Efficient courts? A frontier performance assessment

Gustavo Ferro
Universidad del Centro de Estudios Macroeconomicos de Argentina, Buenos Aires, Argentina and National Council for Scientific and Technical Research, Buenos Aires, Argentina
Carlos A. Romero
Instituto Interdisciplinario de Economía Política de Buenos Aires (IIEP-Baires), Facultad de Ciencias Económicas, Universidad de Buenos Aires (UBA), Buenos Aires, Argentina, and
Exequiel Romero-Gómez
Universidad de Buenos Aires, Buenos Aires, Argentina

Received 21 April 2017
Revised 12 September 2017
Accepted 24 September 2017

Abstract

Purpose – The purpose of this paper is to build performance indicators to assess efficiency for First Instance Federal Courts in Argentina and study the determinants of efficiency in Criminal Instruction Courts.
Design/methodology/approach – The efficiency scores were determined using data envelopment analysis with a database for the period 2006–2010. Then, a search of the efficiency determinants in the Criminal Instruction Courts was performed. Four output-oriented models were developed based on various explanatory and environmental variables.
Findings – Workload is an environmental variable that significantly increased the average levels of efficiency. When analyzing explanatory factors of the efficiency levels of the Criminal Instruction Courts, surrogate judges and temporary staff are more efficient on average than tenured judges and staff.
Research limitations/implications – The method chosen permits flexibility in the analysis. Future research would be interesting to develop the underlying economic model using econometric methods.
Practical implications – This paper's contribution is twofold: first, to estimate the relative efficiency for all First Instance Federal Courts in every jurisdiction; and second, to explain the differences in efficiency in the Criminal Instruction Courts.
Social implications – This study has the potential to greatly impact the discussion of how to structure judicial procedures (from the benchmarking between different branches of Federal justice) and in the design of incentives in a judicial career (e.g. tenured vs temporary judges and clerical employees, the role of seniority of judges and clerical employees and the impact of gender in performance).
Originality/value – To the authors' knowledge, this paper is the first scholarly article to measure efficiency in Argentine justice system using mathematical programming and econometric methods. It has academic interest since it advances on the comprehension of the underlying production function of justice service provision. The paper also has social and practical implications since it permits contributing to the institutional design and opens the discussion for further sequels with other methods and complementary purposes.

Keywords Benchmarking, Data envelopment analysis, Efficiency, Courts

Paper type Research paper

1. Introduction

The purpose of this paper is to build performance indicators to assess efficiency of the First Instance Federal Courts of Argentina. After determining the efficiency scores using data envelopment analysis (DEA) with a database for the period 2006–2010, we conducted a search of the determinants of efficiency in the Criminal Instruction Courts (a subset of the former). This paper's contribution is thus twofold: first, to estimate the relative efficiency for all First Instance Federal Courts in every jurisdiction; and second, to explain the differences in the efficiency in Criminal Instruction Courts. The final goal is to contribute to the judicial statistical system, providing benchmarking tools for decision making and policy design.

Benchmarking: An International Journal
Vol. 25 No. 9, 2018
pp. 3443-3458

1463-5771
DOI 10.1108/BIJ-09-2017-0244

In 2007, Argentina's Supreme Court ruled in favor of creating a National Commission of Judicial Management with the objective of implementing information technology tools to improve accountability and performance. The diagnosis stresses the need of information technology to process data, the convenience of inclusion of incentive schemes for personnel, the importance of performance measurement and benchmarking for accountability reasons and the necessity of discrimination between administrative and judicial tasks to reduce bureaucracy and freeing technical resources, among others.

Thus, a DEA was used to make a detailed efficiency assessment for the First Instance Federal Courts of Argentina at the level of jurisdictions, which facilitate a comparison of the relative performance of the various courts and an analysis of their evolution over time.

The productive efficiency analysis allows us to evaluate the courts' use of production factors as production units. Their output can be measured as the number of legal decisions (e.g. judgments and orders) issued by year. The production process is labor intensive, and to identify the productive factors which are relevant in the analysis, we explore the literature on justice and other service sectors' efficiency.

The paper is structured as follows: Section 2 presents a comprehensive review of the relevant literature on the measuring of the judicial sector's efficiency through efficiency frontier techniques in different countries. Section 3 introduces the methodology and describes the database used. Section 4 explains the results of the frontier studies carried out for the First Instance Federal Courts of Argentina. Section 5 presents some determinants of efficiency in the Criminal Instruction Courts employing econometric techniques. Finally, Section 6 describes the conclusions of the study.

2. Empirical literature survey

The review of literature related to performance measurements in courts serves as a guide to determine variables for the Argentine case, as well as studies on other service sector efficiency.

The "production of justice" is mainly a "labor-intensive" task. The variables most commonly used for inputs are the number of personnel working in the courts, the availability of staff reinforcement in the courts, the educational level of the judges, the prospects for the promotion of judges and clerical personnel, the time devoted by the judges to administrative issues and the task assignment (Kittelsen and Førsund, 1992; Pedraja-Chaparro and Salinas-Jimenez, 1996; Buscaglia and Ulen, 1997; Buscaglia and Dakolias, 1999; Schneider, 2005; Rosales-López, 2008). Although this is not the most relevant aspect, capital and technology have a positive influence on the judicial output (Buscaglia and Dakolias, 1999).

The complexity of the cases admitted to the system, the level of income and the litigious nature of the population have also been key variables to explain judicial performance (Gorman and Ruggiero, 2009). The workload is also included in Lewin *et al.* (1982) and Schneider (2005). The omission of this variable would underestimate productivity in those years in which a court registers a low workload. The increase in public spending has a positive effect on the supply of justice; however, in the medium term, it would stimulate the demand for justice, increasing congestion and making it difficult to weigh its net effect on efficiency (Pastor, 2003; Buscaglia and Ulen, 1997; Rosales-López, 2008).

There are a number of studies that analyze justice performance at a more aggregated level. For example, Lewin *et al.* (1982) evaluate the efficiency of judicial districts in North Carolina. Finocchiaro and Guccio (2012) estimate the Italian judicial districts' technical efficiency. Beenstock and Haitovsky (2004) study Israel's justice system.

Some of the studies that evaluate judicial efficiency have been carried out in the context of judicial reform undertaken by different countries. To measure the impact of the restructuring of the Swedish courts' scale on efficiency, Hagstedt and Proos (2008) analyze the courts' efficiency in two different periods. They show that the levels of efficiency increased on average, concluding that the reform had improved the scale of the courts. Also

in a context of judicial reform, Yeung and Azevedo (2010) suggest that the causes of inefficiency in Brazilian courts could be the lack of material and human resources and the low quality of procedural law. Yeung (2008) also shows that the presence of professional management leads to greater efficiency.

One of the key components of court performance is an analysis of the judicial career. For this reason, different aspects of the incentives scheme, applied to judges and personnel, are commonly addressed in the literature. Some generally used career variables are aging, promotion opportunities, experience and academic record. Schneider (2005) assesses the productivity of civil courts in Germany, arguing that there is an internal labor market in which the main incentive derives from the career development opportunities. He concludes that courts that employ many judges with PhDs are more productive, but they face more reversals by the appeal courts. In the same direction, Dimitrova-Grajzl *et al.* (2011), in a study for Slovenia, find that the most productive judges in Slovenia make lower quality decisions. Hence, the policymaker contemplating a judicial reform aimed at increasing productivity should carefully analyze this effect.

Dimitrova-Grajzl *et al.* (2011) also show the following: the judges with chances of promotion tend to solve a greater number of cases; the life cycle of judicial performance follows an inverted-U pattern; there is a positive relationship between productivity and judicial salaries; and there is no robust evidence that possessing a postgraduate degree, having more experience or the presence of gender differences in staff composition affects productivity. Bhattacharya and Smyth (2001) study the impact of the age of judges in Australia's High Court on their productivity, finding that judges have a lower performance as they get older.

In Brazil, the most efficient courts show a greater interest in the development of their personnel's careers and work motivation (Yeung and Azevedo, 2010). Espasa and Esteller-Moré (2011), for the First Instance Civil Courts of Spain, identify evidence that hiring temporary workers is significantly less effective compared to tenured employees.

Efficiency can be affected by external factors. Gorman and Ruggiero (2009) focus on the staff efficiency of public prosecutors' offices in 26 American states and find that in lower income level counties with a higher percentage of a minority population those are more inefficient. They argue that this is because cases in those counties can be more complicated, and people can cooperate less than in other counties to resolve cases. Kittelsen and Førsund (1992) find that the performance of multifunctional rural courts is different from specialized urban courts in Norway. Espasa and Esteller-Moré (2011) suggest that average efficiency tends to increase with time.

Another important issue is the delay in solving cases. Rosales-López and García-Rubio (2010) estimate the technical efficiency of the First Instance Courts of Andalusia in 2008. They find that the courts could have reduced the pending cases at the end of 2008 by 9.4 percent, on average had the courts operated efficiently. Nevertheless, they conclude that promoting technical efficiency does not necessarily make a significant reduction in the judicial backlog. Therefore, the delay cannot be attributed solely to the courts' low efficiency rate. Dalton and Singer (2008) show that both the number of lawyers and the number of judges per court are critical to explaining the time it takes to resolve a case. Tulkens (1993) focuses on Belgium Peace Courts between 1983 and 1985. He concludes that about 35 percent of the pending cases can be reduced by increasing the personnel in the Peace Courts.

The relationship between efficiency and court size is a critical aspect to explore. Guzowska and Strack (2010) analyze the efficiency of the public prosecutor's offices in Poland with respect to criminal cases in 2007. They find that the inefficiency of several offices can be attributed to the inappropriate scale of operations. de Sousa and Schwengber (2005) find that efficiency seems to be related to the size of the court since efficiency tends to increase with size. Moreover, these authors suggest that economies of scale exist for the court services (attributed to the presence of a minimal staff). Ruíz Hernández (2004) conducts an analysis of

the National Board of Prosecutors in Colombia for the years 1998–2002. He shows that several of the sectionals have an excess of resources used in the productive operation.

There are studies with other objectives, such as Klingner *et al.* (2015), Yasin and Gomes (2010), Andersen *et al.* (2008), Gunasekaran and Putnik (2006), Wynn-Williams (2005), Gunasekaran (2005) on service sector performance, both private and public, which findings are relevant for this paper. In general, age, tenure and skills affect efficiency and productivity in service sectors. Also in judicial sectors, García-Rubio and Elbialy (2011) analyze the performance of the First Instance Courts of Egypt with the aim of differentiating between civil and criminal jurisdictions. They show that the civil courts are relatively inefficient. This result could be influenced by the higher degree of complexity in civil cases. Beenstock and Haitovsky (2004) find that the number of resolved cases does not depend on the number of judges, implying that an increase in the number of judges does not generate a reduction in pending cases. They suggest that the practice of determining the number of judges by fixed coefficients is inappropriate.

3. Data and methodology

Once we identified inputs and outputs, we estimated efficiency frontiers of courts' production using DEA to account for the best performers in the different jurisdictions. Subsection 3.1 is devoted to the method, while Subsection 3.2 deals with the data.

3.1 Data envelopment analysis

DEA is a well-known family of mathematical programming tools for assessing the relative efficiency of a set of comparable decision-making units. It aims to evaluate the relative efficiency of decision units in the public sector and non-profit organizations producing single or multiple outputs with multiple inputs. DEA compares the observed inputs and outputs for all the decision units of a sample, identifies the best practice to define an efficient frontier and measures the degree of inefficiency of each decision unit relative to the frontier.

DEA approaches of technical efficiency can be input-oriented, output-oriented or non-oriented models. In the first case, output is maintained constant, and the model determines which potential proportional reduction in inputs is needed to achieve the frontier result. In the second case, inputs are maintained constant, and the model shows which potential increase in output is needed to achieve the frontier result. In the third case, both inputs and outputs could be adjusted to fill the efficiency gap (in which case the input reduction and output increasing are calculated together).

Mathematically, DEA models are built departing from the maximization of a total factor productivity index such as follows:

$$\text{MAX TFP} = \frac{\sum_{i=1}^{m} v_i y_{ij}}{\sum_{r=1}^{s} u_r x_{rj}}, \tag{1}$$

s.t.:

$$\frac{\sum_{r=1}^{s} v_r y_{rj}}{\sum_{i=1}^{m} u_i x_{ij}} \leqslant 1.$$

Being $u_r, v_i > 0, r = 1, \ldots, \mathrm{s}, i = 1, \ldots, m,$ where y_{rj} are the amount of r outputs (judicial decisions) of the court j; u_i is the weigh given to input r; x_{ij} is the amount of the i input consumed by court j; and v_i is the weigh given to product i.

The former formulation is not used because of the difficulty of managing the properties of linearity and convexity. Instead it is transformed in a linear mathematical program which can be build allowing the maximization of product or the minimization of inputs in the productive process (Ganley and Cubbin, 1993).

Court authorities have a limited control over the inputs. They normally can manage outputs (e.g. sentences and resolved cases). Therefore, we consider that the relevant model is output oriented. Assuming constant returns to scale (CRS), we can characterize the model as follows:

$$\min_{\lambda,\theta}\theta, \tag{2}$$

s.t.:

$$Y\lambda \geqslant y_j; X\lambda \leqslant \theta x_j; Z\lambda \leqslant z_j; \lambda \epsilon R^J_+,$$

where for each court $j = 1, \ldots, N$, there is an output vector y_j, an input vector x_j and z_j an environmental variables vector. Y and X are the corresponding outputs and inputs matrices and Z is the environmental variables matrix, representing the data for the N courts. This problem is solved N times: this means that it is solved for each of the courts in the sample, whereby we can obtain the level of technical efficiency for each court j.

The value of θ obtained will be the efficiency score for the ith court. It will satisfy $\theta < 1$, with a value of 1 indicating a point on the frontier and hence a technically efficient court, according to the Farrell (1957) definition.

3.2 Descriptive statistics of the sample

The data were provided by the National Judicial Power (PJN). The analysis was performed for all the Federal Courts of First Instance. Table I presents the number of courts for the

Type of court	Denomination	Short denomination	Jurisdiction	Penal matter (Yes/No)	Total courts in 2006	Total courts in 2010
Instruction	National Criminal Instruction Court	JNCIO	Criminal	Yes	49	49
Criminal and Correctional	Federal Criminal and Correctional Court	JNCCF	Criminal	Yes	12	12
Economic	National Economic Criminal Court	JNPEO	Penal econ	Yes	8	8
Tributary	National Tributary Criminal Court	JNPTO	Penal econ	Yes	3	3
Juvenile	National Juvenile Criminal Court	JNCMO	Criminal	Yes	7	7
Correctional	National Correctional Court	JNCO	Criminal	Yes	14	14
Interior	Federal Justice of the Interior	JFI	Multijurisdictional	Both	80	85
Social Security	Federal Social Security Court	JFSS	Social Security	No	10	10
Labor	National Labor Court	JNTO	Labor	No	80	80
Fiscal Executions	Federal Contentious Administrative Fiscal Executions Courts	JNCAEFF	Contentious Administrative	No	6	6
Contentious Administrative	Federal Contentious Administrative Courts	JNCAF	Contentious Administrative	No	12	12
	National Commercial Court	JNComO	Commercial	No	26	26
Family	National Family Court	JNCFO	Civil	No	24	24
Patrimonial	National Patrimonial Court	JNCPO	Civil	No	86	86
Civil and Commercial	Federal Civil and Commercial Court	JNCComF	Civil and Commercial	No	11	11

Source: Own elaboration based on "Secretaría de Estadísticas del PJN"

Table I.
First Instance Federal Courts

years 2006 and 2010, as well as the respective matter (i.e. penal and no penal) and jurisdiction (civil, commercial, labor, criminal, etc.). Argentina has a federal organization. There is a federal justice; however, each province has its own provincial justice system. Procedures differ between federal and provincial courts. Civil, Penal, Commercial, Social Security and Labor legislation are of national reach. Federal and national courts include those that are located in the capital city (which are specialized in one jurisdiction) and those in the interior of the country (which are multijurisdictional).

The Statistics Office of the PJN compiles and publishes raw data on the activities (e.g. number of cases admitted, in process and resolved) for each federal court throughout the nation. The budget and personnel information is taken from the Human Resources and Administration offices under the Council of the Judiciary ("Consejo de la Magistratura").

Of all the courts, the more numerous are Patrimonial Courts (86), Interior Courts[1] (85) and Labor Courts (80). The total number of courts grew from 428 in 2006 to 433 in 2010. Five new courts were created between these years, all in the interior of the country. The share of each type of court remains stable over time. Moreover, Table I suggests that the frontier analysis could not be conducted separately for some types of courts (mainly for Juvenile, Fiscal Executions, Tributary Criminal, and Economic Criminal Courts) because of the lack of enough observations. Therefore, to solve the problem of the low number of observations, the analysis was performed at the jurisdiction level. Note that the Federal Courts of the Interior ("Interior"), unlike the rest, are "multijurisdictional" courts. They are considered in the estimates as another jurisdiction, but they cannot be compared with the other courts since they are multijurisdictional ones.

To present the information systematically, variables were grouped into the following categories: files and workload; and structure and typology of personnel. The output of the courts can be measured by the number of closed files (closed cases). The primary information on files is taken from the Yearbooks compiled by the PJN's Statistics Office.

Adjustments in the number of closed files were made to ensure the time consistency of the database. A time-consistent series requires that the sum of the flow variables must equal the difference in the stock variable. The cases admitted, readmitted and resolved by year are flow variables, and cases pending at the beginning or end of each year are stock variables. The time consistency requires that cases at the end of a year equal those existing at the beginning of the next year. The latter is not verified from the available information. This made it necessary to generate a new flow variable that allowed us to take the existing cases at the beginning of each year as a valid quantity[2].

Table II shows the average levels of the variables used to estimate the efficiency for each type of court.

The "Social Security," "Correctional," "Interior" and "Contentious Administrative" courts resolved more than 3,000 cases on average in 2010, whereas the "Economic," "Tributary" and "Labor" courts were unable to exceed 500 closed files. The outputs of the different types of courts have significant differences. This may be due to different formal procedures or to the fact that the difficulty of analysis required for completing files could affect the required time to resolve the cases.

Table II also refers to the workload (i.e. the sum of existing, admitted and readmitted files). The workload behaves similarly to the closed files, where "Contentious Administrative," "Social Security" and "Interior" courts show high values of the two above-mentioned variables. This implies that a significant part of a larger number of resolved cases can be attributed to an increased workload. Interestingly, "Correctional" courts have a relatively low workload and relatively high levels of resolved cases.

At an aggregate level, the average workload grew by 8 percent between 2006 and 2010. The increase could be explained in part by the fall in the resolved cases and in part by the increase in cases admitted. At the court level, a similar behavior is present. Discrepancies

Table II.
Characterization of each type of court (average 2006/2010)

Type of court	Penal matter (Yes/No)	Resolved cases	Workload	Personnel	Proportion of temporary personnel	Proportion of professional personnel	Years for promotion	Personnel age	Personnel seniority
Instruction	Yes	1,370	1,677	11.98	0.53	0.54	5.6	37	11
Criminal and Correctional	Yes	1,459	2,165	42.00	0.35	0.62	4.3	34	10
Economic	Yes	291	636	22.25	0.35	0.57	5.9	38	14
Tributary	Yes	138	454	14.33	0.42	0.66	2.9	32	7
Juvenile	Yes	622	767	26.00	0.36	0.51	7.5	39	15
Correctional	Yes	5,378	6,240	20.14	0.47	0.54	6.8	39	13
Interior	Both	3,157	22,470	34.91	0.23	0.40	7.0	43	14
Social Security	No	4,311	17,086	21.00	0.15	0.59	5.3	37	10
Labor	No	368	929	11.89	0.17	0.60	9.2	44	18
Fiscal Executions	No	660	8,206	3.67	0.04	0.83	2.9	40	11
Contentious Administrative	No	3,806	32,312	20.83	0.37	0.50	6.7	38	14
	No	2,338	15,790	28.08	0.31	0.55	5.7	40	16
Family	No	666	5,105	17.79	0.24	0.67	7.8	44	18
Patrimonial	No	516	3,380	15.60	0.25	0.59	7.0	42	17
Civil and Commercial	No	1,201	6,021	20.64	0.25	0.50	7.8	41	17

Source: Own elaboration based on "Secretaría de Estadísticas del PJN" and "Consejo de la Magistratura"

can also be observed within each type of court. The coefficients of dispersion (i.e. the quotient between standard deviation and mean) in "Interior" and "Contentious Administrative" are visibly higher than the rest, with values close to 0.90 in 2006 and 2010 for both courts[3].

The average number of employees by court for the entire sample is 22 agents. The variability in court employees is low except in the Federal Courts of the Interior. Staff differences were observed by type of court. The personnel's variables did not have significant fluctuations with the time.

We constructed additional explanatory variables based on the available data and the variables that arose from the literature review. These variables are: Seniority (i.e. the average seniority of each court staff measured in years), Age (i.e. the average age of each court staff measured in years), Promotion (i.e. the average time between promotions of the personnel for each court measured in years) and Proportion of Temporary Personnel (i.e. the ratio between temporary agents and the total personnel of each court).

"Interior" shows heterogeneity in terms of the average share of the professional employees in the entire staff of each court (a dispersion of 0.38). The other categories show small deviations.

The ratio of temporary employees and total agents for all courts is 28 percent. This variable contains a high dispersion in different types of courts (minimum 4 percent, maximum 53 percent).

The average seniority is 16 years. The highest values are observed in the "Labor," "Patrimonial" and "Family" courts (with 19 and 18 years, respectively). In contrast, the average "Patrimonial" is 9 years younger. All other types of courts present values near the mean. Dispersion within each type of court is low.

The average age of the agents is 42 years. Most courts present values close to the mean. The time between promotions averages 7 years. The higher average is close to 10 years in "Labor," whereas the lower average in "Patrimonial" is 4 years. Dispersion within each type of court is low.

4. Performance of the First Instance Federal Courts

This section presents the results of alternative models of efficiency. We first discuss the models. Next, we examine the efficiency levels achieved by jurisdiction and type of court.

4.1 Estimated models

Output-oriented models were used for the estimations since it is assumed that judges have limited control over the amount of labor resources, but they can affect the resolution of causes, which is measured in terms of closed files. In line with most of the literature presented in Section 2, we assume that the production function is essentially labor intensive.

The models take the Resolved Cases as an output and the Personnel for each court as inputs, differentiating between Professionals and Nonprofessionals. In Argentina, each First Instance Court has only one judge. Seniority, Age, Promotion and the Proportion of Temporary Personnel are also included in the models. Also, Workload is included as an environmental variable in a couple of specifications.

We considered four alternative models for the analysis (Table III), assuming CRS in four different specifications in terms of the variables included (using notations 1, 2, 3 and 4 for each specification)[4].

Model 1 used Professional agents, Nonprofessional agents and Seniority as explanatory variables. Workload was added as an environmental variable in Model 2. Models 3 and 4 included the other variables to perform a sensitivity analysis, given that they were not statistically significant in econometric regressions. The difference between them is that

Model 1	Model 2	Model 3	Model 4
Dependent variable			
Resolved Cases	Resolved Cases	Resolved Cases	Resolved Cases
Independent variables			
Professional agents	Professional agents	Professional agents	Professional agents
Nonprofessional agents	Nonprofessional agents	Nonprofessional agents	Nonprofessional agents
Seniority	Seniority	Seniority	Seniority
		Age	Age
		Temporary Personnel Ratio	Temporary Personnel Ratio
		Promotion	Promotion
Environmental variable			
	Workload		Workload

Source: Own elaboration

Table III. DEA model specifications

Model 3 does not consider the environmental variable, whereas that variable is included in Model 4.

The nature of the causes of the different courts requires diverse levels of effort to resolve; therefore, it is plausible to assume that the production functions differ between courts. The grouping by jurisdiction makes the outputs (cases) of the courts within the same jurisdiction more homogeneous. In this way, we can ensure a greater comparability of the courts' production functions.

4.2 Estimations of relative efficiency by jurisdiction

The estimations were made for the period 2006–2010. Table IV presents a summary of the results of the estimates of the relative efficiency for each model and for each sample.

Jurisdictions/type	Model 1	Model 2	Model 3	Model 4	Observations (2010)
All	0.477	0.633	0.550	0.685	433
Criminal	0.357	0.805	0.383	0.808	82
Instruction	*0.323*	*0.796*	*0.334*	*0.797*	*49*
Criminal and Correctional	*0.237*	*0.709*	*0.287*	*0.710*	*12*
Juvenile	*0.074*	*0.831*	*0.116*	*0.833*	*7*
Correctional	*0.721*	*0.907*	*0.769*	*0.919*	*14*
Civil	0.478	0.500	0.551	0.565	110
Family	*0.557*	*0.559*	*0.630*	*0.631*	*24*
Patrimonial	*0.456*	*0.484*	*0.529*	*0.546*	*86*
Commercial	0.621	0.642	0.682	0.702	26
Labor	0.658	0.719	0.789	0.794	80
Penal Economic	0.797	0.815	0.875	0.882	11
Economic	*0.827*	*0.848*	*0.901*	*0.908*	*8*
Tributary	*0.717*	*0.727*	*0.806*	*0.813*	*3*
Civil and Commercial	0.624	0.665	0.673	0.702	11
Contentious Administrative	0.508	0.535	0.574	0.607	18
Fiscal Executions	*0.587*	*0.617*	*0.706*	*0.765*	*6*
Contentious Administrative	*0.469*	*0.495*	*0.507*	*0.527*	*12*
Social Security	0.899	0.910	0.946	0.952	10
Interior	0.259	0.514	0.336	0.570	85

Source: Own elaboration

Table IV. Efficiency levels by jurisdiction (average 2006/2010)

Note that the Federal Justice for the Interior is considered a separate jurisdiction, and it is not compared with civil or commercial jurisdictions.

The proportion of efficient courts increases when the environmental variable is included, though proportions differ according to the jurisdiction. Both the Interior and Criminal jurisdictions stand out with lower levels of efficiency in Model 1, without environmental variables (0.36 and 0.26, respectively). If we consider Workload (Model 2), there is a significant increase in the average efficiency (0.8 and 0.51, respectively).

By and large, the use of Workload as the environmental variable significantly increases the average levels of efficiency. There is a significant drop in the variability coefficients when the environmental variable is considered. In addition, we can see that the inclusion of the rest of the explanatory variables (Models 3 and 4) does not generate significant changes in the aggregate results. The largest differences between Models 1 and 2, and 3 and 4 are in the Criminal and Interior jurisdictions.

The choice of the environmental variable is critical to the comparison of the courts. Some types of courts work with a minimum stock of files, and most of their causes are closed within the year. However, other categories of courts (e.g. Fiscal Executions Courts, Contentious Administrative Courts and the Federal Courts of the Interior) have a significant stock. Thus, the Workload variable is useful to compare courts belonging to diverse jurisdictions or when there are several types of courts that differ by the characteristics of the process and disposition of cases within the same jurisdiction.

By including an environmental variable, higher average levels of inefficiency are still found in the Interior Courts. The average efficiency difference between these types of courts and the rest differs because Interior Courts encompass multiple jurisdictions; hence, they have a different production function than Courts in the Capital.

Since Model 1 does not consider environmental variables and Model 3 is used for reference purposes, we take Model 2 as the relevant model for the analysis.

Table IV also shows the average relative efficiency by type of court. The analysis focuses on those jurisdictions with more than one type of court, to identify differences in performance levels within each jurisdiction.

The inclusion of Workload in the model generates significant increases in the levels of efficiency. The Criminal jurisdiction clearly illustrates this effect. The "Correctional" court differs significantly from the rest with an average efficiency of 0.72 followed by the "Instruction" and "Criminal and Correctional" courts with 0.32 and 0.24, respectively. Finally, the estimated average efficiency for the "Juvenile" courts is only 0.07, a tenth of the efficiency of the "Correctional" courts.

These values are indicative of the existence of comparability problems within jurisdictions. Model 2 solves this problem by including the Workload variable. Table IV shows that the relative performance falls within a much more limited range: the maximum value is 0.91 for "Correctional," and the minimum value is 0.71 for "Criminal and Correctional."

The utilization of Model 3 shows that the results are again affected by differences in the class of cases the courts face. This occurs because of the inclusion of variables such as Promotion, Age and Temporary Personnel Ratio does not have the same effect as the environmental variable. Key to this sensitivity analysis is the choice of the additional explanatory variables. Future research would be necessary to develop the economic model (the selection of variables) using econometric methods.

Figures 1–3 show the behavior of the efficiency estimated by each model (the results have been ordered from the more efficient to the less efficient court in Model 2). In Figure 1, it is hard to distinguish the differences between Models 2 and 4 which are virtually coincident. Also, Models 3 and 4 have slight differences between each other. Thus, in Figures 2 and 3, Models 1 and 2 and 3 and 4, respectively, are compared. The bulk of the

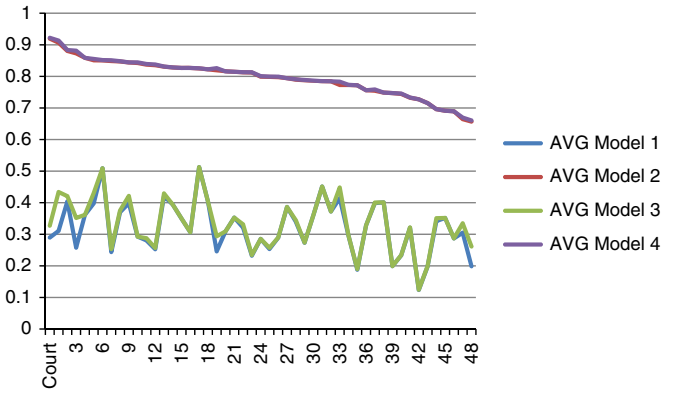

Figure 1.
The efficiency levels of courts in the four models

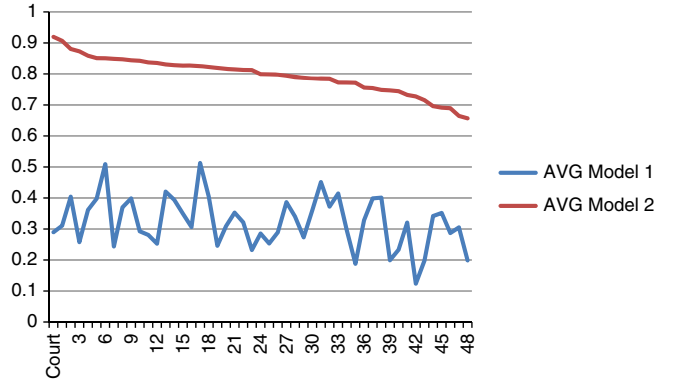

Figure 2.
The efficiency levels of courts in Models 1 and 2

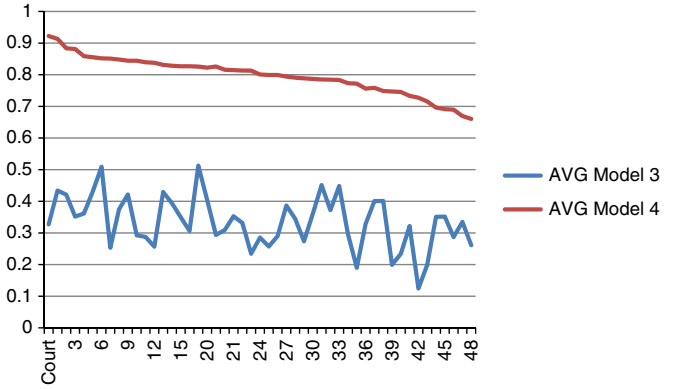

Figure 3.
Efficiency levels of the courts in Models 3 and 4

differences between those pairs of models are attributable basically to the environmental variable "workload." Nevertheless, the slight differences between the pairs of Models 1 and 2, and 3 and 4, respectively, are originated in the other variables which differentiate the models (Age, Temporary Personnel Ratio and Promotion).

5. Determinants of efficiency in Criminal Instruction Courts

This section analyzes the main factors affecting efficiency in the Criminal Instruction Courts applying econometric techniques. First, we present the estimation methodology, including its results for the several models in the period 2006–2010. Next, we interpret the determinants of efficiency in the courts.

The factors affecting the efficiency of the criminal instruction courts were obtained with DEA Model 2. To analyze the correlation between these factors and efficiency measures, we used a fixed-effects panel data model. This model can be represented in the following way:

$$\mathrm{Log}(Y_{it}) = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ki} + \alpha_i + u_t + \varepsilon_{it}, \tag{3}$$

where $\mathrm{Log}(Y_{it})$ represents the logarithms of the efficiency score obtained by utilizing DEA (dependent variable), x_{ki} represents the k explanatory variables (independent variables), α_i is the time invariant fixed effect for each entity, u_t is the control for fixed effects by year and ε_{it} is the error for each entity annually.

The independent variables considered in the different model specifications are Percentage of Female Personnel, Seniority of the Personnel and its square, Age of the Judges and its square, Percentage of Temporary Personnel in the court, Percentage of Tenured Personnel and a dummy for Surrogate Judges. Since the variables (with the exception of the dummies) are expressed in natural logarithms, the estimated coefficients can be interpreted as elasticities.

In Table V, the variables that were utilized are presented in levels. The sample has 245 observations from 2006 to 2010 only corresponding to the Criminal Instruction Courts. As shown, the averages of the variables tend to have a low variability over time, with the Proportion of Surrogate Judges as the exception.

We estimated several models that are presented in Table VI (named with capital letters A, B, C and D to clearly distinguish them from DEA Models 1, 2, 3 and 4). They enable to analyze the significance of the considered variables. Each model represents a different combination of explanatory variables. Employing a fixed-effect model is justified by the need for time and court controls, for improving the estimations since unobserved variables are not affecting the estimated coefficients.

The signs of the coefficients shown in Table VI indicate different effects of variables. The coefficients in all the models keep the same signs, and the significant variables are robust to the specification of the model.

Seniority of the personnel positively affects the efficiency of the criminal instruction courts. This result is expected since the personnel can increase their productivity if they acquire experience. Regarding Age of Judges, the study finds that older judges are more inefficient

	2006	2007	2008	2009	2010	Total
Efficiency	0.4735	0.5159	0.6411	0.6864	0.6247	0.5903
Age of Judges	49	49	50	51	51	50
Age of Personnel	35	34	34	35	35	35
Seniority of Judges	17	18	19	20	20	19
Seniority of Personnel	10	10	10	10	10	10
Proportion of Female Personnel	0.47	0.47	0.5	0.51	0.49	0.49
Proportion of Temporary Personnel	0.28	0.24	0.26	0.24	0.28	0.26
Proportion of Professional Personnel	0.43	0.43	0.45	0.48	0.47	0.45
Proportion of Surrogate Judges	0.16	0.08	0.14	0.27	0.27	0.18

Source: Own elaboration based on "Consejo de la Magistratura"

Table V.
Criminal instruction courts (averages by year)

Variables	Description	Model A	Model B	Model C	Model D
LogSeniority	Seniority of Personnel	3.104* (1.743)	3.030* (1.748)	3.351* (1.733)	3.071* (1.755)
Log^2Seniority	Seniority of Personnel (squared)	−0.709* (0.386)	−0.698* (0.387)	−0.763** (0.383)	−0.703* (0.388)
LogJudgeAge	Age of Judges	−23.61** (11.94)	−22.78* (11.97)	−23.60** (11.91)	−23.52* (11.98)
Log^2JudgeAge	Age of Judges (squared)	2.997* (1.532)	2.895* (1.536)	3.000* (1.528)	2.985* (1.537)
Surrogate	Proportion of Surrogate Judges	0.131* (0.0703)	0.121* (0.0703)	0.129* (0.0718)	0.131* (0.0707)
Temporary	Proportion of Temporary Personnel			0.0998 (0.0642)	
Tenured	Proportion of Tenured Personnel	0.172 (0.109)		0.141 (0.109)	0.173 (0.11)
Female	Proportion of Female Personnel				0.0187 (0.0994)
Constant		42.67* (23.09)	40.98* (23.14)	42.10* (23.03)	42.57* (23.14)
Observations		294	294	283	294
R^2		0.277	0.270	0.275	0.277
Number of courts		49	49	49	49

Notes: Standard errors in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$
Source: Own elaboration

Table VI.
Econometric results

than younger ones. In both cases (i.e. personnel and judges), the effect of Seniority and Age on efficiency is decreasing, as informed by the squared terms. Proportion of Tenured Personnel and Temporary and Female Personnel in each court are not significant in any of the models.

An interesting result is how efficiency is affected by the existence of Surrogate Judges. The sign of this variable is positive. A possible explanation could be that these judges have more incentives to resolve more cases because they seek to be the tenured judge of the court.

6. Conclusions

The purpose of this paper is to build performance indicators to assess efficiency of the First Instance Federal Courts of Argentina. After determining the efficiency scores using DEA with a database for the period 2006–2010, we conducted a search of the determinants of efficiency in the Criminal Instruction Courts (a subset of the former). This paper's contribution is thus twofold: first, to estimate the relative efficiency for all First Instance Federal Courts in every jurisdiction; and second, to explain the differences in the efficiency in Criminal Instruction Courts. The final goal is to contribute to the judicial statistical system, providing benchmarking tools for decision making and policy design.

For analytical purposes, we consider courts as production units whose main output can be measured by the number of legal decisions issued by year. The production process requires a combination of factors as in any production process. Indeed, the "production of justice" is a "labor-intensive task." In this sense, the empirical literature shows that the justice output depends on certain variables, such as the personnel working in the courts, the availability of staff reinforcement in the courts, the educational level of judges and the prospects for the promotion of judges.

The model developed for the Argentine case was an output-oriented model since we assumed that judges have limited control over the amount of labor resources. The Resolved

Cases were used as outputs, and the Personnel of each court were used as input. Workload was included as an environmental variable, which represents the sum of the existing, admitted and readmitted files.

The data cover the period 2006–2010. Four DEA models were developed based on various CRS, and differing by the environmental and explanatory variables considered. The efficiency analysis was performed by subsamples of courts to ensure the comparability of results.

In general, the use of Workload as an environmental variable significantly increases the average levels of efficiency. In particular the Criminal and Interior jurisdictions have the lowest levels of efficiency if the environmental variable is not considered. Otherwise, the average efficiency levels increase significantly. There is also a significant drop in the levels of efficiency score dispersion in these jurisdictions if the environmental variable is accounted for. In other jurisdictions, the inclusion of the environmental variable generates increased levels of efficiency but yields only minor changes in their distribution. The average efficiency difference between the Interior courts and the rest could be based on the fact that the former are multijurisdictional courts deciding over highly heterogeneous causes.

To investigate the factors affecting the efficiency in the Criminal Instruction Courts, we used fixed-effects panel data models. The estimates show some interesting findings. One of them reveals that Surrogate Judges are more efficient than Tenured Judges on average. Another finding shows that courts with Senior Judges have a lower average efficiency.

The work underlying this paper could be seen as a step into the development of consistent statistics and benchmarking indicators and the design of analytical methodologies to estimate relative efficiency for the courts of all federal jurisdictions. Analytical models were applied to explain the differences in efficiency in Criminal Instruction Courts, and this methodology could be extended to other kinds of courts in the future. Future research would be interesting to develop the underlying economic model using econometric methods. Finally, this paper has direct public policy implications, as these methodologies are useful tools to benchmark and set incentive schemes.

Notes

1. "Interior Courts" deal with federal cases in the interior of the country (i.e. provinces outside the city of Buenos Aires).
2. From 2006 to 2016, the difference between the "informed" pending cases at the end of each year and the "adjusted" pending cases represent the 0.17 percent of total pending cases in the same period. The courts with the most adjustments are the National Commercial Courts since the pending cases at the beginning of the year for those courts are not "informed."
3. Standard deviations and dispersion coefficients are not included in the table.
4. Initially, we developed both CRS and VRS models, but no differences were found in efficiency levels. Additionally, we applied a Kolmogorov–Smirnov or K–S test (as developed by Banker and Natarajan, 2004) which allows testing the null hypothesis of constant returns to scale in the sample. The K–S statistic quantifies a distance between the empirical distribution functions of two samples, in this case, efficiency levels under CRS vs efficiency levels under VRS. The low value of the K–S statistic does not allow to reject the null hypothesis.

References

Andersen, B., Henriksen, B. and Spjelkavik, I. (2008), "Benchmarking applications in public sector principal–agent relationships", *Benchmarking: An International Journal*, Vol. 15 No. 6, pp. 723-741.

Banker, R.D. and Natarajan, R. (2004), "Statistical tests based on DEA efficiency scores", in Cooper, W., Seiford, L. and Zhu, J. (Eds), *Handbook on Data Envelopment Analysis*, KAP, Norwell, MA, pp. 299-321.

Beenstock, M. and Haitovsky, Y. (2004), "Does the appointment of judges increase the output of the judiciary?", *International Review of Law and Economics*, Vol. 24 No. 3, pp. 351-369.

Bhattacharya, M. and Smyth, R. (2001), "Aging and productivity among judges: some empirical evidence from the high court of Australia", *Australian Economic Papers*, Vol. 40 No. 2, pp. 199-212.

Buscaglia, E. and Dakolias, M. (1999), *Comparative International Study of Court Performance Indicators: A Descriptive and Analytical Account*, World Bank, Washington, DC.

Buscaglia, E. and Ulen, T. (1997), "A quantitative assessment of efficiency of the judicial sector in Latin America", *International Review of Law and Economics*, Vol. 17, pp. 275-291.

Dalton, T. and Singer, J.M. (2008), "A matter of size: an analysis of court efficiency using hierarchical linear modeling", *Social Science Research Network*.

de Sousa, M.d.C.S and Schwengber, S.B. (2005), "Efficiency estimates for judicial services in Brazil: nonparametric FDH and the expected Ordem-M efficiency scores for Rio Grande do Sul Courts", XXXIII Encontro Nacional da ANPEC, Natal.

Dimitrova-Grajzl, V., Grajzl, P., Sustersic, J. and Zajc, K. (2011), "Judicial incentives and performance at lower courts: evidence from Slovenian judge-level data", *Review of Law and Economics*, Vol. 8 No. 1, pp. 215-252.

Espasa, M. and Esteller-Moré, A. (2011), "Is inefficiency under control in the justice administration?", *Working Paper of Funcas*, Vol. 619, pp. 1-21.

Farrell, M.J. (1957), "The measurement of productive efficiency", *Journal of the Royal Statistical Society: Series A (General)*, Vol. 120 No. 3, pp. 253-290.

Finocchiaro, C. and Guccio, C. (2012), "Searching for the source of technical inefficiency of Italian judicial districts: an empirical investigation", *European Journal of Law and Economics*, Vol. 38, pp. 368-391, doi: 10.1007/s10657-012-9329-0.

Ganley, J.A. and Cubbin, J.S. (1993), *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*, Elsevier, North-Holland.

García-Rubio, M. and Elbialy, N. (2011), "Assessing judicial efficiency of Egyptian first instance courts a DEA analysis", Discussion Paper No. 19-2011, MAGKS, Joint Discussion Paper Series in Economics.

Gorman, M. and Ruggiero, J. (2009), "Evaluating US judicial district prosecutor performance using DEA: are disadvantaged counties more inefficient?", *European Journal of Law and Economics*, Vol. 27 No. 3, pp. 275-283.

Gunasekaran, A. (2005), "Benchmarking in public sector organizations", *Benchmarking: An International Journal*, Vol. 12 No. 4, pp. 401-406.

Gunasekaran, A. and Putnik, G.D. (2006), "Performance measures, benchmarking and best practices in the new economy", *Benchmarking: An International Journal*, Vol. 13 Nos 1-2, pp. 225-230.

Guzowska, M. and Strack, T. (2010), "An examination of the efficiency of polish public sector entities based on public prosecutors offices", *Operations Research and Decisions*, Vol. 2, pp. 41-57.

Hagstedt, K. and Proos, J. (2008), *Has the Recent Restructuring of the Swedish District Courts Improved Efficiency? A DEA Analysis*, Uppsala University, Department of Economics, Bachelor's Thesis (Examensarbete C).

Kittelsen, S.A.C. and Førsund, F.R. (1992), "Efficiency analysis of Norwegian district courts", *The Journal of Productivity Analysis*, Vol. 3 No. 3, pp. 277-306.

Klingner, S., Stephanie, P. and Michael, B. (2015), "Service productivity in different industries – an empirical investigation", *Benchmarking: An International Journal*, Vol. 22 No. 2, pp. 238-253.

Lewin, A., Morey, R. and Cook, T. (1982), "Evaluating the administrative efficiency of courts", *Omega*, Vol. 10, pp. 401-411.

Pastor, S. (2003), "Eficiencia y Eficacia de la Justicia", Papeles de Economía Española No. 95.

Pedraja-Chaparro, F. and Salinas-Jimenez, J. (1996), "An assessment of the efficiency of Spanish courts using DEA", *Journal of Applied Economics*, Vol. 28 No. 11, pp. 1391-1403.

Rosales-López, V. (2008), "Economics of court performance: an empirical analysis", *European Journal of Law and Economics*, Vol. 25, pp. 231-251.

Rosales-López, V. and García-Rubio, M.A. (2010), "Justicia y Economía: Evaluando la Eficiencia Judicial en Andalucía", InDret, Revista para el Análisis del Derecho 4, October.

Ruíz Hernández, A. (2004), "Aplicación del Método de Optimización DEA en la Evaluación de la Eficiencia Técnica de las Seccionales de la Fiscalía", Documentos CEDE 2004, pp. 1-26.

Schneider, M. (2005), "Judicial career incentives and court performance: an empirical study of the German labour courts of appeal", *European Journal of Law and Economics*, Vol. 20, pp. 127-144.

Tulkens, H. (1993), "On FDH efficiency analysis: some methodological issues and applications to retail banking, courts, and Urban transit", *The Journal of Productivity Analysis*, Vol. 4, pp. 183-210.

Wynn-Williams, K.L.H. (2005), "Performance assessment and benchmarking in the public sector: an example from New Zealand", *Benchmarking: An International Journal*, Vol. 12 No. 5, pp. 482-492.

Yasin, M.M. and Gomes, C.F. (2010), "Performance management in service operational settings: a selective literature examination", *Benchmarking: An International Journal*, Vol. 17 No. 2, pp. 214-231.

Yeung, L. (2008), "Measuring the efficiency of Brazilian courts from 2006 to 2008: what do the numbers tell us?", Working Paper No. WPE: 251/2011, Insper.

Yeung, L. and Azevedo, P. (2010), "Beyond conventional wisdom and anecdotal evidence: measuring efficiency of Brazilian courts", Anais do XXXVII Encontro Nacional de Economia 122, ANPEC.

Further reading

Banker, R., Charnes, A. and Cooper, W. (1984), "Some Models for estimating technical and scale inefficiencies in data envelopment analysis", *Management Science*, Vol. 30 No. 9, pp. 1078-1092.

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

Corresponding author
Gustavo Ferro can be contacted at: gferro05@yahoo.com.ar