Efficiency assessment of primary care providers: A conditional nonparametric approach

José Manuel Cordero-Ferrera[1], Edurne Alonso-Morán[2], Regina Sauto Arce[2], Roberto Nuño-Solinis[2], Juan F. Orueta[3]

[1] Universidad de Extremadura, Av Elvas s/n, 06006, Badajoz, Spain

[2] O+berri, Basque Institute for Healthcare Innovation, Plaza Asua 1, 48150 Sondika, Spain

[3] Osakidetza, Basque Health Service, Mezo 35, 48950, Erandio, Spain

Abstract

This paper uses a fully nonparametric approach to estimate efficiency measures for primary care units incorporating the effect of environmental factors. This methodology allows us to account for different types of variables (continuous and discrete) regarding the main characteristics of patients served by those providers. In addition, we use an extension of this nonparametric approach to deal with the presence of undesirable outputs in data, represented by the number of readmissions and hospitalization rates of ambulatory care sensitive condition (ASCS). The empirical results of our application show that all the exogenous variables considered have a significant and negative effect on efficiency estimates.

1.   INTRODUCTION

The combination of growing care demands from an ageing population with the increasing pressures in public budgets as a consequence of the economic crisis has placed the objective of improving the effectiveness and efficiency of health systems at the center of the debate in the future of healthcare in Europe (European Commission, 2010). Within this framework, most of the empirical work has been focused on hospitals (Steinmann *et al.*, 2004), although primary and community-based primary care provision is progressively receiving much attention, since it is considered as a pivotal element in the search for a more efficient organization of healthcare (Amado and Dyson, 2008).

Modeling the production technology of primary care centers (hereafter PCCs) is a difficult task because the final output, which should capture the impact of the services on current and future health status of patients, cannot be measured directly, so intermediate products related with final outputs are normally used as proxies. In addition, these centers provide multiple services that are not clearly connected to those outputs and their performance depends on the health status of the population covered.

In this context, the use of a nonparametric approach, and particularly data envelopment analysis (DEA), has become very popular in empirical studies, since it can easily handle multiple dimensions of performance and is less vulnerable to the misspecification problems that can affect econometric models[1]. Furthermore, multiple extensions of this technique have been developed in the literature to facilitate its adaptation to different frameworks (Emrouznejad *et al.*, 2008).

This paper contributes to the existing literature in the field by applying two extensions of the nonparametric approach to assess the performance of a set of primary care units. Specifically, we combine the use of a transformation approach in the traditional DEA model to incorporate undesirable outputs with the utilization of a conditional approach to include the effect of both continuous and discrete environmental variables.

---

[1]See Hollingsworth (2003, 2008) or Worthington (2004) for an overview of studies on the measurement of efficiency in the health sector.

The potential existence of undesirable (bad) outputs in a production process was already mentioned in the seminal work of Koopmans (1951). These undesirable outputs are prominent in the energy and environmental context in the form of pollution emissions or waste in resources, but they can also appear in health care services (e.g. complications of medical operations). In those cases, the aim of units should be to minimize these outputs instead of maximize them. However, in the standard DEA model, decreases in outputs are not allowed (only inputs are allowed to decrease), thus it would be necessary to transform the original data or change the technology of production in order to take into account the presence of these factors. In the current study, the undesirable factors are represented by the number of readmissions and hospitalization rates of ambulatory care sensitive condition (ASCS). In order to include them in a DEA model we use the extension developed by Seiford and Zhu (2002), since this is the method that fits better with the technology of production assumed. This model has been used previously in the health sector (Hu *et al.*, 2012) as well as in other environmental contexts (Lu and Lo, 2007; Hua *et al.*, 2007).

In addition to the problem of dealing with undesirable outputs in data, we also need to bear in mind that the performance of PCCs can be affected by exogenous or environmental variables, which in the context of our study are represented mainly by the characteristics of the patients served by each unit (Lezzoni, 1994). Those variables, unlike the inputs and the outputs, are not under the control of the Decision Making Units (DMUs), thus we need to include those variables in our efficiency analysis in a different way. For that purpose, many different approaches can be used (see Cordero *et al.*, 2008 for an overview), however the validity of these traditional models is limited, because they need to assume the separability condition between the input-output space and the space of external factors environmental variables. This is difficult to believe in a framework like the health sector, where the characteristics of the population are clearly related to the volume of outcomes and even some inputs. Therefore, in this paper we use the so-called *conditional nonparametric approach* (Cazals, 2002; Daraio and Simar, 2005; Daraio and Simar, 2007a), which avoids the restrictive separability assumption required by traditional approaches in order to provide meaningful results. This method was designed for continuous variables only, but we are interested in considering also discrete variables, so we apply an extension of this methodology developed by De Witte and Kortelainen (2013) to include both types of exogenous variables.

The conditional approach has become very popular in the recent literature on efficiency measurement. Hence, it is possible to find works using this approach to measure the efficiency of units operating in multiple frameworks such as the educational sector (Bonaccorsi *et al*., 2006; Cherchye *et al*., 2010; De Witte and Rogge, 2011; Haelermans and De Witte, 2012; De Witte and Kortelainen, 2013; De Witte *et al*., 2013), banks and mutual funds (Daraio and Simar, 2005; 2006; Daouia and Simar, 2007; Blass *et al*., 2008; Badin *et al*., 2010), post offices (Cazals *et al*., 2008), public libraries (De Witte and Geys, 2011; 2013), regional welfare or environment (Halkos and Tzeremes, 2011a, Halkos and Tzeremes, 2013), local services (Rogge and De Jaeger, 2012; Verschelde and Rogge, 2012) or the water sector (De Witte and Marques, 2010a, 2010b; Carvalho and Marques, 2011; Vidoli, 2011). However, to the best of our best knowledge, this methodology has only been applied previously in the health sector to measure the efficiency of health administrative districts in Greece (Halkos and Tzeremes, 2011b) considering only continuous exogenous variables. Therefore, this is the first study using this approach to estimate efficiency measures of PCCs´ performance and, additionally, we consider both continuous and discrete environmental variables.

In the previous literature, there are few previous studies where the presence of both undesirable outputs and exogenous factors are considered simultaneously in a DEA framework. Hua *et al.* (2007) developed a non-radial output-oriented DEA considering both types of variables in a model to estimate the ecological efficiency of paper mills in China. Yang and Pollitt (2009) assess the performance of Chinese coal-fired power plants combining four different traditional models to incorporate uncontrollable variables into DEA with one approach to deal with undesirable factors. More recently, Halkos and Tzeremes (2013) use a conditional directional distance function approach for measuring regional environmental efficiency, although they opt for transforming the technology of production to adapt it to the presence of undesirable outputs.

The rest of the paper is organized as follows. Next section provides a brief literature review about previous studies attempting to measure efficiency of primary care centers. Section 3 presents the methodology employed in our study with a detailed explanation of the approaches we use to incorporate both undesirable outputs and environmental variables into the estimation of efficiency measures of PCC´s performance. Section 4

describes the dataset and the variables used in our analysis and section 5 contains the main results. Finally, conclusions appear in section 6.

2. LITERATURE REVIEW

As we mentioned in the previous section, most of studies carried out in the context of primary health care have opted to use a nonparametric approach. Since the pioneer work of Nunamaker (1983) on nursing service efficiency, multiple applications used this technique to estimate efficiency measures of primary care units (Huang and McLaughlin, 1989; Pina and Torres, 1992; Szczepura *et al.*, 1993). Later, it was also employed to assess the performance of general practitioners (Bates *et al.*, 1996), physicians (Chilingerian and Sherman, 1996, 1997) and primary care teams (Goñi, 1999).

All these studies use the so-called "*activity-oriented models*" in which the output of primary health care is approximated by the activity levels of the health units being analyzed, in particular, by their recorded number of visits and consultations. However, the use of these *proxies* of primary health care output is clearly prone to criticism, since there is no clear link between the number of visits and the quality of the health care because some patients who attend PCCs may not receive the appropriate attention.

Such criticisms show that it is necessary to consider not just quantitative but also qualitative indicators to properly measure primary health care output. However, the concept of quality is difficult to define, since it encompasses technical aspects reflecting the capacity of medical staff to diagnose and treat medical problems, but also patients´ perceptions about the service delivered (Donabedian, 1980; Campbell *et al*., 2000). Therefore, in the short list of studies using a nonparametric framework to incorporate quality measures into the efficiency analysis it is possible to detect a variety of indicators as proxies to this concept (Salinas-Jiménez and Smith, 1996; García *et al*., 1999; Rosenman and Friesner, 2004; Amado and Santos, 2009).

An indicator widely used as a measure of primary care quality and effectiveness is represented by admission rates of hospitalization for ambulatory care sensitive conditions (ACSC) (Finegan *et al.* 2010; Schiøtz e*t al*., 2011; Pelone *et al.* 2012). The

conditions chosen are those for which timely and effective primary care could be expected to reduce the risk of admission to hospital by preventing the onset of illness, controlling an acute episode of illness or better long term management (Giuffrida *et al*., 1999). Several empirical studies (Kringos *et al.,* 2013) and literature reviews (Rosano *et al*., 2012; Gibson *et al*., 2013) support the (negative) association between access and quality of primary care and ACSC hospitalization rates. The role of primary care in avoiding hospitalizations for ACSCs is supported by experts´ opinions and particularly well suited to the views of primary care professionals as regards to their own role in the healthcare system. Such role would be focused around primary prevention, early detection and monitoring of acute episodes, and monitoring of chronic conditions (Caminal *et al.,* 2004). Likewise, care coordination and integration is a dimension of primary care performance of particular interest in a context where chronic patients represent most of the interactions with the healthcare system (Ham, 2010; Nuño *et al.* 2012). In this sense, previous comparative and evaluation studies have included indicators of integrated care such as unplanned hospital admissions, re-admissions and avoidable hospital admissions (Curry and Ham, 2010), 30-day readmission rates (Schiøtz e*t al*., 2011) and emergency hospital admission (Bardsley *et al*., 2013). The main problem with all these indicators is that they cannot be used in a standard DEA model, because they represent undesirable targets for the evaluated units, so we need to adapt the model used to their presence as we explain in section 3.

Another way of increasing the accuracy of the model specification is to consider the influence that external or environmental variables can have on the performance of health care providers. Particularly, in making an efficiency assessment of primary health care units these variables are mainly represented by the characteristics of the population demanding care. Many previous studies recommend adjusting for case-mix when making comparisons between healthcare organizations on the basis of ACSC hospitalization rates (Finegan *et al*., 2010), which are also found to be influenced by other clinical and socioeconomic factors such as age, health status and co-morbidities, deprivation or income level (Caminal *et al.,* 2004; Rosano *et al*., 2012; Gibson *et al.*, 2013).

However, considering these environmental factors is an approach not much explored in the primary health care performance literature. Actually, most of the applications that

have attempted to incorporate this information have been limited to performing a second-stage analysis in order to identify potential explanatory factors of inefficient behavior, but they do not incorporate the effect of these variables into the efficiency scores (Zavras *et al*., 2002; Kontodimopoulos *et al*., 2007, Ramirez-Valdivia *et al*. 2010).

More recently, Kontodimopoulos *et al.* (2010) and Cordero *et al*. (2010; 2013) estimate corrected efficiency scores including information about the characteristics of the population covered by each PCCs using two traditional approaches such as the three-stage model (Muñiz, 2002) and four-stage model (Fried *et al*., 1999), which adjust original input and output values to obtain a measure of managerial inefficiency that controls for the effect of exogenous factors. However, as we mentioned in the previous section, there have not been previous empirical studies attempting to incorporate this information through a conditional nonparametric model.

3. METHODOLOGY

Following most of previous studies in the health sector, we use a nonparametric approach to measure the efficiency of primary care centers. Introducing the basic notation used in this paper, we consider a production process where units are characterized by a set of inputs x ($x \in \Re_+^p$) and outputs y ($y \in \Re_+^q$). The production technology is the set of all feasible input-output combinations:

$$\psi = \left\{ x, y) \in \Re_+^{p+q} \ \middle| \ x \text{ can produce } y \ \right\} \tag{1}$$

Given that the set ψ cannot be observed as well as the efficiency scores, it has to be estimated from a random sample of production units denoted by $X = \left\{ x_i, y_i) \middle| i = 1,...,n \right\}$. Since the pioneering work of Farrell (1957), multiple approaches have been developed to achieve this goal. In this framework, an observed production unit (x_i, y_i), defines an individual production possibilities set $\psi(x_i, y_i)$, which under the free disposability of inputs and outputs, can be written as:

$$\psi(xi, yi) = \left\{ x, y \right) \in \Re_+^{p+q} \big| x \ge x_i; y \le y_i \right\}$$ (2)

Within this framework, the DEA estimator is the most common in the literature since it does not rely on restrictive hypothesis on the Data Generating Process. Following the notation provided by Daraio and Simar (2007a), this estimator $\hat{\psi}_{DEA}$ can be defined as[2]:

$$\hat{\psi}_{DEA} = \left\{ (x, y) \in \Re_+^{p+q} \Big| y \le \sum_{i=1}^{n} \gamma_i y_i; x \ge \sum_{i=1}^{n} \gamma_i x_i, \text{ for } (\gamma_1, ..., \gamma_n) \right.$$

$$\left. \text{s.t.} \sum_{i=1}^{n} \gamma_i = 1; y_i \ge 0, i = 1, ..., n \right\}$$ (3)

The estimator of the output efficiency score for a given (x_0, y_0) can be obtained by solving a simple linear program:

$$\hat{\lambda}_{DEA}(x_0, y_0) = \sup \left\{ \lambda \big| (x_0, \lambda y_0) \in \hat{\psi}_{DEA} \right\}$$

(4)

This production describes a production process which transforms conventional inputs into conventional outputs. However, in the context of our study the technology must also comprise the production of undesirable factors and the potential effect of external or environmental variables on results. Next subsections are devoted to the analysis of these two aspects within a nonparametric framework.

3.1. Dealing with undesirable outputs

Efficiency measurement usually relies on the idea that inputs have to be minimized and outputs have to be maximized. This means that, for each evaluated unit, more output and less input imply a higher degree of efficiency, which is also implicitly assumed in DEA models. However, in some cases, the production function may also contain

---

[2] This definition represents the case of variable returns to scale (VRS) according to the model introduced by Banker *et al.* (1984).The constant returns to scale model developed by Charnes *et al.* (1978) can also be applied when the equality constrained ($\sum_{i=1}^{n} \gamma_i = 1$) is omitted from the equation.

undesirable that need to be minimized (Chung *et al.*, 1997), which complicates the estimation of the DEA standard efficiency scores using equation (2), because these bad outputs cannot be simply incorporated as another conventional output[3].

In the literature we can find several approaches for handling undesirable factors in DEA models[4], although a general protocol is not clear (Sheel, 2001). These models can be roughly divided into two groups. The first is based on the concept of weak disposability reference technology and allows for using the original data. The second is based on data translation and the utilization of traditional DEA models. It is worth noting that the use of these alternative approaches often leads to different outcomes in terms of the units identified as efficient and in terms of the targets set for inefficient units (Dyson *et al.*, 2001; Sahoo *et al.*, 2011).

Färe *et al.* (1989) proposed the first non-linear DEA program where the desirable outputs are increased and undesirable outputs are decreased using a hyperbolic output efficiency measure. Subsequently, Färe and Grosskopf (2004) suggested an alternative approach in treating the undesirable factors by adopting a directional distance function to estimate efficiency scores based on weak disposability of undesirable outputs[5]. These approaches have been widely applied in the field of environmental performance measurement, where the presence of undesirable outputs is frequent[6]. Actually, the weak disposable reference technology is also referred to as an environmental DEA technology. However, in our study we have decided not to use this methodology because it would imply the introduction of new axioms (weak disposability of outputs) that would be incompatible with the methodology presented above as well as the extension used to deal with external factors[7].

Among the methods transforming the data, which do not need to modify the standard axioms of the technology (free disposability is assumed), there are also different options. The first possibility would be either to treat the negative or undesirable outputs as inputs (Dyckhoff and Allen, 2001, Korhonen and Luptacik, 2004), or to invert the

---

[3] The presence of negative values in data can cause similar problems. See Cheng *et al.* (2013) for details.
[4] SeeHua and Bian (2007) or You and Yan (2011) for an extended review of these methods.
[5] Weak disposability assumes that it is costly to reduce undesirable outputs because it increases the inputs or decreases desirable outputs at the same time (Yang *et al.*, 2008).
[6] See Zhou *et al.* (2008) for a review of studies using this approach.
[7] See Podinovski and Kuosmanen (2011) for details.

value of the original variables (Lovell *et al*., 1995). However, these methods do not truly reflect the real production process or lose an invariant to the data transformation.

To overcome these shortcomings, Seiford and Zhu (2002) developed a methodology based on a monotone decreasing transformation by multiplying each undesirable output by −1 and then find a proper translation vector to let all negative undesirable outputs be positive. In particular, they proposed that a sufficient large positive scalar constant number (K) is added to the reciprocal additive transformation of the undesirable output, so that the final new value would be isotonic.

This methodology can truly reflect the real production process and is invariant to the data transformation within the DEA model (Lovell and Pastor, 1995), so this is the method we apply to treat the undesirable output factors in our study. The problem here arises from the fact that the method is sensitive to the choice of the constant value, thus an over large value can dominate the data and modify the structure of the efficient frontier, while selecting a small value reduces the effect of the translation on results. Therefore, we must take this decision cautiously. Moreover, due to strong convexity constraints, it can only be solved under variable returns to scale (Silva Portela *et al*., 2004), so our empirical analysis should be performed using the BCC model (Banker *et al*., 1984).

3.2. The incorporation of external factors using a conditional approach

Environmental factors that are not under the control of the primary care provider need to be considered in their evaluation since such factors are a potential source of inefficiency. An evaluation of a health care facility should explicitly include this information to ensure that the efficiency score finally assigned to the centre truly reflects the portion of the production process for which that centre is itself responsible (Muñiz, 2002).

Recent years have seen the development of different ways to incorporate the effect of external factors or environmental variables into the production process in estimating

efficiency scores through DEA[8]. The most widely used approach is a two-stage procedure, where initial efficiency scores are estimated using a standard DEA model and then they are regressed on the environmental variables (Simar and Wilson, 2007; McDonald, 2009)[9]. Most studies using this approach in the second stage estimation employed either Tobit regression or ordinary least squares. Unfortunately, usual inference on the obtained estimates of the regression coefficients is not available, so it is necessary to use a bootstrap-based procedure to obtain more reliable results (Simar and Wilson, 2007). However, this two-stage approach still has a major inconvenience, since it requires a restrictive separability condition between the input–output space and the space of external or environmental factors, assuming that those factors do not have an impact on the frontier of the efficiency scores, affecting only the probability of being more or less efficient, which is often unrealistic (Badin *et al*., 2010). One model where the two-stage approach is valid was proposed by Banker and Natarajan (2008), but their model heavily depends on quite restrictive and unrealistic assumptions on the production process, as described in Simar and Wilson (2011).

More recently, the specific literature devoted to exogenous (environmental) factors and their influence on efficiency has advanced significantly with the development of a more general and appealing full nonparametric conditional approach based on a probabilistic definition of the frontier (Daraio and Simar, 2005, 2007a, 2007b). In order to explain the conditional efficiency approach, it is necessary to introduce some basic concepts developed by Cazals *et al*. (2002). These authors proposed an alternative probabilistic formulation to the production process, denoted by $H_{XY}(x, y)$, which represents the probability of dominating a unit operating at level *(x, y)*:

$$H_{XY}(x, y) = \Pr(X \leq x, Y \geq y) \tag{5}$$

This probability function can be further decomposed as follows:

$$H_{XY}(x, y) = \Pr(Y \geq y | X \leq x)\Pr(X \leq x) =$$

$$S_{Y|X}(Y \geq y | X \leq x)F_X(X \leq x) = S_{Y|X}(y|x)F_X(x) \tag{6}$$

---

[8] See Fried *et al*. (2008) for an overview.
[9] In a slight variation on this approach, Fried *et al*. (1999, 2002) regress radial and non-radial slacks on environmental variables.

where $S_{Y|X}(y|x)$ represents the conditional function of Y and $F_X(x)$ the cumulative distribution function of X. Therefore, the output oriented technical efficiency measure can also been defined as the proportionate increase in outputs required for the evaluated unit to have a zero probability of being dominated at the given input level:

$$\hat{\lambda}(x,y) = \sup\{\lambda | S_Y(\lambda y|x) > 0\} = \sup\{\lambda | H_{XY}(x,\lambda y) > 0\} \tag{7}$$

In order to estimate efficiency scores using this probabilistic formulation, the empirical distributions functions $\hat{H}_{XY,n}(x,y)$ and $\hat{S}_{Y,n}(y|x)$ must replace $H_{XY}(x,y)$ and $S_Y(y|x)$ respectively. These empirical analogs are represented by the following expressions:

$$\hat{H}_{XY,n}(x,y) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \le x, y_i \ge y) \tag{8}$$

$$\hat{S}_{Y,n}(y|x) = \frac{\hat{H}_{XY,n}(x,y)}{\hat{F}_{X,n}(x)} = \frac{\hat{H}_{XY,n}(x,y)}{\hat{H}_{XY,n}(x,0)} \tag{9}$$

where *I(-)* is an indicator function. Using the plug-in rule, the conditional DEA estimator (which relies on the convexity assumption of ψ) for the output-oriented efficiency score can be obtained by solving the linear program involved by:

$$\hat{\lambda}_{DEA}(x,y) = \sup\{\lambda | \hat{S}_{Y,n}(\lambda x|y) \in \hat{\psi}_{DEA}\} \tag{10}$$

In this context, Cazals *et al*. (2002) and, subsequently, Daraio and Simar (2005; 2007a; 2007b) suggested that the presence additional external factors $Z \in \mathfrak{R}_+^k$ can be incorporated into the analysis by conditioning the production process to a given value of *Z = z*. This conditional function can be defined as:

$$H_{XY|Z}(x,y|z) = \Pr(X \le x, Y \ge y|Z = z) \tag{11}$$

The function $H_{XY|Z}(x, y|z)$ represents the probability for a unit operating at level (x, y) to be dominated by other units facing the same environmental conditions z. Again, this can also be decomposed into:

$$H_{XY|Z}(x, y|z) = \Pr(Y \geq y|x \leq x, Z = z)\, \Pr(X \leq x, Z = z)$$

$$= S_{Y|X,Z}(Y \geq y|X \leq x, Z = z) F_X(X \leq x; Z = z)$$

$$= S_Y(y|x, z) F_X(x|z) \tag{12}$$

Therefore, the output efficiency measure can be analogously defined as:

$$\hat{\lambda}(x, y|z) = \sup\{\lambda > 0 | S_{Y|XZ}(\lambda y|X \leq x, Z = z) > 0\} \tag{13}$$

However, the estimation of $S_Y(y|x, z)$ is more difficult than the unconditional case, because we need to use smoothing techniques for the exogenous variables in z (due to the equality constraint $Z = z$):

$$\hat{S}_{Y,n}(y|x, z) = \frac{\sum\limits_{i=1}^{n} I(x_i \leq x, y_i \geq y) K_{\hat{h}}(z, z_i)}{\sum\limits_{i=1}^{n} I(x_i \leq x) K_{\hat{h}}(z, z_i)} \tag{14}$$

Therefore, this approach relies on the estimation of a nonparametric kernel function to select the appropriate reference partners and a bandwidth parameter h using some bandwidth choice method[10]. This would be straightforward if all the Z variables are continuous, but it is more complicated if we have mixed data (continuous and discrete variables) as in our empirical study. For that purpose, De Witte and Kortelainen (2013) proposed recently a model to smooth any type of discrete variables (ordered and unordered) along with continuous variables extending the ideas proposed by Racine and Li (2004), Hall *et al.* (2004) and Li and Racine (2007). Basically, this approach consists of multiplying three different multivariate kernel functions (one of each type of

---

[10] The estimation of conditional full frontiers does not depend on the chosen kernel but only on the selected bandwidth.

variable) to obtain a generalized product kernel function ($K_{\hat{h}}$) and substitute it for $K_{\hat{h}}$ in equation 14. Subsequently, the conditional estimator $\hat{\lambda}(x, y|z)$ can be obtained by plugging in the new $\hat{S}_{Y,n}(y|x, z)$ in equation 13.

Given that our dataset only contains continuous and ordered discrete variables, we adapt this methodology to a simpler case with only two multivariate kernel functions. Following De Witte and Kortelainen (2013), we employ the Epanechnikov kernel function ($K_{\hat{h}}(z, z_i) = h^{-1}K((Z - z_i)/h)$ for continuous variables and the Li and Racine (2007) discrete kernel function for ordered discrete variables. Regarding the estimation of the bandwidth parameters, we follow the data-driven selection approach developed by Badin *et al.* (2010), which can be easily adapted to the case of mixed environmental variables[11].

Finally, this conditional approach allows us to evaluate the direction of the effect of exogenous variables on the production process by comparing conditional with unconditional measures. In particular, when Z is continuous and univariate, Daraio and Simar (2005, 2007a) suggest to use a scatter plot of the ratio between those measures ($Q^z = \hat{\lambda}(x, y|z)/\hat{\lambda}(x, y)$) against Z and its smoothed nonparametric regression line. In an output-oriented conditional model, an increasing regression line will indicate that Z is favorable to efficiency whereas a decreasing line will denote an unfavorable effect. In the first case, the environmental variable operates as a sort of *extra* input *freely available*, consequently the value of $\lambda(x, y|z)$ will be much smaller than $\lambda(x, y)$ for small values of Z than for large values of Z. In the second case, the environmental variable can be interpreted as an extra undesired output to be produced, which requires the use of more inputs, thus $\lambda(x, y|z)$ will be smaller than $\lambda(x, y)$ for large values of Z (Daraio and Simar, 2005).

In addition, it is also possible to investigate the statistical significance of Z explaining the variations of Q. For that purpose, we use local linear least squares for regression

---

[11] In the case of an ordered discrete variable, we assure that the performance of each unit is compared only to those in the same category (i.e. the same value of discrete variable) by imposing bandwidth to be zero for that variable in question.

estimation and then we apply the nonparametric regression significance test proposed by Li and Racine (2004) and Racine and Li (2004), which smooths both continuous and discrete variables. Specifically, we test the significance of each of the continuous and discrete variables using bootstrap tests proposed by Racine (1997) and Racine *et al.* (2006), which can be interpreted as the nonparametric equivalent of standard t-tests in ordinary least squares regression.

4. DATA AND VARIABLES

The present paper aims at analyzing the efficiency of PCCs in the public Basque Health Service. In the Basque Country, with a population of approximately 2.2 million inhabitants, a public organization (Osakidetza) provides universal and comprehensive healthcare services (primary, specialized and mental care), free of charge, to all residents. Primary care is provided in group practices, organized in 132 primary care units (with 1,500 registered patients on average per general practitioner) dispersed across the region both in urban and rural areas. Each citizen is registered with a GP, who plays a gate-keeping role regarding access to the rest of the system healthcare services (except for emergency services, which can be directly accessed by patients).

Resources at the disposal of PPCs in the accomplishment of their mission to maintain and improve health status of the population under their responsibility can be divided into four main typologies: labour (health and non-health professionals), pharmaceuticals and other medical products, infrastructure and technology. Assuming that in the public Basque health sector available infrastructure and medical technology are similar across PCCs and that there are few non-health professionals directly involved in patient's care at primary care level, we propose using the following input variables in our empirical study: number of GPs, number of nurses and number of prescriptions per primary care unit which correspond to those most commonly used in the literature (Amado and Dyson, 2008).

Output indicators have been selected following the criterion that they reflect the role of primary care, which directs towards health promotion and education, disease prevention, early diagnosis and timely treatment. Primary care also plays a key role in ensuring the comprehensiveness and continuity of a patient's care, being a patient's first

point of contact with the healthcare system and its guide throughout the different healthcare settings and services. According to this, two output indicators were selected: hospitalisation rates for ACSCs and the rates of re-admissions to hospital 30 days after discharge. Both these indicators are solidly supported by the literature (as seen in section 2) and represent undesirable outputs.

The selection of external variables has been based on available data about several characteristics of population served by PCCs that can have influence on their outcomes. In particular, we use the percentage of population above 65 years old together with a morbidity rate and a deprivation index. The morbidity rate is based on a case-mix system developed by Starfield *et al.* (1991), which is employed to classify annually inhabitants in the Basque Country following an Adjusted Clinical Groups (ACGs) system. This instrument enables the identification of health problems from diagnoses and categorizes prescriptions and individuals into a hundred groups according to their healthcare needs and its costs. In particular, the morbidity rate is defined as the ratio between the observed number of visits and the expected number of visits for each center. The deprivation index is a composed indicator constructed from variables such as the percentage of manual workers, the unemployment and temporary employment rates and low levels of educational attainment for the whole population and also for young people (inhabitants between 16–29 years of age) (Domínguez *et al.*, 2008). This indicator was defined as an ordered measure with five different categories (1-5).

All these variables have been adjusted by 10,000 inhabitants in order to avoid potential distortions due to the existence of significant differences in the size of the PCCs. Likewise, the original values of the (undesirable) output variables have been transformed using the method proposed by Seiford and Zhu (2002). According to this model, the values of both variables are multiplied by -1 and subsequently, we subtracted the value obtained from a large enough parameter, which in our case has been determined by the value K = 1000. Table 1 provides a brief summary of the main descriptive statistics of variables used in the analysis.

Table1. Descriptive statistics of inputs, outputs and exogenous variables

Variables	Role	Mean	SD	Min	Max
Number of GPs	Input	6.53	2.69	3.78	33.08
Number of nurses	Input	6.32	1.26	3.78	12.52
Number of prescriptions	Input	16.65	4.90	6.21	55.52
Hospitalizations due to ACSC	Output	389.95	41.43	213.00	474.22
Re-hospitalizations in 30 days	Output	432.68	23.85	307.30	487.39
Percentage of population above 65	Exogenous	19.55	3.99	4.31	29.24
Morbidity rate	Exogenous	1.03	0.13	0.65	1.38
Deprivation index	Exogenous	3.07	1.25	1.00	5.00

Source: Osakidetza dataset

5. RESULTS

The results of the efficiency estimations for both unconditional and conditional models using VRS and an output orientation are summarized in Table 2. If we focus on the performance of PCC without controlling for exogenous variables (unconditional model), we notice that the mean efficiency score is 1.116, with only 10 PCCs being considered as efficient. Therefore, one could think that inefficient units can still improve their performers almost 12% on average to achieve the efficiency levels of the best practices. However, this initial assessment does not take into account the characteristic of the population served by each center, so those initial efficiency scores could not represent adequately their level of efficiency.

Table 2. Efficiency estimates and bandwidths

	Min	Median	Mean	Max	SD	Efficient units
Unconditional eff.	1.0000	1.1157	1.1160	1.5196	0.0765	10 (7.5%)
Conditional eff.	1.0000	1.0000	1.0251	1.3246	0.0532	74 (56%)
Bandwidth Dep. index	0.0000	0.3171	0.3631	0.7882	0.1897	-
Bandwidth % > 65 years	0.1269	2.9792	8.59E+06	1.84E+08	2.73E+07	-
Bandwidth Morb. index	0.0105	0.1501	3.01E+05	6.25E+06	1.02E+06	-

Once we include the three exogenous indicators (two continuous and one discrete) in the estimation of the conditional efficiency model, more than a half of units become efficient. This result derives from the fact that now each unit is only compared with those operating in a similar environment, so the reference set is smaller. Likewise, the margin for improvement is reduced notably, since the average level of efficiency decreases to 1.0251.

Table 2 also includes the summary statistics for bandwidths estimates, which present a reasonable average value for the discrete variable (deprivation index) and very high average and maximum values for the two continuous variables. At this point, it is worth noting that those high values can be explained due to the some outlying maximum scores. Nevertheless, those values do not imply that the influence of the variables is not significant as we demonstrate above.

In order to test the influence of the exogenous variable, we regress the ratio between conditioned and unconditioned efficiency scores on the environmental variables using the local linear estimator described in section 3.2. Table 3 presents the p-values of the significance tests proposed by Li and Racine (2004) and Racine and Li (2004), which allow us to detect that all the variables have a significant impact on PCC´s performance, although the level of significance is slightly low for the deprivation index.

Table 3. Nonparametric significance test

Variable	p-value
Deprivation index	0.084*
Percentage of people over 65 years	< 2E-16***
Morbidity index	0.035**
R^2	0.309

** denotes statistical significance at 5%
 * denotes statistical significance at 10%

As we are interested in the identification of the effect of the exogenous variables, we analyze the values of the ratio against those Z variables. Hence, following the principles established by Daraio and Simar (2005, 2007), firstly we observe the partial regression scatter plots for both continuous variables, on their median value and, respectively, on

their first and third quartile to capture the heterogeneity among units (Figure 1). Since we are examining an output oriented case, a decreasing regression line indicates that the environmental variable is unfavorable to PCCs´ efficiency. This evidence confirms the result we expected and is in line with some previous studies (see Cordero *et al.*, 2013), since the performance of units is frequently worse when the levels of morbidity and the percentage of old population (proxies for a bad health condition) are higher.

Figure 1. Effects of continuous Z variables on PCCs´ efficiency scores
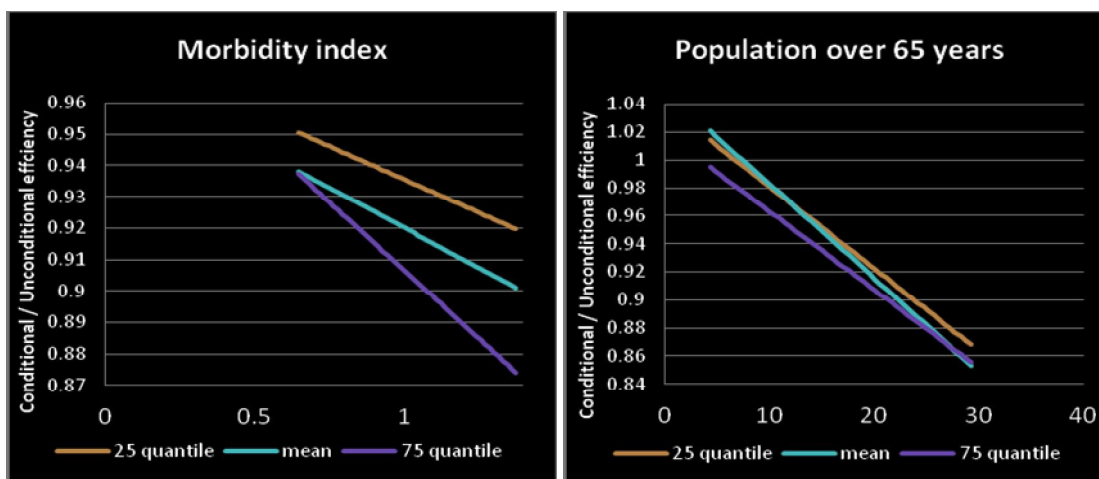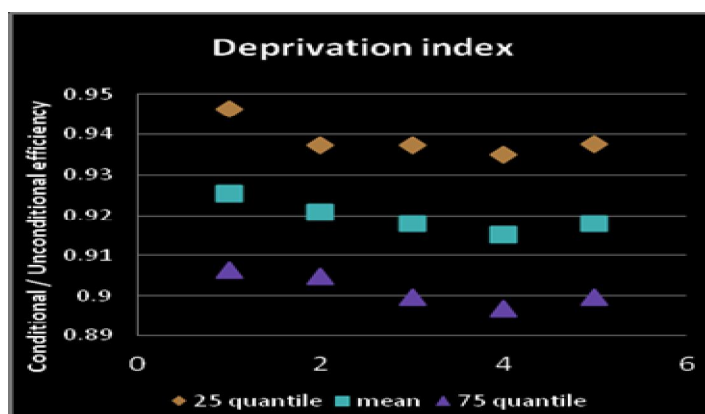

Figure 2 also illustrates the partial regression plot for the discrete ordered variable (Deprivation index). In this case, the result observed is similar, since the variable has an unfavorable effect on the level of efficiency, although there is an unexpected favorable increase for the highest level (5). This can be explained by the existence of a low number of units with these values, which implies that most of them are considered as efficient in the conditional model, where the reference set has a reduced size.

Figure 2. Effect of the discrete variable on PCCs´ efficiency score

6. CONCLUSIONS

This paper uses a recently developed conditional nonparametric approach to estimate efficiency measures for a set of Spanish primary care units incorporating the effect of different types of environmental factors (continuous and discrete) representing the characteristics of the population served by those providers. This method allows us to avoid the restrictive separability assumption between the input-output space and the space of external factors environmental variables required by traditional approaches in order to provide meaningful results. In addition, this methodology makes it possible to determine the statistical significance and the direction of the effect of those exogenous variables. To the best of our knowledge, this is the first empirical study using this method in the health sector.

Moreover, the variables selected to represent the outcomes of the evaluated units can be interpreted as undesirable outputs according to the structure of the production process, thus we have had to use an extension of the traditional DEA model to transform the original values of those variables and estimate valid measures of performance for the evaluated units.

The empirical results show that all the environmental variables considered has a significant and negative effect on the performance of primary health care providers, which is in line with the results obtained in the scarce previous literature using traditional semi-parametric approaches (Cordero *et al.*, 2013). At this point, it is worth noting that in this paper we have only focused on exploring the potential influence of environmental variables related to the characteristics of the population served by primary care providers, but there are other potential factors that can explain differences in ACSC hospitalisation rates such as different quality of specialist care or different admission policies between hospitals (Muecke, 2010).

References

Amado, C.A. and Dyson, R. (2008): On comparing the performance of primary care providers, *European Journal of Operational Research*, 185, 915-932.

Amado, C.A. and Santos, S. (2009): Challenges for performance assessment and improvement in primary health care: The case of the Portuguese health centres, *Health Policy*, 91, 43-56.

Badin, L., Daraio, C. and Simar, L. (2010): Optimal bandwidth selection for conditional efficiency measures: A data-driven approach, *European Journal of Operational Research*, 201, 633-640.

Badin, L., Daraio, C. and Simar, L. (2012): How to measure the impact of environmental factors in a nonparametric production model?,*European Journal of Operational Research,* 223, 818-833.

Banker, R, D., Charnes, A. and Cooper, W. W. (1984): Models for estimating technical and scale efficiencies in data envelopment analysis, *Management Science*, 30(9), 1078-92.

Banker, R. D. and Natarajan, R. (2008): Evaluating contextual variables affecting productivity using data envelopment analysis, *Operations Research*, 56(1), 48–58.

Bardsley, M., Steventon, A., Smith, J. and Dixon, J. (2013): Evaluating integrated and community-based care: How do we know what works? Research summary, London: Nuffield Trust.

Bates J, Baines D, and Whynes, D. (1996): Measuring the efficiency of prescribing by general practitioners, *Journal of the Operational Research Society*, 47, 1443–51.

Blass, R., Da Silva, G.and Miranda, B. (2008): A Probabilistic Approach for Assessingthe Significance of Contextual Variables in Nonparametric Frontier Models: An Application for Brazilian Banks,*Brazilian Review of Econometrics*, 28(1), 111-125.

Bonaccorsi, A., Daraioc C. and L. Simar (2006): Advanced Indicators of Productivity ofUniversities, An Application of Robust Nonparametric Methods to Italian Data, *Scientometrics*, 66 (2), 389-410.

Caminal, J., Starfield, B., Sánchez, E., Casanova, C. and Morales, M. (2004): The role of primary care in preventing ambulatory care sensitive conditions, *European Journal of Public Health*, 14(3), 246–51.

Campbell, S.M., Roland, M.O. and Buetow, S.A. (2000): Defining quality of care, *Social Science & Medicine*, 51, 1611-1625.

Carvalho, P. and Marques, R.C. (2011): The influence of the operational environment on the efficiency of water utilities, *Journal of Environmental Management*, 92 (10), 2698-2707.

Cazals, C., Florens, J.P.andSimar, L. (2002): Nonparametric frontier estimation: A robust approach, Journal of Econometrics, 106, 1–25.

Cazals, C., Dudley, P., Florens, J.P., Patel, S. and Rodriguez, F. (2008): Delivery OfficesCost Frontier: A Robust Non Parametric Approach with Exogenous Variables, *TheReview of Network Economics*, 7 (2), 294-308.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978): Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 2 (6), 429-444.

Cheng, G., Zervopoulos, P. and Qian, Z. (2013): A variant of radial measure capable of dealing with negative inputs and outputs in data envelopment analysis, *European Journal of Operational Research*, 225, 100-105.

Cherchye, L., De Witte, K. Ooghe, E. andNicaise, I. (2010): Efficiency and equity in private and public education: A nonparametric comparison, *European Journal of Operational Research*, 202, 563-573.

Chilingerian, J. A. (1995): Evaluating physician efficiency in hospitals: a multivariate analysis of best practices, *European Journal of Operational Research*, 80, 548–574.

Chilingerian, J. and H. Sherman (1996): Benchmarking physician practice patterns with DEA: A multi-stage approach for cost containment,*Annals of Operations Research*, 67, 83-116.

Chung, Y.H., Fare, R. and Grosskopf, S. (1997): Productivity and undesirable outputs:a directional distance function approach, *Journal of Environmental Management*, 51, 229-240.

Cordero, J.M., Pedraja, F. and Salinas, J. (2008): Measuring Efficiency in Education: an Analysis of Different Approaches for Incorporating Non-Discretionary Inputs, *Applied Economics*, 40 (10), 1323-1339.

Cordero, J.M., Crespo, E. y Murillo, L.R. (2010): Measuring Efficiency in Primary Health Care: The Effect of Exogenous Variables on Results, *Journal of Medical Systems*, 35, 545-554.

Cordero, J.M., Crespo, E. y Murillo, L.R. (2013): The effect of quality and socio-demographic variables on efficiency measures in primary health care, *The European Journal of Health Economics, in press*.

Curry, N. and Ham, C. (2010): *Clinical and service integration: The route to improved outcomes*, London: The King´s Fund.

Daouia, A. and Simar, L. (2007): Nonparametric Efficiency Analysis: A MultivariateConditionalQuantileApproach, *Journal of Econometrics*, 140, 375-400.

Daraio, C.andSimar, L. (2005): Introducing environmental variables in nonparametric frontier models: A probabilistic approach, *Journal of Productivity Analysis*, 24 (1), 93–121.

Daraio, C.andSimar, L. (2006):A robust nonparametric approach to evaluate and explain the performance of mutual funds, *European Journal of Operational Research,* 175 (1), 516–542.

Daraio, C. and Simar, L. (2007a): *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*, Springer, New York.

Daraio, C. and Simar, L. (2007b): Conditional nonparametric Frontier models for convex and non convex technologies: A unifying approach, *Journal of Productivity Analysis*, 28, 13–32.

De Witte, K. and Geys, B. (2011): Evaluating efficient public good provision: Theory and evidence from generalized conditional efficiency model for public libraries, *Journal of Urban Economics*, 69 (3), 319-327.

De Witte, K. and Geys, B. (2013): Citizen co-production and efficient public good provision: Theory and evidence from local public libraries, *European Journal of Operational Research*, 224 (3), 592-602.

De Witte, K. and Kortelainen, M. (2013): What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables, *Applied Economics*, 45, 2401-2412.

De Witte, K. and Marques, R.C. (2010a): Incorporating heterogeneity in non-parametric models: a methodological comparison, *International Journal of Operational Research*, 9 (2), 188-204.

De Witte, K. and Marques, R.C. (2010b): Designing performance incentives, an international benchmark study in the water sector, *Central European Journal of Operations Research*, 18 (2), 189-220.

De Witte, K. and Rogge, N. (2011): Accounting for exogenous influences in performance evaluations of teachers, *Economics of Education Review*, 30, 641-653.

De Witte, K., Rogge, N., Cherchye, L. and van Puyenbroeck, T. (2013): Economies of scope in research and teaching: A non-parametric investigation, *Omega*, 41, 305-314.

Domínguez, M.F., Borrell, C., Cano, G., Esnaola, S., Nolasco, A., Pasarín, M.I., Ramis, R., Saurina, C. and Escolar, A. (2008): Construcción de un índice de privación a partir de datos censales en grandes ciudades españolas: (Proyecto MEDEA), *Gaceta Sanitaria*, 22:179–187.

Donabedian, A. (1980): *The Definition of Quality and Approaches to its Assessment*, vol. 1, Health Administration Press, Ann Arbor, Michigan.

Dyckhoff, H. and Allen, K. (2001): Measuring ecological efficiencywith data envelopment analysis (DEA),*European Journal of Operational Research*, 132,312–325.

Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V., Sarrico, C.S. and Shale, E.A. (2001): Pitfalls and protocols in DEA, *European Journal of Operational Research*, 132,245-259.

Emrouznejad, A., Parker, B.R. and Tavares, G. (2008): Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA, *Socio-Economic Planning Sciences*, 42 (3), 151-157.

European Commission (2010): Joint Report on Health Systems. European Economy, Occassional Papers 74.

Färe, R., Grosskopf, S., Lovell, C.A.K. and Pasurka, C. (1989): Multilateral productivity comparisons when some outputs are undesirable: a nonparametricapproach, *The Review of Economics and Statistics*, 71, 90–98.

Färe, R.andGrosskopf, S. (2004): Modeling undesirable factors inefficiency evaluation: Comment, *European Journal of Operational Research*, 157 (1), 242–245.

Farrell, M.J. (1957): The measurement of productive efficiency, *Journal of the Royal Statistics Society, Serie A*, 120 (3), 253-281.

Finegan, M. S., Gao, J., Pasquale, D. and Campbell, J. (2010): Trends and geographic variation of potentially avoidable hospitalizations in the veteran's health-care system, *Health Service Management Research*, 23(2):66-75.

Fried, H., Lovell, C.A.K. and Schmidt, S. (2008): The *measurement of Productive Efficiency and Productivity Growth,* Oxford University Press.

Fried, H., Schmidt, S. and Yaisawarng, S. (1999): Incorporating the Operating Environment into a Nonparametric Measure of Technical Efficiency, *Journal of Productivity Analysis*, 12, 249-267.

Fried, H.O., Lovell, C.A.K., Schmidt, S.S. andYaisawarng, S. (2002): Accounting for environmental effectsand statistical noise in data envelopment analysis,*Journal of Productivity Analysis*, 17, 157–174.

García, F. Marcuello, C. Serrano, D. and O. Urbina (1999): Evaluation of efficiency in primary healt care centres: An application of data envelopment analysis, *Financial Accountability and Management*, 15(1), 67-83.

Gibson, O.R., Segal, L. and McDermott, R.A. (2013): A systematic review of evidence on the association between hospitalization for chronic disease related ambulatory care sensitive conditions and primary health care resourcing, *BMC Health Services Research*, 13(1):336.

Giuffrida, A., Gravelle, H. and Roland, M. (1999): Measuring quality of care with routine data: avoiding confusion between performance indicators and health outcomes, *BMJ*, 319, 94-98.

Goñi, S. (1999): An analysis of the effectiveness of the Spanish primary health care teams, *Health Policy*, 48, 107-117.

Haelermans, C. and De Witte, K. (2012): The role of innovations in secondary school performance: Evidence from a conditional efficiency model, *European Journal of Operational Research*, 223, 541-549.

Hall, P., J.S. Racine and Li, Q. (2004): Cross-validation and the estimation of conditional probability densities, *Journal of the American Statistical Association*, 99 (468), 1015-1026.

Halkos, G. and Tzeremes, N. (2011a): Modelling regional welfare efficiency applying conditional full frontiers, *Spatial Economic Analysis*, 6(4), 451-471

Halkos, G. and Tzeremes, N. (2011b): A conditional nonparametric analysis for measuring the efficiency of regional public healthcare delivery: An application to Greek prefectures, *Health Policy*, 103(1), 73-82.

Halkos, G. and Tzeremes, N. (2013): A conditional directional distance function approach for measuring regional environmental efficiency: Evidence from UK regions, *European Journal of Operational Research*, 227, 182-189.

Ham, C. (2010): The ten characteristics of the high-performing chronic care system, *Health economics, policy, and law*, 5(1), 71-90.

Hollingsworth, B. (2008): The measurement of efficiency and productivity of health care delivery, *Health Economics*, 17, 1107-1128.

Hu, H., Qi, Q. and Yang, C.H. (2012): Evaluation of China´s regional hospital efficiency: DEA approach with undesirable output, *Journal of the Operational Research Society*, 63, 715-725.

Hua, Z. and Bian, Y. (2007): DEA with undesirable factors, in Zhu, J. and Cook, W.D. (eds): *Modeling Data Irregularities and StructuralComplexities in Data Envelopment Analysis*, Springer Science, Boston, MA, 103–121.

Hua, Z., Bian, Y. and Liang, L. (2007): Eco-efficiency analysis of paper mills along the HuaiRiver:Anextended DEA approach, *Omega*, 35, 578-587.

Huang, Y-G. L. and McLaughlin, C. P. (1989): Relative efficiency in rural primary health care: an application of data envelopment analysis, *Health Services Research*, 24, 143-158.

Koopmans, T.C. (1951): An analysis of production as an efficient combination of activities*, in Koopmans, T.C. (Ed.): *Activity Analysis of Production and*

*Allocation,*Cowles Commission for Research in Economics*, Wiley, New York Monograph, 13.*

Kontodimopoulos, N., Papathanasiou, N., Tountas, Y. and Niakas, D. (2010): Separating Managerial Inefficiency from Influences of the Operating Environment: An Application in Dialysis, *Journal of Medical Systems*, 34 (3), 397-405.

Korhonen, P. and Luptacik, M. (2004): Eco-efficiency analysis of power plants: Anextension of data envelopment analysis, *European Journal of Operational Research*, 154, 437–446.

Kringos, D.S., Boerma, W., Van der Zee, J. and Groenewegen, P (2013): Europe's Strong Primary Care Systems Are Linked To Better Population Health But Also To Higher Health Spending, *Health Affairs*, 32 (4), 686-694.

Lezzoni, L. (1994): *Risk Adjustment for Measuring Health Care Outcomes*, Ann Arbor, Michigan: Heath Administration Press.

Li, Q. and Racine, J.S. (2004): Cross-Validated Local Linear Nonparametric Regression, *Statistica Sinica*, 14(2), 485-512.

Li, Q. and Racine, J. (2007): *Nonparametric econometrics: theory and practice.* Princeton: Princeton UniversityPress.

Lovell, C.A.K. and Pastor, J. (1995): Units invariant and translation invariant DEA models, *Operations Research Letters*, 18, 147–151.

Lovell, C. A. K., Pastor, J. and Turner, J. (1995): Measuring Macroeconomic Performance in the OECD: A Comparison of European and Non-European Countries, *European Journal of Operational Research*, 87, 507-518.

Lu, W.M. and Lo, S.F. (2007): A closer look at the economicenvironmentaldisparities for regional development in China. *European Journal of Operational Research*, 183, 882–894.

McDonald, J. (2008): Using least squares and tobit in second stage DEA efficiencyanalyses,*European Journal of Operational Research*, 197 (2), 792-798.

Muecke, S. (2010): *Avoiding hospitalisation: ambulatory care sensitive conditions,* Research Roundup, 12.

Muñiz, M. (2002): Separating Managerial Inefficiency and External Conditions in Data, *European Journal of Operational Research*, 143-3, 625-643.

Nunamaker, T.R. (1983): Measuring routine nursing service efficiency: A comparison of cost per patient day and data envelopment analysis models, Health Services Research, 18: 183-205.

Nuño, R., Coleman K., Bengoa R. and Sauto R (2012): Integrated care for chronic conditions: the contribution of the ICCC Framework, *Health Policy*, 105(1):55-64.

Pelone, F., Kringos, D. S., Valerio, L., Romaniello, A., Lazzari, A., Ricciardi, W. and Giulio de Belvis, A. (2012): The measurement of relative efficiency of general practice and the implications for policy makers, *Health policy*, 107(2-3), 258–68.

Pina, V. and L. Torres (1992): Evaluating the efficiency of non-profit organizations: An application of data envelopment analysis to the public health services, Financial Accountability and Management, 8, 213-225.

Podinovski, V. and Kuosmanen, T. (2011): Modelling weak disposability in data envelopment analysis under relaxedconvexity assumptions, *European Journal of Operational Research*, 211, 577-585.

Racine, J.S. (1997): Consistent significance testing for nonparametric regression, *Journal of Business and Economic Statistics*, 15, 369–379.

Racine, J.S. and Li, Q. (2004): Nonparametric Estimation of Regression Functions with both categorical and continuous Data, *Journal of Econometrics*, 119 (1), 99-130.

Racine, J.S., Hart, J.D. and Li, Q. (2006): Testing the significance of categorical predictor variables in nonparametric regression models, *Econometric Reviews*, 25, 523–544.

Rogge, N. and De Jaeger, S. (2012): Measuring and explaining the cost efficiency of municipal solid waste collection and processing services, *Omega*, 41 (4), 653-664.

Rosano, A., Abo Loha, C., Falvo, R., Van der Zee, J., Ricciardi, W., Guasticchi, G. and De Belvis, A.G. (2012): The relationship between avoidable hospitalization and accessibility to primary care: a systematic review, *European Journal of Public Health*, 23(3):356-60.

Rosenman, R. and D. Friesner (2004): Scope and scale efficiencies in physician practices, *Health Economics*, 13: 1091-1116.

Sahoo, B.K., Luptacik, M. and Mahlberg, B. (2011): Alternative measures of environmental technology structure in DEA: An application, *European Journal of Operational Research*, 215, 750-762.

Salinas-Jiménez, J. and Smith, P.C. (1996): Data envelopment analysis applied to quality in primary health care, *Annals of Operations Research*, 67, 141-161.

Scheel, H. (2001): Undesirable outputs in efficiency valuations, *European Journal of Operational Research*, 132, 400-410.

Schiøtz, M., Price, M., Frølich, A., Søgaard, J,.Kristensen, J.K., Krasnik, A., Ross, M.N., Diderichsen, F. and Hsu, J. (2011): Something is amiss in Denmark: a comparison of preventable hospitalisations and readmissions for chronic medical

conditions in the Danish Healthcare system and Kaiser Permanente, *BMC Health Services Research,* 11:347.

Seiford, L. M., and Zhu, J. (2002): Modeling undesirable factors in efficiency evaluation,*European Journal of Operational Research*, 142, 16–20.

Silva Portela, M.C.A., Thanassoulis, E. and Simpson, G. (2004): Negative data in DEA: a directional distance function approach applied to bank branches, *Journal of the Operational Research Society*, 55, 1111-1121.

Simar, L. and Wilson, P. W. (2007): Estimation and inference in two-stage, semi-parametric models of productionprocesses, Journal of Econometrics, 136(1), 31–64.

Simar, L. and Wilson, P. W. (2011): Two-stage DEA: caveat emptor,*Journal of Productivity Analysis*, 36,205–218.

Starfield B., Weiner J., Mumford, L. and Steinwachs, D. (1991): Ambulatory care groups: a categorization of diagnoses for research and management, *Health Services Research*, 26:53-74.

Steinmann, L., Dittrich, G., Karmann, A. and Zweifel, P. (2004): Measuring and comparing the (in)efficiency of German and Swiss hospitals, *The European Journal of Health Economics*, 5 (3), 216-226.

Szczepura, A., Davies, C. and Fletcher, J. (1993): Efficiency and Effectiveness in General Practice, *Journal of Management in Medicine*, 7 (5), 36-47.

Verschelde, M. and Rogge, N. (2012): An environment-adjusted evaluation of citizen satisfaction with local police effectiveness: Evidence from a conditional Data Envelopment Analysis approach, *European Journal of Operational Research*, 223, 214-215.

Vidoli, F. (2011): Evaluating the water sector in Italy through a two-stage method using the conditional robust nonparametric frontier and multivariate adaptive regression splines, *European Journal of Operational Research*, 212 (3), 583-595.

Worthington, A.C. (2004): Frontier Efficiency Measurement in Health Care: A Review of Empirical Techniques and Selected Applications, *Medical Care Research and Review*, 61 (2), 135-170.

Yang, C., Hsiao, C. and Yu, M. (2008): Technical efficiency and impact of environmental regulations in farrow-to-finish swine production in Taiwan, *Agricultural Economics*, 39, 51–61.

Yang, H. and Pollitt, M. (2009): Incorporating both undesirable outputs and uncontrollable variables into DEA:The performance of Chinese coal-fired power plants, *European Journal of Operational Research*, 197, 1095-1105.

You, S. and Yan, H. (2011): A new approach in modelling undesirable outputin DEA model, *Journal of the Operational Research Society*, 62, 2146-2156.

Zavras, A., Tsakos, G. Economou, C. and Kyriopoulos, J. (2002): Using DEA to Evaluate Efficiency and Formulate Policy Within a Greek National Primary Health Care Network, *Journal of Medical Systems*, 26 (4), 285-292.

Zhou, P., Ang, B.W. and Poh, K.L. (2008): A survey of data envelopment analysis in energy and environmental studies, *European Journal of Operational Research*, 189 (1), 1-18.