Technical efficiency in primary health care: does quality matter?

Murillo-Zamorano, Luis R. and Petraglia, C.

September 2008

TECHNICAL EFFICIENCY IN PRIMARY HEALTH CARE: DOES QUALITY MATTER? *

Luis R. Murillo-Zamorano§
Department of Economics, University of Extremadura

Carmelo Petraglia
Department of Economics, University Federico II di Napoli

Abstract

The accuracy required in the measurement of output is an issue that has as yet still not been satisfactorily addressed in empirical research on efficiency in primary health care. We exploit information retrieved from a newly constructed database (APEX06) for the Spanish region of Extremadura. The richness of our dataset allows us to consider original synthetic measures of output that take into account both the quantity and the quality of services provided by 85 primary care centres (PCCs) in 2006. We provide evidence that neglecting the issue of properly accounting for the quality of health services can lead to misleading results. Our main finding is that adjusting output for quality influences efficiency analysis in three senses. First, inefficiency now explains relatively more of the deviation from the potential output. Second, the average technical efficiency in the sector is lower, while its dispersion among PCCs is significantly higher. And third, the efficiency ranking of the PCCs is also affected.

Keywords: Primary Health Care, Stochastic Frontier Analysis, Technical Efficiency, Quality.
JEL-codes: I12, C10.

---

* An earlier version of this paper "Efficiency in Primary Health Care" was presented at the conference "XXVIII Jornadas de Economía de la Salud" (Salamanca, 28-30 May 2008). We are grateful to the participants in the session "Análisis económico de la sanidad (II)" for useful comments. We also thank Marcelino Sánchez Rivero for his helpful comments and to the Consejería de Sanidad y Dependencia de la Junta de Extremadura and Programa "José Castillejo" del Ministerio de Ciencia e Innovación for its financial support. The usual disclaimer applies.

§ *Corresponding author*: Department of Economics, University of Extremadura, Avenida de Elvas s/n, 06071 Badajoz, Spain; Phone: 00 34 924289520. Fax: 00 34 924272509. email: lmurillo@unex.es

1. Introduction

According to the activity analysis developed in Koopmans (1951), a producer is said to be technically inefficient when it can produce the same amount of output with less than at least one input, or can use the same package of inputs to produce more than at least one output. This definition establishes the twofold orientation –output augmentation and input reduction– of the technical component of economic efficiency. Either makes technical efficiency a very attractive concept with which to investigate a productive sector as sensitive to demand as health care. Indeed, since the seminal study of Huang and McLaughlin (1989), there has been a numerous and wide ranging collection of papers devoted to the measurement of technical efficiency in health care.

Even though the initial contribution of Huang and McLaughlin (1989) was concerned with the evaluation of relative efficiency in primary health care, most subsequent investigations focused on hospitals, with primary health care receiving far less attention. As Amado and Dyson (2008) point out, while a hospital is an organization with clear boundaries, where patients are admitted and discharged, primary health care delivery can be thought of as an open community-based system with unclear boundaries. This introduces greater complexity when it comes to the economic modeling of the sector, especially with respect to the appropriate definition of primary care providers' output.

By definition, any measure of primary care output should capture the impact of the services on the current and future health status of patients. Unfortunately, the lack of adequate information on this aspect means that observation of the causality connection between the provision of health services and the health status of the population served is far from straightforward. This has led scholars to the adoption of so-called "activity-oriented" models where the primary care output is proxied by activity levels of the health care units under analysis, i.e., by the number of registered visits or consultations. Examples of this approach can be found in Pina and Torres (1992), Chilingerian and Sherman (1996, 1997), Ozcan (1998), and Goñi (1999).

However, as Amado and Dyson (2008) and Puig-Junoy and Ortún (2004) indicate, the use of such quantity proxies is subject to major criticisms. First, the number of visits or consultations provided by a given health centre is likely to be affected by factors that are beyond their control, for instance, the socio-demographic characteristics of the population it serves. Second, there is no clear relationship between the visits, which are largely determined by the physician, and the quality of primary care. Third, how much visits contribute to health improvement depends on their effectiveness. Fourth, it assumes that all the patients accessing health care centres are receiving the necessary services. And fifth, it assumes that the services provided during the visits are appropriate and of similar quality.

These criticisms make it necessary to consider not only quantitative but also qualitative indicators in order to provide an adequate measure of primary care output. However, although it is a well recognized problem, the accuracy required in the measurement of output is an issue that has as yet still not been satisfactorily addressed in empirical research on efficiency in primary health care. Hence, the main objective of the present study was to shed some light on this controversial and challenging issue by means of assessing whether adjusting output for quality affects the measurement of technical efficiency in the primary health care sector, and if so in which direction.

In practice, the inclusion of output quality information inolves the consideration of multi-output technologies. This introduces further difficulty into the analysis, and indeed there have been only a few studies taking this line, such as Salinas-Jiménez and Smith (1996), García *et al*. (1999), and, more recently, Rosenman and Friesner (2004). All of them calculate technical efficiency by means of Data Envelopment Analysis (DEA), i.e., by means of a non-parametric and deterministic model. None, however, have attempted to implement a parametric and stochastic approach so as to take advantage of the inherent characteristics of models of this type.

Deterministic models, such as those obtained from DEA, envelop all the observations, and identify technical inefficiency as being the distance between the observed production and the maximum production as defined by the frontier and the available technology. DEA is well suited to working with multiple output scenarios, but presents the drawback that it does not permit one to distinguish between technical efficiency and statistical noise, so that the entire deviation from the frontier is artificially considered to be inefficiency. Parametric and stochastic models not only deal with this problem, but they allow one to test for the statistical significance of alternative hypotheses and to estimate an appropriate functional form for the productive technology under consideration[1].

The intention with the present work is to contribute to extending the existing literature on technical efficiency in primary health care in three ways. Firstly, and in line with the above theoretical underpinnings, it defines an appropriate measure for primary care output by combining both activity and quality indicators. Secondly, it deals with multiple output technologies by using multivariate data techniques and generating a set of synthetic indices for quality and primary care output. And thirdly, it estimates technical inefficiency by means of a parametric and stochastic frontier production model.

To the best of our knowledge, the present paper describes the first empirical study on technical efficiency in the primary health care sector allowing for the quality adjustment of health output in a parametric frontier production model. The novelty of the analysis was made it possible by the richness of information retrieved from a newly constructed database for the Extremadura Primary Health System (APEX06). As a main result, we provide evidence that not properly adjusting for the quality of health services leads to misleading results in both the average technical efficiency of the sector and the efficiency ranking of health service providers.

The remainder of the paper is organized as follows. Section 2 briefly reviews the techniques employed to measure technical efficiency. Section 3 provides a discussion of the dataset and the variables involved in the estimations. Section 4 presents the results, and Section 5 gives the conclusions.

2. Methods

The measurement of productive efficiency by means of parametric techniques requires the specification of a particular frontier function. This specification can be either deterministic or stochastic. Deterministic models envelop all the observations, identifying the distance between the observed production and the maximum production,

[1] The reader is referred to Murillo-Zamorano (2004) for a detailed discussion of parametric and non-parametric techniques for the measurement of economic efficiency.

defined by the frontier and the available technology, as economic inefficiency. Stochastic approaches, on the other hand, permit one to distinguish between technical inefficiency and statistical noise.

The stochastic frontier production function model originally proposed by Aigner, Lovell and Schmidt (1977), Battese and Corra (1977) and Meeusen and van den Broeck (1977) to account for random errors and non-negative inefficiency effects can be expressed as follows:

$$Y_i = f(x_i;\beta)\exp(v_i - u_i) \tag{1}$$

where Y_i denotes production in the *i*-th cross-section unit; x_i is a $(1 \times k)$ vector of values of known functions of inputs of production and other explanatory variables associated with the *i*-th cross-section unit; and β is a $(k \times 1)$ vector of unknown parameters which are to be estimated. The v_i's are assumed to be iid N $(0,\sigma_V^2)$ random variables, distributed independently of the u_i's.

With respect to the one-sided (inefficiency) error u_i, a number of distributions have been assumed in the literature. The most frequently used are the half-normal and the truncated half-normal distributions. In the former case, the u_i's are assumed to be iid $N(0,\sigma_U^2)$, while in the latter case, the distribution of u_i is obtained by truncation at zero of the normal distribution with mean m_i and variance σ_U^2.

If the two error terms are assumed independent of each other and of the input variables, and one of the two aforementioned distributions is used, then the likelihood functions can be defined and maximum likelihood estimates made. Estimating the model using maximum likelihood techniques, one obtains a fitted value for the composed error term $(v_i - u_i)$. For efficiency measurement, these two error terms need to be separated. Jondrow *et al.* (1982) proposed one way to do this. They derived an explicit formula for the expected value of u_i conditional on the composed error term $(E(u_i|v_i-u_i))$ in the half-normal case:

$$E[u_i|v_i] = \frac{\sigma\lambda}{(1+\lambda^2)}\left[\frac{\phi(v_i\lambda/\sigma)}{\Phi(-v_i\lambda/\sigma)} - \frac{v_i\lambda}{\sigma}\right] \tag{2}$$

where $\phi(.)$ is the density of the standard normal distribution and $\Phi(.)$ the cumulative density function.

For the truncated case, Greene (1993) shows that the conditional technical inefficiencies are obtained by replacing $v_i\lambda/\sigma$ in the expression for the half-normal case, with

$$u_i^* = \frac{v_i\lambda}{\sigma} + \frac{u_i}{\sigma\lambda} \tag{3}$$

Finally, measures of technical efficiency relative to the production frontier in (1) are obtained as

$$TE_i = e^{-E[u_i|v_i]} \tag{4}$$

where TE_i takes a value between zero and one. If TE equals one, no technical inefficiency is observed and all the deviation from the optimum frontier is due to random noise. The greater the value of TE in general, the lower the level of technical inefficiency of the health units under consideration.

3. Data

Our estimations were made for a cross-section of 85 primary care centres (PCCs) located in the Spanish region of Extremadura, and observed in the year 2006. Because of the extension of its territory (41,634 km^2 in area) and its low population density (26.18 inhab/km^2), the Extremadura primary health system is structured around two territorial administrative levels of aggregation: Health Areas and Basic Health Zones. The system is divided into 8 Health Areas (Badajoz, Mérida, Don Benito-Villanueva, Zafra-Llerena, Cáceres, Coria, Plasencia, and Navalmoral de la Mata), each consisting of a number of Basic Health Zones. Covering a total population of 1,081,845 inhabitants, in 2006 there were 104 operating Basic Health Zones, each organized around a PCC as the main provider of primary health care services in the zone. Figure 1 illustrates the distribution of the Basic Health Zones and Health Areas of Extremadura.

Figure 1. Extremadura primary care map.

We retrieved data from APEX06[2], an integrated information system for Primary Care in Extremadura that provides detailed information for each one of the aforementioned Health Areas and Zones on a number of variables, including the population covered, human resources, activity levels, costs, accessibility indicators, quality indicators, and environmental indicators. Table 1 presents descriptive statistics of the variables included in the estimated model, together with their role in the productive process of a PCC.

As was indicated in the Introduction, the most commonly used indicator for health care output is the number of visits or consultations by primary care professionals. For each PCC considered in the study, the variables FREQUENCYGP, FREQUENCYP, FREQUENCYN, and FREQUENCYU indicate the number of visits or consultations per capita[3] with general practitioners (GPs), paediatricians, nurses, and emergency units, respectively. For the inputs the PCCS employ to provide their health care services, we used labour, capital, and prescription data: LAB is the per capita total number of

[2] Vega, Murillo et al. (2007).

[3] When applicable, the variables are expressed in per capita terms to avoid the scale effect associated with the heterogeneous sizes of the Basic Health Zones considered in our study.

(equivalent) personnel, including both medical and non-medical staff, i.e., GPs, pædiatricians, nurses, nursing assistants, emergency GPs, emergency nursing assistants, administrative staff, and porters; CAPITAL is a proxy for the capital structure of the care centre, expressed in terms of per capita PCC area measured in squared meters; and PHARMA is the per capita number of prescriptions.

Table 1. Main descriptive statistics.

Variable	Role	Mean	Std. Dev.	Min.	Max.
Frequencygp	Activity	10.7361	3.4825	0.7970	18.9386
Frequencyp	Activity	6.5765	2.9187	0.9500	17.0800
Frequencyn	Activity	7.7829	3.1290	2.5500	19.1400
Frequencyu	Activity	1.6478	0.6716	0.3948	3.2700
Lab	Input	0.0037	0.0020	0.0007	0.0110
Capital	Input	0.1351	0.0885	0.0133	0.4686
Pharma	Input	21.6089	4.4121	11.6037	32.5179
Dayvisitsgp	Quality1	43. 6103	12.1248	10.6700	70.3100
Dayvisitsp	Quality1	19.4026	10.4988	1.17004	45.2700
Dayvisitsn	Quality1	31.8411	9.5633	14.2700	85.6000
Experience	Quality2	5,244	703.175	2,216	5,741
Hostests	Quality2	0.4674	0.1885	0.0641	0.9749
Healthtarget	Quality2	57.8383	15.4451	35.3100	89.8100
Questions	Quality2	5.4823	1.3505	2	10
Indact	Activity Index	50.2906	29.6064	3.3442	98.8323
Indqua1	Quality1 Index	48.7984	33.0134	1.8895	98.8105
Indqua2	Quality2 Index	50.6329	24.4078	6.1950	93.1340
Indquat	Total Quality Index	50.5186	29.7056	6.2224	95.5539
Indout	Output Index	50.0725	32.1066	6.0309	95. 4037

To measure primary care quality, we used seven variables capturing quality from different points of view. A first set – DAYVISITSGP, DAYVISITSP, and DAYVISITSN – represents the daily number of visits or consultations per GP, pædiatrician, and nurse in each PCC. It is assumed that, *ceteris paribus*, personnel with fewer visits or consultations per day are able to provide better care services, thus increasing the quality of the PCC. The variable EXPERIENCE is a proxy for the experience of GPs and pædiatricians working in each PCC, measured in days of work during the previous 15 years. HOSTESTS represents the number of per capita diagnostic test requests from each PCC to the zone's reference hospital. HEALTHTARGET is another proxy of the quality of care provided by each health centre, designed to capture the extent to which the PCC is able to fulfil certain specific health targets. It is an average of the coverage ratios of each of the programs implemented within the PCC's portfolio of services, expressed in terms (percentage) of the share of effective population served over the potential population to be served. Finally, QUESTIONS indicates the number of affirmative answers to a ten-item questionnaire distributed to the managers of the PCCs. This questionnaire was based on some of the standards considered in the more general model of total quality elaborated for the Extremadura Health Service (Servicio Extremeño de Salud, SES), and was designed to give information on three categories of quality: the medical personnel's continued education, health management skills, and patient satisfaction.

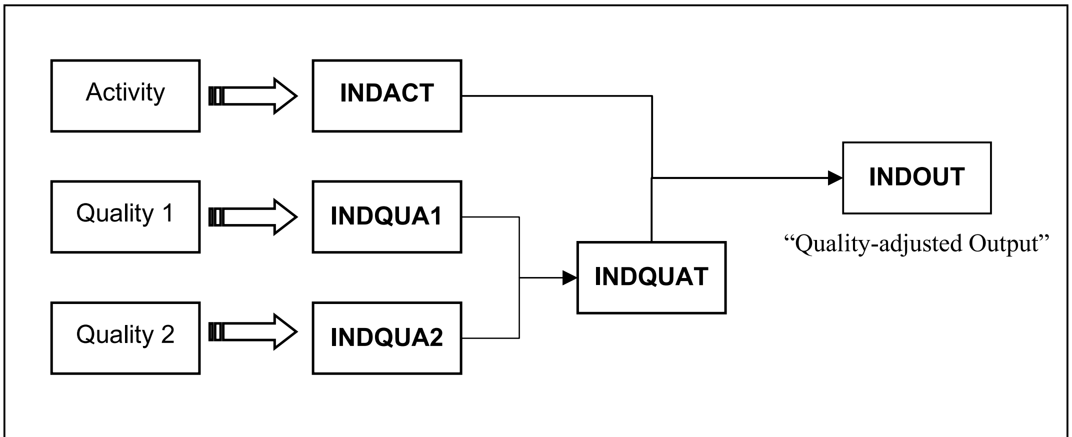
The last group of variables listed in Table 1 are synthetic indices generated by means of the multivariate data technique of principal component analysis (PCA). The goal of PCA is to decompose a data table with correlated measurements into a new set of uncorrelated (i.e., orthogonal) variables[4]. Depending upon the context, these variables are known as principal components, factors, eigenvectors, singular vectors, or loadings. Each factor or principal component is calculated as a linear combination of the standardized values of the original variables used for the definition of the index. The weight given to each of these variables corresponds to its statistical correlation with the latent dimension that the synthetic index attempts to measure.

How many factors to retrieve depends on the correlation of the initial variables. If they are strongly correlated with each other, one factor will be sufficient to explain most of their variance. However, if the correlation is weak, several factors will be required in order to explain a significant percentage of their variance. In this case, one will get a set of intermediate indicators, as many as there were common factors, and the final synthetic index will be calculated as their weighted sum. The importance of each factor is given by the proportion of the total variance explained.

With this methodological approach[5], five synthetic indices were calculated for each PCC of the sample. The relationships among them, the correlation levels [...], and the common factors (F) involved in their definitions are presented in Figure 2.

[4] See Abdi (2003) for a detailed discussion of PCA.
[5] The reader is referred to Jobson (1992) for a detailed explanation of multivariate data techniques.

Figure 2. Synthetic indices and multivariate data analysis.



INDACT = [0.642]*F1 + [0.358]*F2

F1 = [0.890]*FREQUENCYGP + [0.025]*FREQUENCYP + [0.865]*FREQUENCYN + [0.412]*FREQUENCYU
F2 = [0.060]*FREQUENCYGP + [0.938]*FREQUENCYP + [0.170]*FREQUENCYN + [0.410]*FREQUENCYU

INDQUA1 = [0.838]*DAYVISITSGP + [0.779]*DAYVISITSP + [0.593]*DAYVISITSN

INDQUA2 = [0.543]*F1 + [0.457]*F2

F1 = [0.768]*EXPERIENCE + [0.640]*HOSTESTS + [0.011]*HEALTHTARGET + [0.447]*ITEMS
F2 = [0.173]*EXPERIENCE + [0.050]*HOSTESTS + [0.962]*HEALTHTARGET + [0.237]*ITEMS

INDQUAT = [0.761]* (INDQUA1+INDQUA2)

INDOUT = [0.804]* (INDQUAT+INDACT)

INDACT is a quantity-output index that integrates information on each PCC's activity by combining the per capita number of visits or consultations with each of the types of primary care professionals. INDQUA1, INDQUA2, and INDQUAT are synthetic quality indices. INDQUA1 is associated with the daily numbers of visits or consultations per GP, pædiatrician, and nurse. INDQUA2 is asociated with the experience of the medical personnel, the number of diagnostic tests, the coverage ratio of the portfolio of services, and the affirmative answers to the questionnaire sent to the PCC managers. INDQUAT is an overall quality index constructed from INDQUA1 and INDQUA2. Finally, INDOUT is the quality-adjusted measure for the health output of any given PCC constructed by combining the activity index, INDACT, and the overall quality index, INDQUAT.

4. Results

The empirical application described in this section will be considered in three stages. The first deals with the definition of the estimated model and its correct specification by means of determining the set of relevant inputs and control factors to take into account, the functional form for the productive technology, and the distribution to use for the one-sided (inefficiency) error term component. In the second, we shall estimate the frontier production function in two scenarios – with and without adjusting output for quality – and then calculate the individual technical efficiency scores for the PCCs. And the third will present a graphical and comparative analysis of the efficiency scores obtained in the two scenarios.

4.1 Model specification

Since our main objective is to assess the incidence of quality on the measurement of technical efficiency in the primary health care sector, we considered two stochastic frontier productive scenarios. The first is defined by the use of a purely quantitative measure of output (Model 1), while the second adjusts the output to account for quality (Model 2). We will then be able to compare the results in the two scenarios, assessing whether they are sensitive to this quality adjustment of the output. The specifications of Models 1 and 2 are as follows:

Model 1: Half-normal Cobb-Douglas with no quality-adjustment of the output.

$$\ln OUT_i = \alpha + \beta_1 \ln LAB + \beta_2 \ln PHARMA + \beta_3 RURAL + v_i - u_i$$
$$v_i \sim N(0, \sigma_v^2)$$
$$u_i \sim \mathrm{N}^+(0, \sigma_u^2)$$

Model 2: Half-normal Cobb-Douglas with quality-adjusted output.

$$\ln OUTQ_i = \alpha + \beta_1 \ln LAB + \beta_2 \ln PHARMA + \beta_3 RURAL + v_i - u_i$$
$$v_i \sim N(0, \sigma_v^2)$$
$$u_i \sim \mathrm{N}^+(0, \sigma_u^2)$$

where i=1,…85 indexes the PCCs, α is the intercept, β_1, β_2, β_3 are output elasticities, OUT is the total per capita number of GP, pædiatrician, nurse, and emergency visits or consultations, OUTQ the synthetic quality-adjusted output index (INDOUT), LAB the per capita total number of (equivalent) personnel, PHARMA the per capita number of prescriptions, and RURAL is a dummy variable which takes the value of 1 for PCCs located in rural zones and 0 otherwise.

In order to check various alternative hypotheses on the above half-normal Cobb-Douglas models, we carried out a number of generalized log-likelihood ratio (LR) tests.[6] The results are summarized in Table 2. With respect to the functional form, the implementation of an LR test at a 1% significance level suggested that the Cobb-Douglas specification was to be preferred over more flexible technologies, in particular, different versions of the translog production function.[7]

With respect to other regressors to consider in the definition of the frontier production function, we also estimated alternative specifications for Models 1 and 2 by including capital. In both cases, the value of the elasticity of output with respect to capital was not found to be statistically significant, and the null hypothesis of not including capital as a relevant input could not be rejected. Giuffrida (1999) argues that the use of physical capital is quite limited in the provision of primary care, and much less important than in other health services such as hospitals. This view seems to be confirmed with our dataset.[8]

A further test was of the statistical relevance of including RURAL as a dummy variable in the regression, thus controlling for the heterogeneity across PCCs related to their location in either rural or urban zones, which in turn is a proxy of their specific economic and socio-demographic characteristics. As was to be expected, the specification without RURAL could not be accepted for either model.

---

[6] A detailed set of estimates for each of the alternative hypotheses tested is available from the authors on request.

[7] Without the factor share equations, the estimation of these translog functions seems to be hampered by an important problem of multicollinearity. According to *Klein's rule of thumb,* multicollinearity is a problem if max $R_j^2 > R^2$, where R_j^2 is the R^2 statistic from the OLS estimation of the auxiliary regression of the j-th regressor on the other regressor and the intercept term. Several auxiliary regressions were estimated, and this condition was found in all of them.

[8] It is worth mentioning that, on average, capital accounted for only 6% of overall costs in our sample.

Table 2. Generalized likelihood ratio tests of hypothesis for the regressors, functional form, and inefficiency error term distribution of the stochastic frontier production function.

Regressors			
Alternative hypothesis	LR ratio test	Critical value for χ^2	Decision
H_1: Model 1 with Capital	0.44	$\chi^2_{(1)} = 6.63$	Can not reject H_0
H_1: Model 2 with Capital	1.64	$\chi^2_{(1)} =6.63$	Can not reject H_0
H_1: Model 1 with Rural	24.18	$\chi^2_{(1)} =6.63$	Reject H_0
H_1: Model 2 with Rural	15.06	$\chi^2_{(1)} =6.63$	Reject H_0
Functional Form			
Alternative hypothesis	**LR ratio test**	**Critical value for χ^2**	**Decision**
H_1: Translog version of Model 1	0.20	$\chi^2_{(3)} = 11.34$	Can not reject H_0
H_1: Translog version of Model 1 without quadratic terms	0.04	$\chi^2_{(1)} = 6.63$	Can not reject H_0
H_1: Translog version of Model 1 without interaction term	0.10	$\chi^2_{(2)}= 9.21$	Can not reject H_0
H_1: Translog version of Model 2	3.26	$\chi^2_{(3)}= 11.34$	Can not reject H_0
H_1: Translog version of Model 2 without quadratic terms	0.52	$\chi^2_{(1)} = 6.63$	Can not reject H_0
H_1: Translog version of Model 2 without interaction term	2.12	$\chi^2_{(2)} = 9.21$	Can not reject H_0
Inefficiency error term			
Alternative hypothesis	**LR ratio test**	**Critical value for χ^2**	**Decision**
H_1: Truncated-normal in Model 1	4.90	$\chi^2_{(1)}= 6.63$	Can not reject half-normal
H_1: Truncated-normal in Model 2	0.04	$\chi^2_{(1)} = 6.63$	Can not reject half-normal

Finally, we checked the statistical significance of the assumption made concerning the distribution of the inefficiency error component. As in most of the frontier literature, we tested two possible distributions against each other: the half-normal and the truncated half-normal. The assumption of a half-normal distribution could not be rejected for either model. In our sample therefore, the average pure managerial efficiency of the PCCs was not statistically significantly different from zero.

The parameters of the final estimates for Models 1 and 2 that resulted from the hypothesis tests described above are listed in Table 3. The LR test on $\sigma_u = 0$ shows that the null hypothesis that the variance of the inefficiency error term is zero can be rejected in Model 1. This is evidence that there is room for a contribution from technical inefficiency in the sample data, which makes it preferable to estimate a frontier production function rather than an average production function. In other words, this result suggests that it is necessary to assume that the deviations from optimal output are due to both random noise and inefficiency. Indeed, the estimated value of λ – the ratio between the standard errors of the two components of the composed error term – is

2.011, i.e., inefficiency contributes twice as much as random noise to determining the deviations from the potential output.

Table 3. Estimated parameters for the full Cobb-Douglas half-normal model.

	Model 1 DEPENDENT VARIABLE: OUT			**Model 2 DEPENDENT VARIABLE: OUTQ**		
	Coefficient	t-stat	p-value	Coefficient	t-stat	p-value
CONSTANT	2.548	3.65	0.000	3.999	2.34	0.019
LAB	0.130	2.16	0.030	0.441	2.57	0.010
PHARMA	0.491	3.31	0.001	0.744	2.21	0.027
RURAL	0.165	2.73	0.006	0.679	3.99	0.000
Wald χ^2	135.32		0.000	65.02		0.000
Log-likelihood	27.148			-59.201		
LR test $\sigma_u = 0$	6.39		0.006	6.74		0.005
σ_u	0.229			0.814		
σ_v	0.114			0.174		
$\lambda = \sigma_u / \sigma_v$	2.011			4.669		
Observations	85			85		

For Model 2, the evidence for technical inefficiency is again confirmed by the result of the LR test. Moreover, the value of λ (4.669) is now much higher than that estimated in Model 1 . This leads to a first major deduction: that adjusting the output for quality seems to increase the importance that must be assigned to inefficiency in determining the deviations from the frontier.

With respect to the estimation results for Model 1 presented in Table 3, the coefficients for the two inputs and the RURAL dummy variable are overall quite significant, and show the expected positive signs. In particular, the elasticity of labour is estimated to be lower than that of prescriptions. The technological dummy variable RURAL is positive, thus confirming the positive impact played by economic and socio-demographic factors in shifting the productive frontier upward for PCCs located in rural zones. In other words, the specified production frontier differs across zones such that in rural zones the potential output that can be reached by PCCs is higher than in urban zones.

Adjusting the output for quality in Model 2 leads to similar signs and significance of the coefficients. In magnitude, however, all the estimates in Model 2 are greater than in

Model 1. For the inputs, this implies drawing different conclusions regarding the return to scale: decreasing returns in Model 1, increasing returns in Model 2. For the dummy RURAL, the technological gap is now even greater than before, i.e., the upward shift of the frontier for rural PCCs relative to those in urban zones is greater when the output is adjusted for quality.

4.2 Efficiency scores

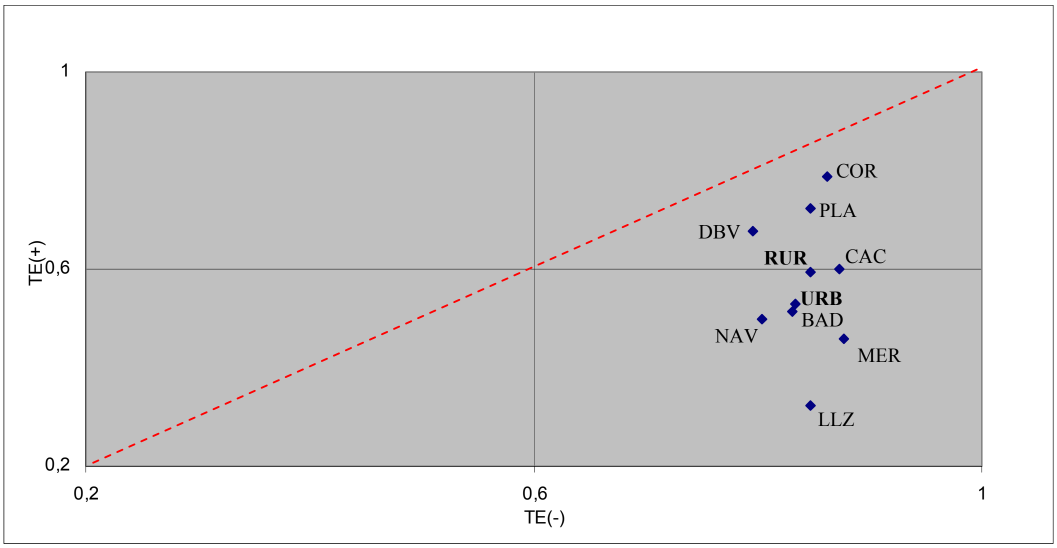
The result described above on the greater contribution of inefficiency in explaining deviations from the frontier when the estimations are performed using the "quality-adjusted" measure of health services will be further explored in this section. One observes in Table 4 that, with the quality adjustment, the average technical efficiency decreases, while the dispersion among the PCCs rises significantly. In particular, the average technical efficiency decreases from 84% to 58%, implying that not adjusting for quality leads to overestimating the efficiency. The dispersion around the average efficiency in the entire sample as measured by the standard deviation increases from 0.0868 to 0.2226, i.e., the variability of efficiency across PCCs is underestimated if output is not adjusted for quality.

Table 4. Average efficiency scores.

	Model 1 DEPENDENT VARIABLE: OUT				**Model 2 DEPENDENT VARIABLE: OUTQ**			
	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.
All (85)	0.8437	0.0868	0.5140	0.9637	0.5786	0.2226	0.1680	0.9455
Rural (65)	0.8471	0.0725	0.5565	0.9637	0.5940	0.2240	0.1680	0.8923
Urban (20)	0.8329	0.1243	0.5140	0.9437	0.5283	0.2162	0.2422	0.9455
Badajoz (15)	0.8314	0.0869	0.5755	0.9437	0.5150	0. 2218	0.2405	0.9455
Mérida (9)	0.8763	0.0337	0.8209	0.9198	0.4591	0.2423	0.1706	0.8266
Don Benito-Villanueva (10)	0.7954	0.1708	0.5140	0.9637	0.6780	0.2377	0.1680	0.8844
Llerena-Zafra (7)	0.8467	0.0300	0.7819	0.8763	0.3244	0.1520	0.1745	0.5567
Cáceres (18)	0.8722	0.0425	0.7954	0.9413	0.6004	0.1609	0.2383	0.8291
Coria (7)	0.8619	0.0725	0.7644	0.9498	0.7892	0.0931	0.6449	0.8782
Plasencia (11)	0.8469	0.0786	0.7198	0.9486	0.7222	0.1467	0.4417	0.8923
Navalmoral (8)	0.8039	0.0936	0.6113	0.8817	0.4990	0.2122	0.2423	0.8341

The same pattern holds when the sample is split into rural and urban zones, or into the eight Health Areas into which the Health Map of Extremadura is organized: adjusting for quality lowers the average efficiency and increases the dispersion. This is also illustrated in Figure 3, in which the vertical axis is the average technical efficiency level including quality adjustment for the output, and the horizontal axis the average technical efficiency level without quality adjustment. Both for the urban and rural zones, and for the eight Health Areas, the observations lie below the 45° bisectrix, indicating a decline in technical efficiency when output is adjusted for quality.

Figure 3. Average technical efficiency with (+) and without (-) adjusting for quality



BAD: Badajoz. MER: Mérida. DBV: Don Benito-Villanueva. LLZ: Llerena-Zafra. CAC: Cáceres. COR: Coria. PLA: Plasencia. NAV: Navalmoral de la Mata. RUR: Rural Health Zones. URB: Urban Health Zones

4.3 Ranking

The next step was to see whether the quality adjustment also affects the rank of the PCCs in terms of technical efficiency. Figure 4 plots the ranks of the ten most efficient PCCs (ranked 1-10) in Model 1 against their respective ranking after adjusting for quality. It is interesting to note that all the PCCs experience a – sometimes substantial – change. With only two exceptions, the common feature is a negative change in the ranking.

Figure 4. Changes in ranking after adjusting for quality for ten most efficient PCCs

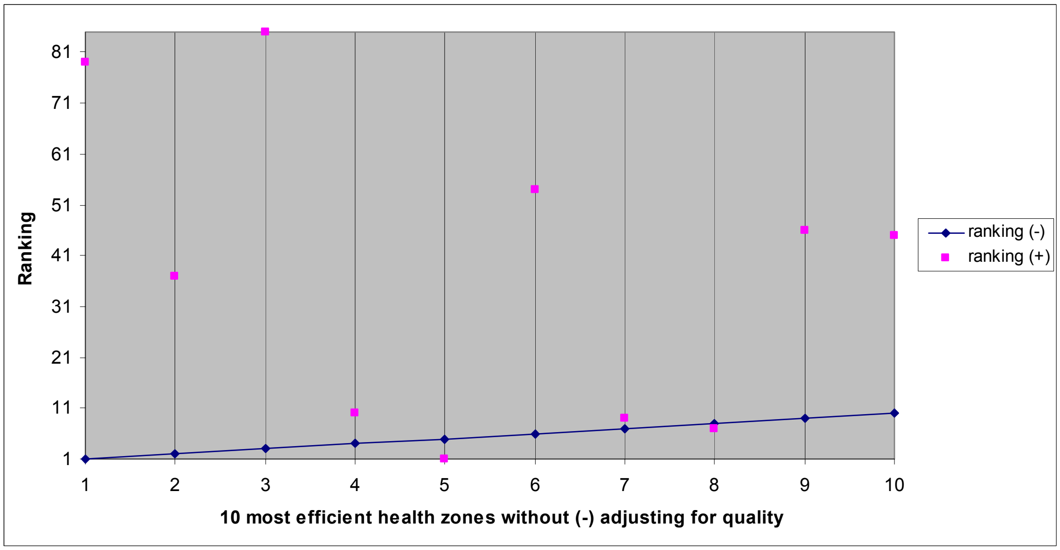
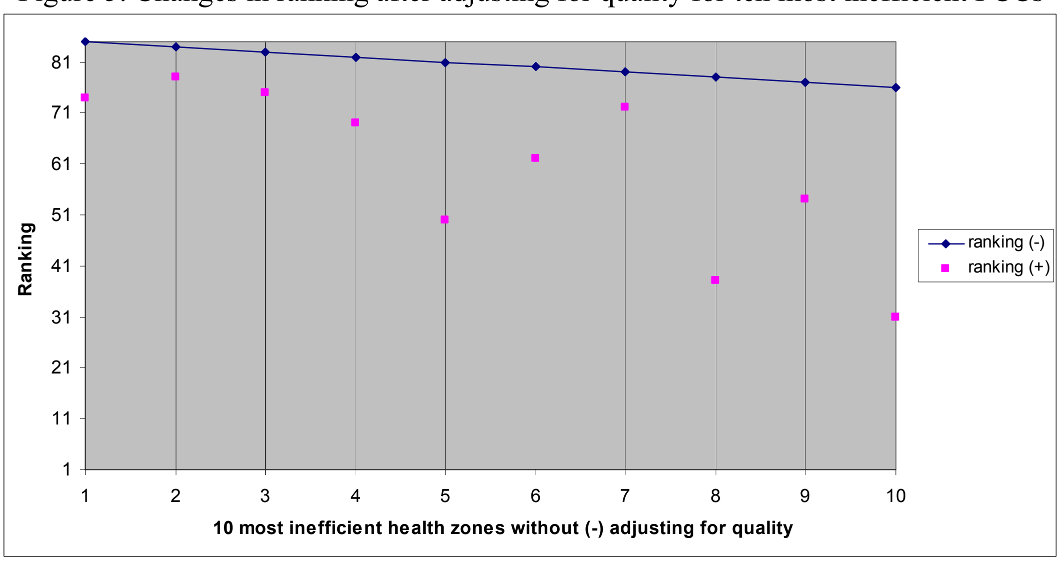


Figure 5 is the analogue of Figure 4 for the 10 least efficient PCCs (ranked 76-85) in Model 1. One observes that they all obtain a higher ranking after adjusting for quality.

Figure 5. Changes in ranking after adjusting for quality for ten most inefficient PCCs



This evidence is further corroborated by the pairwise Spearman rank correlation coefficients of the PCCs' efficiency scores (Table 5) and the respective *p* values. For the full sample, the positive correlation between the two rankings is quite low (about 20%), and significant only at 10%. For the rural and urban sub-samples, in the former the correlation coefficient is less than 30%, while in the latter the two sets of rankings are not significantly correlated.

Table 5. Pairwise Spearman rank correlation coefficients of the PCCs' efficiency scores.

	All (85)	Rural (65)	Urban (20)
Spearman rho	0.2002	0.2816	0.1519
P value	0.0662	0.0231	0.5227

In sum, the above graphical analysis and the Spearman correlation coefficient evidence show how misleading an efficiency ranking of PCCs based on measures of health services not adjusted for quality might be. One must conclude that the use of quality-adjusted measures of PCC output is important in order both to provide reliable measures of average technical efficiency and to rank health providers. Indeed, the results of neglecting quality would be biased, and even simply incorrect.

5. Conclusions

Using a newly constructed dataset for the Extremadura Primary Health System (APEX06), we have estimated a stochastic production model for a representative sample of PCCs observed in 2006. While the existing literature on the topic has dealt with non-parametric and deterministic techniques of estimation, we here provided the first empirical study on technical efficiency in the primary health care sector taking into account the quality adjustment of health output within a parametric and stochastic frontier production model.

As a second aspect departing from the previous literature, we were able, given the richness of the information of our dataset, to construct a synthetic measure of output taking both the activity level and the quality features of health services into account, and hence to overcome some of the criticisms leveled at the use of purely quantitative indicators of output.

Our main objective was to assess how, to what extent, and in which direction disregarding the quality features of health services might affect the results of efficiency analysis. In order to do so, we estimated two models. The first was defined by the use of a purely quantitative measure of output, and the second included an adjustment of output in order to account for quality. By comparing the estimations obtained with the two models, we found that adjusting output for quality influences efficiency analysis in three senses. First, a greater proportion of the deviations from potential output are explained by inefficiency. Second, the average technical efficiency in the sector is lower, while its dispersion among the PCCs is significantly greater. And third, the efficiency ranking of the PCCs is also affected.

In view of the present set of results, one must conclude that great accuracy should always be required in the definition of the output in the primary health care sector.

References

Abdy, H. (2003). Multivariate Analysis, in M. Lewis-Beck, A. Bryman, T. Futing (Eds): *Encyclopedia for research methods for social sciences*.

Aigner, D.J., C.A.K. Lovell and P.J. Schmidt (1977). "Formulation and estimation of stochastic frontier production function models". *Journal of Econometrics*, 6: 21-37.

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

Battese G. and G. Corra (1977). "Estimation of a production frontier model with application to the pastoral zone of Eastern Australia". *Australian Journal of Agricultural Economics*, 21(3): 167-179.

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.

Chilingerian, J. and H. Sherman (1997). "DEA and primary care physicians report cards: Deriving preferred practice cones from managed care service concepts and operating strategies". *Annals of Operations Research*, 73: 35-66.

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.

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

Greene, W.M. (1993). "The econometric approach to efficiency analysis". In *The Measurement of Productive Efficiency: Techniques and Applications*, Harold O. Fried, C.A.K. Lovell and S.S. Schmidt (Eds.), Oxford: Oxford University Press: 68-119.

Giuffrida, A. (1999). "Productivity and efficiency changes in primary care: a Malmquist index approach". *Health Care Management Science*, 2: 11-26

Jobson, J.D. (1992). Applied Mulitvariate Analysis. Volume II: Categorical and Multivariate Methods. Editorial Springer-Verlag, New York, pp: 388-426.

Jondrow, J, C.A.K. Lovell, I.S. Materov and P. Schmidt (1982). "On the estimation of technical inefficiency in the stochastic frontier production function model". *Journal of Econometrics*, 23: 269-274.

Koopmans, T.C. (1951). "An analysis of production as efficient combination of activities". In *Activity Analysis of Production and Allocation*, Koopmans, T.C., eds, Cowles Commission for Research in Economics, Monograph no. 13. New York.

Meeusen, W. and J. van den Broeck (1977). “Efficiency estimation from Cobb-Douglas production functions with composed error”. *International Economic Review*, 18: 435-444.

Murillo-Zamorano, L.R. (2004). “Economic efficiency and frontier techniques”. *The Journal of Economic Surveys*, 18 (1): 33-77.

Ozcan, Y. A. (1998). “Physician benchmarking: measuring variation in practice behavior in treatment of otitis media”. *Health Care Management Science*, 1: 5-17.

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.

Puig-Junoy, J. and V. Ortún (2004). “Cost efficiency in primary care contracting: a stochastic frontier cost function approach”. *Health Economics*, 13: 1149-1165.

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

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

Vega, J., Murillo, L., De Miguel, F., Morillo, J. and M. Sánchez (2007). “Un sistema integrado de información para la atención primaria en Extremadura, APEX06”. Consejería de Sanidad y Consumo, Junta de Extremadura.