Length of trials in the Italian Judicial System: An Efficiency Analysis by Macro-Area

Elisa Fusco, Martina Laurenzi, Bernardo Maggi*

Department of Statistical Science, University of Rome La Sapienza, Italy
Insight service, Logista Italia
Department of Statistical Science, Faculty of Engineering of Information, Informatics and Statistics, University of Rome La Sapienza, Italy

Abstract

In recent years, the Italian judicial system has been at the center of both the political debate and policy actions aiming at modifying the territorial structure and the organization of the courts as well as the procedural processes. The measures adopted concerned the reorganization of the magistrates career and the reform of judicial districts.

Despite the several reforms, the Italian judicial system does not reach yet the European standards, principally for the so called *magistrate-duration procedures binomial*, according to which the number of magistrates is above the European average level and the time of legal trials is too long compared with most European countries. Hence, performance and efficiency are worthy of attention in this field.

Here, the territorial displacement of efficiency for the Italian judicial districts is studied using a *DEA* approach followed by a spatial analysis consisting in a PCA for macro-area where the problem of long trials is taken into account. We assess on the geographical characterization of both productivity and expenses of the judicial sector and provide indications in order to improve efficiency according to the dimensions of the courts.

Keywords: Judicial efficiency, Data envelopment analysis, Length of trials, Geographical Analysis, Principal Component Analysis

---

*Corresponding author at bernardo.maggi@uniroma1.it.

Email addresses: efusco86@gmail.com (Elisa Fusco), martina.laurenzi@gmail.com (Martina Laurenzi), bernardo.maggi@uniroma1.it (Bernardo Maggi)

Length of trials in the Italian Judicial System: An Efficiency Analysis by Macro-Area

1. Introduction

In recent years, the Italian judicial system has been at the center of both the political debate and administrative choices aiming at modifying the territorial structure and the organization of the courts as well as the procedural processes. Important measures have regarded both the reorganization of the magistrates career (Mastella's reform - Law n. 111 of July 30, 2007) and the reform of judicial districts (Decree Law of 7 September 2012, n. 155 and 156) that reconsidered the territorial organization of the judicial offices through the abolition of 31 tribunal and prosecutor offices, 220 separate offices and 667 peace courts.

However, despite the consistent measures adopted, the Italian judicial system does not reach yet the European standards, principally for the so called *magistrate-duration procedures binomial*. In fact, as pointed out by the European Commission for the Efficiency of Justice (CEPEJ) and according to the European Judicial System database (see Bianco *et al.*, 2007), Italy has a number of magistrates well above the European average as well as the length of trials. Then, performance and efficiency are worthy of attention in such a context.

Our paper analyzes the efficiency of Italian judicial districts, using a *Data Envelopment Analysis* (DEA - Charnes *et al.*, 1978) approach followed by a spatial analysis consisting in a *Principal Component Analysis* (PCA-Manly, 1994) by macro-area where the problem of long trials is taken into account.

Although the methodology proposed fits particularly with the Italian judicial system inasmuch trials are highly length conditioned, it is also extendable to the other judicial systems when affected by such a problem, or even more in general when trials may be categorized according to other specific cogent problems.

The quality and the quantitative amount of statistical information about the Italian judicial system are extremely poor and the studies till now developed recurred usually both to proxies and *ad hoc* methods in order to account for relevant variables such as the length of trials (see Giacomelli & Menon, 2013 and Peyrache & Zago, 2016). Our database, provided by the Italian Ministry of Justice (Ministero di Grazia e Giustizia) and by the Higher Judiciary Council (Consiglio Superiore della Magistratura), allows to shed lights on the judicial efficiency related to this aspect. In particular, we assess on the structure and expenses of the judicial sector across space taking into account of the lengths of trials. Moreover, we evaluate efficiency according to the dimensions of the courts in order to draw implications for enlarging offices or dividing the existing ones in the macro-areas considered.

The paper unfolds as follows. In Section 2, we explain our contribution to the literature of the efficiency analysis for the justice sector. In Section 3 we present our approach from a technical point of view. Section 4, describes data and variables. Section 5 comments on the empirical models adopted. In Section 6, results concerning the Italian Ordinary courts efficiency are provided. In Section 7, the efficiency analysis of the Appeals courts completes the study of the Italian judicial system for the two degrees available. Section 8 concludes.

2. Theoretical foundations

As extensively surveyed in Liu *et al.*, 2013, efficiency analysis has been widely applied to public sector in a broad sense. However, in the justice sector few studies has been done. This is particularly true for Italy where, as generally recognized, the justice sector is strategic for the improvement of productivity and growth at national level. On this point, the "Doing Business" report, World Bank, 2010 is particularly effective in emphasizing the positive correlation between inefficiency and the length of trials and, yet, its negative effect on the level of productivity. Such an analysis is prevalently descriptive but, nonetheless, calls for a more rigorous treatment of this problem.

With this research we intend to close the gap in the literature of efficiency in the justice sector concerning the interaction between inefficiency and an excessive length of trials: the study of courts inefficiency implies *per se* a geographical analysis, here we compare courts of different areas in order to understand if inefficiency is due to *scale* problems or to *technical* inefficiency which, in its turn, may be due to a wrong distribution of resources, performing in a different way depending on the length of trials, or to a low productivity by length of trials. Hence, we perform comparisons among courts at length of trials parity conditions and assess in favor of a change in the court dimensions once there emerges that there is little to do in terms of inefficiency pertaining inputs (basically magistrates).

The geographical comparisons we do are at regional and macro level in order to reach some conclusions for an improvement of the efficiency in the justice sector also at national level, which is crucial for determining the country's capacity of enhancing investments and increasing production, as argued in the above mentioned report. More specifically, Ordinary and Appeals courts will be examined not only with DEA

techniques, but we also disentangle the courts productivity (overall productivity) into ones (ratios) specialized in the several lengths of trials and analyze the criticality of the overall productivity due to inefficiency of the specific ratios or to a wrong distribution of resources through trials characterized by different lengths.

Lewin *et al.* (1982) conducted for first in the sector under question a research on the state of North Carolina. They examine the criminal sector using the CRS hypothesis. Kittelsen & Førsund (1992) analyze scope economies of the Norwegian courts for searching optimal size. Pedraja-Chaparro & Salinas-Jimenez (1996) examine the judicial courts of Spain, and, analogously to Yeung & Azevedo (2011), evaluate the personnel efficiency. Deyneli (2012) adopts a two stage analysis in order to evaluate how efficiency affects the magistrates compensation. Santos & Amado (2014) consider the largest number of outputs in this field for Portugal. They control such a variety with appropriate weights calculated on the base of the trade-off between outputs.

As for the Italian case, Antonelli & Marchesi (1999) pioneered the efficiency problem of justice and, with a translog parametric approach, find that the dimension of the judicial districts is suboptimal compared with the number of judges and courts. Then, they conclude that an efficiency improvement would need some combination of small courts compatibly with the economies of specializations. Also Finocchiaro Castro & Guccio (2014) use judicial districts data, developing a two stage procedure - DEA plus bootstrap - corroborated by the Wagner & Shimshak (2007) methodology. They conclude that demand factors are relevant for improving judicial efficiency and suggest to specialize the judicial system in the fields mostly requested by trials. In both cases the geographical dimension is the judicial district, which helps evaluate the judicial system efficiency at macro level. Peyrache & Zago (2016) perform a macro analysis as well, though moving from micro data appositely clustered at different aggregated levels. They find the directional distance function (DDF) to be an appropriate methodology in that allows for different aggregation levels of inefficiency sources for the whole industry. More specifically, they adopt discretionary weights on outputs in order to condition the efficiency results on the time required to resolve the trials.[1]

3. Strategy of Analysis

In the present research we intend to study the efficiency of the Italian judicial system at both judicial district and macro-area level. We use judicial district data in line with the strategy of Finocchiaro Castro & Guccio (2014) and take into special account the time required to resolve trials since it is one of the main causes of the claimed inefficiency in the judicial sector.

Our distinguishing feature depends on the peculiarity of our database, which is detailed by the length of each trial considered (civil or penal). From one hand, this gives the possibility to conduct a DEA analysis that accounts properly for such a characteristic by weighting each process (output variable) using the time employed for its resolution, instead of adopting discretionary weights. In practice, the output represented by the number of trials is penalized by weights representing the lengths of trials on the presumption that the longer the time required to resolve the trial the less efficient is the production process and, consequently, the lower should be the evaluation of output when compared with the other DMUs for assessing efficiency (like in the DDF case). From the other hand, the length of trials information allows for addressing geographical comparisons of productivities at parity conditions in such a respect. In particular, we study the effect of resources distribution dedicated to outputs with different duration, which qualifies more into deep the results on the offices dimensions, and indicates the policy actions to undertake in terms of offices restructuring, such as dividing or creating offices in the macro-areas considered.

Another peculiarity of our approach is that we take as inputs also the interception expenses besides, as usual, the magistrates.

More in details, we start with a more traditional first step, consisting in the application of the DEA methodology to judicial districts (our decisional making units, DMUs), and check the robustness of the results obtained.

As a second step, we consider two kinds of indexes representing productivity by magistrate and expenditure with the aim of deepening the spatial comparison between geographical macro-areas (NUTS1) with respect to the problem of the lengths of trials. These indexes are obtained from the related DMUs (NUTS2). Of course, the deriving implications of such an analysis, in terms of mergers of diversifications of judicial offices, have valuable consequences also at national level (NUTS0) in terms of growth and production.

In order to do comparisons at parity conditions of lengths of trials, we conceive overall indexes for each macro-area as weighted average of productivity, or expenditure, ratios specified for each time required to

---

[1] Also Falavigna *et al.* (2015) use a DDF approach, however they consider the case of tax courts, which is quite different from that one of Ordinary and Appeal tribunals for civil and penal procedures.

conclude the process (the typical output) with weights given by the percentage of magistrates (the typical input), or interception expenses, employed for that length of trial.

As for the first step, the DEA methodology (Charnes *et al.*, 1978) is a non-parametric frontier estimation for the efficiency analysis of a decisional making unit through the comparison with a non-parametric production boundary (*best-practice frontier*), using suitable linear programming techniques.

The shape of the frontier is related to the concept of *returns to scale* deriving from the restrictions assumed on the characteristics of the production set, mainly the convexity necessary for identifying a frontier.

Technically, in a cross-sectional setting, we have the case of N DMUs (public authorities), $i = 1, ..., N$, using P inputs, $\mathbf{x}_i = \{x_1, x_2, ..., x_P\}'$, $p = 1, ..., P$, to provide Q services, $\{y_1, y_2, ..., y_Q\}'$, $q = 1, ..., Q$. The data sample can be indicated with $\chi = (\mathbf{x}_i, \mathbf{y}_i)$ and all the combinations of inputs and outputs - such that $\mathbf{y}_i$ can be produced with the input vector $\mathbf{x}_i$ - represent the *production set* $\Psi(\mathbf{x}, \mathbf{y}) = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \in \mathbb{R}_+^P, \mathbf{y} \in \mathbb{R}_+^Q$ is feasible$\}$. With DEA approach the production set $\Psi(\mathbf{x}, \mathbf{y})$ is assumed to be free disposal and convex.

The first proposed DEA model was input-oriented -aiming at minimizing inputs in order to obtain a given quantity of output- with *constant returns to scale* (CRS). For the generic DMU and for each input and output, the attainable set is given by $\Psi_{DEA} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{P+Q} | x_p \geq \sum_{i=1}^{N} \lambda_i x_{i,p}, y_q \leq \sum_{i=1}^{N} \lambda_i y_{i,q}$ for λ such that $\lambda_i \geq 0\}$ and the efficiency score can be obtained by solving the following linear program:

$$\min_{\theta, \lambda} \quad \theta$$
$$\text{s.t.} - y_q + \sum_{i=1}^{N} \lambda_i y_{i,q} \geq 0$$
$$\theta x_p - \sum_{i=1}^{N} \lambda_i x_{i,p} \geq 0 \tag{1}$$
$$\lambda_i \geq 0, \forall q = 1, ...Q, p = 1, ..., P$$

where $\theta \leq 1$ is the efficiency score to be obtained (equal to 1 for the DMUs lying on the frontier) and λ is the optimal weights vector.

An important extension of this approach was proposed in 1984 by Banker *et al.* (1984) that generalized the original DEA model allowing for the production function to exhibit increasing, constant, or diminishing returns to scale. The so called *variable returns to scale* model (VRS) modifies the CRS model by adding a further constraint on weights:

$$\min_{\theta, \lambda} \quad \theta$$
$$\text{s.t.} - y_q + \sum_{i=1}^{N} \lambda_i y_{i,q} \geq 0$$
$$\theta x_p - \sum_{i=1}^{N} \lambda_i x_{i,p} \geq 0 \tag{2}$$
$$\sum_{i=1}^{N} \lambda_i = 1$$
$$\lambda_i \geq 0, \ \forall q = 1, ...Q, \ p = 1, ..., P$$

VRS provides a decomposition of CRS Farrell efficiency into two components that we are going to examine: one related to scale efficiency and the other one to technical efficiency.

Another aspect we take into consideration is the detection of slacks. In fact, it often happens that the piece-wise linear form of the frontier determined with DEA might generate weakly efficient DMUs (units located in sections of frontier parallel to axes) causing multiple optimal solutions and so categorizing some units as benchmark of efficiency even if this is not true (further inputs reduction is possible at a given output). In such a case a second stage (see Ali & Seiford, 1993), optimizing the slack variables[2] (Harrison *et al.*, 2004), is required in order to determine the possible non-zero slacks and the quantity of inputs to be reduced to reach the frontier. A more rigorous definition of technical efficiency could be provided basing on Koopmans (1951): a firm is technically efficient only if operates on the frontier, and so all associated slacks are zero (Coelli *et al.*, 2005; Cooper *et al.*, 2007).

---

[2]Slacks describe the magnitude of inefficiency and outlines the overuse of inputs or the underproduction of output.

Using DEA leads to some notable advantages: the possibility to work well with small samples and to handle with multiple inputs and outputs without assuming a functional form of the frontier;[3] the ease identification of the scale efficiency; the possibility to compare DMUs against an observed peer or a peer group (benchmark units) in the dataset.

After having performed the DEA approaches (1) and (2), we apply the Order-m method in order to check the robustness of our results. In fact, the main disadvantage of the DEA method is the high influence of outliers and extreme data points on the best-practice frontier enveloping all the observations in the sample. Cazals *et al.* (2002), with the aim to overcome this drawback, proposed a more robust nonparametric estimator of the frontier sensitive to the magnitude of the extreme observations but, simultaneously, resistant to their influence in case they are possible outliers.[4] In accordance with such a methodology, we compare a DMU with a set of size $m < N$, given a level of $\mathbf{Y}_i$, and re-sample for several times.[5] The random set of order m can be defined as follows:[6]

$$\Psi_m = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{P+Q} | \mathbf{X}_i \leq \mathbf{x},\ \mathbf{Y}_i \geq \mathbf{y},\ i = 1, ..., m\}, \tag{3}$$

Then, the benchmark φ^m is identified with the expected minimal value of inputs for m units producing at least $\mathbf{Y}_i \geq \mathbf{y}$:

$$\varphi^m = \mathbb{E}\left[min(\mathbf{X}_1, ..., \mathbf{X}_m | \mathbf{Y}_i \geq \mathbf{y})\right] \tag{4}$$

Consequently, the obtained frontier envelopes a sub-sample of data keeping out the "super-efficient" units, and, in practice, the unit efficiency score θ is calculated as the arithmetic average of the ones relative to the numbers of the re-samplings.

From the application of this method we validate the DEA methodology used.[7]

Once concluded the first step of our analysis, we go on with the second step for better qualifying comparisons across space and to corroborate further the findings of the first step. It consists in conceiving the overall indexes of productivity (expenditure) as weighted averages of ratios specialized in the different lengths of trials for $t = 1, ..., T$, where T is the maximum time required to conclude the trials, and the weights are the percentages of inputs (expenditure) specific for the same t

$$Index^{AREA_j} = \frac{\sum_{t=1}^{T} y_t^{AREA(j)}}{\sum_{t=1}^{T} x_t^{AREA(j)}} = \sum_{t=1}^{T} \left(\frac{y_t^{AREA(j)}}{x_t^{AREA(j)}}\right) \frac{x_t^{AREA(j)}}{\sum_{t=1}^{T} x_t^{AREA(j)}}. \tag{5}$$

We consider two different strategies for comparing areas at *parity conditions* per each length of trial: 1) either using the same weights, PE_t (*population effect*), 2) or the same specific ratios, CE_t (*coefficient effect*). In the first case, differences in performance between areas - in terms of productivity or expenditure index - depend on the specific ratios, whist, in the second case, they are due to the distribution of magistrates, or interception expenses, on processes with different lengths.

For the $AREA_j$ the two effects are represented by:

$$PE_t^{AREA(j)} = \frac{x_t^{AREA(j)}}{\sum_{t=1}^{T} x_t^{AREA(j)}}; CE_t^{AREA(j)} = \frac{y_t^{AREA(j)}}{x_t^{AREA(j)}}, \tag{6}$$

The indexes corresponding to expression (6) are:

Population index

$$Index_{PE}^{AREA_j} = \sum_{t=1}^{T} PE_t^{AREA(\hat{j})} \cdot \frac{y_t^{AREA(j)}}{x_t^{AREA(j)}} \tag{7}$$

Coefficient effect index

$$Index_{CE}^{AREA_j} = \sum_{t=1}^{T} CE_t^{AREA(\hat{j})} \cdot \frac{x_t^{AREA(j)}}{\sum_{t=1}^{T} x_t^{AREA(j)}} \tag{8}$$

---

[3]The same advantage with the possibility of an inferential analysis is also possible with a semi-parametric approach. However, as argued by Maggi & Guida (2011), due to the large sample requirements such an approach is not applicable here.

[4]Such an estimator is termed as to "*partial frontier*" and is well near the upper (or lower) boundary. Daraio & Simar (2005) extended this method to the multivariate case.

[5]In other words, after fixing the level of output $\mathbf{Y}_i$ we consider m i.i.d. random variables $\mathbf{X}_i, i = 1, ...m$ generated by the conditional distribution of $\mathbf{x}_i$.

[6]The convexity is not assumed since it is derived from the Free Disposable Hull (FDH) approach (Deprins *et al.*, 1984).

[7]Results available upon request.

where $\hat{j}$ indicates the geographic area with weights (population) or specific ratios (coefficient effects) kept equal across all the j areas for making comparisons at parity conditions. As said above, in order to have a more representative term of comparison at macro level, we use national weights and specific ratios.

We apply these steps in order to answer the questions if the long legal trials in Italy depend on inputs or scale dimensions inefficiency (first step), and if there might exist improvements in productivity due either to resources distribution or to specific productivity ratios by lengths of trials, which may support appropriate policies of industrial configuration in terms of efficiency of offices dimension in the justice sector (second step). Furthermore, the second step allows evaluating how productivities by duration and the corresponding distribution of resources affect, in a separate way, the overall productivity. Therefore, from the DEA analysis one can assign IRS or DRS technologies and with the PCA on the above indicators one can judge about the organization of resources and on the impact of the specific productivities.

A relevant aspect of such an analysis is that it is of general validity irrespective of the judicial system of reference, and also that it may be exploited for other possible characterizations of output and resources of interest, different from that one of lengths of trials considered here.[8]

4. The data

Data on the Italian justice system have been provided by the Ministry of Grace and Justice-D.G. Statistics and Organizational Analysis (*Ministero di Grazia e Giustizia - Direzione Generale di Statistica e Analisi Organizzativa*, https://webstat.giustizia.it/_layouts/15/start.aspx#/SitePages/Home.aspx). The dataset is subdivided in Ordinary courts (*Tribunali Ordinari*) and Appeals courts (*Corti d'Appello*), for the period 2005-2011. However, after a descriptive analysis, we apply the methodology previously expounded to the more restricted period 2009-2011 when data are available for all the geographical specifications considered. A valuable contribution in collecting data for this sub-period has been given by the Higher Judiciary Council (*Consiglio Superiore della Magistratura*), which provided us with the series of magistrates, *judges* or *prosecutors*, offices and territorial districts.

The deployment of judicial offices in Italy is rather capillary. Actually, there exist 179 magistrates of peace (for minor litigations), 136 Ordinary courts, 26 districts with an Appeals court per each, and one Supreme court placed in Rome.

In order to frame the production process of the judicial system, we first define three kinds of outputs according to the time dedicated to them and to the category of trial to which they belong, civil or criminal. In particular, we distinguish between *enrolled* trials, those just began, *unresolved* trials, those not yet ended and *resolved* trials, those concluded. Secondly, as input variables, we consider the number of magistrates (according to the functional task) and the wiretapping expenses.

Actually, the Italian Institute for Political, Economic and Social Studies surveyed that the 62% of the Italians complain about the long time required for the trials resolution. EURISPES (2009) and World Bank (2010)) indicated Italy as the country with the longest time of trials resolution within OCSE. For this reason, the output variables will be also considered as weighted by the time necessary to conclude the trial. In particular, the longer is the time required the more the number of resolved trials will be "reduced" by the weight adopted.

A relevant question is if such long processes of justice are due to an inefficiency problem of inputs management of the judicial system or rather to the distribution of magistrates (or interception expenses) over the processes with different durations. We cope with such a question by considering the indicators of input-efficiency and returns to scale together with the geographical PCA analysis for the index productivity ratios.

Tables 1-3 report the descriptive statistics for the two degrees of judgment and Prosecutor's office.

---

[8]Furthermore, it may be conceived as a way to cope with the DEA analysis when outputs and inputs are too many, as it would be in case of explicit consideration of such variables by lengths of trials. For alternative DEA dealing with this problem see Doyle & Rodney (1994), Andersen & Christian (1993), Mehdi *et al.* (2015).

Table 1: Ordinary courts descriptive statistics

Year	Number (National)	Mean	Median	Min	Max	Std.Dev.
		Input				
		Magistrates (total)				
2009	**6,774**	260.54	188.5	42	814	208.91
2010	**6,553**	252.04	165	42	817	209.8
2011	**6,554**	252.08	165	42	817	209.77
		Judges				
2009	**4,897**	188.35	136.5	28	610	158.26
2010	**4,910**	188.85	136.5	28	613	158.64
2011	**4,911**	188.88	136.5	28	613	158.61
		Prosecutors				
2009	**1,877**	72.19	52.5	14	204	51.41
2010	**1,883**	72.42	53.5	14	204	51.35
2011	**1,883**	72.42	53.5	14	204	51.35
		Output				
		Civil resolved trials				
2009	**2,800,435**	107,709.04	84,092.00	15,129	333,878	85,748.69
2010	**2,742,081**	105,464.65	81,250.00	14,515	325,187	85,172.17
2011	**2,702,744**	103,951.69	78,053.00	13,222	333,126	86,554.30
		Weighted				
2009	**2,638,701**	99,097.36	79,762.41	15,059	289,254	77,963.06
2010	**2,584,345**	97,104.21	77,350.60	14,456	284,615	77,582.73
2011	**2,552,871**	98,187.34	74,723.51	13,160	304,158	78,943.68
		Penal resolved trials				
2009	**1,282,581**	49,330.04	37,029.50	7,325	170,558	38,386.30
2010	**1,293,001**	49,730.81	37,117.00	8,592	167,522	39,182.95
2011	**1,265,022**	48,654.69	35,498.50	8,336	149,959	39,090.62
		Weighted				
2009	**1,196,007**	44,893.83	35,633.94	7,306	142,336	32,779.09
2010	**1,208,552**	45,574.15	35,900.69	8,573	144,029	34,450.00
2011	**1,184,284**	45,549.40	34,108.87	8,305	129,533	34,181.84

The statistics are calculated over 26 districts, physical data are in units, monetary data are in euro

Table 2: Appeal courts descriptive statistics

Year	Number (National)	Mean	Median	Min	Max	Std.Dev.
		Input				
		Magistrates (total)				
2009	**1,569**	60.35	43	15	196	48.3
2010	**1,573**	60.5	43	15	196	48.22
2011	**1,574**	60.54	43	15	196	48.2
		Judges				
2009	**1,279**	49.19	34.5	11	171	41.94
2010	**1,281**	49.27	34.5	11	171	41.85
2011	**1,282**	49.31	34.5	11	171	41.87
		Prosecutors				
2009	**290**	11.15	8.5	4	26	6.82
2010	**292**	11.23	9	4	26	6.78
2011	**292**	11.23	9.5	4	26	6.8
		Output				
		Civil resolved trials				
2009	**141,112**	5,427.38	3,991.00	1,695	26,348	5,561.05
2010	**149,838**	5,763.00	3,957.50	1,234	29,288	6,374.90
2011	**148,839**	5,724.58	4,058.00	1,020	27,574	5,798.10
		Weighted				
2009	**128,125**	4,927.90	3,932.53	1,688	19,691	4,200.96
2010	**135,681**	5,218.49	3,890.97	1,229	22,183	4,893.22
2011	**135,583**	5,214.71	3,926.03	1,015	20,629	4,452.26
		Penal resolved trials				
2009	**76,751**	2,951.96	2,198.00	284	10,217	2,370.90
2010	**81,014**	3,115.92	2,593.00	380	10,067	2,267.46
2011	**79,171**	3,045.04	2,839.00	359	8,029	2,080.04
		Weighted				
2009	**72,914**	2,804.38	2,048.81	281	9,399	2,191.53
2010	**76,672**	2,948.91	2,468.90	377	9,190	2,059.02
2011	**74,726**	2,874.07	2,715.44	359	7,100	1,870.66

The statistics are calculated over 26 districts, physical data are in units.

Table 3: Prosecutor's Office

Year	Number (National)	Mean	Median	Min	Max	Std.Dev.
			Interception expenses			
2009	255,059,249	9,809,971.12	26,532,808.00	358,908	47,926,659	13,240,640.93
2010	237,041,484	9,116,980.15	25,122,030.00	374,359	39,670,400	11,007,736.30
2011	225,987,187	8,691,814.88	20,993,098.00	239,723	36,279,033	10,245,836.63

The statistics are calculated over 26 districts and are referred to Ordinary courts, Juvenile courts and Appeal courts, monetary data are in euro

As it is easy to see from the comparison between the mean and the median in these tables, a large proportion of magistrates is located in few big districts (the skewness is to the right hand side of the distribution since media>median). Instead, interception expenses are also concentrated in small districts. This is because interception expenses require fixed costs, which are independent of the districts dimension (the skewness is to the left hand side of the distribution since media<median). The differences among districts are quite substantial as described by the consistent dimension of the standard deviation.

As for output, we have reported the resolved trials both weighted by the time necessary to conclude the process and not. Also in this case data reveal more concentration in the big districts. However, the asymmetry towards the right hand side of the distribution (media > then median) diminishes for the weighted variable, which describes that big districts suffer more from long time for trials resolution than small ones.

5. DEA empirical models

The choice of an input oriented model, and so the identification of a benchmark based on both the minimum number of magistrates and wiretapping expenditures, is due to the exogenous nature of the number of processes.

The first step of our efficiency analysis consists in a DEA, separated for Ordinary and Appeals courts, carried out in different consequential phases.[9] (*i*) In a first phase, a specific analysis on the returns to scale has been conducted by means of three inferential tests (see Banker, 1993, 1996). (*ii*) In a second phase, having as a term of comparison the VRS hypothesis, a CRS-DEA model (equations (1) and (2) in Section 3) has been chosen in order to measure the global efficiency and to identify benchmark districts. (*iii*) In a third phase, we investigate on the presence of slacks. A DEA slack model has been used to find possible non-zero slacks and to measure the additional inefficiency and so the further reduction of inputs needed to reach the optimal mix.

Three different inputs and outputs approaches have been investigated. In particular, *Model I* considers as inputs the total number of magistrates and the interception expenses and as outputs the number of civil and penal procedures (2 inputs and 2 outputs); *Model II* splits the number of magistrates into *judges* and *prosecutors* (3 inputs and 2 outputs); in *Model III* the number of civil and penal procedures is weighted for the time average stock in order to take workload and processing times into account (3 inputs and 2 weighted outputs).

6. Ordinary courts analysis

6.1. Ordinary courts DEA results

In this subsection the results for Ordinary courts are presented. Firstly, the Kolmogorov-Smirnov test and the test on the equality of the VRS and the CRS average scores are reported. In particular, in Table 4 the null hypothesis of constant returns to scale is tested against the alternative hypothesis of variable returns to scale by using the non-parametric test of Kolmogorov-Smirnov for the three years and models considered. In all cases the null hypothesis is accepted (p-values > 0.05) and so constant returns to scale setting is the most appropriate for representing technology.

Table 4: Ordinary courts efficiency scale test by year and model

	Kolmogorov-Smirnov test					
	CRS 09 vs VRS 09		CRS 10 vs VRS 10		CRS 11 vs VRS 11	
	Test statistic	P-value	Test statistic	P-value	Test statistic	P-value
Model I	0.231	0.493	0.192	0.722	0.269	0.303
Model II	0.231	0.493	0.192	0.722	0.192	0.722
Model III	0.269	0.303	0.231	0.493	0.192	0.722

---

[9] As mentioned above, the robustness of this analysis has been confirmed by the Order -*m* approach.

This finding is confirmed in Table 5 by a second test on the equality of the VRS and the CRS average scores (test statistics lower than 1.960).

Table 5: Ordinary courts VRS vs CRS score mean test by year and model

	Test of mean equality		
	CRS 09 vs VRS 09	CRS 10 vs VRS 10	CRS 11 vs VRS 11
	Test statistic	Test statistic	Test statistic
Model I	1.118	0.979	1.077
Model II	1.039	0.907	0.932
Model III	1.292	1.095	1.074

Finally, Table 6 shows the results of the test on the changes in CRS efficiency scores over time using a Friedman's test. For all models the test accepted the null hypothesis, meaning that the average efficiency of Ordinary courts does not differ significantly over years, remaining around 70%.

Table 6: Ordinary courts efficiency changes test by model - Friedman test

	Friedman test	
	Test statistic	P-value
Model I	1.432	0.489
Model II	1.435	0.488
Model III	2.193	0.334

Table 7 summarizes the yearly efficiency scores obtained with CRS-DEA method for *Models I*, *II*, *III*.[10]

Table 7: Ordinary courts CRS efficiency score by model - years 2009-2010-2011

District	*Model I*			*Model II*			*Model III*		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
ANCONA	0.906 s	1.000	0.915	0.933 s	1.000	0.937 s	0.953 s	1.000	0.964 s
BARI	0.942 s	0.900	0.914	0.944 s	0.910 s	0.932 s	0.963 s	0.905 s	0.814 s
BOLOGNA	0.918 s	1.000	1.000	0.940 s	1.000	1.000	0.881 s	1.000	1.000
BRESCIA	0.867 s	0.832	0.848	0.873 s	0.865 s	0.889 s	0.890 s	0.904 s	0.917 s
CAGLIARI	0.755 s	0.928	0.762 s	0.757 s	0.952 s	0.876 s	0.754 s	0.958 s	0.851 s
CALTANISSETTA	0.244 s	0.211 s	0.225 s	0.256 s	0.221 s	0.232 s	0.259 s	0.221 s	0.233 s
CAMPOBASSO	0.938 s	1.000	1.000	0.959 s	1.000	1.000	1.000	1.000	1.000
CATANIA	0.526 s	0.394 s	0.400 s	0.535 s	0.398 s	0.409 s	0.527 s	0.392 s	0.409 s
CATANZARO	0.444 s	0.486	0.435	0.449 s	0.495 s	0.436 s	0.441 s	0.491 s	0.425 s
FIRENZE	0.868 s	0.778	0.739	0.873 s	0.783 s	0.751 s	0.852 s	0.780	0.763 s
GENOVA	0.646 s	0.663	0.668	0.659 s	0.665 s	0.669 s	0.667 s	0.700	0.693 s
L'AQUILA	1.000	1.000	0.996 s	1.000	1.000	1.000	1.000	1.000	1.000
LECCE	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MESSINA	0.432 s	0.475	0.532	0.454 s	0.497 s	0.550 s	0.458 s	0.504 s	0.552 s
MILANO	0.480 s	0.503	0.524	0.529 s	0.578 s	0.609 s	0.537 s	0.574 s	0.601 s
NAPOLI	0.671 s	0.672	0.588	0.718 s	0.744 s	0.656 s	0.615 s	0.673 s	0.586 s
PALERMO	0.291 s	0.304 s	0.315 s	0.308 s	0.320 s	0.340 s	0.298 s	0.313 s	0.331 s
PERUGIA	0.921 s	0.791	0.789	0.972 s	0.811 s	0.814 s	0.992 s	0.835 s	0.848 s
POTENZA	0.624 s	0.692	0.544	0.630 s	0.698 s	0.562 s	0.633 s	0.701 s	0.579 s
REGGIO CALABRIA	0.298 s	0.302 s	0.298 s	0.303 s	0.311 s	0.317 s	0.297 s	0.305 s	0.310 s
ROMA	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.894	0.880	0.827 s	0.896 s	0.896 s	0.870 s	0.936 s	0.934 s	0.894 s
TORINO	0.535 s	0.712	0.669	0.552 s	0.715 s	0.683 s	0.551 s	0.725 s	0.703 s
TRENTO	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TRIESTE	0.571 s	0.651	0.612	0.586 s	0.671 s	0.627 s	0.593 s	0.687 s	0.643 s
VENEZIA	0.808 s	0.778	0.811	0.821 s	0.780 s	0.816 s	0.826 s	0.799 s	0.831 s

Note: s=Presence of positive slacks.

Table 7 shows that in the considered years, under the hypothesis of constant returns to scale and for all models, 4 districts display as benchmarks in 2009 - Trento, L'Aquila, Lecce and Rome - and 6 in 2010-2011 with the inclusion of Bologna and Campobasso. The worst districts for all years and models are Reggio Calabria, Palermo and Caltanissetta with an efficiency score approximately equal to 0.3.

For the inefficient units, a further analysis has been carried out in order to investigate the possible presence of slacks. The "*s*" denotes the inability of districts to reach the efficient frontier with only an equiproportional reduction of the two inputs considered, which implies a necessary reorganization of activities and functions performed by the judicial offices. As an example, in 2009 almost all DMUs present inefficiencies and Ancona, among others, still in the presence of a high efficiency score, could even more increase it by contracting only one input.

---

[10]The Appendix A reports DEA-VRS efficiency score per each district and an Order-m test for evaluating the presence and the effect of extreme values. Also in this case the robustness of our results is confirmed.

The overall result of CRS would suggest that on average -or at national- level, and compared to output, the size of the judicial offices is appropriate. However, in order to single out the configuration of the individual DMUs, we compare the CRS results with those of VRS and perform, in the following Table 8, a return to scale analysis by district.

Table 8: Ordinary courts returns to scale by model - years 2009-2010-2011

District	Model I			Model II			Model III		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
ANCONA	IRS	**MPSS**	IRS	IRS	**MPSS**	IRS	IRS	**MPSS**	IRS
BARI	DRS	DRS	IRS	DRS	DRS	IRS	DRS	DRS	IRS
BOLOGNA	DRS	**MPSS**	**MPSS**	DRS	**MPSS**	**MPSS**	DRS	**MPSS**	**MPSS**
BRESCIA	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
CAGLIARI	IRS	DRS	DRS	IRS	DRS	DRS	IRS	DRS	DRS
CALTANISSETTA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
CAMPOBASSO	IRS	**MPSS**	**MPSS**	IRS	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**
CATANIA	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
CATANZARO	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
FIRENZE	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
GENOVA	DRS	IRS	IRS	DRS	IRS	IRS	DRS	IRS	IRS
L'AQUILA	**MPSS**	**MPSS**	DRS	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**
LECCE	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**
MESSINA	IRS	DRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
MILANO	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
NAPOLI	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
PALERMO	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
PERUGIA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
POTENZA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
REGGIO CALABRIA	IRS	IRS	IRS	IRS	DRS	DRS	DRS	DRS	DRS
ROMA	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**
SALERNO	IRS	IRS	IRS	IRS	IRS	IRS	DRS	DRS	IRS
TORINO	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
TRENTO	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**	**MPSS**
TRIESTE	DRS	IRS	IRS	DRS	IRS	IRS	DRS	IRS	IRS
VENEZIA	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS

Note: IRS=Increasing returns to scale, DRS=Decreasing returns to scale, MPSS=Most Productive Scale Size.

In Figure 1 we report for simplicity sake only the efficiency map of the most complete *Model III*.[11]

Figure 1: Ordinary courts returns to scale, Model III

(a) 2009   (b) 2010   (c) 2011
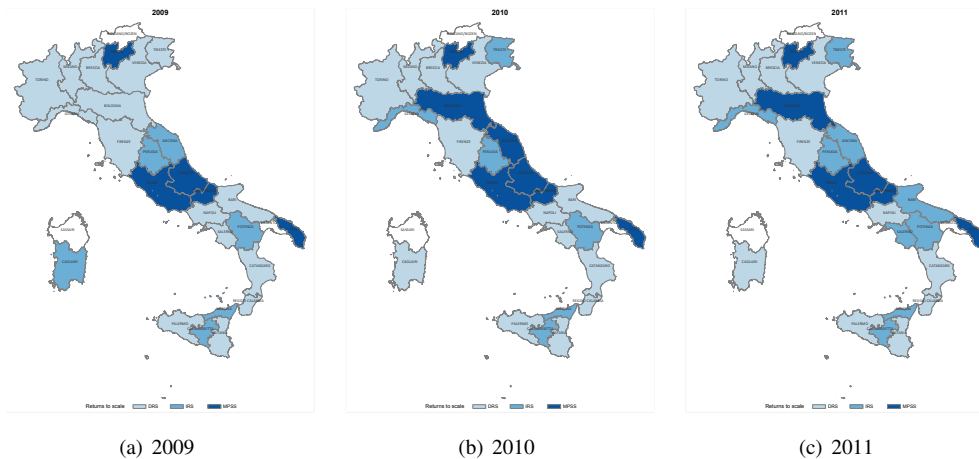
From Table 8 there emerges that the *most productive scale size* (MPSS) districts -in terms of resources optimization and size of activities- are not located in a specific area of Italy, like Trento, Rome, L'Aquila, Campobasso and Lecce. Furthermore, there is a diffused presence of DRS or IRS and, for some of them, technical efficiency increases from Models I and II to Model III, where output is reduced by weights that account for the lengths of trials.[12] Hence Model III exhibits a rise in the efficiency score for those districts that are less penalized, compared to the others, by amounts of output with large duration.[13]

---

[11]The others efficiency maps are reported in Appendix B.

[12]As an example for DRS districts see Milano and Torino whilst, for IRS districts, Napoli and, for some years, Firenze and Catanzaro.

[13]In such cases, it is likely that the large duration is due to a large scale demand problem which affects productivity ratio, rather

A straightforward conclusion would be that, in case of districts with an improvement in efficiency in Model III and DRS technology, a division should be suggested whilst, if districts are with IRS technology, larger offices should be required.[14] However, in order to support further such policies, we are going to perform another comparative evaluation of DMUs recurring to a PCA on indexes (8) and (7), by choosing national average of weights and specific coefficients. We presume that index (8) performs better for the DRS offices by denoting an increase in the overall productivity with respect to the real data, and that the same occurs for index (7) with IRS offices. In fact, in the above mentioned PCA the national average specific ratios are expected to be greater than those of DRS offices thus causing an increase in the productivity with respect to real data, and for the latter PCA the national average weights would be larger than the resources employed by the IRS offices, which brings about an index (7) larger than the real productivity.

6.2. Implications by macro-area

The presence of both DRS and IRS technologies for several Ordinary courts and the overall optimality of CRS asks for further analysis on productivity in order to reach more robust conclusions on the policy implications for the industrial configurations of offices.

In particular, we recur to a Principal Component Analysis (PCA - Manly 1994) appositely suited for underlying geographical differences in productivities of districts in different macro-areas. We initially do this with the actual variables and then use data derived by formulae (7) and (8).

The first PCA has the aim to propose a general view of the phenomenon and so it has been applied on total magistrates, interception expenses, civil and penal procedures, weighted and not. The results are similar for the three years as it can be seen in Figure 2. The first component depends positively on both output variables and total magistrates with a percentage of explained variance of about $77\% - 79\%$, against a $19\% - 22\%$ variance explained by the second component representing the interception costs.

---

than to an inefficiency problem.

[14]Of course in this reasoning it is supposed that the *scale* and *scope* economies, concerning the dimensions and the amount of output of the courts, should be exploited after understanding the problem of the inefficiency pertaining inputs *(technical* inefficiency) net of the length of trials effect.

Figure 2: PCA on total magistrates, interception expenses, civil and penal procedures defined and weighted

(a) 2009

(b) 2010

(c) 2011
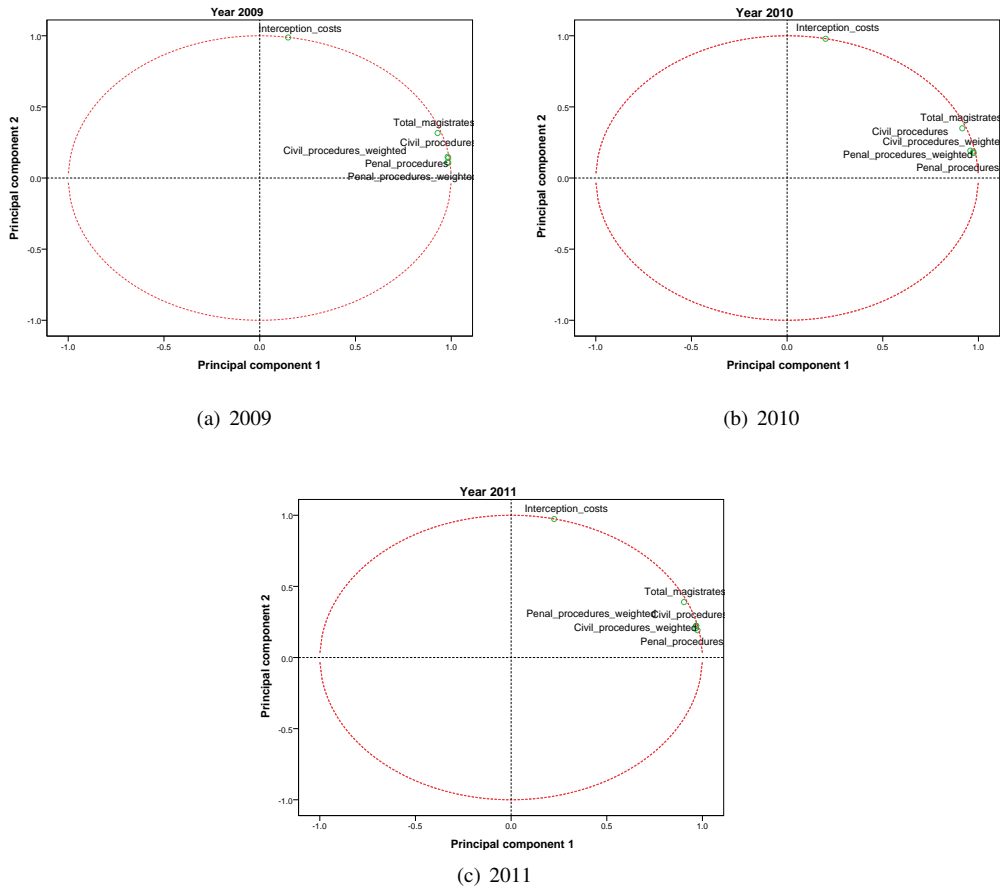
Moreover, by analyzing the individuals charts in Figure 3, it can be seen that most courts are distributed below the average level associated with the first component and around the average for interception expenses. Outliers are the tribunals of Reggio Calabria, Palermo, Milan, Naples and Rome. In particular, Reggio Calabria and Palermo exhibit high levels of interception expenditures and a number of penal and civil procedures below the average level. Differently, the districts of Milan and Naples are characterized by a high level of interception expenses and a level of penal and civil defined procedures and of magistrates employed above the average. Noteworthy, the representation of the Rome tribunal shows a level of interception expenditures and a number of defined procedures below and above the average of the other courts, respectively.

Figure 3: Individuals charts

(a) 2009

(b) 2010

(c) 2011
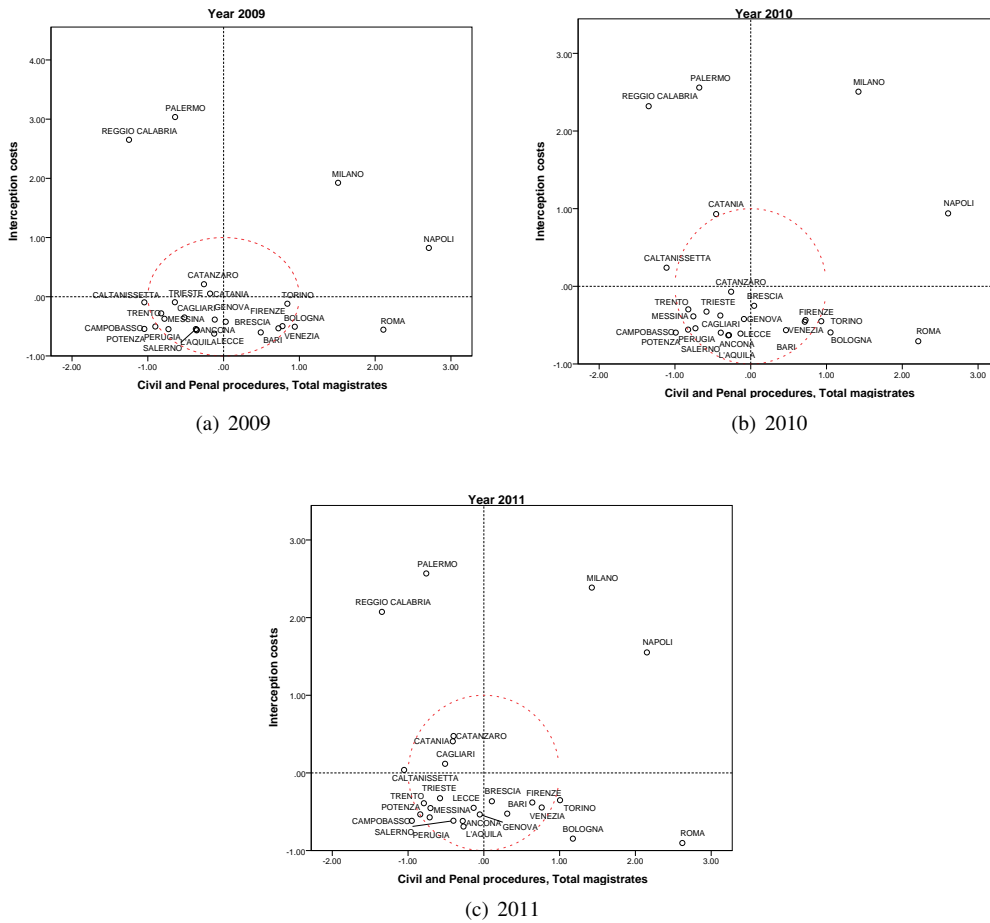
The second PCA focuses on measures of inputs performance (productivity) inside geographical areas. Specifically, we define the ratios of output variables (defined procedures in civil and penal cases) out of the total number of magistrates (assigned to each court) as structural indexes. An expenditure index is also defined as the ratio between outputs out of the interception expenditure.

We find very similar results for the geographical areas in the three years considered. Hence, for sake of simplicity, we present in Table 9 the last one. [15] The first component is linked to expenditures indexes and the second one to structural indexes.

The North-East of Italy is affected positively by the structural indexes and negatively by the spending indexes. This underlines that from one side the productivity of magistrates is high, compared to the average level (particularly due to Trento Ordinary court, coherently with the good performances identified by the DEA method), from the other side such good productivity is supported by high expenditures per unit of output. North-West and Center's geographic areas exhibit a medium-level of structural indexes but different scenarios in terms of spending indexes. In particular, North-West has a poor performance in terms of expenses whilst the Center has an output per expenditure well above the average. South shows a good performance of expenditures per unit of output and a ratio of output per magistrate below the mean. The geographic area of Isles has the worst output per expenses ratio with the most critical structural indexes. Then, Isles have the worst performance for both inputs (magistrates and expenses for interception) and North-East the best one followed by Center, North-West and South.

---

[15] More specifically, for major precision, we adopt the "rotated matrix".

Table 9: PCA rotated matrix (2011)

	Components (difference from the mean)	
Indexes	First	Second
Civil procedures/Total magistrates	0.287	0.887
Penal procedures/Total magistrates	0.137	0.964
Weighted Civil procedures/Total magistrates	0.238	0.920
Weighted Penal procedures/Total magistrates	0.105	0.963
Civil procedures/Interception expenditures	0.962	0.146
Penal procedures/Interception expenditures	0.967	0.148
Weighted Civil procedures/Interception expenditures	0.965	0.147
Weighted Penal procedures/Interception expenditures	0.955	0.169
Territorial area	First	Second
North-East	-0.264	0.622
North-West	-0.192	-0.054
Center	0.515	-0.027
South	0.367	-0.146
Isles	-0.497	-0.319

Finally, the PCA has been replicated in two different ways in order to account for the geographical effect of the areas considered on the DMUs and to make comparisons at parity conditions. As stated in formula (5), we consider per each geographic area the structure and expenditure indexes in terms of weighted average of the specific ratios referred to the several classes of time (days of average duration) required to complete the trials ($t = 1,..T$), indicated in the following Table 10. Then, in accordance with formula (6), we identify a *population effect* (PE_t) in order to control for the effect of the weights - *i.e.*, the geographical distributions of inputs - and to evaluate the role of specific ratios across areas, and *vice-versa* for the *coefficient effect* (CE_t).

Table 10: Population and Specific effects

	Component	
Time to complete the trials	Civil trials	Penal trials
very low	≤ 250	≤ 150
low	≤ 450	≤ 250
medium	≤ 650	≤ 350
high	≤ 850	≤ 450
very high	≥ 850	≥ 450

Therefore, starting with the population effect, we analyze if, *net* of the inputs distribution, some geographical differences with the PCA performed with actual data there occur. In that case, the only remaining explanatory effect would be that one of specific productivity.

From Figure 4, based on formula (7), it is possible to discover that isolating the effect of the specific productivity does not improve the performance condition of North East, North West improves little, Center and South improve, whilst Isles remain definitely in the same condition. Hence, differently from Center and South, what matters for a better performance of North East and North West is an appropriate distribution of resources (magistrates and interception expenses) according to the lengths of the trials. In other words, once blocking such an effect, the sole effect of specific productivity does not suffice to ensure a better performance of these areas compared with Center and South. As for Isles, their performance remains unaltered, meaning that the negative performance registered by the actual data is due to an overall poor productivity.

14

Figure 4: Individuals charts - population effect

(a) 2009

(b) 2010

(c) 2011
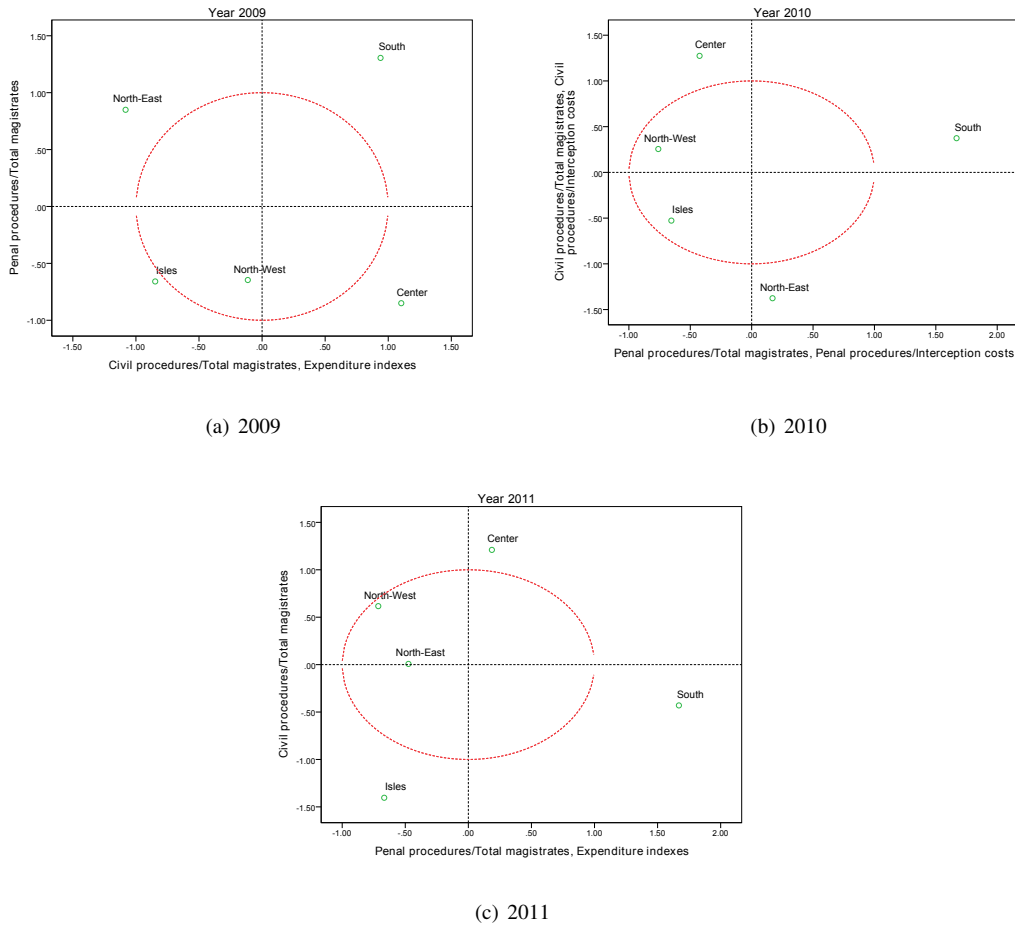
The results obtained by blocking the specific ratios in formula (8) are represented in Figure 5 and confirm those reached with the previous PCA. Specifically, compared with the results obtained with actual data, North East improves its performance, North West improves little, South worsens, Center turns into worse or better condition according to the year under consideration and the Isles seem to improve but only in the first two years.

Therefore, we may observe that the good performance of North East is essentially due to an efficient distribution of the resources according to the time necessary to conclude the trials and not much to the specific productivity of the used inputs. The same conclusion is valid also for North West, though in a more tenuous way. Differently, South is characterized by a good performance of the inputs but with an inefficient repartition through the lengths of the trials. Analogous situation is that one of Center, though less quantitatively relevant.

Figure 5: Individuals charts - coefficient effect

(a) 2009

(b) 2010

(c) 2011
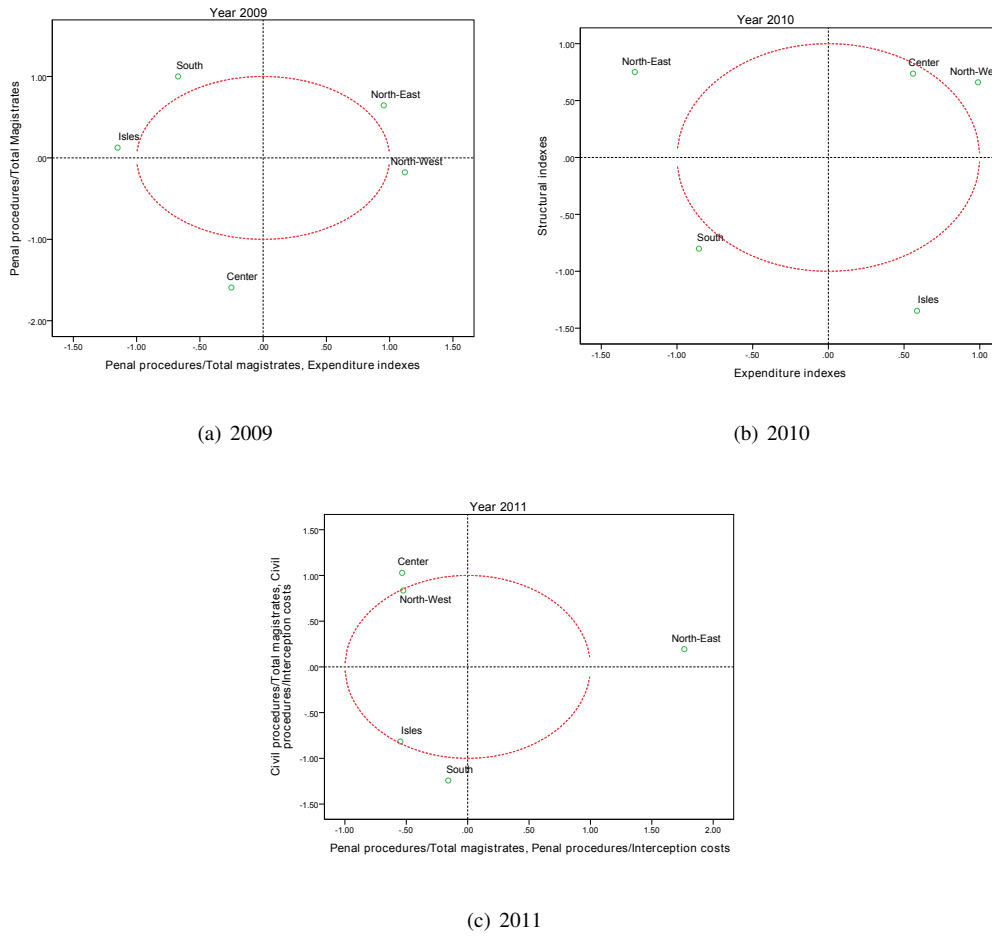
Here below, Table 11 reports the improvement (plus sign) or the worsening (minus sign) due to the population or the coefficient specific effects in comparison with the actual data as indicated by the first and second principal components. As said above, the first component (*PC*1) represents the expenditure index (interception expenditure productivity) and the second one (*PC*2) the structural index (magistrate productivity) .

Table 11: Geographic comparisons at parity conditions

Territorial area	Actual values		Population effect					
			2009		2010		2011	
	*PC*1	*PC*2	*PC*1	*PC*2	*PC*1	*PC*2	*PC*1	*PC*2
North East	-	+	-	+	+ ($\simeq 0$)	-	-	0
North West	-	-	- ($\simeq 0$)	-	-	+ ($\simeq 0$)	-	+
Center	+	-	+	-	-	+	+	+
South	+	-	+	+	+	+	+	-
Isles	-	-	-	-	-	-	-	-
	Actual values		Coefficient specific effect					
			2009		2010		2011	
	*PC*1	*PC*2	*PC*1	*PC*2	*PC*1	*PC*2	*PC*1	*PC*2
North East	-	+	+	+	-	+	+ ($\simeq 0$)	+
North West	-	-	+	- ($\simeq 0$)	+	+	-	+
Center	+	-	-	-	+	+	-	+
South	+	-	-	+	-	-	-	-
Isles	-	-	-	+ ($\simeq 0$)	+	-	-	-

Conclusively, from the analysis developed it seems that, looking at the South, and in minor part at the Center, and given the IRS technology, the remedy for improving the performance should be to increase resources with the scope of improving the distribution of magistrates and interception expenses by lengths of trials rather than merging offices, as at a first sight it might seem. As for the North (and some big cities of the South), since the distribution of inputs is efficient but the dimensions of offices are larger than the optimal level, it seems advisable to divide offices in order to face the large demand for justice but, at the

same time, to preserve the actual organization of resources through lengths of trials. The Isles suffer from a low productivity and also from a not efficient distribution of the inputs, which are both problems to face.

7. Appeals courts analysis

7.1. DEA Efficiency results

In order to analyze the two degrees of the Italian justice, the efficiency has also been studied for the Appeals courts (second degree of judgment).

In such a case, the Kolmogorov-Smirnov test in Table 12 seems to indicate the validity of the CRS hypothesis. However, this is not obtained for models I and II in 2010, which is also confirmed in Table 13 where the Test statistics are greater than 1.960.

Table 12: Appeals courts efficiency scale test by year and model

| | | *Kolmogorov-Smirnov test* (alternative hypothesis: TWO-SIDES) | | | | |
| | | CRS 09 vs VRS 09 | | CRS 10 vs VRS 10 | | CRS 11 vs VRS 11 | |
		Test statistic	P-value	Test statistic	P-value	Test statistic	P-value
Model I		0.308	0.171	0.423	0.019	0.308	0.171
Model II		0.308	0.171	0.385	0.043	0.308	0.171
Model III		0.308	0.171	0.346	0.089	0.269	0.301

Table 13: Appeals courts CRS vs VRS score mean test by year and model

| | *Test of mean equality* | | |
| | CRS 09 vs VRS 09 | CRS 10 vs VRS 10 | CRS 11 vs VRS 11 |
	Test statistic	Test statistic	Test statistic
Model I	1.59	2.634	1.894
Model II	1.462	2.56	1.753
Model III	1.596	1.473	1.573

To verify if changes in efficiency exist over years, a Friedman test has been set up coherently with the hypotheses of returns to scale obtained in tables 12 and 13. For Appeals courts we find significant changes between 2009 and 2011 as shown in Table 14.

Table 14: Appeals courts efficiency changes test by model - Friedman test

| | *Friedman test* | |
	Test statistic	P-value
Model I	26.547	0.0000
Model II	20.191	0.0000
Model III	20.242	0.0000

Finally, to prove if some differences exist between specific pairs of years a rank-Wilcoxon test has been implemented. Table 15 shows that there are no significant differences between the performance of the 26 Appeals courts of the first model in the 2009-2010 period. Likewise, the reduction in efficiency observed in years 2010-2011 for the second and third model is not statistically significant. However, from the comparison between 2009 and 2011, the significant reduction in efficiency, which involves all three models, is evident. No less relevant is the loss of efficiency observed in the 2009-2010 and 2010-2011 in models II-III and Model I, respectively.

Table 15: Appeals courts efficiency changes test by model - Wilcoxon test

| | | *Wilcoxon test* | | |
		CRS_09 vs VRS_10	VRS_10 vs CRS_11	CRS_09 vs CRS_11
Model I	Test statistic	-0,614b	-3,878a	-4,158a
	P-value	0.539	0.000	0.000
		CRS_09 vs NIRS_10	NIRS_10 vs CRS_11	CRS_09 vs CRS_11
Model II	Test statistic	-3,057a	-1,354a	-4,106a
	P-value	0.002	0.176	0.000
		CRS_09 vs CRS_10	CRS_10 vs CRS_11	CRS_09 vs CRS_11
Model III	Test statistic	-3,957a	-0,292a	-3,632a
	P-value	0.000	0.77	0.000

a. Based on positive ranks
b. Based on negative ranks

Under the assumptions of constant returns to scale, Table 16 points out the efficient Appeals courts districts as Rome, Lecce and Bari in 2009 and, in 2010-2011, Rome, Lecce and L'Aquila. Analyzing the

Italian average efficiency, in the Appeals courts framework there appear a decreasing trend over years, with a total gap of about 14%.

Table 16: Appeals courts CRS efficiency score by model - years 2009-2010-2011

District	Model I			Model II			Model III		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
ANCONA	0.944 s	0.790 s	0.884	0.974 s	0.753 s	0.894 s	0.935 s	0.746 s	0.906 s
BARI	1.000	0.835 s	0.988 s	1.000	0.858 s	1.000	1.000	0.866 s	1.000
BOLOGNA	0.777 s	0.600 s	0.425 s	0.778 s	0.571 s	0.433 s	0.749 s	0.554 s	0.423 s
BRESCIA	0.603 s	0.512 s	0.543 s	0.622 s	0.501 s	0.575 s	0.583 s	0.487 s	0.569 s
CAGLIARI	0.620 s	0.726 s	0.604 s	0.674 s	0.760 s	0.669 s	0.678 s	0.773 s	0.678 s
CALTANISSETTA	0.562 s	0.509 s	0.436 s	0.620 s	0.544 s	0.484 s	0.632 s	0.558 s	0.493 s
CAMPOBASSO	0.902 s	0.611 s	0.826 s	0.989 s	0.648 s	0.826 s	1.000	0.841 s	1.000
CATANIA	0.614 s	0.479 s	0.402 s	0.621 s	0.462 s	0.413 s	0.600 s	0.457 s	0.406s
CATANZARO	0.513 s	0.461 s	0.458 s	0.607 s	0.546 s	0.543 s	0.591 s	0.539 s	0.539 s
FIRENZE	0.722 s	0.567 s	0.454 s	0.739 s	0.557 s	0.471 s	0.748 s	0.561 s	0.470 s
GENOVA	0.731 s	0.604 s	0.607	0.740 s	0.572 s	0.620 s	0.731 s	0.582 s	0.635 s
L'AQUILA	0.969 s	1.000	1.000	0.970 s	1.000	1.000	0.994 s	1.000	1.000
LECCE	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MESSINA	0.694 s	0.568 s	0.615 s	0.718 s	0.588 s	0.642 s	0.709 s	0.586 s	0.652 s
MILANO	0.330 s	0.273 s	0.301 s	0.365 s	0.290 s	0.312 s	0.360 s	0.291 s	0.305 s
NAPOLI	0.770 s	0.707 s	0.540 s	0.816 s	0.803 s	0.685 s	0.801 s	0.765 s	0.589s
PALERMO	0.553 s	0.523 s	0.435 s	0.587 s	0.543 s	0.469 s	0.588 s	0.548 s	0.470 s
PERUGIA	0.878 s	0.624 s	0.809	0.926 s	0.659 s	0.865 s	0.939 s	0.674 s	0.900s
POTENZA	0.931 s	0.672 s	0.503 s	1.000	0.730 s	0.548 s	1.000	0.793 s	0.572 s
REGGIO CALABRIA	0.752 s	0.491 s	0.389 s	0.762 s	0.501 s	0.417 s	0.765 s	0.499 s	0.412 s
ROMA	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.669 s	0.683 s	0.638	0.676 s	0.683 s	0.673 s	0.713 s	0.749 s	0.833 s
TORINO	0.615 s	0.427 s	0.389 s	0.687 s	0.423 s	0.406 s	0.663 s	0.426 s	0.404 s
TRENTO	0.630 s	0.479 s	0.452 s	0.652 s	0.495 s	0.469 s	0.666 s	0.509 s	0.483 s
TRIESTE	0.734 s	0.495 s	0.400 s	0.763 s	0.488 s	0.424 s	0.772 s	0.506 s	0.432 s
VENEZIA	0.920 s	0.508 s	0.519 s	0.934 s	0.509 s	0.531 s	0.905 s	0.493 s	0.520 s

Note: s=Presence of positive slacks.

7.2. Implications for offices dimensions

As in the case of the Ordinary courts, by comparing the hypotheses of CRS and VRS, we try to assess if the inefficiency is the result of an inadequate management of productive factors or the effect of an inadequate dimensional scale by which districts are operating. In Table 17 it can be seen how the number of districts operating at an optimal size is very small (Rome, Lecce, L'Aquila and Bari) and that the number of districts with IRS is considerably higher than those characterized by DRS, which are, in their turn, represented by the major cities. Figure 6 confirms the same results for Model III which is, as said above, the most complete in terms of variables involved.

Table 17: Ordinary courts returns to scale by model - years 2009-2010-2011

District	Model I			Model II			Model III		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
ANCONA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
BARI	MPSS	DRS	IRS	MPSS	DRS	MPSS	MPSS	DRS	MPSS
BOLOGNA	DRS	DRS	IRS	DRS	DRS	IRS	DRS	DRS	IRS
BRESCIA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
CAGLIARI	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
CALTANISSETTA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
CAMPOBASSO	IRS	IRS	IRS	IRS	IRS	IRS	MPSS	IRS	MPSS
CATANIA	DRS	DRS	IRS	DRS	DRS	IRS	DRS	DRS	IRS
CATANZARO	IRS	IRS	IRS	DRS	IRS	IRS	DRS	IRS	IRS
FIRENZE	DRS	DRS	IRS	DRS	DRS	IRS	DRS	DRS	IRS
GENOVA	DRS	DRS	IRS	DRS	DRS	IRS	DRS	DRS	IRS
L'AQUILA	IRS	MPSS	MPSS	IRS	MPSS	MPSS	IRS	MPSS	MPSS
LECCE	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS
MESSINA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
MILANO	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
NAPOLI	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
PALERMO	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
PERUGIA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
POTENZA	IRS	IRS	IRS	MPSS	IRS	IRS	MPSS	IRS	IRS
REGGIO CALABRIA	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
ROMA	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS	MPSS
SALERNO	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
TORINO	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS	DRS
TRENTO	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
TRIESTE	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS	IRS
VENEZIA	DRS	IRS	IRS	DRS	IRS	IRS	DRS	IRS	IRS

Note: IRS=Increasing returns to scale, DRS=Decreasing returns to scale, MPSS=Most Productive Scale Size.

Figure 6: Appeals courts returns to scale, Model III

(a) 2009   (b) 2010   (c) 2011
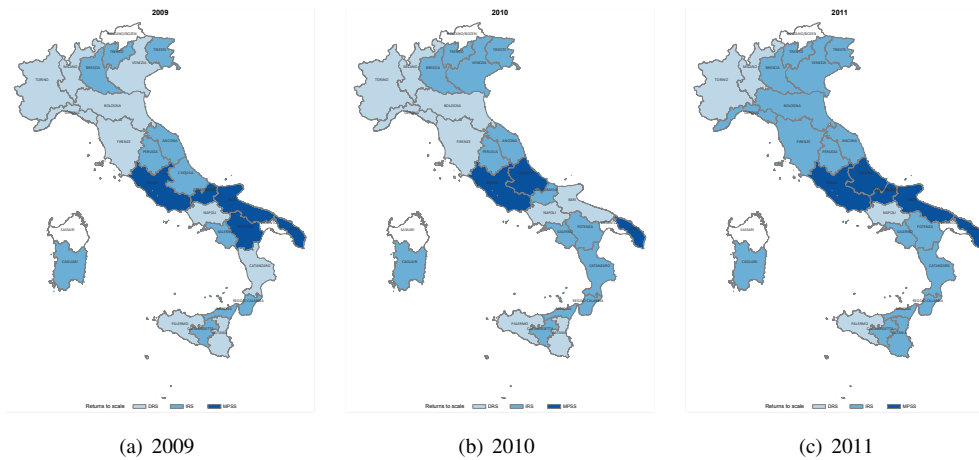
Therefore, the result reached here is that the major part of the Italian Appeals courts should change size by increasing activities. In particular, Milan's Appeals court for the whole period exhibits an efficiency score around 35%. An opposite situation is given by the Venice's Appeals court in 2009, that operates in an over-dimensional scale condition with an inefficiency level of about 8%.

For the Appeals courts we do not perform the PCA second-step basically because there is no ambiguity on the dimensions of these offices which have been discovered to be mainly with IRS technology.[16]

Actually, in Italy, the raise in the demand for outputs of the judicial sector is becoming an urgent problem to cope with the huge number of laws and decrees always increasing. This is basically due to the increasing number of local authorities, and of their competences, which caused an always more increasing amount of regulations and practices. Alternatively, a more uniform legislative system should help control for this problem.

8. Conclusions and further remarks

Being strategic for the improvement of productivity and growth at national level, this research investigates, for the first two degrees of judgment, on the Italian judicial efficiency over some years, taking into account of the lengths of trials.

As regards the Ordinary courts, the technical inefficiency found is consistent and stable over years. On average 20 districts out of 26 are inefficient, with an average inefficiency score of 27.2%. The analysis developed consisted in two steps with a DEA followed by a PCA, on the productivity ratios.

The second step has been motivated by the necessity of qualifying more into deep the dysfunctions detected for many Ordinary courts represented by DRS and IRS technologies. This method, here developed, allows for studying differences in productivity between macro-areas characterized by DMUs with distributions of resources performing in a different way through processes with different lengths, or with different specific productivities specialized in the lengths of trials.

In order to improve efficiency and to satisfy a growing demand for justice, a division of offices would be requested in the North, being the most part of them characterized by DRS technology, whilst more resources, together with a better repartition of the inputs through offices, would be necessary in the South where there is a prevalence of IRS districts. The Isles suffer from both a low productivity and an inefficient distribution of inputs on which should be relevant to concentrate the efforts for improving efficiency.

As for Appeals courts, on average, 23 districts out of 26 were found inefficient in the case of CRS hypothesis and 20 out of 26 under the VRS hypothesis, with a relative inefficiency of 28.8% and 25%, respectively. For such a degree level, the majority of courts are IRS. Hence, an increase in the inputs would be desirable for processing trials more efficiently and in a shorter time.

A promising field for further research would be that of studying and formalizing the endogenous relationship between inefficiency and length of trials, which is a peculiarity of efficiency of the judicial system. This would constitute an alternative, if not a step forward, to that one developed here, where the kind of

---

[16]Furthermore, the Italian law prescribes a single Appeals court per each district, thus allowing only for the enlargement or contraction of such offices without the possibility of merging.

inefficiency possibly inherent to high length of trials has been controlled for either by appropriate weights (applied to output) or by conducting the productivity analysis at parity conditions.

We also stress the general validity of the analysis employed which is, in principle, independent of the judicial system of reference, and so may be of a more general use for possible characterizations of output and resources of interest different from that of lengths of trials considered here.

Acknowledgments

A. Appendix

A.1. Ordinary courts

Table 18: Ordinary courts efficiency score by method and model - year 2009

District	Model I		Model II		Model III	
	DEA CRS	DEA VRS	DEA CRS	DEA VRS	DEA CRS	DEA VRS
ANCONA	0.906	0.909	0.933	0.936	0.126	0.411
BARI	0.942	1.000	0.944	1.000	1.000	1.000
BOLOGNA	0.918	1.000	0.940	1.000	0.766	0.837
BRESCIA	0.867	0.918	0.873	0.943	0.108	0.332
CAGLIARI	0.755	0.760	0.757	0.766	0.198	0.607
CALTANISSETTA	0.244	0.422	0.256	0.428	0.022	0.436
CAMPOBASSO	0.938	1.000	0.959	1.000	0.026	1.000
CATANIA	0.526	0.567	0.535	0.571	0.222	0.358
CATANZARO	0.444	0.509	0.449	0.519	0.265	0.398
FIRENZE	0.868	0.949	0.873	0.963	0.369	0.460
GENOVA	0.646	0.663	0.659	0.672	0.123	0.303
L'AQUILA	1.000	1.000	1.000	1.000	0.154	0.403
LECCE	1.000	1.000	1.000	1.000	0.665	0.935
MESSINA	0.432	0.539	0.454	0.550	0.174	0.491
MILANO	0.480	0.902	0.529	0.917	0.216	0.281
NAPOLI	0.671	1.000	0.718	1.000	1.000	1.000
PALERMO	0.291	0.338	0.308	0.371	0.134	0.239
PERUGIA	0.921	1.000	0.972	1.000	0.126	0.637
POTENZA	0.624	0.712	0.630	0.715	0.143	0.643
REGGIO CALABRIA	0.298	0.302	0.303	0.303	0.089	0.342
ROMA	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.894	0.899	0.896	0.900	0.251	0.464
TORINO	0.535	0.647	0.552	0.672	0.169	0.241
TRENTO	1.000	1.000	1.000	1.000	0.046	1.000
TRIESTE	0.571	0.574	0.586	0.590	0.034	0.357
VENEZIA	0.808	0.901	0.821	0.906	0.246	0.338

Table 19: Ordinary courts efficiency score by method and model - year 2010

District	Model I		Model II		Model III	
	DEA CRS	DEA VRS	DEA CRS	DEA VRS	DEA CRS	DEA VRS
ANCONA	1.000	1.000	1.000	1.000	0.072	0.379
BARI	0.900	1.000	0.910	1.000	1.000	1.000
BOLOGNA	1.000	1.000	1.000	1.000	0.439	0.524
BRESCIA	0.832	0.876	0.865	0.908	0.039	0.279
CAGLIARI	0.928	0.938	0.952	0.967	0.249	0.634
CALTANISSETTA	0.211	0.422	0.221	0.427	0.016	0.435
CAMPOBASSO	1.000	1.000	1.000	1.000	0.016	1.000
CATANIA	0.394	0.475	0.398	0.487	0.172	0.307
CATANZARO	0.486	0.487	0.495	0.495	0.276	0.404
FIRENZE	0.778	0.810	0.783	0.818	0.230	0.327
GENOVA	0.663	0.679	0.665	0.686	0.060	0.254
L'AQUILA	1.000	1.000	1.000	1.000	0.087	0.347
LECCE	1.000	1.000	1.000	1.000	0.561	0.860
MESSINA	0.475	0.536	0.497	0.546	0.207	0.520
MILANO	0.503	0.895	0.578	0.898	0.121	0.187
NAPOLI	0.672	1.000	0.744	1.000	0.694	1.000
PALERMO	0.893	1.000	0.320	0.395	0.098	0.206
PERUGIA	0.791	0.871	0.811	0.878	0.050	0.590
POTENZA	0.692	0.762	0.698	0.773	0.115	0.632
REGGIO CALABRIA	0.302	0.308	0.311	0.313	0.107	0.356
ROMA	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.880	0.892	0.896	0.904	0.380	0.592
TORINO	0.712	0.750	0.715	0.765	0.085	0.164
TRENTO	1.000	1.000	1.000	1.000	0.011	1.000
TRIESTE	0.651	0.676	0.671	0.688	0.023	0.354
VENEZIA	0.778	0.895	0.780	0.899	0.151	0.264

Table 20: Ordinary courts efficiency score by method and model - year 2011

District	Model I		Model II		Model III	
	DEA CRS	DEA VRS	DEA CRS	DEA VRS	DEA CRS	DEA VRS
ANCONA	0.915	0.944	0.937	0.956	0.082	0.383
BARI	0.914	0.918	0.932	0.934	1.000	1.000
BOLOGNA	1.000	1.000	1.000	1.000	0.422	0.510
BRESCIA	0.848	0.895	0.889	0.926	0.042	0.279
CAGLIARI	0.762	0.963	0.876	1.000	0.400	0.741
CALTANISSETTA	0.225	0.422	0.232	0.427	0.020	0.434
CAMPOBASSO	1.000	1.000	1.000	1.000	0.034	1.000
CATANIA	0.400	0.413	0.409	0.425	0.246	0.376
CATANZARO	0.435	0.484	0.436	0.492	0.328	0.450
FIRENZE	0.739	0.777	0.751	0.804	0.296	0.389
GENOVA	0.668	0.688	0.669	0.702	0.063	0.255
L'AQUILA	0.996	1.000	1.000	1.000	0.136	0.392
LECCE	1.000	1.000	1.000	1.000	0.529	0.826
MESSINA	0.532	0.582	0.550	0.589	0.225	0.528
MILANO	0.524	0.882	0.609	0.888	0.133	0.196
NAPOLI	0.588	1.000	0.656	1.000	0.763	1.000
PALERMO	0.924	1.000	0.340	0.388	0.125	0.231
PERUGIA	0.789	0.855	0.814	0.857	0.093	0.614
POTENZA	0.544	0.655	0.562	0.675	0.155	0.654
REGGIO CALABRIA	0.298	0.312	0.317	0.317	0.135	0.380
ROMA	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.827	0.881	0.870	0.904	0.550	0.713
TORINO	0.669	0.704	0.683	0.732	0.075	0.153
TRENTO	1.000	1.000	1.000	1.000	0.010	1.000
TRIESTE	0.612	0.632	0.627	0.642	0.040	0.363
VENEZIA	0.811	0.882	0.816	0.890	0.203	0.304

A.2. Appeal courts

Table 21: Appeals courts efficiency score by method and model - year 2009

District	Model I		Model II		Model III	
	DEA CRS	DEA VRS	DEA CRS	DEA VRS	DEA CRS	DEA VRS
ANCONA	0.944	1.000	0.974	1.000	0.618	0.940
BARI	1.000	1.000	1.000	1.000	0.464	0.515
BOLOGNA	0.777	0.844	0.778	0.868	0.426	0.453
BRESCIA	0.603	0.645	0.622	0.669	0.626	0.662
CAGLIARI	0.620	0.777	0.674	0.829	0.184	0.562
CALTANISSETTA	0.562	0.885	0.620	0.938	0.016	0.668
CAMPOBASSO	0.902	1.000	0.989	1.000	0.056	1.000
CATANIA	0.614	0.626	0.621	0.631	0.291	0.384
CATANZARO	0.513	0.529	0.607	0.613	0.115	0.368
FIRENZE	0.722	0.802	0.739	0.819	0.078	0.316
GENOVA	0.731	0.741	0.740	0.753	0.168	0.436
L'AQUILA	0.969	1.000	0.970	1.000	1.000	1.000
LECCE	1.000	1.000	1.000	1.000	0.186	0.639
MESSINA	0.694	0.746	0.718	0.748	0.128	0.621
MILANO	0.330	0.379	0.365	0.386	0.166	0.200
NAPOLI	0.770	0.924	0.816	0.935	0.703	0.772
PALERMO	0.553	0.573	0.587	0.617	0.054	0.229
PERUGIA	0.878	0.989	0.926	0.993	0.116	0.814
POTENZA	0.931	1.000	1.000	1.000	0.018	0.805
REGGIO CALABRIA	0.752	0.776	0.762	0.789	0.188	0.582
ROMA	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.669	0.752	0.676	0.758	0.185	0.533
TORINO	0.615	0.668	0.687	0.744	0.266	0.320
TRENTO	0.630	0.915	0.652	1.000	0.001	1.000
TRIESTE	0.734	0.881	0.763	0.906	0.037	0.670
VENEZIA	0.920	1.000	0.934	1.000	0.381	0.472

Table 22: Appeals courts efficiency score by method and model - year 2010

District	Model I		Model II		Model III	
	DEA CRS	DEA VRS	DEA CRS	DEA VRS	DEA CRS	DEA VRS
ANCONA	0.790	0.909	0.753	0.934	1.000	1.000
BARI	0.835	0.882	0.858	0.891	0.847	1.000
BOLOGNA	0.600	0.928	0.571	0.951	0.694	0.826
BRESCIA	0.512	0.600	0.501	0.582	1.000	1.000
CAGLIARI	0.726	0.821	0.760	0.845	0.150	0.532
CALTANISSETTA	0.509	0.833	0.544	0.845	0.021	0.671
CAMPOBASSO	0.611	1.000	0.648	1.000	0.059	1.000
CATANIA	0.479	0.486	0.462	0.470	0.393	0.429
CATANZARO	0.461	0.503	0.546	0.587	0.129	0.381
FIRENZE	0.567	0.708	0.557	0.726	0.111	0.322
GENOVA	0.604	0.689	0.572	0.670	0.205	0.441
L'AQUILA	1.000	1.000	1.000	1.000	0.854	0.968
LECCE	1.000	1.000	1.000	1.000	0.230	0.654
MESSINA	0.568	0.720	0.588	0.729	0.135	0.624
MILANO	0.273	0.297	0.290	0.292	0.466	0.528
NAPOLI	0.707	0.879	0.803	0.891	0.988	1.000
PALERMO	0.523	0.588	0.543	0.625	0.063	0.231
PERUGIA	0.624	0.915	0.659	0.925	0.311	0.844
POTENZA	0.672	0.957	0.730	0.973	0.017	0.804
REGGIO CALABRIA	0.491	0.576	0.501	0.580	0.184	0.578
ROMA	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.683	0.748	0.683	0.748	0.322	0.550
TORINO	0.427	0.532	0.423	0.520	0.798	1.000
TRENTO	0.479	0.887	0.495	1.000	0.002	1.000
TRIESTE	0.495	0.755	0.488	0.730	0.051	0.672
VENEZIA	0.508	0.532	0.509	0.529	0.817	0.910

Table 23: Appeals courts efficiency score by method and model - year 2011

District	Model I		Model II		Model III	
	DEA CRS	DEA VRS	DEA CRS	DEA VRS	DEA CRS	DEA VRS
ANCONA	0.884	0.964	0.894	1.000	1.000	1.000
BARI	0.988	0.994	1.000	1.000	0.633	0.661
BOLOGNA	0.425	0.437	0.433	0.441	0.458	0.486
BRESCIA	0.543	0.612	0.575	0.635	0.598	0.635
CAGLIARI	0.604	0.755	0.669	0.807	0.069	0.511
CALTANISSETTA	0.436	0.795	0.484	0.818	0.039	0.679
CAMPOBASSO	0.826	1.000	0.826	1.000	0.050	1.000
CATANIA	0.402	0.436	0.413	0.443	0.328	0.409
CATANZARO	0.458	0.508	0.543	0.593	0.108	0.361
FIRENZE	0.454	0.467	0.471	0.483	0.121	0.323
GENOVA	0.607	0.631	0.620	0.635	0.130	0.425
L'AQUILA	1.000	1.000	1.000	1.000	0.342	0.871
LECCE	1.000	1.000	1.000	1.000	0.222	0.662
MESSINA	0.615	0.727	0.642	0.737	0.142	0.629
MILANO	0.301	0.367	0.312	0.377	0.349	0.351
NAPOLI	0.540	0.852	0.685	0.866	1.000	1.000
PALERMO	0.435	0.452	0.469	0.490	0.054	0.231
PERUGIA	0.809	1.000	0.865	1.000	0.384	0.878
POTENZA	0.503	0.859	0.548	0.860	0.023	0.806
REGGIO CALABRIA	0.389	0.529	0.417	0.543	0.161	0.468
ROMA	1.000	1.000	1.000	1.000	1.000	1.000
SALERNO	0.638	0.761	0.673	0.762	0.281	0.575
TORINO	0.389	0.401	0.406	0.407	0.871	0.972
TRENTO	0.452	0.913	0.469	1.000	0.002	1.000
TRIESTE	0.400	0.674	0.424	0.704	0.035	0.670
VENEZIA	0.519	0.540	0.531	0.549	0.672	0.676

B. Appendix

Figure 7: Ordinary courts returns to scale, Model I

(a) 2009   (b) 2010   (c) 2011
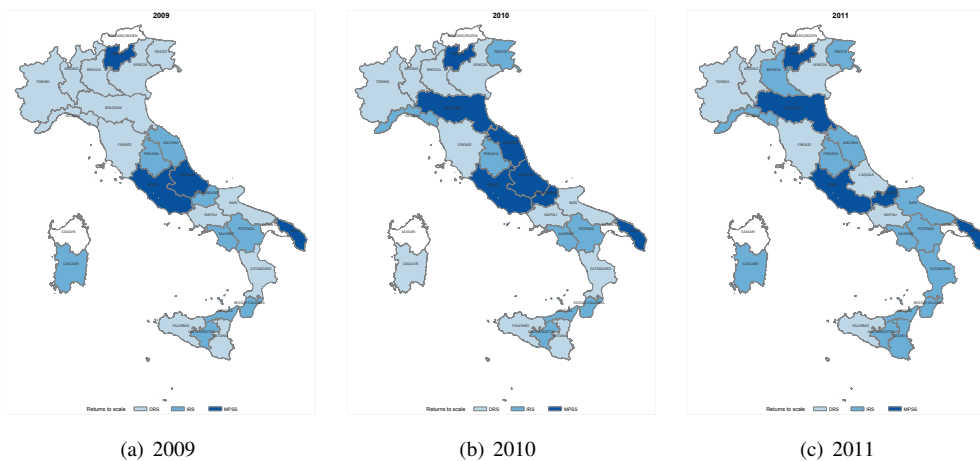
Figure 8: Ordinary courts returns to scale, Model II

(a) 2009   (b) 2010   (c) 2011
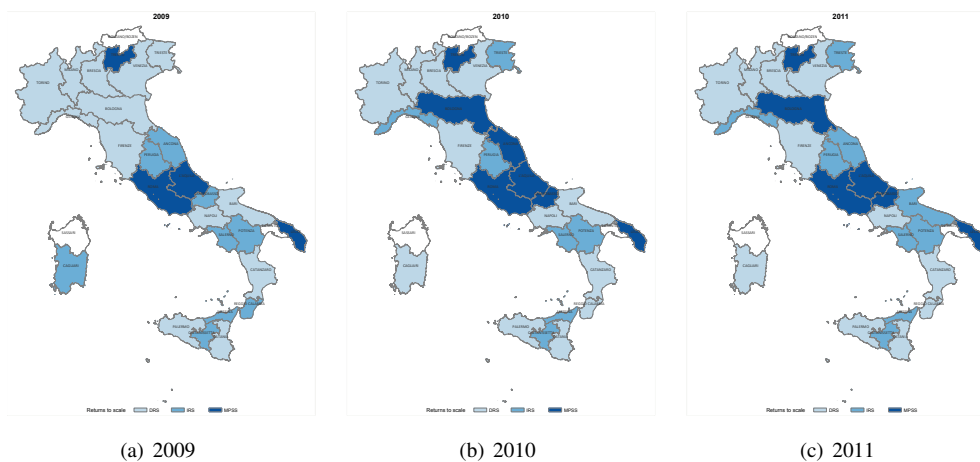
Figure 9: Appeals courts returns to scale, Model I

(a) 2009 (b) 2010 (c) 2011
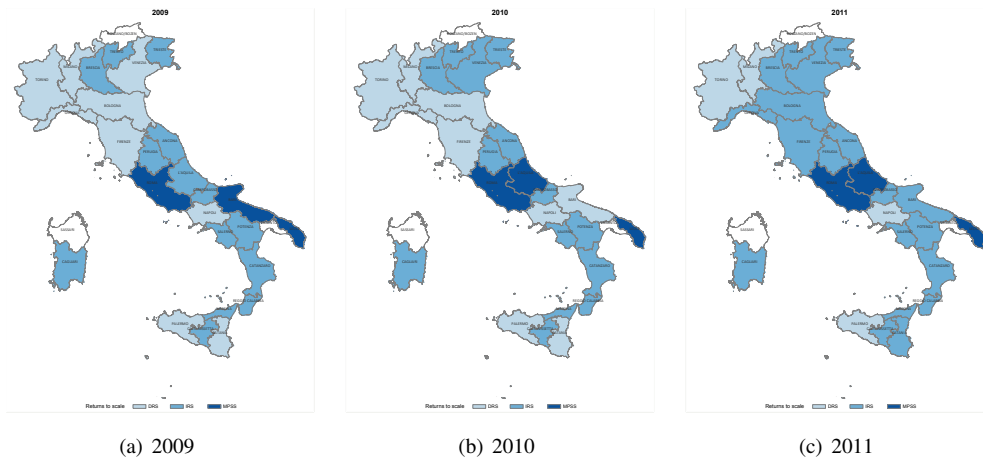
Figure 10: Appeals courts returns to scale, Model II

(a) 2009 (b) 2010 (c) 2011
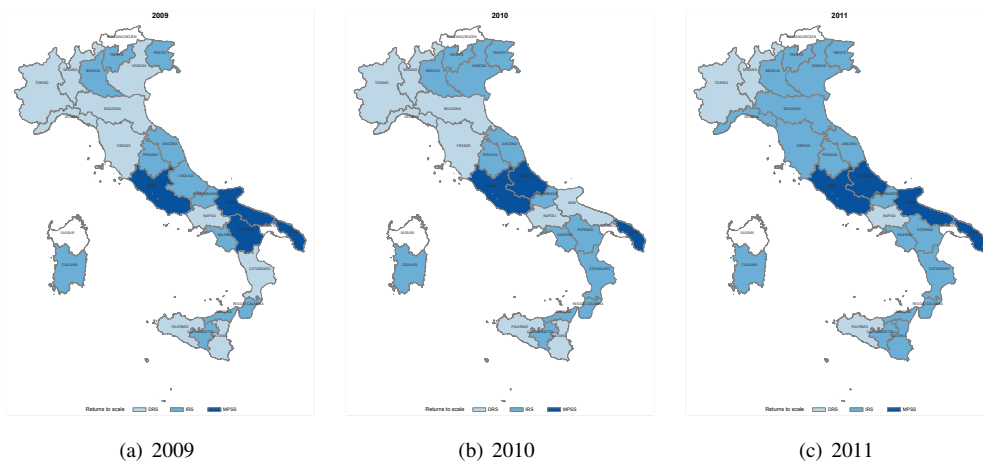
Table 24: Wilcoxon test of ranks (DEA vs Order-*m*) - Appeals courts

	2009	2010	2013
Model I	0.993	0.978	0.964
Model II	0.993	0.964	1.000
Model III	0.993	0.964	1.000

Table 25: Wilcoxon test of ranks (DEA vs Order-*m*) - Ordinary courts

	2009	2010	2013
Model I	0.978	0.993	0.993
Model II	0.964	0.964	1.000
Model III	0.993	1.000	0.993

References

Ali, A.I., & Seiford, L.M. 1993. *The mathematical programming approach to efficiency analysis*. Oxford U.K. Chap. in Fried HO and SS Schmidt (eds.) the Measurement of Productive Efficiency: Techniques and Applications, pages 120–159.

Andersen, Per Petersen, & Christian, Niels. 1993. A procedure for ranking efficient units in data envelopment analysis. *Management science*, **39**(10), 1261–1264.

Antonelli, A., & Marchesi, D. 1999. Dimensione e composizioni ottimali degli uffici giudiziari e effetti dell'introduzione del giudice unico. *Commissione tecnica della spesa pubblica, ricerche, rome, italy*, **19**.

Banker, Rajiv D. 1993. Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. *Management science*, **39**(10), 1265–1273.

Banker, Rajiv D. 1996. Hypothesis tests using data envelopment analysis. *Journal of productivity analysis*, **7**(2), 139–159.

Banker, R.D., Charnes, A., & Cooper, W.W. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, **30**, 1078.

Bianco, M., Giacomelli, S., Giorgiantonio, C., & Palumbo, G.and Szego, B. 2007. La durata (eccessiva) dei procedimenti civili in italia: offerta, domanda o rito? *Rivista di politica economica*, **97**(5), 3–54.

Cazals, C., Florens, J., & Simar, L. 2002. Nonparametric frontier estimation: A robust approach. *Journal of econometrics*, **106 (1)**, 1–25.

Charnes, A., Cooper, W.W., & Rhodes, E. 1978. Measuring the efficiency of decision making units. *European journal of operational research*, **2**, 429–444.

Coelli, T., Rao, D.S., O'Donnell, C., & Battese, G. 2005. *An introduction to efficiency and productivity analysis*. Springer US.

Cooper, W.W., Seiford, L.M., Tone, K., & Zhu, J. 2007. Some models and measures for evaluating performances with dea: Past accomplishments and future prospects. *Journal of productivity analysis*, **28**, 151–163.

Daraio, C., & Simar, L. 2005. Introducing environmental variables in nonparametric frontier models:a probabilistic approach. *Journal of productivity analysis*, **24(1)**, 93–121.

Deprins, D., Simar, L. L., & Tulkens, H. 1984. *The performance of public enterprizes: Concepts and measurements*. North-Holland. Chap. Measuring Labor Inefficiency in Post Offices, pages 243 – 267.

Deyneli, Fatih. 2012. Analysis of relationship between efficiency of justice services and salaries of judges with two-stage DEA method. *European journal of law and economics*, **34**(3), 477–493.

Doyle, John, & Rodney. 1994. Efficiency and cross-efficiency in dea: Derivations, meanings and uses. *The journal of the operational research society*, **45**(5), 567–578.

EURISPES. 2009. *21° rapporto italia*. Tech. rept. EURISPES.

Falavigna, Greta, Ippoliti, Roberto, Manello, Alessandro, & Ramello, Giovanni B. 2015. Judicial productivity, delay and efficiency: A directional distance function (ddf) approach. *European journal of operational research*, **240**(2), 592 – 601.

Finocchiaro Castro, Massimo, & Guccio, Calogero. 2014. Searching for the source of technical inefficiency in italian judicial districts: an empirical investigation. *European journal of law and economics*, **38**(3), 369–391.

Giacomelli, S., & Menon, C. 2013. Firm size and judicial efficiency: evidence from the neighbour's court. *Temi di discussione*, 4–45.

Harrison, J., Coppola, M.N., & Wakefield, M. 2004. Efficiency of federal hospitals in the united states. *Journal of medical systems*, **28**, 411–422.

Kittelsen, Sverre A. C., & Førsund, Finn R. 1992. Efficiency analysis of norwegian district courts. *Journal of productivity analysis*, **3**(3), 277–306.

Koopmans, T.C. 1951. *Activity analysis of production and allocation*. New York: Wiley. Chap. An Analisys of Production as an Efficient Combination of Activities, page monograph no.13.

Lewin, Arie Y, Morey, Richard C, & Cook, Thomas J. 1982. Evaluating the administrative efficiency of courts. *Omega*, **10**(4), 401 – 411.

Liu, John S., Lu, Louis Y.Y., Lu, Wen-Min, & Lin, Bruce J.Y. 2013. A survey of DEA applications. *Omega*, **41**(5), 893–902.

Maggi, Bernardo, & Guida, Marco. 2011. Modelling non-performing loans probability in the commercial banking system: efficiency and effectiveness related to credit risk in italy. *Empirical economics*, **41**, 269–291.

Manly, B.F.J. 1994. *Multivariate statistical methods: A primer*. Chapman and Hall.

Mehdi, Toloo, Mona, Barat, & Atefeh, Masoumzadeh. 2015. Selective measures in data envelopment analysis. *Annals of operations research*, **226**(1), 623–642.

Pedraja-Chaparro, Francisco, & Salinas-Jimenez, Javier. 1996. An assessment of the efficiency of spanish courts using dea. *Applied economics*, **28**(11), 1391–1403.

Peyrache, Antonio, & Zago, Angelo. 2016. Large courts, small justice!: The inefficiency and the optimal structure of the italian justice sector. *Omega*, **64**, 42 – 56.

Santos, Sérgio P., & Amado, Carla A.F. 2014. On the need for reform of the portuguese judicial system – does data envelopment analysis assessment support it? *Omega*, **47**, 1 – 16.

Wagner, J. M., & Shimshak, D. G. 2007. Stepwise selection of variables in data envelopment analysis: procedures and managerial perspectives. *European journal of operational research*, **180**, 57—-67.

World Bank. 2010. *Doing business*. Tech. rept. World Bank.

Yeung, Luciana L., & Azevedo, Paulo F. 2011. Measuring efficiency of brazilian courts with data envelopment analysis (dea). *Ima journal of management mathematics*, **22**(4), 343–356.