A DATA ENVELOPMENT ANALYSIS OF ITALIAN COURTS EFFICIENCY

Eugenia Nissi[1]

Economic Department, "G. d'Annunzio" University of Chieti-Pescara, Italy

Agnese Rapposelli[2]

Department of Quantitative Methods and Economic Theory, "G. d'Annunzio" University of Chieti-Pescara, Italy

Abstract. *During the last two decades public sector activities at the regional or national level have been involved in a worldwide trend for improved performance. The concept of efficiency has also penetrated the legal and judicial systems at all levels, since courts around the world are in a state of crisis, dealing with backlogs and inefficiencies. Moreover, political authorities place emphasis on courts delay, which represents a widely observed phenomenon and undermines the functioning of court systems around the world. Hence, nowadays there is an increasing interest in determining whether courts and judicial systems operate efficiently. In this context, the aim of this study is to analyse the productive efficiency of the Italian Courts of Appeal for the year 2008 by using a non-parametric technique, Data Envelopment Analysis.*

Keywords: *Accountability, Public sector performance assessment, Courts, Technical efficiency, Data Envelopment Analysis.*

1. INTRODUCTION

During the 1980's, in the United Kingdom, accountability and performance measurement in the public sector received increased attention. Nowadays, all over the world it is common belief that public sector organisations should account for the services they provide, so the need for the application of improved productivity measurement in the public sector appears indisputable. Efficiency evaluation breaks in these circumstances and a number of empirical techniques have been applied to public sector performance assessment.

In this paper, an analysis of the efficiency of Courts is carried out. There is a voluminous literature in law journals interested in courts performance, since there is an increasing body of evidence that the efficiency of the court system is important

[1] Eugenia Nissi, email: nissi@dmqte.unich.it
[2] Agnese Rapposelli, email: a.rapposelli@unich.it

to a well-functioning economy. Lately, there has been some discussion as to the causes of courts delay, which represents a widely observed phenomenon: it is frequently cited as a source of inefficiency that undermines the courts' very essence[3] (Vereeck and Muhl, 2000) and is attributed to a substantial growth in caseloads and to understaffed courts adopting poor management practices (Torre, 2003). However, except for a few studies – Lewin *et al.* (1982) in the USA, Kittelsen and Forsund (1992) in Norway, Tulkens (1993) in Belgium, Pedraja-Chaparro and Salina-Jiménez (1996) in Spain, Schneider (2005) in Germany and Azevedo and Yeung (2011) in Brazil – the problem of measuring the efficiency of courts has remained relatively unexplored.

The main objective of the present paper is to examine the productive efficiency of Italian Courts of Appeal in terms of dispute resolution. Our interest in this subject was not only awakened by the observation of enormous backlogs at courts in Italy[4] (Italian Ministry of Justice, 2009), but also by the current belief that, in addition to being held accountable for judicial decisions, courts, like other public agencies, must compete for resources in a period dominated by resource scarcity (Lewin *et al.*, 1982). In many courts the growing backlogs of cases has led to considerable waiting periods for those seeking justice and a number of contributions have pointed out that the increasing caseload of Italian courts may put at risk the performance of their judicial duties. Even the Council of Europe has urged Italy for the adoption of legislative measures necessary to shorten the trials duration, both civil and criminal. In this context, the situation is particularly dramatic, though improving, in the Courts of Appeal[5], where the average court delay in the last three years was over 1100 days. This workload problem only seems to worsen, due to the number of new cases that markedly increases.

It is well known that economic theory has defined many concepts of efficiency: in this study we focus only on the technical component of the efficiency of the

---

3 Court delay also comes into conflict with art. 6 of the European Convention of Human Rights requesting a case to be treated within a reasonable period of time.

4 The structure of the Italian judiciary system is divided into three tiers: inferior courts of original and general jurisdiction; intermediate courts which hear cases on appeal from lower courts; Courts of last resort which hear appeals from lower appellate courts on the interpretation of law.

5 The Court of Appeal represents the second level in the Italian judicial system: decisions of the first level, both civil and criminal, may be appealed against at one of the Courts of Appeal, geographically identified. With the entry into force of the reform about the single judge of first instance, the Court of Appeal became the competent office for labour, social security and assistance matters, too.

Courts to coordinate the actors involved in the production of their decisions. The service provided by the judicial system at the level of judicial offices is likened, in fact, to a production process, whose technical conditions are determined by procedural rules governing the exercise of the judicial function and whose inputs consist of the persons employed. Our goal is to verify whether the system's malfunctions, represented by the extent of backlog, the length of trials and the level of costs, can be traced back to a non-optimal court size and a wrong inputs combination.

To this purpose, according to existing empirical literature, we apply two different Data Envelopment Analysis (DEA) models, accounting for different returns to scale assumptions, in order to evaluate the performance of the Italian Courts of Appeal system. In particular, the analysis focuses on the 26 Italian districts of Courts of Appeal for the year 2008, the most recent for which all required data are available. All data required have been obtained from the Judicial Organization Department of the Italian Ministry of Justice, and they are available online.

The paper is organised as follows. Section 2 introduces the concept of technical efficiency and presents in detail the DEA method for the evaluation of efficiency in organisations, Section 3 presents the data used and discusses the results obtained, Section 4 contains the main implications for future research.

2. METHOD

Modern efficiency measurement began with Farrell (1957), who drew upon the work of Debreu (1951) and Koopmans (1951) and introduced a measure for technical efficiency. According to them, a technically efficient organization is one that produces the maximum possible outputs from a given set of inputs, or one that produces a certain level of outputs with the minimum amount of inputs. Farrell (1957) suggested measuring the efficiency of a unit in terms of distance from the best unit on the production frontier, that is represented by the production function of the efficient units. This frontier, also called best practice or efficiency frontier, specifies for a unit the maximum quantities of outputs it can produce given any level of inputs or, for any level of outputs, the minimum quantities of inputs needed for producing the outputs.

This efficiency framework was reformulated as a mathematical programming problem by Charnes *et al.* (1978), thus initiating the linear programming approach to efficiency measurement known as Data Envelopment Analysis (DEA). DEA is a non-parametric method of frontier estimation. It estimates technical efficiency by first constructing the production possibility set assumed to contain all input-output correspondences which are feasible and then estimating the maximum feasible expansion of the output (output orientation) or the maximum feasible contraction of the input levels (input orientation) of the decision making units (DMUs) within the set.

To present formally the basic DEA model, known as CCR (Charnes, Cooper and Rhodes), consider a set of *n* DMUs, each consuming different amounts of *m* inputs to produce ***s*** outputs, and let x_{ij} denote the amount of input *i* ($i = 1,\ldots,m$) and y_{ri} the amount of output level *r* ($r = 1,\ldots,s$) for DMU *j* ($j = 1,\ldots,n$). The objective is to measure the efficiency of one of the set of DMUs, unit j_0, relative to the best observed practice in the sample. It is possible to obtain a measure of relative efficiency of unit j_0 that is defined by the ratio of a weighted sum of its outputs to a weighted sum of its inputs. To this end, weights are not defined a-priori, but they are chosen in order to maximise the efficiency ratio of the analysed unit j_0 so that they are shown in the best possible light. Thus, the relative efficiency of DMU j_0 is obtained by treating weights as variables and by maximising the efficiency ratio of the unit subject to the efficiencies of all the units being constrained to be less than an arbitrary limit such as 1:

$$e_0 = \max \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}} \tag{1}$$

$$s.t. \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1, \qquad j = 1,\ldots,n \tag{2}$$

$$u_r, v_i \geq 0 \qquad \forall\, r, i$$

where u_r and v_i are the weights for the *r*-th output and the *i*-th input, respectively.

This fractional model must be solved *n* times, once for each unit in the sample, for obtaining an efficiency score, which is bounded between zero and one, for each DMU. DMU j_0 is said to be efficient (has a score of unity) if no other unit or combination of units can produce more than DMU j_0 on at least one output without producing less in some other output or requiring more of at least one input.

Note that the CCR model is built on the assumption of constant returns to scale[6] of activities (CRS): that is, if an activity (x, y) is feasible, then, for every positive scalar t, the activity (tx, ty) is also feasible. However, it must be remembered that this assumption is only appropriate when all units are operating at an optimal scale. The use of CRS specification when not all units are operating at the optimal scale, will result in measures of technical efficiency which are confounded by scale efficiencies. Banker *et al.* (1984), in fact, have shown that the efficiency score generated by the CCR model is a composite total efficiency score that can be decomposed into two components, one due to scale efficiency and one due to pure technical efficiency, devoid of these scale efficiencies effects. They have also suggested an extension of the original CCR model to account for variable returns to scale (VRS) situations. Their model, known as BCC (Banker, Charnes and Cooper), is obtained by adding the convexity constraint in the previous formulation:

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{3}$$

The presence of the convexity constraint in the BCC model reduces the feasible region for DMUs, which, in general, results in an increase of efficient units; otherwise CRS and VRS models work in the same way. Hence, the efficiency score of a unit under VRS is not less than the efficiency score under CRS since the imposition of the convexity constraint makes the feasible region of the BCC model a subset of that of the CCR model.

The technical efficiency of the analysed DMU j_0 can be determined either under input reduction or output expansion orientations. Due to the CRS assumption, the relative efficiency score of the output-orientated model relate to that of the input-orientated model via $e_0 = 1/h_0$. Differently from the CRS model, under VRS assumption the model orientation (input or output) affects the projection point on the frontier and the resulting efficiencies may not be the same. Thus, for inefficient DMUs we may have $e_0 \neq 1/h_0$, although the subset of efficient DMUs is the same irrespective of model orientation.

The CCR model yields an evaluation of overall technical efficiency. The BCC model, on the other hand, can distinguish between technical and scale inefficiencies by estimating the pure technical efficiency at the given scale of operation for each unit. Hence, the divergence between the CRS and VRS efficiency score captures the

---

6 Returns to scale of a point on the production frontier are defined as the amount by which all the outputs will increase for a proportionate increase in all inputs.

impact of scale size on the performance of the unit concerned but not the nature of scale inefficiency.

DEA has several advantages over other methodologies for performance evaluation with regard to the study of the productive behaviour of public or not-for-profit organizations (Ganley and Cubbin, 1992). First of all, differently from the alternative approach for estimating the efficiency of organizational units, the parametric one represented by SFA, DEA does not require the specification of an explicit functional form for the production frontier. SFA on the other hand, requires assumptions about the production function form and the distribution of the random error and inefficiency terms, hence it does not avoid the danger of imposing a-priori wrong functional forms. Besides, since DEA makes few assumptions about the technology structure, it can easily adjust to the uncertainty surrounding public production technology (Pedraja-Chaparro and Salina-Jiménez, 1996). Another virtue of DEA is that it can easily handle multiple inputs and outputs as opposed to the usual stochastic frontier formulation. Moreover, the DEA methodology does not depend on *a priori* weights or prices for the inputs or the outputs and it can handle factors outside the control of the unit under analysis (Lewin *et al.*, 1982). Finally, DEA gives extra information in the form of peers units, i.e. the reference set of efficient DMUs, and targets, which are the optimal values of the inputs and outputs that the units under assessment should be able to achieve in order to become efficient (Coli *et al.*, 2011). It must be remembered that, in spite of the benefits offered by DEA method, SFA has the advantage that it allows random noise to be incorporated into the model, whilst DEA is deterministic, which means that any statistical noise, measurement errors, omitted variables and other misspecifications are counted as inefficiency.

3. MAIN RESULTS

The first step in applying the DEA method is to identify the set of inputs and outputs to be included in the analysis. However, the non-parametric approach to efficiency measurement does not offer any tools that can aid researchers to specify the most appropriate model. To deal with this drawback, a lot of attention has to be paid to the selection of the input-output set, giving emphasis to what is postulated by efficiency theory and to what is appropriate in the particular context under investigation.

Our DEA analysis is therefore performed with the following input variables: the number of judges employed, the number of new cases filed during the year, which is a factor outside the control of the unit being evaluated and which represents the justice demand, and the number of pending cases, which indicates the inefficiency degree of justice in relation to social expectations. The sum of the two latter variables represents the caseload. It is important to control for the caseload because judges cannot provide their services unless lawsuits are filed. Therefore, omitting the caseload would imply that productivity is underestimated for those years in which a court is charged with a small caseload (Schneider, 2005)[7]. The output of a court in terms of dispute resolution is captured by the number of cases finished during the year, so our model includes as output the total number of dispositions[8]. With regard to the input side of the model, we have not considered inputs reflecting the capital, due to data limitations. However, the results are still relevant if one takes into account that Courts are labour-intensive organisations (Pedraja-Chaparro and Salina-Jiménez, 1996).

Efficiency results are computed for each courts using input-orientation, so their objective is to minimise inputs while producing at least the given output levels. The measures of input-efficiency, listed in Table 1, have been obtained by using the DEA-Solver software developed by Kaoru Tone (Cooper *et al.*, 2000).

We can observe that three courts – Perugia, Reggio Calabria and Trento – are identified, with a score of one, as being fully CCR-efficient in the provision of judicial services, i.e. essentially court's decisions. The remaining units are sub-efficient but only three – Bari, Brescia and Campobasso – show quite low ratings. The mean CCR-efficiency score is 0.8209. With regard to the results provided by the BCC model, we may note that twelve courts – Perugia, Reggio Calabria and Trento plus Bolzano, Campobasso, L'Aquila, Lecce, Milano, Napoli, Roma, Torino and Trento – form the best practice frontier. Besides, most units register high efficiency scores, in fact the average efficiency is 0.9070. Only Brescia is at the bottom of the BCC ranking, with efficiency rating equal to 0.6650. In addition, we can note a substantial difference between the Campobasso ratings: it shows the worst scores (0.6070) in the CRS model, whilst it is fully efficient in the VRS specification. Napoli and Roma show a similar situation, too, with quite low scores in the first model, but the maximum score in the second one.

---

7 We have considered the question as to which, if any, demographic variables should be included as inputs. One might reason that variables such as per capita income and population size may help explain the variation in outputs. But it should be noted that the size of the caseload has already been included among the inputs (Lewin *et al.*, 1982).

8 We have not distinguished the output according to the type of case, because a problem would be faced: due to the high number of variables included in the DEA analysis with respect to the number of units, there would not be a reasonable level of discrimination between DMUs evaluated.

Table 1: DEA efficiency scores by Courts of Appeal, 2008

DMU	CRS score	VRS score	Scale eff.	DMU	CRS score	VRS score	Scale eff.
Ancona	0.8949	0.8958	0.9990	Messina	0.7063	0.7063	1.0000
Bari	0.6626	0.7888	0.8400	Milano	0.7431	0.9379	0.7923
Bologna	0.9397	1.0000	0.9397	Napoli	0.7723	1.0000	0.7723
Brescia	0.6650	0.6650	1.0000	Palermo	0.8448	0.9740	0.8673
Cagliari	0.8103	0.8224	0.9853	Perugia	1.0000	1.0000	1.0000
Caltanissetta	0.9425	1.0000	0.9425	Potenza	0.8436	1.0000	0.8436
Campobasso	0.6070	1.0000	0.6070	R. Calabria	1.0000	1.0000	1.0000
Catania	0.8039	0.8552	0.9399	Roma	0.7312	1.0000	0.7312
Catanzaro	0.7006	0.7839	0.8937	Salerno	0.7557	0.8023	0.9419
Firenze	0.8611	1.0000	0.8611	Torino	0.9721	1.0000	0.9721
Genova	0.7023	0.7730	0.9086	Trento	1.0000	1.0000	1.0000
L'Aquila	0.9984	1.0000	0.9984	Trieste	0.9124	0.9643	0.9462
Lecce	0.7500	0.8475	0.8890	Venezia	0.7206	0.7647	0.9423

Table 2: Summary statistics for DEA efficiency scores

	CCR efficiency	BCC efficiency	Scale efficiency
Mean	0.8209	0.9070	0.9082
Minimum	0.6070	0.6650	0.6070
Maximum	1.0000	1.0000	1.0000
Standard deviation	0.1191	0.1080	0.0966

When specifying a VRS frontier the question of the most efficient scale of the units also arises. The ratio between CRS and VRS efficiency scores provides a measure of scale efficiency: when it is equal to one, it means that the unit operates at an optimal size (MPSS, most productive scale size); on the contrary, if it is lower than one unit inefficiency also depends on scale factors (Ganley and Cubbin, 1992). We may note that a large number of courts are found efficient when the assumption of CRS has been relaxed to VRS. This indicates the presence of diseconomies of scale in the operation of individual courts. Our findings also show that Messina, Perugia, Reggio Calabria and Trento are fully efficient according to scale (Tab. 1). On average, the scale efficiency score is higher than the CCR and BCC scores (Tab. 2), whereas the score variability is lower.

Table 3: Frequency in reference set

Peer set CCR model	**Frequency to other DMUs**
Perugia	23
Reggio Calabria	9
Trento	8
Peer set BCC model	Frequency to other DMUs
Bologna	6
Firenze	5
L'Aquila	3
Napoli	1
Perugia	12
Reggio Calabria	1
Torino	4

Table 4: Reference sets – BCC model

DMU	**Reference sets (lambda)**							
Ancona	L'Aquila	0.5296	Perugia	0.4704				
Bari	Bologna	0.0830	L'Aquila	0.5150	Napoli	0.1368	Perugia	0.2650
Brescia	Bologna	0.0008	Perugia	0.6758	RCalabria	0.3234		
Cagliari	Bologna	0.0663	Firenze	0.0171	Perugia	0.9166		
Catania	Bologna	0.6382	Firenze	0.0647	Perugia	0.2972		
Catanzaro	Bologna	0.1529	Firenze	0.1270	Perugia	0.5409	Torino	0.1790
Genova	Firenze	0.2576	Perugia	0.5599	Torino	0.1825		
Lecce	L'Aquila	0.3058	Perugia	0.0923	Torino	0.6019		
Messina	Perugia	0.9059	R. Calabria	0.0941				
Milano	Napoli	0.1321	Torino	0.8679				
Palermo	Firenze	0.2843	Perugia	0.0369	Torino	0.6788		
Salerno	Bologna	0.0950	Firenze	0.1557	Perugia	0.7492		
Trieste	Perugia	0.5106	Trento	0.4894				
Venezia	Bologna	0.4935	L'Aquila	0.1377	Perugia	0.3688		

With respect to each inefficient unit DEA also identifies its reference set of efficient DMUs (peer units). Focusing on efficient units, the number of citations in peer groups can be interpreted as a measure of the "robustness" of best practice units. Tab. 3 displays the frequency with which efficient courts appear in the peer group of the inefficient ones. Perugia is the most CCR and BCC robustly efficient unit (23 and 12 times, respectively). The efficient courts Napoli and Reggio Calabria appear only in their own peer group, therefore their efficiency could be considered questionable. In particular, in the BCC model we can observe that Perugia is used as an efficient peer by all the inefficient courts, except Milano (Tab. 4). It also represents the main reference DMU for seven inefficient courts (Brescia, Cagliari, Catanzaro, Genova, Messina, Salerno and Trieste).

4. CONCLUDING REMARKS

This paper focused on performance measurement in a very traditional area of the public sector, i.e. the administration of justice. In particular, we have evaluated the performance of Italian Courts of Appeal for 2008 by means of the non-parametric approach to efficiency measurement, represented by Data Envelopment Analysis. To this purpose, we obtained measures of input-oriented technical efficiency from CRS and VRS production frontiers. Furthermore, we have decomposed the total efficiency score into two components, one due to scale efficiency and one due to pure technical efficiency, devoid of these scale efficiency effects.

Both CCR and BCC results point out that DMUs are operating at a quite high level of efficiency in the provision of judicial services, i.e. essentially court decisions. In particular, the results provided by the first DEA model show that out of the 26 units analysed three (Perugia, Reggio Calabria and Trento) are technically efficient. When in the second step we introduced the VRS specification, the number of efficient units became twelve, with the arrival of Bologna, Caltanissetta, Campobasso, Firenze, L'Aquila, Napoli, Potenza, Roma and Torino). In addition, many Courts are operating at a high level of efficiency. Based on this evidence it seems appropriate to say that there are some problems related to inadequate operational dimension.

This application has helped us to identify benchmarking courts so that the best practices can be implemented to become efficient. Hence, we can conclude that the present study represents an additional source of useful information to policy makers for future policy actions and that DEA can be useful as a diagnostic tool for distinguishing between most and least efficient courts in public sector efficiency measurement.

However, it should be taken into account that the efficiency analysis applied in this work can be improved. This study, in fact, suggests three main directions for future research. First, court's performances can be studied over time by applying the analysis for different years. Second, when data become available, an obvious way to improve the model considered would be to take into account the different complexity of the trials, and consequently their different length. Finally, the examined model could be applied to different judicial systems, or to other justice courts, too.

References

Azevedo, P.F. and Yeung, L.L. (2011). Measuring efficiency of Brazilian courts with Data Envelopment Analysis. *Journal of Management Mathematics*, (22): 343-356.

Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, (30): 1078-1092.

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

Coli, M., Nissi, E. and Rapposelli, A. (2011). Efficiency evaluation in an airline company: Some empirical results. *Journal of Applied Sciences*, (11): 737-742.

Cooper, W.W., Seiford, L.M. and Tone, K. (2000). *Data Envelopment Analysis, A Comprehensive Test with Models, Applications, References and DEA-Solver Software*. Kluwer Academic Publishers, Boston.

Debreu, G. (1951). The measurement of productive efficiency. *Econometrica*, (19): 273-292.

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society A*, (120): 253-290.

Ganley, J.A. and Cubbin, J.S. (1992). *Public Sector Efficiency Measurement – Applications of Data Envelopment Analysis.* North Holland, Amsterdam.

Italian Ministry of Justice (2009). *Intervento del Ministro della Giustizia, On. Angelino Alfano, presso la Corte Suprema di Cassazione. Inaugurazione anno giudiziario 2009.*

Kittelsen, S.A.C. and Forsund, F.R. (1992). Efficiency analysis of Norwegian district courts. *Journal of Productivity Analysis*, (3): 277-306.

Koopmans, T.C. (1951). An analysis of production as an efficient combination of activities. In: T.C. Koopmans, ed., *Activity Analysis of Production and Allocation*. Cowles Commission for Research in Economics, Monograph n.13, Wiley, New York.

Lewin, A.Y., Morey, R.C. and Cook, T.J. (1982). Evaluating the administrative efficiency of Courts. *Omega - International Journal of Management Science*, (10): 401-411.

Pedraja-Chaparro, F. and Salina-Jiménez, J. (1996). An assessment of the efficiency of Spanish courts using DEA. *Applied Economics*, (28): 1391-1403.

Schneider, M.R. (2005). Judicial career incentives and court performance: an empirical study of the German Labour Courts of Appeal. *European Journal of Law and Economics*, (20): 127-144.

Torre, A. (2003). The impact of court delays on the prosecutor and the defendant: An economic analysis. *European Journal of Law and Economics*, (16): 91-111.

Tulkens, H. (1993). On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts, and urban transit. *Journal of Productivity Analysis*, (4): 183-210.

Vereeck, L. and Muhl, M. (2000). An economic theory of court delay. *European Journal of Law and Economics*, (10): 243-268.