Accepted Version of:

SARRICO CS, ROSA MJ, COELHO IP (2010). The Performance of Portuguese Secondary Schools. An exploratory study. *Quality Assurance in Education* 18: 4, 286-303. https://doi.org/10.1108/09684881011079143

The performance of Portuguese secondary schools – an exploratory study

The purpose of this exploratory study is to measure and compare the performance of Portuguese secondary schools. Some data on the schools was collected via an electronic questionnaire sent to 103 secondary schools of the Centre Region of Portugal; other needed data was available through the Centre Regional Education Authority. Of the 33 schools that participated in the study, there was complete data for 29. We first use a non-parametric technique, data envelopment analysis, to assess the sample of schools. Then we compare the results obtained with some preliminary results of a national evaluation programme of schools. From our findings, we tentatively conclude that most schools are following national education policy priorities relating to decreasing dropout rates and increasing completion rates. School performance does not seem to relate to geographic location, size of the school, typology of the school or rotation of its executive committee. The article concludes by discussing the need for metric benchmarking exercises of the type proposed, to inform schools, evaluators and policy decision-makers. Finally, the complementarities between metric and practice benchmarking exercises are argued for.

Keywords: performance; Portuguese secondary schools; data envelopment analysis

Introduction

If one assumes, like Drucker (1999), that we live in a time of rapid change and profound transition, where society faces great and new challenges, it is easy to understand the importance of education in the context of today's world and its impact on the economic, cultural and social development of countries, irrespective of their stage of development.

Education should, then, be a priority area for Portugal, especially because its educational results, in spite of the significant changes in the last 30 years, are still well below the majority of other developed nations (OECD, 2007). In this light, it is important to know to what extent schools are achieving their stated mission, assuming that it is only possible to define educational policy and improve the quality of the education system if adequate models of institutional performance measurement are available.

Portugal faces serious problems with enrolment, progress and completion rates at secondary school level, which are significantly below all other EU countries. Even those students that make it into higher education, frequently, find it hard to finish their degrees. According to a recent official study the average survival rate for state higher education is around 65% (Portugal, Observatório da Ciência e do Ensino Superior, 2007). One of the reasons put forward is a poor secondary education.

Since the school year 2000/01, the results of national exams at the end of secondary school, which are used as entry requirements for higher education, have been made publicly available. These have given rise to the construction of secondary school rankings by several newspapers, raising questions about school performance, school choice, but also the perverse effects that these simplistic, output-based rankings have.

Analysis of performance in Portugal has to be contextualized with the relative maturity of the concept for the country and level of understanding implicit for the stakeholders.

Our study focuses on the development of a model to measure and compare the performance of secondary schools in Portugal. A performance measurement model for Portuguese secondary schools, using data envelopment analysis (DEA), was developed and tested, in order to answer several fundamental questions. Firstly the work explores what are the expectations of school performance for the different stakeholders in the system. Belying this is an analysis of how expectations are formed and thus relative judgements made by different stakeholders in society. A second objective of the work is to understand the socio-economic factors which significantly determine the performance of the schools. Thirdly, definition of the performance measurement model in the Portuguese case requires an understanding of the country specific problems and the generalized objectives for a western-European developed economy. A final question drawn out by this work is to what extent performance differences are attributable to socio-economic input factors or managerial factors. This question helps us to understand the effectiveness of the system in a more objective way.

Evaluating public sector efficiency and performance in Portugal is generally hampered by lack of data or problems with the quality of the data. One problem is that the official statistics agency for education uses the school as the statistic unit, rather than the pupil. Thus no data on pupil socio-economic background, such as education and employment of parents is generally available. To surpass this problem, a questionnaire was developed and administered electronically to all secondary schools in the central region of Portugal, in cooperation with the Centre Regional Authority for Education, to obtain the necessary data. Having obtained the necessary data a metric benchmarking exercise, using the developed DEA model was conducted.

Meanwhile, the Ministry of Education has launched an external evaluation exercise of all schools in Portugal based on peer review, conducted by a panel of inspectors from the General Inspectorate for Education and lay members. A comparison between the results of these evaluations and the results of our model will be discussed, namely their consistency and the complementarities between metric and practice benchmarking exercises.

It is hoped that our study will contribute towards a better understanding of the Portuguese situation in terms of secondary school performance and to inform the debate on school evaluation, performance improvement and policy setting.

The development of a model of performance measurement for Portuguese schools

Education is an important area for the application of techniques to the evaluation and improvement of public sector performance (Pollitt, 2003). The literature on school evaluation reports several investigations that centre their analysis on regression methods to compare expected performance, given a set of independent variables that are believed to have an impact on school performance, and actual performance (see for instance, Jesson & Gray, 1991; Jesson, Mayston, & Smith, 1987; D. Mayston & Jesson, 1988). More recently, there are reports on the use of a mathematical programming technique, data envelopment analysis (DEA), for measuring educational performance (see the review of Worthington, 2001). This technique has gained favour due to some advantages that it presents relative to regression techniques (Thanassoulis, 1993).

DEA was developed by Charnes, Cooper & Rhodes (1978) to evaluate public sector and not-for-profit organisations. Subsequently, it has also been extensively used in for-profit service organisations. DEA compares each decision-making unit (DMU) with all other DMUs in a set of DMUs, and calculates an aggregate performance measure based

on a ratio of outputs and inputs. DEA can deal with multiple inputs and multiple outputs for measuring the performance of each DMU. With this information the DEA model determines the observed frontier of performance, based on the units with 100% performance – all those that perform better relative to all others in the set of units. DEA can also identify areas for improvement of performance for each unit (relative to the observed performance at the frontier), comparing the practice of the units on the frontier with the practice of the units below it. The sharing of the best practice of the units on the frontier with the units below it allows an opportunity for the units with worse performance to improve and, thus, improve the performance of the whole set. The repeated use of DEA helps to establish an environment of organisational learning.

The development of appropriate techniques of performance measurement in education, such as DEA, and its application to available data, raises a number of questions that deserve being researched (D. J. Mayston, 2003).

The first issue that merits being investigated is the relationship of education performance evaluation to the underlying objectives of the educational system. If this relationship is weak, there is a risk of sub-optimisation of the educational outcomes achieved, through concentration of managerial attention and resources on targets that do not, in fact, measure what is intended (Dyson, 2000).

Another important property that any model of performance measurement should have, and that should be considered, is the need to adequately account for the differences in resources and in the characteristics of the cohorts of students that each school faces. The utilisation of aggregate exam results, not adjusted for these differences, will unduly favour schools that receive pupils from less privileged socio-economic backgrounds. The utilization of more sophisticated methods of performance measurement, such as regression and DEA, might be able to deal with some of these problems (see for

instance, Arnold, Bardhan, Cooper, & Kumbhakar, 1996). However, its use is highly dependent on the availability of data that, frequently, is not available.

Finally, it is important that any performance measurement model will assure the consistency and comparability of any standards used: the values of the educational performance measures for each school need to be validly compared across schools and time.

Evaluation of schools in Portugal is not a totally new experience, as several projects on school evaluation have taken place over the last fifteen years (Coelho, Sarrico, & Rosa, 2007):

Observatório da Qualidade da Escola (1992-1999)

Projecto Qualidade XXI (1999-2002)

Programa Avaliação Integrada das Escolas (1999-2002)

Projecto Melhorar a Qualidade (2000-2004)

Programa AVES – Avaliação de Escolas Secundárias (2000-…)

Projecto de Aferição da Efectividade da Autoavaliação das Escolas (2004-2006)

Although some of these studies highlighted the preoccupation with contextualising the results obtained for each school, with the socio-economic context of its pupils (see for instance, Portugal, Inspecção Geral de Educação, 2002), they all suffered from a lack of data both in quantity and quality terms, to be able to deal properly with the issue. Especially no socio-economic data had been collected systematically and publicly available at pupil level.

On the other hand, there was a clear lack of continuity from one project to the next, due, above all, to changes in government and/or minister of education: Portugal has had 25

different education ministers since the first free elections in 1976. From a recent study one can conclude that all these experiences have not promoted a culture of self-evaluation and continuous improvement at schools (Oliveira et al., 2006). Meanwhile, and in an attempt to change things, national exams at the end of each cycle of study are now being generalised. This will sure increase pressure of school rankings published by the media: not just at the end of secondary education, but also at other school levels. At the same time, a new universal programme of school evaluation has been launched, *Programa de Avaliação Externa das Escolas*, initiated in 2006, based on peer review by an external panel of assessors composed of educations inspectors and lay members.

Secondary education in Portugal starts with the 10th year of study and lasts for three years (see Figure 1). One of the reasons for studying this level of study was the possibility of gathering data for entry and exit qualifications, which would allow a value-added approach to be taken, i.e. the higher the input levels of a school, the higher the output levels should be.

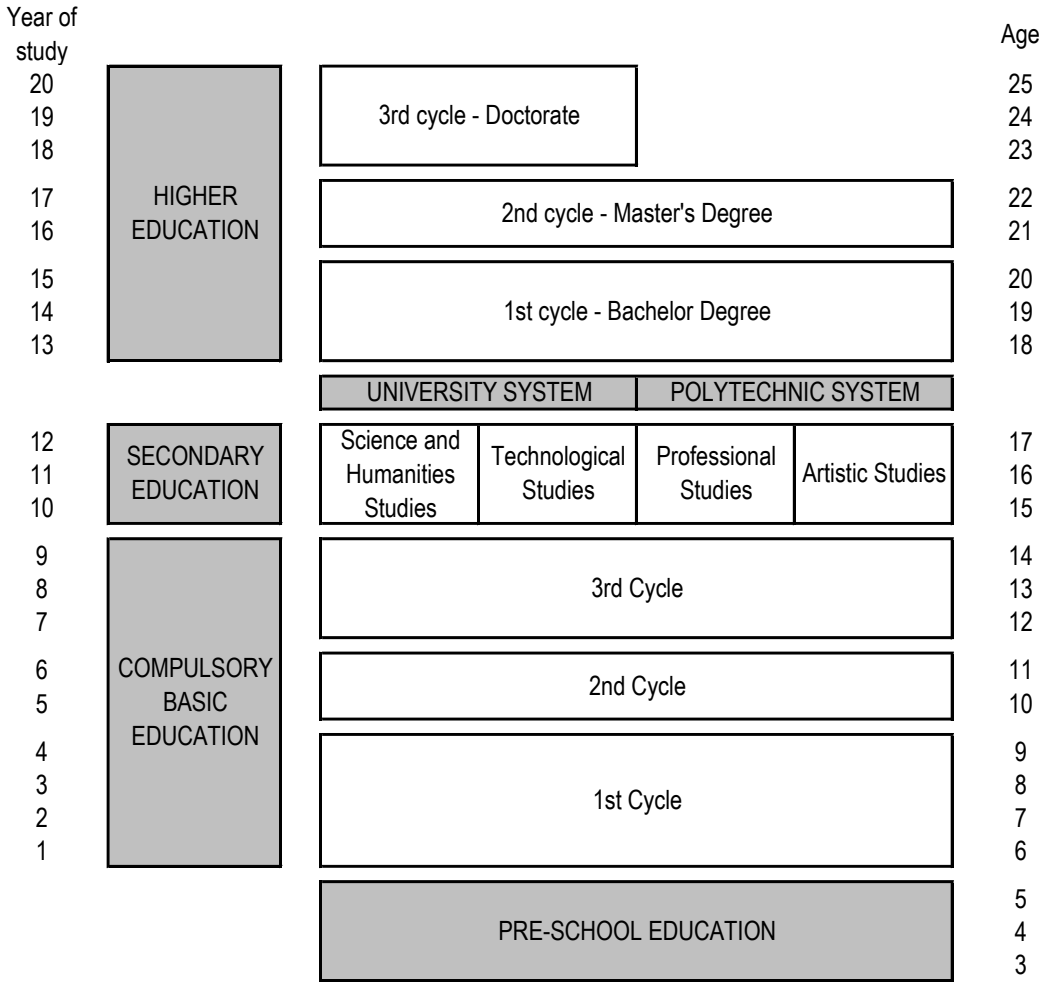

Figure 1. Portuguese education system.

Taking a value-added approach a model specification can be conceived, which will call the School Choice Model (see Table 1). This will take into consideration one type of stakeholder of the education system: students and their families. Given the ability on entry and socio-economic characteristics of the student (which are not under the control of the secondary school), how much value can the school add to its students in terms of non-dropout rate, completion rate and examination success? A similar model was used in the context of university selection by applicants in Sarrico, Hogan, Dyson, & Athanassopoulos (1997).

Table 1. School Choice Model.

Inputs	Outputs
Ability on entry	Examination success
Socio-economic characteristics	Completion rates
	Non-dropout rates

One can envisage another type of model, taking into account another stakeholder perspective, that of the government and ultimately of the tax-payer citizen, which will call School Management Model (see Table 2) (see Sarrico & Dyson, 2000 for a similar approach using universities). This model will take into consideration not only student characteristics, but also levels of school resourcing, both in terms of quantity and quality, which again, in the case of Portugal, are not under the control of the school, but only of the ministry and its education authorities.

Table 2. School Management Model.

Inputs	Outputs

Ability on entry	Examination success
Socio-economic characteristics	Completion rates
Measures of resourcing	Non-dropout rates
Quality of the teaching staff	

A similar study to ours, meanwhile, has been developed with data from one of the programmes mentioned above, *Programa AVES*, which covers however a different sample (Portela & Camanho, 2007).

Sample and variables

Since there was no data readily available on a number of factors in the models, a questionnaire was designed to collect data on school resources, socio-economic characteristics of the student body, ability on entry, and classifications on exit. All other data was available through the Centre Regional Authority for Education. The questionnaire was administered electronically to all 103 secondary schools in the Centre region of Portugal. We have received data for 33 schools (32% return rate), but there was missing data for 4 of the schools. Data refers to the 2005/06 school year, with the exception of 12^{th} grade classification, which refers to the end of the 2004/05 school year. Obviously, a strong assumption is being taken here: that average school 12^{th} grade classifications do not vary greatly from one year to the next.

The final variables for the two models can be found in Table 3 and Table 4.

Table 3. Variables for School Choice Model.

Inputs	Outputs
Mean 9th grade classification Portuguese + Mathematics	Mean 12th grade classification of the school
Mean school years of parents	% of pupils that complete 12th grade
% pupils without income support	% of pupils that do not abandon

Table 4. Variables for School Management Model.

Inputs	Outputs
Mean 9th grade classification Portuguese + Mathematics	Mean 12th grade classification of the school
Mean school years of parents	% of pupils that complete 12th grade
% pupils without income support	% of pupils that do not abandon
Ratio of teachers per student	
Mean number of years of teachers at the school	

In Table 5 descriptive statistics for every variable can be found, and a detailed discussion of each of them follows.

Table 5. Descriptive statistics.

	9th grade classification	school years of parents	% pupils withouth income support	ratio of teachers per pupil	number of years of teachers at the school	12th grade classification of the school	% of pupils that complete 12th grade	% of pupils that do not abandon
Mean	3.36	7.44	81.61	0.32	14.52	9.99	49.98	96.87
Median	3.37	7.37	83.53	0.27	14.70	9.99	47.14	97.24
Standard Deviation	0.19	1.09	9.04	0.13	3.76	0.86	14.63	3.18
Minimum	3.04	5.60	58.79	0.15	9.19	7.91	27.27	87.50
Maximum	3.76	9.49	95.97	0.60	24.49	11.71	92.34	100.00

Inputs

Mean 9th grade classification Portuguese + Mathematics. For every pupil in the 12th grade, the Mathematics and Portuguese 9th grade classification was collected, and the mean for each secondary school was calculated. The scale goes from 1 to 5, 1 and 2 represent failure, 3 is a pass, 4 is good and 5 very good. The mean of this variable for all schools involved is not very high (3.36) and the standard deviation is low, meaning that there is not a great deal of variability between schools on this factor.

Mean school years of parents. Data was collected for the school attainment of each parent of each pupil in the 12th grade, and the mean for the school was calculated. On average, for the schools concerned, the parents hold 7.44 years of study, and there is more variability between schools than with the previous factor, with a minimum of 5.60 and a maximum of 9.49 years of study. It is important to note that basic compulsory education in Portugal is the 9th grade (9 years of study) or until the pupil reaches 16 years of age, whatever happens first. It is, thus, easy to see that a considerable number of parents have not completed basic compulsory education.

% pupils without income support. This variable represents the percentage of pupils in the school that do not receive income support from the state to help them with expenses (from school materials, transport, meals, etc.). On average 81.61% of pupils do not receive income support, but the figure ranges from 58.79% to 95.97%.

Ratio of teachers per student. This represents the level of resourcing, measured by the number of teachers of the school by the number of pupils. The variability on this variable is quite high, ranging from 0.15 to 0.60, with a mean of 0.32. This is probably due to the fact that the number of teachers at the school reflects historical numbers of students rather than the actual numbers. There is a significant time lag between a decrease in the numbers of pupils at a school (due to demographics, for instance) and its effect on the number of teachers at that school.

Mean number of years of teachers at the school. This variable attempts to measure, to some degree, the quality of the teaching staff, but also the stability of the teaching staff at the school. More problematic schools tend to have a high rotation of teaching staff. It varies from 9.19 to 24.49 years at the school.

Outputs

Mean 12th grade classification of the school. This represents the average of the final grades of the pupils of the school (these are composed of a classification obtained in a national exam, and the internal classification given to the pupils by their teachers). The classification scale goes from 0 to 20, 10 representing a pass. As it can be seen, some schools have a negative (below 10) average classification, the minimum being 7.91 and the maximum 11.71, which is quite low, given the scale.

% of pupils that complete 12th grade. A significant number of pupils do not complete 12th grade, because they fail to pass the year, given the negative classifications, as discussed in the previous variable. On average, on our sample, 49.98% complete, ranging from a minimum of 27.27% to a maximum of 92.34%.

% of pupils that do not abandon. Some pupils do not even reach the end of the academic year; they abandon their studies altogether. The problem does not seem to be

particularly acute in our sample, with an average of 96.87 not abandoning, ranging from a minimum of 87.50% and a maximum of 100%.

Results

An output-oriented CCR model of DEA (Charnes et al., 1978), which assumes constant returns to scale (CRS), was used for both the school choice model and the school management model. There are some indications that variable returns to scale (VRS) are present in assessments of Portuguese secondary schools (Pereira & Moreira, 2007). However, since our models use ratios, and thus scale of the school is not accounted for in any of the variables, a decision was made to use CCR. This raises the problem of having targets for non-abandon rate higher than 100%, so improvement rates in this case might be inflated. The problem does not arise with completion rates. An output orientation is used since the objective here is not one of resource minimisation, but a value-added approach: given a school resources, of which it has little control, how far in terms of outputs can it go. Results are shown in Table 6 and Table 7.

Table 6. Results for School Choice Model.

Unit	Efficiency score	Reference count	Potential improvement (%)		
			12th grade classification	Completion rate	Non abandon rate
EB23SALC	100	0	0%	0%	0%
EB23SSC	100	5	0%	0%	0%
ES3ESM	100	5	0%	0%	0%
ES3GAB	100	4	0%	0%	0%
ES3JC	100	6	0%	0%	0%
ES3JLB	100	0	0%	0%	0%
ES3JMF	100	6	0%	0%	0%
ES3PM	100	2	0%	0%	0%
ES3VNP	100	11	0%	0%	0%
ESDD	100	5	0%	0%	0%
ESMC	100	4	0%	0%	0%
ESMV	100	5	0%	0%	0%
ESSERT	100	3	0%	0%	0%
ESCANT	99.69	-	0%	0%	2%
EB23SMRF	98.65	-	2%	61%	0%
ESARG	98	-	6%	2%	2%
ES3JCCG	96.62	-	14%	4%	3%
ES3FV	96.24	-	8%	34%	0%
EB23SPEN	95.83	-	4%	4%	7%
ES3JE	95.8	-	5%	11%	2%
ES3MC	94.7	-	5%	6%	5%
ES3MA	92.78	-	9%	8%	2%
ES3FA	92.74	-	8%	20%	3%
EB23SPAC	91.66	-	9%	9%	0%
ES3FCR	90.93	-	27%	20%	7%
ES3BAT	90.84	-	10%	10%	9%
ESSEIA	89.68	-	12%	12%	2%
ES3JML	89.29	-	12%	16%	8%
EB23SPJM	87.21	-	14%	15%	3%

As it would be expected with a relatively small sample, there is not much discrimination between the schools' performance, for the school choice model (Table 6): efficiency scores are generally high, ranging from 87.21 to 100, with 13 out of the 29 schools being in the frontier of observed performance, 11 of which are peers to inefficient schools, with school ES3VNP being a reference to 11 other schools, and thus a good candidate for further investigation. Another important consideration is the fact that these schools represent a self-selected sample: the schools in the sample chose to answer the questionnaire, and might not be representative of the rest of the schools. In fact, by showing a willingness to participate, they might be more pro-active in taking an

interest in their performance and, consequently, have better performance than the rest of the schools which have not participated in the study.

As to potential improvements to the outputs, it must be noted that the non-abandon rate improvement takes into account the fact that the target cannot be greater than 100%, and thus the improvement was calculated taking that into consideration, when the model gave a target greater than 100%. The most dramatic potential improvements, as expected, relate to completion rates, although there is quite a lot of scope for improvement, for some schools, also in the grade classifications.

Table 7. Results for School Management Model.

Unit	Efficiency score	Reference count	Potential improvement (%)		
			12th grade classification	Completion rate	Non abandon rate
EB23SALC	100.00	0	0%	0%	0%
EB23SMRF	100.00	2	0%	0%	0%
EB23SSC	100.00	3	0%	0%	0%
ES3BAT	100.00	0	0%	0%	0%
ES3ESM	100.00	2	0%	0%	0%
ES3FV	100.00	1	0%	0%	0%
ES3GAB	100.00	2	0%	0%	0%
ES3JC	100.00	3	0%	0%	0%
ES3JLB	100.00	0	0%	0%	0%
ES3JMF	100.00	1	0%	0%	0%
ES3MA	100.00	0	0%	0%	0%
ES3PM	100.00	0	0%	0%	0%
ES3VNP	100.00	6	0%	0%	0%
ESARG	100.00	5	0%	0%	0%
ESCANT	100.00	1	0%	0%	0%
ESDD	100.00	3	0%	0%	0%
ESMC	100.00	1	0%	0%	0%
ESMV	100.00	4	0%	0%	0%
ESSERT	100.00	6	0%	0%	0%
ESSEIA	99.77	-	21%	18%	0%
ES3JCCG	97.68	-	9%	2%	2%
ES3JE	96.14	-	5%	4%	2%
EB23SPEN	95.83	-	4%	4%	7%
ES3FA	95.33	-	5%	11%	3%
EB23SPJM	94.88	-	5%	5%	3%
ES3MC	94.70	-	5%	6%	5%
EB23SPAC	93.42	-	7%	7%	0%
ES3FCR	92.37	-	23%	57%	7%
ES3JML	89.73	-	12%	12%	8%

When it comes to the results for the school management model (Table 7), as expected, given the increased number of variables, results improve, with 19 schools now at the frontier of observed performance. This means that some schools might be able to be more attractive to pupils and their families should they be better resourced. This is an important message for the education ministry in terms of its policy for allocating resources. For instance, the school ES3BAT, who had an efficiency score of 90.84% in the first model, has now an efficiency score of 100%, which means that it might not be particularly attractive from a pupils and their families' perspective, but is doing a good job, given its allocated resources by the school authorities. The scope for improvement

in terms of classifications and, especially, completion rates remains high for some of the inefficient schools.

Current national policy for education emphasises the fight against abandoning school without qualifications, and thus an effort is being put on making pupils complete the year successfully (instead of having to repeat the year due to failure). There is some anecdotal evidence that some schools, in order to improve their position in the league tables published by the media, using 12th grade exam results, might neglect completion and dropout rates, escalating the problem. In order to assess the issue with the sample of our schools, the average virtual weights assigned to the three outputs of our models were calculated in Table 8.

Table 8. Distribution of virtual weights for the outputs.

	Average virtual weights allocation		
	12th grade classification	12th grade completion	Non-abandon rate
SCM	33.6	14.9	51.5
SMM	35.6	13.8	50.7

At least, for the sample being study, the improvement of grades at the expense of completion and dropout rates does not seem to be widespread, although there are some

extreme examples of schools putting all their weight (100%), or close, into 12th grade classifications.

In order to avoid this behaviour, national policy can be translated into the models by means of weights restrictions where the virtual weight assigned to the percentage of pupils that do not abandon school should be greater than the virtual weight assigned to the percentage of pupils that complete 12th grade, which in turn should be greater than the weight assigned to the mean 12th grade classification of the school. Virtual weights restrictions were then added to the models (see Sarrico & Dyson, 2004, for more on virtual weights restrictions).

The results of the models with virtual weights restrictions can be found in Table 9 and Table 10.

Table 9. Results for School Choice Model with virtual weights restrictions.

Unit	Efficiency score	Reference count	Potential improvement (%)		
			12th grade classification	Completion rate	Non abandon rate
EB23SALC	100.00	4	0%	0%	0%
EB23SSC	100.00	5	0%	0%	0%
ES3JMF	100.00	9	0%	0%	0%
ES3VNP	100.00	18	0%	0%	0%
ESMC	100.00	10	0%	0%	0%
ES3JC	100.00	10	0%	0%	0%
ES3JLB	100.00	0	0%	0%	0%
ESDD	100.00	8	0%	0%	0%
ES3GAB	100.00	3	0%	0%	0%
ESMV	99.48	-	-7%	9%	0%
ES3PM	98.47	-	-12%	14%	0%
ESARG	98.00	-	6%	2%	2%
ES3JCCG	96.62	-	14%	4%	3%
ES3FV	96.24	-	8%	34%	0%
ESSERT	95.93	-	-16%	24%	4%
ESCANT	95.50	-	-12%	21%	5%
ES3MC	94.70	-	5%	6%	5%
EB23SMRF	93.16	-	-2%	55%	0%
EB23SPEN	93.01	-	-6%	21%	8%
ES3JE	92.63	-	3%	14%	2%
ES3FA	91.71	-	1%	17%	3%
EB23SPAC	91.09	-	4%	16%	0%
ES3FCR	90.93	-	27%	20%	7%
ES3MA	90.68	-	9%	8%	2%
ES3BAT	90.48	-	5%	16%	9%
ES3ESM	90.04	-	-10%	32%	4%
ESSEIA	89.68	-	12%	12%	2%
ES3JML	86.72	-	9%	22%	8%
EB23SPJM	86.29	-	4%	28%	3%

Given the weights restrictions, which translate current national education policy, the number of units who appear efficient in the school choice model decreases, from 13 to 9, as expected. The schools ES3ESM, ES3PM, ESMV, and ESSERT are no longer efficient: they need to improve completion rates, possibly at the expense of average classification results. ES3ESM, especially, had a significant drop in efficiency to 90.04, putting it at the bottom part of the table. The schools ES3VNP, ESMC and ES3JC continue to be pointed out as good comparators for underperforming schools.

Table 10. Results for School Management Model with virtual weights restrictions.

Unit	Efficiency score	Reference count	Potential improvement (%)		
			12th grade classification	Completion rate	Non abandon rate
EB23SALC	100.00	1	0%	0%	0%
EB23SMRF	100.00	0	0%	0%	0%
EB23SSC	100.00	4	0%	0%	0%
ES3GAB	100.00	3	0%	0%	0%
ES3JLB	100.00	0	0%	0%	0%
ES3MA	100.00	0	0%	0%	0%
ES3VNP	100.00	11	0%	0%	0%
ESARG	100.00	7	0%	0%	0%
ESDD	100.00	2	0%	0%	0%
ESMC	100.00	5	0%	0%	0%
ESMV	100.00	4	0%	0%	0%
ESSERT	100.00	3	0%	0%	0%
ES3JC	100.00	8	0%	0%	0%
ES3JMF	100.00	0	0%	0%	0%
ESCANT	100.00	1	0%	0%	0%
ES3FV	100.00	2	0%	0%	0%
ESSEIA	99.77	-	21%	18%	0%
ES3PM	98.47	-	-12%	14%	0%
ES3JCCG	97.68	-	9%	2%	2%
ES3FA	95.09	-	2%	8%	3%
ES3MC	94.70	-	5%	6%	5%
ES3BAT	94.63	-	-4%	16%	6%
ES3JE	93.42	-	-3%	17%	2%
EB23SPEN	93.01	-	-6%	21%	8%
EB23SPAC	92.41	-	4%	13%	0%
ES3FCR	92.37	-	23%	57%	7%
EB23SPJM	91.49	-	3%	15%	3%
ES3ESM	90.06	-	-10%	32%	4%
ES3JML	87.01	-	4%	26%	8%

When the same weights restrictions are applied to the school management model, as expected the number of efficient schools increases to 16. However, relative to the school management model without weights restrictions 3 schools are no longer efficient: ES3BAT, ES3ESM and ES3PM. Again, ES3SM had the most significant drop. ES3VNP remains the most used peer for inefficient schools.

Discussion of results

Although our sample is quite small, and self-selected, some interesting results can be discussed. Although some schools might be emphasising school classifications at the expense of increasing completion and decreasing dropout rates, most seem to be 'good students' when it comes to national policy. More appealing would be to try and explain why some schools obtain higher efficiency scores than others. We had information on the geographic location of each school – namely its rural or urban location –, its typology – exclusively secondary school or covering other cycles of study –, its size, in terms of the number of pupils taught, and the rotation rate of the executive committee (elected governing body) in the previous 10 years. We could find no relationship between efficiency scores and any of these variables. The need for a practice benchmarking following our proposed metric benchmarking using DEA seemed obvious. This is intended to be undertaken as future research.

However, the fact that the first results of the *Programa de Avaliação Externa das Escolas (Programme for the External Evaluation of Schools)* have recently become available, will allows us to do a practice benchmark «by proxy». The program started in 2006 with the appointment of a working group for school evaluation by the Ministry for Education. The group developed a model for external evaluation of schools, which covers five domains: results (including, but not only, academic results), service delivery; organisation and management; leadership; and self-regulation and progress (see Oliveira et al., 2006). Each domain is given a classification on a scale of insufficient, sufficient, good, and very good. Since data has become available on 24 schools evaluated in 2006, and 102 in 2007; all on a voluntary basis. The assessment has mostly a qualitative nature: the schools are visited for 2 to 3 days by a panel of three evaluators: two education inspectors (belonging to the General Inspectorate for

Education) and a lay member. The panel of assessors interviews several representative panels of school stakeholders, such as executive committee, pedagogical council, school assembly, student representatives, heads of departments, class directors, parents, non-teaching staff, etc. The programme foresees that each school should be evaluated every four years, as a necessary condition for giving more autonomy to schools. All evaluation reports are made public, being available on the Internet.

Given this context, we decided to compare the results of our exploratory study with the results of the qualitative peer assessment. Only two schools are common to both samples: ESMC and ES3JC. Both schools are efficient in all models: school choice and school management models, with and without weights restrictions, and are used as peers for inefficient schools, particularly in the models with weights restrictions. Their results in the national evaluative exercise are as in Table 11.

Table 11. Classifications for selected schools on the national evaluation program.

	ESMC (2007)	ES3JC (2006)
Results	Good	Very good
Service delivery	Good	Very good
Organisation and management	Good	Good
Leadership	Very good	Very good
Self-regulation and progress	Good	Very good

The results in our benchmarking exercise do, to some extent, match the results of our practice benchmarking by proxy. If one believes that the results attained by the school

(academic or otherwise) are a consequence of good service delivery, organisation and management, leadership and self-regulation, than the good results these two schools obtain in the DEA models can be explained by good or very good practices in these fields.

Moreover, it is our belief that an external peer review exercise of school evaluation could benefit from a metric benchmarking exercise to help evaluators classify academic results (albeit not non-academic results). At present, we are inclined to think that evaluators might tend to over rate the good results of schools with high levels of inputs (being it pupils' background and/ or resourcing levels) and under rate the apparently bad results of schools with low levels of inputs.

Conclusions

In this paper we have undertaken an exploratory study of Portuguese secondary schools' performance, using DEA. This study is part of a research program on school performance, which wants to use, in a first instance, the results of a metric benchmarking exercise to, in another instance, develop practice benchmarking exercises to understand school performance management, in order to inform education policies.

School rankings based on just exam classifications are not particularly useful. However, metric benchmarking exercises, provided that they are done with adequate data and techniques, would provide a systematic diagnostic tool for the whole country. Metric benchmarking that takes into account not only measures of output, such as exam classifications, completion and dropout rates, but also indicators of input, such and pupils' previous performance, socio-economic characteristics and school resources, would be fundamental to both schools' self-evaluation but also to external evaluation.

At the level of self-evaluation, metric benchmarking, via DEA, allows to inform each school of its positioning in relation to other schools in the region and the country, pointing to appropriate schools as peers, who have similar levels of input but better performance. From here, the school assessing itself can set more realistic targets, because they are observed attainable targets, identify better practices at the reference schools, which it can emulate or adapt to its own context, in order to improve its performance. Repeated benchmarking exercises, over the years, would allow for the best practices to be disseminated across the school network.

At the level of external evaluation, a metric benchmarking exercise of the kind proposed in this study, in its capacity to digest data, which is far away from any human mind, is a source of objective information, which can complement the value judges formed by a panel of external assessors, who will always present a degree of subjectivity.

Furthermore, it can centre the energy of the administration in the follow-up and support to under-performing schools, instead of splitting equally the efforts of external evaluation by all schools, independently of their performance.

This paper reports on an exploratory study, and thus it contains several limitations which need to be overcome by further work. The sample used is small and self-selected; in the future we intend to use the whole school population for the country. The DEA models used are quite simple; more sophisticated models can and will be used, namely non-discretionary- and bounded-variable models. Also, we would like to compare the results of a non-parametric technique, such as DEA, with parametric techniques that have been used to evaluate school performance.

References

Arnold, V. L., Bardhan, I. R., Cooper, W. W., & Kumbhakar, S. C. (1996). New uses of DEA and statistical regressions for efficiency evaluation and estimation - With an illustrative application to public secondary schools in Texas. *Annals of Operations Research, 66*(4), 255-277.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444.

Coelho, I. P., Sarrico, C. S., & Rosa, M. J. (2007). Avaliação de escolas em Portugal: que futuro? *Forthcoming in Revista Portuguesa e Brasileira de Gestão*.

Drucker, P. F. (1999). *Management Challenges for the 21st Century*. New York: Harper Business.

Dyson, R. G. (2000). Performance measurement and data envelopment analysis, ranks are rank! *OR Insight, 13*(4), 3-8.

Jesson, D., & Gray, J. (1991). Slants on slopes: using multi-level models to investigate differential school effectivenss and its impact on pupils' examination results. *School Effectiveness and School Improvement, 2*(3), 230-247.

Jesson, D., Mayston, D., & Smith, P. (1987). Performance assessment in the education sector: educational and economic perspectives. *Oxford Review of Education, 13*(3), 249-266.

Mayston, D., & Jesson, D. (1988). Developing models of educational accountability. *Oxford Review of Education, 14*(3), 321-339.

Mayston, D. J. (2003). Measuring and managing educational performance. *Journal of the Operational Research Society, 54*(7), 679-691.

OECD. (2007). *Programme for International Student Assessment 2006*. Paris: OECD.

Oliveira, P. G., Clímaco, M. C., Carravilla, M. A., Sarrico, C. S., Azevedo, J. M., & Oliveira, J. F. (2006). *Relatório Final da Actividade do Grupo de Trabalho para a Avaliação das Escolas*. Lisboa: Ministério da Educação.

Pereira, M. C., & Moreira, S. (2007). Eficiência das escolas secundárias portuguesas: uma análise de fronteira de produção estocástica. *Boletim Económico, Banco de Portugal, 13*(1), 101-118.

Pollitt, C. (2003). Public sector, private sector – where would we be without a few good stereotypes? . In C. Pollitt (Ed.), *The Essential Public Manager* (pp. 1-25). Maidenhead: Open University Press.

Portela, M. C., & Camanho, A. S. (2007). Performance assessment of Portuguese secondary schools. *working paper 07/2007, Faculty of Economics and Management, Portuguese Catholic University, Oporto, Portugal*.

Portugal, Inspecção Geral de Educação. (2002). *Avaliação Integrada das Escolas: Relatório Nacional, Ano Lectivo 2000-2001*. Lisboa: IGE.

Portugal, Observatório da Ciência e do Ensino Superior. (2007). *Sucesso Escolar no Ensino Superior: Diplomados em 2004-2005*. Lisboa: OCES.

Sarrico, C. S., & Dyson, R. G. (2000). Using DEA for planning in UK universities - an institutional perspective. *Journal of the Operational Research Society, 51*(7), 789-800.

Sarrico, C. S., & Dyson, R. G. (2004). Restricting virtual weights in data envelopment analysis. *European Journal of Operational Research, 159*(1), 17-34.

Sarrico, C. S., Hogan, S. M., Dyson, R. G., & Athanassopoulos, A. D. (1997). Data envelopment analysis and university selection. *Journal of the Operational Research Society, 48*(12), 1163-1177.

Thanassoulis, E. (1993). A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessment. *Journal of the Operational Research Society, 44*(11), 1129-1144.

Worthington, A. C. (2001). An empirical survey of frontier efficiency measurement techniques in education. *Education Economics, 9*(3), 245-268.