20$) by reducing their consumed resources. 45 % of the sample was identified as congested. From these, 65 % were weakly congested and the remaining 35 % strongly congested. As before, we test hypothesis $H_{0(2)} : \Lambda_0 = \Lambda_{\text{weak},0}$ and $H_{0(3)} : f_m \left(1 - \frac{t_i^{-*}}{x_{i0}}\right) = \Lambda_0 \cdot \theta_0^* + \frac{\Lambda_0 \cdot p_r^{+*} + t_r^{+*}}{y_{r0}}$ by using bootstrap and, once again, p-values are close to zero, which allows us to conclude that weak congestion is effectively lower than the strong congestion and it is statistically significant. Accordingly, very low weak congestion coefficients were obtained, revealing that weakly congested ICUs show considerable congestion levels and, as consequence, there is a huge room for inputs management improvement and costs savings. This fact is consistent with the average DSE value found for weakly congested ICUs, −3.5178 (Table 7, Appendix E), i.e., about 7 new discharges could have occurred by an average decreasing of inputs of 2,000€ (on costs) and/or 2 hospital days.

4.1.3 Determinants on congestion levels

Categorical and discrete environment variables The ICU specialty is clearly a determinant on both technical efficiency and congestion. As a matter of fact, the splitting of the sample on those specialties has proven to return different efficiency and congestion distributions by the Kruskal-Wallis test. In view of that, we conclude that Cardiology ICUs are the DMUs presenting the lowest both strong and weak congestion levels. They are followed by Polyvalent ICUs and Surgical ICUs, *ex aequo*.

Time did not show to be a determinant on congestion, either strong or weak, at least at the 1 % significance level. That means that congestion distributions do not significantly change over time and we can expect that those significant congestion levels may keep it up in present days.

As the time, the legal status of the hospital where the ICU(s) is(are) placed has no significant impact on congestion distributions, so it is not a determinant of congestion. Likewise, neither the hospital type nor the hospital merging status is congestion determinant. As in the efficiency case, these Government reforms have not produced yet the desired outcomes, in terms of resources wastefulness reduction and as expected the New Public Management system, the philosophy under which those reforms were made.

SMI Figure 1 shows the Nadaraya-Watson nonparametric regression (and the 95 %CIs) of $\Gamma_0(z_6) = \Lambda_0(z)/\Lambda_0(z \backslash z_6)$ and $\Gamma_{\text{weak},0}(z_6) = \Lambda_{\text{weak},0}(z)/\Lambda_{\text{weak},0}(z \backslash z_6)$, so as to study the impact of the SMI on both strong and weak congestion measures.

It is evident (and statistically significant) that the higher the ICU complexity, the lower the strong congestion, which means that these services (e.g. transplants and burn units) seem to present a better resource allocation and management than other, less complex, ICU services. On the other hand, the SMI has little (though statistically non-null) impact on weak congestion.

GSI Appendix E contains the Nadaraya-Watson regressions against the continuous environment variables, h = 7...10, cf. Figs. 2, 3, 4 and 5. Hospital specialty degree (GSI) is negatively related to strong congestion, as the higher that degree is, the higher the strong congestion level (this variables seems to have no meaningful impact on weak congestion levels). This means that highly differentiated hospitals are more prone to strong congestion, as we can state by the comparison among undifferentiated hospitals and oncology centers (Table 4), such that the latter (more differentiated) present slightly higher congestion levels than the former, though those differences are not statistically significant by the Kruskal-Wallis test, which can be attributed to the fact that undifferentiated hospitals present a considerable range on GSI. For instance, general hospitals deal with the most complex cases (in terms of illness) and present the highest services complexity (given by the SMI), but still they are quite undifferentiated (low GSI), and then they are less congested than differentiated hospitals. That is, both results from SMI and GSI are consistent.

Population density The population density only affects the weak congestion. The higher that environment variable is, the higher the weak congestion levels. Rural regions usually present low population density and their hospitals typically present less complex services (low SMI). On the other hand, urban regions, like Lisbon and Oporto, have a set of highly differentiated hospitals (e.g. oncology centers), Table 6 (Appendix E), which, by the previous subsection, are prone to congestion. These results seem to be consistent with previous findings.

Wealth index Like the SMI, the higher the wealth index, the lower the congestion levels. The fact that this index and the population density are positively correlated, Table 6, could lead to different conclusions. However, by the Grossman model, [48], we can expect that wealthier populations have higher education levels, so their health relies on prevention rather than on treatment. In other words, it is expected that their illness and probability of unexpected mortality is lower than those ones on poorer people, and *a priori* consume less resources and require lower levels of treatments on differentiated care (say higher prevention, less cancers), which are more prone to congestion.

Aging index The aging index, as the SMI and the wealth index, does impact on both strong and weak congestion levels. The higher that index is, the lower the congestion in ICUs. As a matter of fact, the aging index and the population density are negatively correlated, see Table 6, i.e., aging populations tend to be located in rural areas, so the hospitals where they are treated do not have highly complex ICU services such as transplants and burned units, and then are more congested as predicted by the SMI impact. When these (elderly) patients need highly complex ICU treatments, they are transported to urban general hospitals.

4.2 Identifying Congestion main sources

So as to identify the congestion sources, we compute the MP for each input. As there is only one output, Eq. (8) reduces to $MP(x_{i0}, y_{10}) = y_{10} \cdot \tilde{v}^*_{i0}$, $\mathrm{i} = 1 \ldots m$, a quantity that is negative when $\exists_{i=1 \ldots m} \tilde{v}^*_{i0} < 0$, revealing the existence of congestion. Table 8 (Appendix E) summarizes the MP values (10 % trimmed means) by different categories. Note that under the strong congestion, all inputs are strong congestion sources, so the results in Table 8 regard only the weak congestion cases. These results shall be interpreted as follows: keeping the remaining inputs unchanged, the increasing of 1 HospDay or 1,000€ on a specific cost-related input will raise (resp. Decrease) the no. discharges by a ratio equal to $\mathrm{MP}(x_{i0}, y_{10})$, if it is a positive (resp. Negative) quantity. For instance, and taking the global results, an increase of 1,000€ on medicines and clinical stuff (CGSC) will increase, on average, the number of discharges by about 18 inpatients. Likewise, the decrease of 2 hospital days will likely increase discharges by about 3 inpatients. This shows that the length of stay in ICU is its main congestion source, being relevant in ICU specialties such as Cardiology and Surgical ICUs. In these cases, the decrease of 1 HospDay would increase the number of discharges by 4-5. This is an expected result as the input HospDays strongly depends on the beds availability and the average delay, which assumes considerable values in ICUs due to the required level of care. Accordingly, we recall and confirm our assumption that HospDays contributes to the services congestion as it likely potentiates the appearance of other diseases, like nosocomial infections.

Costs with staff also exhibit a role on congestion, although it is weak on average as it would be necessary an average decrease of about 37,600€ on StaffC to increase only one InpD, which in practice is not likely. However, its role becomes meaningful on Polyvalent ICUs and Pediatric, Gynecology, Obstetrics and Neonatology ICUs, on Oncology centers and on EPE hospitals, which came into force in 2005. We also observe that from 2005 onwards, StaffC becomes the most important congestion source instead of HospDays. According to Ferreira and Marques [28], the corporatization reform (legal status) "*refers to the application of private management tools to the public sector*", so EPE hospitals are more autonomous on human resources contracting. This freedom is perhaps the reason why StaffC is a congestion source. On the other hand, HospDays is no longer a major source of congestion in ICUs of EPE hospitals, which means that the introduction of such private management rules on corporatized hospitals may have imposed the reduction of the average delay, even in ICUs. So, unnecessary and even harmful HospDays were reduced to a minimum.

Furthermore, StaffC and HospDays are highly correlated ($\tau = 0.5067$), so the higher the inpatients illness, complexity and severity, the higher the hospitalization time and the costs with doctors and nurses. But as there is a surplus on HospDays, the same applies to StaffC. This also seems to justify why StaffC is a congestion source in hospital centers, as the (horizontal) merging reform did not change the staffing quantity; rather some services were closed so as to explore potential scope and scale economies. In ICUs, this appears to show a perverse effect, exhibiting some staff surplus.

Finally, it is worth to mention that CGSC, SEServ and CapC are not meaningful sources of congestion in ICUs, revealing a good management of clinical stuff (including medicines), outsourcing and capital investments. These three variables may also contribute to the improvement of ICUs production, due to their considerable positive MPs.

5 Concluding remarks

The main objective of the paper is threefold: firstly, to achieve the bias- and environment-corrected congestion of ICUs; secondly, to check whether any of those environment variables, either internal or external, are determinants of congestion; and thirdly, to verify which (if any) of the inputs exhibits signs of congestion source. Considerable and statistically significant congestion levels were identified, meaning that ICUs management shall be careful in resource allocation so as to prevent the congestion phenomenon and improve the production (in this case, the number of alive inpatients leaving the ICU to other hospital services). This resource allocation shall focus on the dimensions identified as congestion sources: costs with staff and hospital days. The remaining inputs seem to be well managed in terms of congestion, but still they can exhibit some other signs of inefficiency. Nevertheless, they do not seem to negatively affect the number of discharges of ICUs. Clearly, both congestion levels and sources are dependent on some determinants, including the ICU specialty, the ICU complexity, the hospital differentiation degree and the demographic patterns of the population. This means that these factors shall also be taken into account by the ICU management on such a resource allocation process. Despite the relative database seniority, it was shown that time is not a determinant of congestion levels, meaning that it is expected that congestion levels remain slightly unchanged at present and some results can be inferred regarding the current days. This clearly must be confirmed by using a more recent database.

Finally, it shall be mentioned that this study, like any other, is not absolutely flawless. An important issue that is left for further research includes the adjustment of inpatients by their probability of death at the ICU entrance. However, our model is adjusted for environment which, according to Ferreira and Marques [27], mimics the inpatients complexity and avoids the heterogeneity among ICUs. Needless to say that for a more recent database this kind of data shall be included and this hypothesis confirmed. Additionally, no quality data has been considered in this study. Unfortunately such information does not exist neither for the ICU services nor for the considered time period (2002-2009). Nevertheless, we believe that quality assumes a really important role in ICU services (and hospitals, in general) performance; therefore, the inclusion of quality data (such as in-ICU mortality rates) shall be a hot topic for further research.

Acknowledgments We would like to thank to three anonymous referees who kindly and significantly have improved this paper's quality, clarity and structure, due to their beneficial comments. We also acknowledge the financial support of the Portuguese Foundation for Scientific and Technology (FCT): SFRH/BD/113038/2015. The second author thanks the FCT (Portuguese national funding agency for science, research and technology) for the possibility of being under sabbatical leave in the University of Cornell in the USA for the period when part of this research took place.

Appendix

A. Literature review

Please, see Table 5

Table 5 Literature review on congestion measurement and/or ICU performance evaluation

Author(s)	Level of analysis (DMUs)	Model(s) specification(s)	Variable(s)	Main conclusion(s)	Comment(s)
Puig-Junoy [8]	993 critically ill patients (treated on 16 ICUs); Catalonia (Spain), 1991-1992	1st stage: DEA with non-discretionary and categorical variables (as either inputs or outputs) and multiplier restrictions 2nd stage: Regression of efficiency results against some environment variables	Inputs: Survival probability at admission; Mortality risk level at admission; Weighted patient days in ICU and hospital; Availability of resources during the hospital stay; Average FTE** no. of nurses and physicians; Technological availability Outputs: No. days surviving in the hospital; Surviving discharge status	Efficiency significantly varies across different mortality risk groups: "*higher risk patients are tendentiously managed less efficiently than lower risk ones – a large amount of resources was being devoted to more severe patients who died*", [8], which reveals the existence of some congestion on these services. Environment variables assume a prominent role on efficiency of critically ill patients management.	Congestion is not evaluated. There is not a bias-correction of results. The separability condition between environment and the efficiency model is intrinsically assumed.*
Tsekouras et al. [9]	39 ICUs; Greece	1st stage: Bias-corrected DEA (with bootstrap) 2nd stage: Regression of efficiency results against some environment variables	Inputs: Equipment acquired before 1992; Equipment acquired between 1993 and 2006; No. available beds, doctors and nurses Outputs: Death-rate adjusted length of stay at ICUs (days of treatment)	New medical technology allows the improvement of ICUs technical efficiency. Asymmetric information, ICUs proximity to pools of knowledge and the composition of staff seem to be relevant factors to improve efficiency.	Congestion is not evaluated. The separability condition between environment and the efficiency model is intrinsically assumed.*
Dervaux et al. [10]	15,029 critically ill patients (treated on 25 hospitals); Paris region (France), 2000	Robust Free Disposal Hull (order-*m*, [39], with a directional input distance function)	Inputs: Length of stay; Omegas 1, 2 and 3 as measures of resources usage per type of care provided Outputs: Probability of death minus the discharge status Control variables (environment): 10 categories of diagnoses; SAPSII score; Mode of entry in ICU; Treatment: surgical or medical	Medical and economic inefficiencies are not significantly correlated, so both must be carefully analyzed by hospital management.	Congestion is not evaluated. The inclusion of control variables is not properly done.
Clement et al. [11]	667 hospitals; USA, 2000.	1st stage: DEA (radial), all outputs under strong disposability 2nd stage: DEA (radial), undesirable outputs under weak disposability	Inputs: Registered and licensed practical FTE** nurses; Other staff (FTE**); Staffed beds Outputs: No. births; No. outpatient surgeries; No. emergency room visits; No.	Significant congestion levels were identified. Lower technical efficiency and poorer risk-adjusted quality outcomes are associated.	Only the congestion over undesirable outputs is analyzed. There is not employed a congestion measurement over inputs. There is neither a bias- nor an environment-correction of results.

Table 5 (continued)

Author(s)	Level of analysis (DMUs)	Model(s) specification(s)	Variable(s)	Main conclusion(s)	Comment(s)
			outpatient visits; No. case-mix adjusted admissions; Risk-adjusted mortality rate (%) for a set of diseases as acute myocardial infarction		As mentioned by Kuntz and Sülz [49], there is the possibility of wrong results in Clement et al. [11] due to some models mistakes.
Valdmanis et al. [12]	1377 urban hospitals; USA, 2004	Identical to Clement et al. [11]	Inputs: Bassinets; Registered and licensed practical FTE** nurses; Other staff (FTE**); Interns/ residents FTE**; Acute and other beds Outputs: No. births; No. total surgeries; No. outpatient visits; No. adjusted admissions; No. other patient days Undesirable outputs: Failure to rescue; Infection due to medical care; Postoperative respiratory failure; Postoperative sepsis Environmental and organizational variables: Total expenses; No. high-tech services; Births by admission; Emergency room visits by outpatient visit; Outpatient surgeries by outpatient visit; % Medicare and % Medicaid; Occupancy rate; Cost/ adjusted admission; Average length of stay; Herfindahl-Hirschman index *inter alia*	Only 3 out of 26 % of inefficiency was attributed to the congestion of undesirable outputs. Quality of care could be improved by increasing labor inputs on low-quality hospitals; still, high quality hospitals show some staff excess, which suggests a resources reallocation.	Only the congestion over undesirable outputs is analyzed. There is not employed a congestion measurement over inputs. The separability condition between environment and the efficiency model is intrinsically assumed.*
Ferrier et al. [13]	170 short-term, general, community hospitals; Pennsylvania (USA), 2002.	Identical to Clement et al. [11] and Valdmanis et al. [12].	Inputs: No. beds; Registered and licensed practical FTE** nurses; Interns/ residents FTE**; Other staff (FTE**) Outputs: No. inpatient surgeries; No. outpatient surgeries; No. emergency visits; No. outpatient visits; No. adjusted inpatient days; Uncompensated care	Outputs could be increased by ~ 7 % if units were technically efficient; uncompensated care reduced the production of hospitals by ~ 2 %.	Only the congestion over undesirable outputs is analyzed. There is not employed a congestion measurement over inputs. There is neither a bias- nor an environment-correction of results. **As** mentioned by Kuntz and Sülz [49], there is the possibility of wrong results in Clement et al. [11] due to some model's mistakes.
Arrieta and Guillén [14]	35 neonatal care units; Peru, 2008-2009	Identical to Clement et al. [11], Valdmanis et al. [12] and Ferrier et al. [13], but with the corrections mentioned by Kuntz and Sülz [49]	Inputs: No. physicians; No. nurses; No. incubators; No. phototherapy equipment; No. monitors	*"About half of hospital's neonatal care units, technical efficiency is affected by output congestion. For those hospitals, patient safety is*	Only the congestion over undesirable outputs is analyzed. There is not employed a congestion measurement over inputs.

Table 5 (continued)

Author(s)	Level of analysis (DMUs)	Model(s) specification(s)	Variable(s)	Main conclusion(s)	Comment(s)
			Outputs: No. neonatal care unit admissions; Risk-adjusted mortality rate; Risk-adjusted near-miss rate	*being compromised by receiving too many patients. […] Most congested hospitals are located in the capital city and suburban areas, and are more likely to be hospitals with the lowest and the highest level of care"* [14]	There is neither a bias- nor an environment-correction of results.
Matranga and Sapienza [15]	116 short-term, acute-care hospitals; Sicily (Italy), 2009	Theoretically, identical to Clement et al. [11], Valdmanis et al. [12] and Ferrier et al. [13].	Inputs: No. inpatient beds; No. medical staff; No. nursing staff; No. other staff Outputs: Case-mix adjusted ordinary discharges; Day hospital admissions; No. inappropriate ordinary discharges; No. inappropriate day hospital admissions	"*Most of the measured overall inefficiency of Sicilian hospitals could be attributed to congestion and pure technical inefficiency and that congestion was statistically different among hospital trusts, local public hospitals and for-profit hospitals and along the provinces*" [15]	Matranga and Sapienza [15] point out that the difference between a strong- and a weak-disposability of outputs-based model is the change from an inequality to an equality on undesirable outputs, which is not entirely true, [49], and may introduce some misleading results and conclusions. Only the congestion over undesirable outputs is analyzed. There is not employed a congestion measurement over inputs. There is neither a bias- nor an environment-correction of results.
Simões and Marques [16]	68 Public hospitals; Portugal, 2005	1st stage: Bias-corrected DEA (with bootstrap) 2nd stage: Regression of efficiency results against some environment variables (with double bootstrap) 3rd stage: Congestion analysis through three different models (based on radial and non-radial	Inputs: Capital expenses; No. full-time employees (staff); Other operational expenses (OPEX-Staff costs) Outputs: No. inpatients; No. emergency visits; No. outpatients' visits	Considerable levels of technical inefficiency, scale inefficiency and congestion were found.	There is not a distinction between strong and weak congestion, which is a relevant issue on congestion assessment. The separability condition between environment and the efficiency model is intrinsically assumed.* The effect of the environment on congestion levels is not analyzed.
Mathews and Long [17]	36 ICU and 15 step-down units	Queuing model	Patient acuity; Arrival rate; Unit length of stay	"*Hospital queuing and simulation modeling with empiric data inputs can evaluate how changes in ICU bed assignment could impact unit occupancy levels and patient wait times.*" [17]	This is not a congestion measurement on its economic meaning.

*Regressing efficiency scores against environment variables without a proper a priori adjustment on the efficiency model presupposes that the separability condition holds (i.e., the environment variables only impacts on efficiency distributions, not on the frontier shape), which is not true in practice. That is, efficiency results are not adjusted for the environment.

**FTE – full-time equivalent.

Service-Mix Index

The inefficiency-corrected SMI for the ICU service (DMU_0) defined by the pair $(x_{i0}, y_{r0}) \in \mathbb{R}_+^{5+1}$ can be computed as follows:

Use Eq. (19) to obtain the optima $\left\{\theta_0^*, p_1^{-*}, \ldots, p_5^{-*}, p_1^{+*}\right\}_{SDH}$ and project $(\boldsymbol{x}_{i0}, \boldsymbol{y}_{r0})$ in the SDH frontier using the transformation $(x^{\vee}{}_{i0}, y^{\vee}{}_{10}) = \left(\theta_0^* \cdot x_{i0} - p_i^{-*}, y_{10} + p_1^{+*}\right) \in \mathbb{R}_+^{5+1}$. This projection removes the technical inefficiency of units. To obtain Ω we follow the strategy proposed by subsection 3.3, considering all environment variables but Z_6 (by obvious reasons).

$$\{\theta_0^*, p_1^{-*}, \ldots, p_5^{-*}, p_1^{+*}\}_{SDH}$$

$$= \min_{\theta_0, \lambda_1, \ldots, \lambda_n, p_1^-, \ldots, p_5^-, p_1^+} \left\{ \theta_0 - \varepsilon\left(p_1^+ + \sum_{i=1}^{5} p_i^-\right) \middle| \begin{array}{c} \sum\limits_{\substack{j=1 \\ j \in \Omega'}}^{n} \lambda_j x_{ij} + p_i^- = \theta_0 x_{i0} \\ \sum\limits_{\substack{j=1 \\ j \in \Omega'}}^{n} \lambda_j y_{1j} - p_1^+ = y_{10} \\ \sum\limits_{\substack{j=1 \\ j \in \Omega'}}^{n} \lambda_j = 1 \\ \lambda_j, p_1^+, p_i^- \geq 0 \\ j \in \Omega'; i = 1, \ldots, 5 \end{array} \right\} \tag{19}$$

1) Re-run the step 1) for all DMUs and obtain the set $\nabla = \left\{ \left(\sum_{i=1}^{4} x^{\vee}{}_{ij}, y^{\vee}{}_{1j} \right) = \left(\theta_0^* \cdot \sum_{i=1}^{4} x_{ij} - \sum_{i=1}^{4} p_i^{-*}, y^{\vee}{}_{1j} \right) \in \mathbb{R}_+^2, \right.$ $\left. j = 1, \ldots, n \right\}$. ∇ only contains data from the single output and the first 4 inputs (monetary resources).
2) Compute the ratio for the j-th unit $\boldsymbol{\xi_j} = \left(\sum_{i=1}^{4} \boldsymbol{x}^{\vee}{}_{\boldsymbol{ij}} \right) / \boldsymbol{y}^{\vee}{}_{1\boldsymbol{j}}$, $\boldsymbol{j} = 1, \ldots, \boldsymbol{n}$' which represents the *efficient* unitary cost of such DMU_j.
3) Compute $\mathscr{B} = \prod_{j=1}^{n} \left(\boldsymbol{\xi_j}^{1/\boldsymbol{n}} \right)$, which represents the unitary costs' national average (baseline).
4) The SMI for the DMU_0 is then $\boldsymbol{SMI}_0 = \boldsymbol{\xi}_0 / \mathscr{B}$, where $\boldsymbol{\xi}_0 = \left(\sum_{i=1}^{4} \boldsymbol{x}^{\vee}{}_{i0} \right) / \boldsymbol{y}^{\vee}{}_{10}$

Gini's Specialization Index

The GSI_k for the hospital k is computed as follows, [45, 46]:

1) Let $\mathscr{L}$ be the number of Disease Related Groups (DRG);
2) Sort DRGs by discharges treated, in ascending order;
3) Let D_w^k be the number of the w-th DRG group discharges;
4) Let $\boldsymbol{q_i^k}, \boldsymbol{i} = 1, \ldots, \mathscr{L}-1$, be the ratio of total discharges treated by the first i DRGs, i.e., $\boldsymbol{q_i^k} = \Sigma_{\ell=1}^{i} D_\ell^k / \sum_{w=1}^{\mathscr{L}-1} D_w^k$
5) Compute $\boldsymbol{GSI_k} \in [0; 1]$ using Eq. (20).

$$GSI_k = \left\{ \sum_{i=1}^{L-1} \left(\frac{i}{L} - q_i^k \right) \Big/ \sum_{i=1}^{L-1} \left(\frac{i}{L} \right) \right\} = \left\{ \sum_{i=1}^{L-1} \left(\frac{i \cdot \sum_{w=1}^{L-1} D_w^k - L \cdot \sum_{l=1}^{i} D_l^k}{L \cdot \sum_{w=1}^{L-1} D_w^k} \right) \Big/ \sum_{i=1}^{L-1} \left(\frac{i}{L} \right) \right\} \tag{20}$$

Bootstrapping

Based on Simar and Wilson [47] and Daraio and Simar [38], the output-oriented bootstrap algorithm is as follows:

1) Compute the n output-oriented DEA efficiency scores, under the strong or the weak disposability assumption, Eqs. (6) and (7), respectively; for the sake of generality, let's suppose we obtain the set of efficiency scores, $\Phi = \{\theta_j, j = 1, \ldots, n\}$, with a standard $\sigma_{\Phi^{"}}$ deviation and an interquartile range $r_{\Phi^{"}}$.
2) Reflect Φ and obtain the $2n$-length set $\Phi' = \{2-\theta_1, 2-\theta_2, \ldots, 2-\theta_n, \theta_1, \theta_2, \ldots, \theta_n\}$.
3) Consider only those p DMUs such that $\theta_b > 1$, $b = 1, \ldots, p < n$; from Φ', create the $2p$-length set $\Phi'' = \{2-\theta_1, 2-\theta_2, \ldots, 2-\theta_p, \theta_1, \theta_2, \ldots, \theta_p\} \subset \Phi'$; Φ'' has a standard deviation $\sigma_{\Phi''}$ and an interquartile range $r_{\Phi''}$.
4) Compute a bandwidth $d \approx \left(1.06 \cdot \sigma_{\boldsymbol{\Phi}} \cdot \min\left\{\sigma_{\boldsymbol{\Phi}''}, \frac{r_{\boldsymbol{\Phi}''}}{1.34}\right\} (2p)^{4/5}\right) / (\boldsymbol{n} \cdot \sigma_{\boldsymbol{\Phi}''})$.

5) Randomly (with reposition) draw a n-length sample from Φ' (step 2)) and obtain the set $\Phi^{\star} = \left\{\tilde{\theta}_j^{\star}, \mathrm{j} = 1, \ldots, n\right\}$, with a standard deviation σ ⋆ and an arithmetic mean m ⋆.

6) Use a perturbation $\chi_j = d \cdot \zeta_j$, where $\zeta_j \sim N(\mu = 0, \sigma = 1)$, to obtain the set

$$\Phi^{\star\star} = \left\{ \tilde{\theta_j^{\star\star}} = \frac{\tilde{\theta_j^{\star}} + \chi_j - m^{\star}}{\sqrt{1 + \left(\frac{d}{\&^{\star}}\right)^2}} + m^{\star}; \mathrm{s.t.}\chi_j = d \cdot \zeta_j; \zeta_j \sim N(\mu = 0, \sigma = 1); \mathrm{j} = 1, \ldots, n; \tilde{\theta_j^{\star}} \in \Phi^{\star} \right\} \tag{21}$$

7) Reflect those $\boldsymbol{n}$ units from $\boldsymbol{\Phi}^{\star\star}$, as follows:

$$\theta_j^{\star\star} = \begin{cases} 2 - \tilde{\theta_j^{\star\star}} & \text{if } \tilde{\theta_j^{\star\star}} < 1 \\ \tilde{\theta_j^{\star\star}} & \text{otherwise} \end{cases} \tag{22}$$

8) Create the set $\mathcal{T}^{\star\star} = \left\{ \left(x_{ij}^{\star}, y_{rj}^{\star} \right) \in \mathbb{R}_{+}^{m+s} : \ x_{ij}^{\star} = x_{ij} \cap y_{rj}^{\star} = y_{rj} \frac{\theta_j}{\theta_j^{\star\star}}, \mathrm{j} = 1, \ldots, n \right\}$, and re-run Eqs. (6) and (7) to project units in the new frontier and to obtain the bootstrap-based efficiency scores, under the strong or the weak disposability assumption, resp.

9) Repeat steps 5)-8) B times, where B is large, say $B \sim 1{,}000$ iterations.

10) Let $m_{\boldsymbol{Bj}}$ and $\sigma_{\boldsymbol{Bj}}$ be the arithmetic mean and the standard deviation of those B bootstrap-based efficiency score for unit j. Bias is then $bias_j \approx m_{Bj} - \theta_j$ and the bias-corrected DEA efficiency score is $\hat{\theta_j} = \theta_j - bias_j \approx 2 \cdot \theta_j - m_{Bj}$. Still, this bias correction shall not be performed if $|bias_j| \leq \sigma_{Bj}/4$.

Some additional graphics and tables

Please check Figs. 2, 3, 4 and 5, as well as Tables 6, 7 and 8.

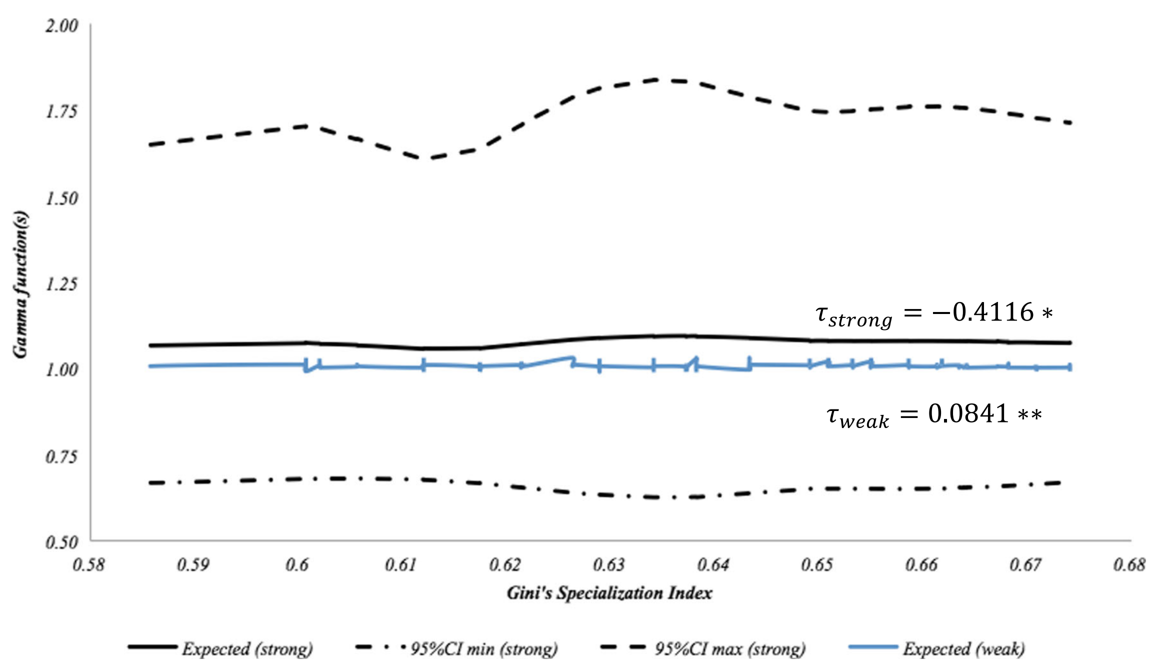

Fig. 2 Nadaraya-Watson regression of Gamma functions, $\Gamma_0(z_7)$ and $\Gamma_{weak,0}(z_7)$, against the Gini's Specialization Index, z_7. τ is the Kendall's correlation coef$z_9$ficient, * (resp. **) indicates significance (resp. Lack of significance) at 5 %

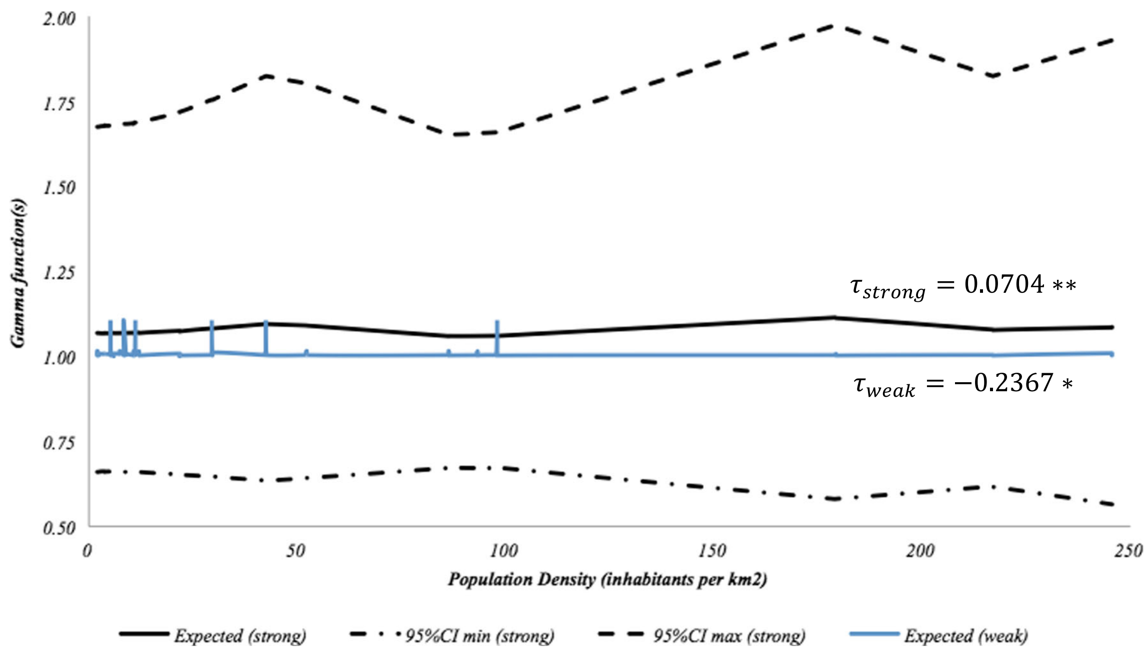

Fig. 3 Nadaraya-Watson regression of Gamma functions, $\Gamma_0(z_8)$ and $\Gamma_{weak,0}(z_8)$, against the Population Density, z_8. τ is the Kendall's correlation coefficient, * (resp. **) indicates significance (resp. Lack of significance) at 5 %

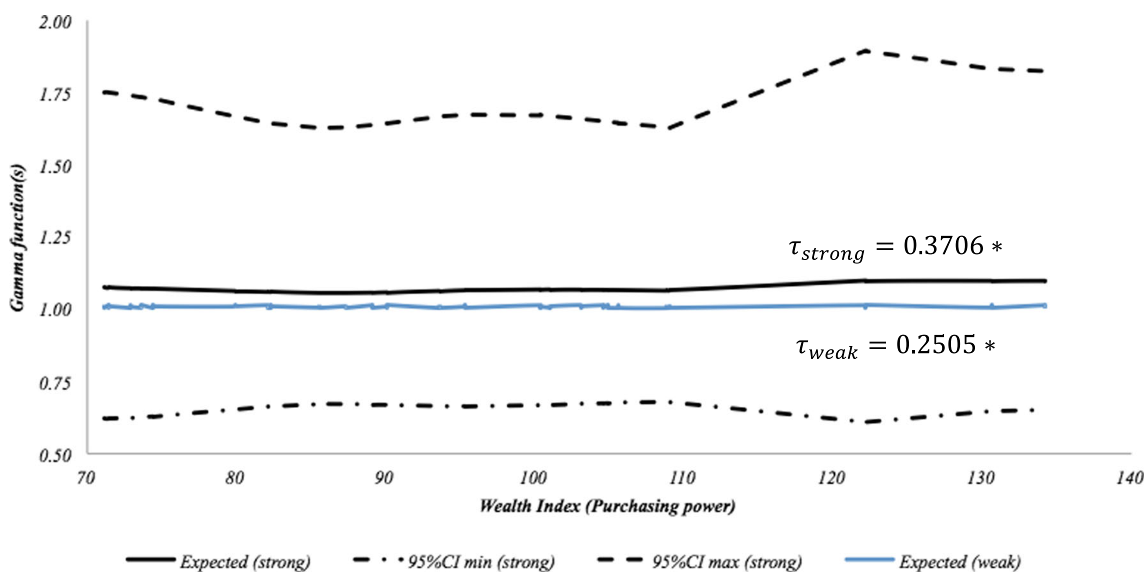

Fig. 4 Nadaraya-Watson regression of Gamma functions, $\Gamma_0(z_9)$ and $\Gamma_{weak,0}(z_9)$, against the Wealth Index, z_9. τ is the Kendall's correlation coefficient, * indicates significance at 5 %

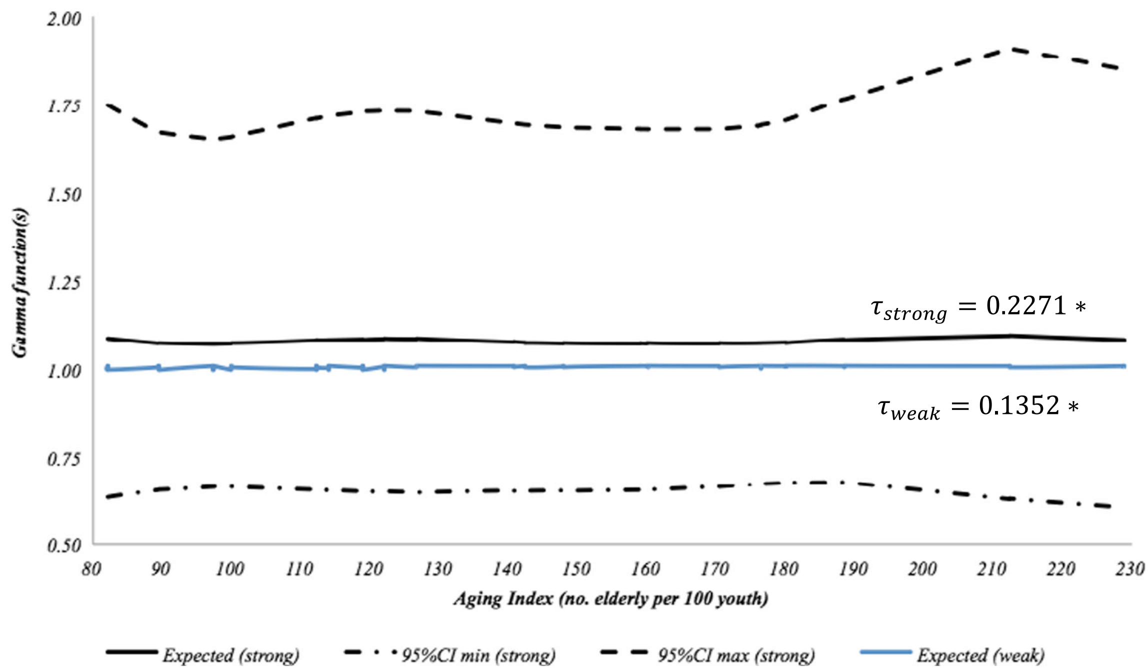

Fig. 5 Nadaraya-Watson regression of Gamma functions, $\Gamma_0(z_{10})$ and $\Gamma_{weak,0}(z_{10})$, against the Aging Index, z_{10}. τ is the Kendall's correlation coefficient, * indicates significance at 5 %

Table 6 Pearson's correlation coefficients for the non-categorical environment variables, Z_2 up to Z_{10}

	Z_2	Z_6	Z_7	Z_8	Z_9	Z_{10}
Z_2	1.0000					
Z_6	0.0490	1.0000				
Z_7	0.0315	-0.0741	1.0000			
Z_8	-0.0121	-0.0052	0.1547[a]	1.0000		
Z_9	-0.0079	-0.0055	0.1295[a]	0.6105[a]	1.0000	
Z_{10}	0.0092	0.0100	-0.3171[a]	-0.4899[a]	-0.5157[a]	1.0000

[a] Statistically non-zero Pearson's correlation coefficient, 5% level

Table 7 Degree of Scale Economies. Values are presented in the form "average ± std. deviation"

	Category	Not congested DMUs*	Strongly congested DMUs*	Weakly congested DMUs**
ICU Specialty	Polyvalent ICUs	2.3527 ± 2.3984	-1.6245 ± 2.8567	-3.0537 ± 3.0707
	Cardiology ICUs	2.2062 ± 2.8155	-2.2600 ± 1.5423	-1.5933 ± 2.9749
	Pediatric, Gynecology, Obstetrics and Neonatology ICUs	3.4541 ± 3.1206	-3.5000 ± 3.0358	-5.8421 ± 7.4004
	Surgical ICUs	2.5456 ± 4.3539	-3.8729 ± 4.5665	-6.4562 ± 23.3604
	Kruskal-Wallis p-value	*0.0683*	*0.0721*	$\sim 10^{-3}$ ***
Year	2002	2.4536 ± 2.0036	-0.4400 ± 0.0424	-3.6634 ± 5.7779
	2003	2.5416 ± 2.2321	-2.1100 ± 2.3218	-2.0748 ± 3.1509
	2004	2.2580 ± 1.9033	-0.7550 ± 0.8503	-2.8544 ± 5.6561
	2005	1.7234 ± 1.4755	-0.7675 ± 0.5864	-3.1132 ± 6.2122
	2006	2.6471 ± 3.1901	-3.5222 ± 2.8134	-8.4795 ± 25.7100
	2007	2.9541 ± 4.0797	-3.6225 ± 3.7837	-3.6577 ± 20.6786
	2008	2.6162 ± 3.4313	-3.3150 ± 2.1626	-3.6737 ± 5.3991

Table 7 (continued)

	Category	Not congested DMUs*	Strongly congested DMUs*	Weakly congested DMUs**
	2009	2.5879 ± 3.0552	-2.5850 ± 3.9369	-3.4676 ± 8.3689
	Kruskal-Wallis p-value	*0.4204*	*0.1534*	*0.9201*
Legal Status	SPA hospitals	2.6527 ± 2.6341	-2.3453 ± 2.9201	-2.9949 ± 11.0768
	SA hospitals	2.5153 ± 2.1841	-1.6367 ± 1.7198	-3.3890 ± 5.9274
	EPE hospitals	2.1520 ± 3.1824	-2.5743 ± 3.3199	-3.9511 ± 14.6341
	Kruskal-Wallis p-value	*0.0271******	*0.8771*	*0.9428*
Hospital Type	Undifferentiated hospitals	2.3672 ± 2.9018	-2.4515 ± 3.1038	-3.8671 ± 13.1227
	Maternities	2.6775 ± 2.3256	N.A.	-1.4938 ± 1.0863
	Oncology centers	3.7150 ± 1.5699	-0.3400 ± 0	-1.4443 ± 1.8188
	Kruskal-Wallis p-value	*0.1309*	*0.2615*	*0.1248*
Merging Status	Singular Hospitals	2.5763 ± 2.6920	-2.0912 ± 3.0119	-3.3894 ± 13.0565
	Hospital Centers	1.9682 ± 3.1379	-3.4069 ± 3.2601	-3.7319 ± 12.0284
	Local Health Units	2.8890 ± 4.1808	-0.1700 ± 0	N.A.
	Kruskal-Wallis p-value	*0.0049****	*0.1009*	*0.6153*
Global results		**2.3834 ± 2.8745**	**-2.4075 ± 3.0927**	**-3.5178 ± 12.6430**

*As computed by Eqs. (9)-(10); ** As computed by Eq. (12); *** Reject the null hypothesis (equal distributions) at both 5 % and 1 % levels; **** Do not reject the null hypothesis at the 1 % level

Table 8 Congestion sources (10 % trimmed mean of Marginal Products for each input, given a single output (InpD))

Category		MP(*CGSC*, *InpD*)	MP(*SEServ*, *InpD*)	MP(*StaffC*, *InpD*)	MP(*CapC*, *InpD*)	MP(*HospDays*, *InpD*)
Global results		18.1447	91.1657	**-0.0266**	64.2085	**-1.5426**
ICU Specialty	Polyvalent ICUs	0.0573	58.9931	**-4.2376**	38.4514	0.4776
	Cardiology ICUs	34.0936	129.6380	9.7884	94.7073	**-4.1157**
	Pediatric, Gynecology, Obstetrics and Neonatology ICUs	23.4305	15.1942	**-3.0018**	72.0594	**-0.1652**
	Surgical ICUs	30.0661	272.9375	2.1015	56.5756	**-5.1755**
	Kruskal-Wallis p-value	*0.0186***	*0.3342*	*0.0072**	*0.6750*	*0.4312*
Year	2002	6.6223	120.3139	21.6736	116.8436	**-7.7294**
	2003	40.8448	49.6468	3.0453	59.0633	**-6.0143**
	2004	5.5706	39.1087	18.4888	96.4398	**-7.5332**
	2005	23.4032	192.0016	**-4.4267**	130.2893	**-2.1478**
	2006	10.5027	64.6359	**-0.9001**	51.7045	**-0.3385**
	2007	19.3308	81.7728	**-5.8501**	31.5438	1.3036
	2008	31.4362	90.1111	**-6.5503**	16.3908	1.3850
	2009	20.3511	127.7868	**-13.6509**	74.4281	2.5048
	Kruskal-Wallis p-value	*0.8588*	*0.6507*	*0.0076**	*0.1684*	*0.0501*
Legal Status	SPA hospitals	5.2989	117.6459	6.3082	77.2840	**-4.1357**
	SA hospitals	54.3361	12.1758	0.5617	72.0957	**-6.3242**
	EPE hospitals	24.3196	93.1856	**-4.7562**	55.6677	0.6684
	Kruskal-Wallis p-value	*0.2260*	*0.0370***	*0.0400***	*0.2654*	*0.0279***
Hospital Type	Undifferentiated hospitals	17.1078	87.9799	0.0380	59.9188	**-1.1974**
	Maternities	**-0.6160**	**-1.4894**	10.5318	82.7765	**-4.1285**
	Oncology centers	215.6307	328.2780	**-52.7463**	146.8278	**-11.0725**
	Kruskal-Wallis p-value	*0.6507*	*0.5530*	*0.0162***	*0.2608*	*0.7010*

Table 8 (continued)

Category		MP(*CGSC*, *InpD*)	MP(*SEServ*, *InpD*)	MP(*StaffC*, *InpD*)	MP(*CapC*, *InpD*)	MP(*HospDays*, *InpD*)
Merging Status	Singular Hospitals	17.1573	62.4545	0.3335	79.2229	**-2.3999**
	Hospital Centers	20.7799	194.5607	**-0.9292**	27.8207	**-0.1896**
	Local Health Units	N.A.***				
	Kruskal-Wallis p-value	*0.8428*	*0.3324*	*0.9215*	*0.0055**	*0.7000*

Bold entries identify the main source(s) of congestion

* Reject the null hypothesis (equal distributions) at both 5 % and 1 % levels; ** Do not reject the null hypothesis at the 1 % level;

***N.A. – not applicable

References

1. Cooper WW, Seiford LM, Zhu J (2000) A unified additive model approach for evaluating inefficiency and congestion with associated measures in DEA. Socio-Econ Plan Sci 34:1–25
2. Nayar P, Ozcan YA, Yu F, Nguyen AT (2013) Benchmarking urban acute care hospitals: Efficiency and quality perspectives. Health Care Manag Rev 38(2):137–145
3. Hsieh HM, Clement DG, Bazzoli GJ (2010) Impacts of market and organizational characteristics on hospital efficiency and uncompensated care. Health Care Manag Rev 35(1):77–87
4. Kahn JM, Goss CH et al (2006) Hospital volume and the outcomes of mechanical ventilation. N Engl J Med 355(1):41–50
5. Khandelwal N, Benkeser DC, Coe NB, Curtis JR (2016) Potential Influence of Advance Care Planning and Palliative Care Consultation on ICU Costs for Patients with Chronic and Serious Illness. Crit Care Med 44(8):1474–1481
6. Halpern NA, Pastores SM, Greenstein RJ (2004) Critical care medicine in the United States 1985-2000: An analysis of bed numbers, use, and costs. Crit Care Med 32(6):1254–1259
7. Barrett, M.L., Smith, M.W., Elixhauser, A., Honigman, L.S., Pines, J.M., 2014. Utilization of Intensive Care Services, 2011. HCUP Statistical Brief #185. Agency for Healthcare Research and Quality, Rockville. Retrieved from https://www.hcup-us.ahrq.gov/reports/statbriefs/sb185-Hospital-Intensive-Care-Units-2011.jsp, 15th September 2016.
8. Puig-Junoy J (1998) Technical Efficiency in the Clinal Management of Critically Ill Patients. Health Econ 7:263–277
9. Tsekouras K, Papathanassopoulos F, Kounetas K, Pappous G (2010) Does the adoption of new technology boost productive efficiency in the public sector? The case of ICUs system. Int J Prod Econ 128:427–433
10. Dervaux B, Leleu H, Minvielle E, Valdmanis V, Aegerter P, Guidet B (2009) Performance of French intensive care units: A directional distance function approach at the patient level. Int J Prod Econ 120: 585–594
11. Clement J, Valdmanis V, Bazzoli G, Zhao M, Chukmaitov A (2008) Is more, better? An analysis of hospital outcomes and efficiency with a DEA model of output congestion. Health Care Manag Sci 11:67–77
12. Valdmanis V, Rosko M, Muller R (2008) Hospital quality, efficiency, and input slack differentials. Health Serv Res 43:1830–1848
13. Ferrier GD, Rosko M, Valdmanis V (2006) Analysis of uncompensated hospital care using a DEA model of output congestion. Health Care Manag Sci 9:181–188
14. Arrieta A, Guillén J (2015) Output congestion leads to compromised care in Peruvian public hospital neonatal units. Health Care Manag Sci. doi:10.1007/s10729-015-9346-y
15. Matranga D, Sapienza F (2015) Congestion analysis to evaluate the efficiency and appropriateness of hospitals in Sicily. Health Policy 119(3):324–332
16. Simões P, Marques RC (2011) Performance and Congestion Analysis of the Portuguese Hospital Services. CEJOR 19:39–63
17. Mathews KS, Long EF (2015) A Conceptual Framework for Improving Critical Care Patient Flow and Bed Use. Ann Am Thorac Soc 12(6). doi:10.1513/AnnalsATS.201409-419OC
18. Chan CW, Farias VF, Escobar GJ (2016) The Impact of Delays on Service Times in the Intensive Care Unit. Manag Sci. doi:10.1287/mnsc.2016.2441
19. Farrell MJ (1957) The measurement of productive efficiency. J R Stat Soc Ser A 120:253–281
20. Charnes A, Cooper WW, Rhodes E (1978) Measuring efficiency of decision-making units. Eur J Oper Res 2:429–444
21. Ozcan YA (2008) Health Care Benchmarking and Performance Evaluation: An Assessment using Data Envelopment Analysis (DEA). Springer, New York
22. Hollingsworth B (2003) Non-parametric and parametric applications measuring efficiency in health care. Health Care Manag Sci 6:203–218
23. Ruggiero J (2007) A comparison of DEA and the stochastic frontier model using panel data. Int Trans Oper Res 14:259–266
24. De Witte K, Marques R (2009) Capturing the environment, a metafrontier approach to the drinking water sector. Int Trans Oper Res 16:257–271
25. Jacobs R, Smith PC, Street A (2006) Measuring Efficiency in Health Care: Analytic Techniques and Health Policy. Cambridge University Press, Cambridge
26. Hollingsworth B (2008) The Measurement of Efficiency and Productivity of Health Care Delivery. Health Econ 17(10):1107–1128
27. Ferreira D, Marques RC (2016) Should inpatients be adjusted by their complexity and severity for efficiency assessment? Evidence from Portugal. Health Care Manag Sci 19(1):43–57
28. Ferreira D, Marques RC (2015) Did the corporatization of Portuguese hospitals significantly change their productivity? Eur J Health Econ 16(3):289–303
29. Marques RC, Carvalho P (2012) Estimating the efficiency of Portuguese hospitals using an appropriate production technology. Int Trans Oper Res 20(2):1–17
30. Cooper WW, Seiford LM, Tone K (2007) Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-solver Software. Kluwer Academic Publishers, Boston
31. Sueyoshi T, Sekitani K (2009) DEA congestion and returns to scale under an occurrence of multiple optimal projections. Eur J Oper Res 194(2):592–607

32. Färe R, Grosskopf S, Lovell CAK (1985) The measurement of efficiency of production. Kluwer-Nijhoff Publishing
33. Färe R, Grosskopf S, Lovell CAK (1994) Production Frontiers. Cambridge University Press, Cambridge
34. Färe R, Svensson L (1980) Congestion of production factors. Econometrica 48:1745–1753
35. Sueyoshi T (2003) DEA Implications of Congestion. Asia Pacific Manag Rev 8(1):59–70
36. Tone K, Sahoo BK (2004) Degree of scale economies and congestion: a unified DEA approach. Eur J Oper Res 158:755–772
37. Ferreira, D., Marques, R.C. 2016. Malmquist and Hicks-Moorsteen Productivity Indexes for Clusters Performance Evaluation. *International Journal of Information Technology & Decision Making*, forthcoming.
38. Herr A (2008) Cost and technical efficiency of German hospitals: does ownership matter? Health Econ 17(9):1057–1071
39. Daraio C, Simar L (2007) Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications. Springer Science + Business Media, LLC, New York
40. Ferreira, D., Marques, R.C., Pedro, I., 2016. Comparing efficiency of holding business model and individual management model of airports. *Journal of Air Transport Management, forthcoming*
41. Rego G, Nunes R, Costa J (2010) The challenge of corporatisation: the experience of Portuguese public hospitals. Eur J Health Econ 11:367–381
42. Ozcan YA, Luke RD (1993) A national study of the efficiency of hospitals in urban markets. Health Serv Res 28(6):719–739
43. Ozcan YA (1993) Sensitivity analysis of hospital efficiency under alternative output/input and peer groups: a review. Knowl Policy 5(4):1–29
44. Lobo MSC, Ozcan YA, Lins MPE, Silva ACM, Fiszman R (2014) Teaching hospitals in Brazil: findings on determinants for efficiency. Int J Healthcare Manag 7(1):60–68
45. Daidone S, D'Amico F (2009) Technical efficiency, specialization and ownership form: evidences from a pooling of Italian hospitals. J Prod Anal 32:203–216
46. Lindlbauer I, Schreyögg J (2014) The relationship between hospital specialization and hospital efficiency: do different measures of specialization lead to different results? Health Care Manag Sci 17:365–378
47. Simar L, Wilson PW (1998) Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. Manag Sci 44:49–61
48. Grossman M (1972) On the Concept of Health Capital and the Demand for Health. J Polit Econ 80(2):223–255
49. Kuntz L, Sülz S (2011) Modeling and notation of DEA with strong and weak disposable outputs. Health Care Manag Sci 14(4):385–388