ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

LIV REUNIÓN ANUAL | NOVIEMBRE DE 2019

Benchmarking Labor Courts:
An Efficiency Frontier Analysis

Ferro, Gustavo Adolfo
Oubiña, Victoria
Romero, Carlos A.

ISSN 1852-0022 / ISBN 978-987-28590-7-7

Benchmarking Labor Courts: an Efficiency Frontier Analysis

Gustavo Ferro[1], Victoria Oubiña[2] and Carlos A. Romero[3]

Abstract

Our aim is to answer: What are the average efficiency levels of the Argentine labor courts and how do they behave individually with respect to the average? Which are the determinants of the relative efficiency levels of those courts? In doing so we estimate a Data Envelopment Analysis efficiency frontier, and in a second stage we analyze the efficiency scores drivers and the Judiciary career incentives. Our sample are 80 courts during the period 2006-2012.Our findings are of average high levels of efficiency, nonetheless with room for improvements of 9 to 12 percent the output with the same inputs on average, whether variable or constant returns are considered, respectively. The efficiency results take into account caseload and backlog. No measure of capital input is used, thus the analysis is for the short run. The analysis of the efficiency scores against its determinants shows that more variables, outside our sample, are playing a role to explain the variance of the scores.

Resumen

Nuestra meta es responder: ¿Cuál es el nivel de eficiencia relativa promedio de los tribunales laborales en Argentina? ¿Cuáles son sus determinantes? Para contestarlas, estimamos fronteras de eficiencia mediante Análisis Envolvente de Datos y en una segunda etapa analizamos los determinantes de los niveles de eficiencia relativos y los incentivos de la carrera judicial. Nuestra muestra es de 80 tribunales durante el período 2006-2012. Nuestros hallazgos son de un nivel promedio de eficiencia alto, con posibilidades de mejorar el producto entre un 9 y un 12% con los mismos insumos según se considere el modelo de rendimientos variables a escala o el de rendimientos constantes. Los resultados de eficiencia toman en cuenta los casos acumulados y los tiempos promedio para resolver cada pleito. Ninguna medida de capital físico se utiliza, por lo cual el análisis es de corto plazo. El análisis de los determinantes de los niveles relativos de eficiencia indica que las variables de nuestra base de datos explican poco de la varianza de los resultados.

Keywords: Courts, Efficiency, Benchmarking, Data Envelopment Analysis

JEL Codes: C65, D24

[1] Universidad del CEMA (UCEMA) and CONICET. Córdoba 374, (1054) Buenos Aires, Argentina. Telephone: +54-11-63143000, extension 780, email: gaf97@ucema.edu.ar and gferro05@yahoo.com.ar. Corresponding Author.

[2] Universidad de Buenos Aires (UBA). Email: v.oubina@gmail.com

[3] Instituto Interdisciplinario de Economía Política, Universidad de Buenos Aires (UBA). Email: carlos.adrian.romero@gmail.com

1. Introduction

The efficiency of the judiciary is a first order concern for citizenship worldwide. Length of trials, costs of the system and difficulty in the procedures rise concerns on the efficacy and efficiency of justice in many places. It is relevant to analyze the differences in efficiency of individual courts and to identify best practices for benchmarking. That tasks demands to measure output, inputs, quality and environmental conditions.

Court efficiency can be defined as maximizing court output produced by court inputs holding court output quality constant. In the same vein, best practices can be defined as the fastest method that uses the least inputs to produce the highest quality output (Djafari, 2012). Additionally, performance can be measured over time to incorporate the dynamics and/or to evaluate reforms (such as mergers), according Matsson et al. (2018).

Efficiency Frontier Analysis can yield some clues on how to improve relative efficiency. Therapies need a prior diagnosis, and these methods are appropriate to deal with questions such as: what are the average efficiency levels of the Argentine labor courts and how do they behave individually with respect to the average? Which are the determinants of the relative efficiency levels of those courts? The choice and implementation of remuneration or other rewards is expected to influence results. Even when procedures (process technology) and career paths are regulated out of the courts, this kind of study provides the opportunity to discuss incentives “out of the box” of the judiciary. In this paper, we aim to respond those questions.

In doing so, we estimate technical efficiency of labor courts of Argentina by means of a Data Envelopment Analysis which uses a database produced by the Judicial System of Argentina (“Poder Judicial de la Nación” or PJN for short). The PJN compiles statistical information of all their branches and some recent efforts have been made to improve the quantity, quality and availability of the information. That information is comprehensive, with respect to personnel and its characteristics (age, seniority, education, tenure and genre), cases (incoming, being resolved and resolved), etc. That is, the available information permits identifying some inputs and their qualitative attributes, outputs as well as quality and environmental (contextual) information on some peculiarities of every court.

Unfortunately, it is not easy to work on some jurisdictions (for instance Civil Courts) since the diversity of cases’ nature, which difficult the efficiency comparison, and also the dissimilar procedures of the different jurisdictions in a federal country which introduce noise in the analysis. For these reasons, we focus our study to Labor Courts, where the “output” is more standardized and thus it is more comparable.

We have information on 80 Labor Courts for the period 2006-2012 in a homogeneous database. Each court is taken as a decision unit which produces certain number of sentences each year and in doing so uses inputs (mainly labor, including judges, degreed and not degreed employees which work as secretaries, clerical employees, and the like). We do not have information concerning physical inputs, such as computers, square meters of offices, software, thus our analysis assumes they are constant in the period of analysis, implying that ours is a short run assessment. Courts cannot control the inflow of the cases and the extent of disposable resources. Rules on staff appointment, remuneration and promotion (the incentive structure) are in a great extent exogenous to the courts. Thus judges have limited ability to incentivize the staff to increase efficiency and productivity, at least, they cannot change the career rules, exogenously determined.

The paper organizes as follows: Section 2 accounts for the literature review and court efficiency, being its main purpose to detect the variables used in efficiency frontier studies of the judiciary; the Section 3 explains the methodology and data employed; the Section 4 is for results; it shows the different models that were estimated, the numerical results and inferences; and finally, the Section 5 presents the conclusions of the study.

2. Literature Review

The purpose of this section is to provide background and context to the study.

The results or outcome of the service of justice provision (outputs) are resolved cases, that is, decisions which award sentences and other forms of resolution to disputes (agreements, mediations, transferring of decisions to other jurisdiction, etcetera). It is worth recognizing that different jurisdictions or different cases within the same jurisdiction demand more or less resources (staff time, procedures, etcetera), and because of that reason the heterogeneity should be recognized. We address it focusing in the Labor jurisdiction, more standardized cases than the existent in other jurisdictions such as criminal or civil courts. Efficiency can be affected by external factors. Gorman and Ruggiero (2009) focus on the staff efficiency of public prosecutors' offices 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 due partly 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.

The production of justice services is labor-intensive. The variables most commonly used for inputs (see Table 1) are the number of personnel working in the courts -distinguishing into different types, if data is available-, some indication of other resources devoted to the service (square meters of office, computers, etcetera, can be included also if data is disposable), and the caseload (stock from precedent periods plus new arrived pending cases), which is the raw material of the task, as well as can be understood as the demand of justice.

Judiciaries can be considered as internal labor markets in which the main incentives derives from career opportunities. Age, seniority, tenure and gender of staff or judges can drive court efficiency. Espasa and Esteller-Moré (2011) suggest that average efficiency tends to increase with time, and identify evidence that hiring temporary workers is significantly less effective compared to tenured employees. Some qualitative aspects can enrich the interpretation of the value added by each input, such as the educational level of the judges (Schneider, 2005), the prospects for the promotion of judges and staff, the time devoted by the judges to administrative issues (Yeung and Azevedo, 2011), and beyond the labor inputs, another important issue is the delay in solving cases. The backlog and the speed or timeliness of the decisions, give an idea of the diligence and quality of the task assignments' completion (Rosales-López and García-Rubio, 2010) as well as the proportion of appeals give clues of the quality of the judges' sentences.

To the output, inputs and quality variables (the latter being in some respects under control of the courts) environmental drivers can be add. The distinguishing feature of the latter is their condition of uncontrollable by the courts.

The units of measure can be physical (as number of full time equivalent staff, or square meters of office space) or monetary (as expenditures in salaries or in other resources). Some of the studies that evaluate judicial efficiency have been carried out in the context of judicial reform undertaken by different countries, such as Sweden, the Netherlands or Italy (Hagsted and Proos, 2008, Djafari, 2012, Falavigna. et al., 2015, Finocchiaro Castro and Guccio, 2018). The relationship between efficiency and court size is a critical aspect to explore. Gusowska and Strak (2010) find that the inefficiency of several offices can be attributed to the inappropriate scale of operations.

We examine in detail the literature, and besides some conceptual papers, we concentrated in 30 empirical ones. First, we develop chronologically its purpose and main findings, and then we present in the Table 1 an exhaustive summary of the empirical literature on court efficiency.

The first article we detected is Lewin et al (1982) who study efficiency of criminal courts and judicial districts in North Carolina (US), finding a fairly inefficient judiciary, both at the district and county level. Kittelsen and Forsund (1992) explore Norwegian district courts' efficiency, finding that most of inefficiency was scale inefficiency with inefficient courts being smaller on average. Tulkens (1993) evaluates courts productivity and backlogs in Belgium, determining that courts efficiency have room for improvements. Chaparro and Jiménez (1996) provide a measure of technical efficiency of the administrative litigation division of Spanish high courts. Both the efficiency and avoidable delays have been calculated. Pedraja-Chaparro and Salinas-Jiménez (1996) assess the efficiency of the administrative litigation division of Spanish high courts, finding considerable scope for improvement. Bhattacharya and Smyth (2001) study the relationship between aging and productivity for a sample of retired judges, they achieve that productivity increase, reach a peak and thus declines closing to retirement. Beenstock and Haitovsky (2004) study efficiency of courts in Israel. They find that judges complete more sentences when facing heavy caseload and courts complete less cases when new judges are appointed. Marselli and Vannini (2004) examine the efficiency of the Italian district courts. They observe low efficiency and productivity increase across time, mostly due to technical change. Schneider (2005) analyzes how judges' education and careers affect courts productivity and pointed out that caseload inclusion is needed to avoid underestimation of productivity. They find positive and significant relationship between judicial efficiency and salaries. Also, judges with ex ante high likelihood of promotion are less productive.

Recently some Judiciary systems were reformed, giving room for more efficiency analysis, in this case of the before and after the court mergers. Hagstedt and Proos (2008) develop an efficiency analysis after restructuring in Swedish district courts to study whether efficiency improved after court reduction, finding overall increasing in efficiency, with many units operating at decreasing returns to scale. Gorman and Ruggiero (2009) study efficiency of prosecutor offices in Italy, with part time magistrates. Prosecutor offices in more socio-economically disadvantaged counties are more inefficient. Guzowska and Strack (2010) research on efficiency of public prosecution organizational units, as well as returns to scale. Potential savings are determined. Nissi and Rapposelli (2010) analyze the productive efficiency of the Italian Courts of Appeal. Best practices are identified; diseconomies of scale are documented. Elbialy and García Rubio (2011) analyze the performance of first instance courts 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.

The incentives to staff and careers for judges are also a promising field to apply efficiency studies. Espasa and Esteller-Moré (2011) address inefficiency of the justice administration in Spain and its drivers. They observe that the greater the percentage of temporary judges, the lower is the efficiency of the courts. Yeung and Azevedo (2011) evaluate efficiency across Brazilian state courts, finding considerable efficiency variations across courts, depending on their internal management and organization. Deyneli (2012) determines the relationship between justice efficiency and judges salaries in 22 different European countries. Positive and significant relationship exist between justice efficiency and judge salaries. Dimitrova-Grajzl et al (2012) examine the performance of lower court judges in Slovenia, documenting possible tradeoffs between quantity and quality of judicial case resolution.

Heterogeneous matters and court practices can influence efficiency. Ferrandino (2014) estimate the efficiency of criminal courts in the US, finding only part of the courts acts efficiently. Odhiambo (2014) track the technical efficiency of the Kenyan Judiciary (first instance courts) after a reform, observing an improvement in technical efficiency. Santos and Amado (2014) measure the incidence of scale factors on efficiency and the need of reform in

Portugal. They show size affects efficiency: small courts are inefficient, suggesting the closing of small courts. Falavigna et al., (2015) analyze the impact of structural changes in Italian district courts. They determine that the role of judges is correlated with court productivity and efficiency. Major (2015) addresses efficiency of Polish courts in Cracow, worried by backlogs and determined that there is room for shorten the queue of pending cases. Ippoliti and Ramello (2016) measure efficiency of courts system in Italy and try to identify main drivers of efficiency, in tax (part time) judges. Find that with judges who maximize utility, their opportunity cost determine the time devoted to Judiciary and to private activity.

Among the more recent studies, Ferro et al. (2018) estimate efficiency of criminal courts and drivers of the judiciary careers in Argentina's federal courts. They find that caseload is an important environmental variable. Also, surrogate judges and temporary staff are more efficient on average than tenured judges and staff. Finocchiaro Castro and Guccio (2018) assess the financial efficiency gains after the merging of Italian courts to increase its size. They find low efficiency and small number of courts fully efficient. Fusco et al (2018) analyze the efficiency of Italian judicial districts, observing technical efficiency consistent and stable. Mattsson et al. (2018) study TFP in Swedish courts, showing considerable scope for improvements and pointing that less efficient courts can catch up more efficient ones. Rushid (2018) measures Swedish district courts' technical efficiency in a merger context. The majority of the shut courts were among the lowest efficiency scores. Finally, Yeung (2018) estimates efficiency of Brazilian judiciary and its dynamism in recent years. Results seem useful for evaluating recent discussions on judiciary efficiency.

Table 1: Summary of Literature Review

Author	Method	Yearly sample	Period	Place	Outputs	Inputs	Environmental	Results
Beenstock and Haitovsky (2004)	Regression	25	1964-95	Israel	Case completion	Cases lodged, cases pending, number of judges		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.
Bhattacharya and Smyth (2001)	Regression	28	1999	Australia	Productivity measured as citations	Age, age square		Productivity of judges has a profile similar to academia careers
Chaparro and Jiménez (1996)	DEA	21	1991	Spain	Cases ending in Judgement; cases ending in another resolution	Judges, Staff		77% of the courts were inefficient
Deyneli (2012)	DEA, two stages	22	2006	22 European Countries	Resolved cases	Judges, Staff	Judge salaries, judge education and number of courts	Efficiency is identified at the level of national judiciary systems
Dimitrova-Grajzl et al (2012)	Regression	360	2008	Slovenia	Resolved cases	Judges, staff	Education, experience, salary, gender, recent promotion, imminent	Judges with imminent promotion tend to be less productive

Author	Method	Yearly sample	Period	Place	Outputs	Inputs	Environmental	Results
							promotion	
Elbialy and García Rubio (2011)	DEA	22	2010	Egypt	Resolved cases	Judges, staff, computers per court		Courts could increase their number of solved civil cases by 25.8% and criminal cases by 33.6%
Espasa and Esteller-Moré (2011)	SFA	119	2005-2009	Catalonia, Spain	Resolved cases	staff	Various considering vacancies, temporary workers, vacancies, pending cases, trends, immigrants percentage	Efficiency of the courts tend to remain low
Falavigna. Et al. (2015)	DEA, directional function	309	2009-2011	Italy	Resolved cases	Judges	Court delays	Low productivity increase across time, mostly technical change
Ferrandino (2014)	DEA	20	1993-2008	Florida, USA	Resolved cases per judge	Judges		Wide variability in criminal courts efficiency within and between circuits
Ferro et al. (2018)	DEA, two stages	82	2006-2010	Argentina	Resolved cases	Professional staff, non-professional staff, age, seniority	Workload	Senior Judges have a lower average efficiency. Interior, less specialized courts are less efficient.
Finocchiaro Castro and Guccio (2018)	DEA, two stages			Italy	Resolved civil cases, resolved criminal cases	Judges, Staff, Civil caseload and Criminal caseload		Italian district courts on average largely technically inefficient
Fusco et al (2018)	DEA	26	2005-2011	Italy	Enrolled trials, unresolved trials, resolved trials	Magistrates, Wiretapping expenses		On average, 20 of 26 districts are inefficient with an average efficiency score of 27.2%
Gorman and Ruggiero (2009)	DEA	151	2001	US	Felony jury verdicts, felony cases closed, misdemeanor cases closes, population	Prosecutors, other staff	Median income, minority population, bachelor's degree	Average technical efficiency of 0.68. Only 14 % of units were efficient

Author	Method	Yearly sample	Period	Place	Outputs	Inputs	Environmental	Results
Guzowska and Strack (2010)	DEA	45	2007	Poland	Completed criminal cases, criminal cases discontinued , refusal to institute proceedings	Specialists, auxiliary staff, other inputs, various costs measures		Expenditure could been reduced in 12.6 percent in 36 non efficient regions
Hagstedt and Proos (2008)	DEA	21	2006-2007	Sweden	Case settled per 1000 inhabitants	Wages, administrative costs		Important increases in efficiency and evidence of economies of scale achieved.
Ippoliti and Ramello (2016)	DEA, two stages	103	2009-2011	Italy	Resolved cases	Judges, Pending cases, Incoming cases	Technical efficiency score	The stock of accumulated cases has an impact on the timing affecting the resolution of new incoming cases
Kittelsen and Forsund (1992)	DEA, Malmquist	107	1983-86	Norway	Seven crime categories	Judges, Staff		Average efficiency of 90%, scale problems explain inefficiency. 6% of growth in productivity on average for the whole period
Lewin et al (1982)	DEA	30		North Carolina, USA	Number of judgements, number of cases pending less than 90 days	Caseload, Number of district attorneys and staff, Days of a court held, Number of misdemeanors within the caseload, Size of population		11 out of 30 district courts were deemed as inefficient
Major (2015)	DEA	26	2013	Poland	Resolved cases	Judges, Assistants, Officials		Capacity of the courts allows them to shorten the caseload
Mattsson et al. (2018)	DEA, Malmquist, Bootstrap	48	2012-2015	Sweden	Resolved civil cases, resolved criminal, real estate and environmental cases and resolved matters	Judges, Clerks, Other personnel, Office space		1.7% average decline (geometric mean)
Nissi and Rapposelli (2010)	DEA	26	2008	Italy	Finished cases	Number of Judges, New cases filed during the year, Number of pending cases (filed+pending=caseload)		Avg. Efficiency 0,82 (CRS) to 0,9082 (VRS)

Author	Method	Yearly sample	Period	Place	Outputs	Inputs	Environmental	Results
Odhiambo (2014)	FDH	119	2014-2016	Kenya	Civil cases criminal cases	Filed cases, staff		Courts resolved more criminal than civil cases
Pedraja-Chaparro and Salinas-Jiménez (1996)	DEA	21	1991	Spain	Cases fully resolved, other resolved cases	Number of Judges, Number of staff		Average efficiency of 77.4%. Only 23.8% of courts were technically efficient
Rushid (2018)	Stochastic Distance Function	102	2000-2016	Sweden	Resolved criminal cases, resolved civil cases, resolved matters	Judges, Clerks, Other personnel (full time equivalent)		93% average efficiency, great variation across the country
Santos and Amado (2014)	DEA	223	2007-2011	Portugal	43 different case types (civil jurisdiction)	Judges, Staff		Less of 16% courts are efficient
Schneider (2005)	DEA, two stages	9	1980-1998	Germany	Number of cases	Judges, Caseload	Judges salaries, training and number of courts	Average efficiency score of 88.8%. Promotion probabilities have a negative impact on the DEA scores
Tulkens (1993)	Free Disposal Hull	187	1983-85	Belgium	Civil and commercial settled cases, Family arbitration sessions held, minor offense cases settled	Staff		Percentage of inefficient courts of 82% to 87%. Only 35% of backlog solvable by productivity increase.
Yeung (2018)	DEA	27	2009-2015	Brazil	Number of decisions, weighted by workload	Judges, Staff		Low scores for most states. Similar rankings than in Yeing and Azevedo (2011)
Yeung and Azevedo (2011)	DEA	27		Brazil	Number of first and second instance judgements	Judges, Staff		Average efficiency of 42%

Source: Own Elaboration

3. Method, Database and Models

There are two families of techniques for measuring relative efficiency through frontier estimation: parametric (regression based), and non-parametric (mathematical programming). Efficiency refers to the ability of decision units (courts for instance) in maximizing output (sentences and other forms of dispute settlement) given inputs (human and non-human resources), as well as in minimizing costs, maximizing revenues, or maximizing profits –if applicable-, conditional on the existing technology.

The non-parametric methods model the productive process, but they do not estimate parameters of a function, being the most common the Data Envelope Analysis (DEA) method. This seeks to determine which units form an envelope surface with respect to the sample, and characterizes the set of efficient producers as those on the frontier. Inefficiency is measured in terms of how far each observation deviates from the most efficient "peers". DEA yields scores between 0 (totally inefficient) and 1 (completely efficient). As a deterministic model, all distance from every decision unit to the frontier is deemed as inefficiency, not considering randomness neither statistical noise separately.

There are two main different kinds of envelopment surfaces: one assumes constant returns to scale (Charnes et al., 1978) while the other supposes variable returns to scale (Banker et al., 1984). Technical efficiency DEA models can be also input-oriented, output-oriented, or not oriented at all. They differ in terms in which direction the distance of each unit from the frontier is measured.

Productivity refers to changes in technology over time, such that a decision unit can generate more output with a given amount of inputs -technical progress- or less output with a given amount of inputs -technical regress-. When efficiency is studied in different periods, the change in productivity of each unit can be decomposed as catching up to the frontier plus the shifting of the latter. The Malmquist index measures productivity change, technical change and efficiency change. But it assumes Constant Returns to Scale, which can be a restrictive assumption.

3.1 Method

We employ DEA efficiency models to assess efficiency levels. DEA evaluates the relative efficiency of N decision units. Each one has m inputs and r outputs (Charnes et al, 1978). The input vector of decision unit i, x_i, is thus m dimensional, and its output vector, y_i, is r dimensional. For each i, an optimization problem is solved to find the optimal weights of the inputs, v_i and the optimal weights of the outputs, u_i, all non-negative, which maximize the ratio between the sum of the weighted output and the sum of the weighted inputs, while securing that all relative efficiencies are less than or equal to 1.

The models can be formulated as N linear programming problems. In the output oriented specification, the sum of the weighted outputs is maximized by one unit of weighted inputs; while in input oriented approach, the sum of the weighted inputs is minimized per one unit of weighted outputs.

Given that court authorities have a limited control over the inputs and, instead, they normally manage outputs, we made use of an output oriented model. Assuming constant returns to scale (CRS), the model can be characterized as follows:

$$max\ _{\lambda,\theta}\ \delta\ subject\ to{:}\,Y\lambda \geq \delta y_i\,; X\lambda \leq x_i\,; \lambda_i \geq 0 \qquad (1)$$

Where for each court i=1,…, N there is a vector y_i for outputs and a vector x_i for inputs. Y and X are the corresponding matrices for output and inputs representing the data for the N courts. This problem is solved N times: this means that it solved for each of the courts in the sample, whereby we can obtain the level of technical efficiency for each court i.

The technical inefficiency obtained is measured as the proportional increase of output possible while maintaining fixed levels of inputs. The δ variable is greater than or equal to 1 and $1/\delta = \theta$ is the efficiency score.

Assuming the existence of variable returns to scale (VRS), enables us to incorporate the possibility of distinguishing between pure technical inefficiency and inefficiency that could be due to differences in the scale of production in the DEA. Adding a convexity restriction, $e\lambda = 1$, to the CRS problem, we obtain the following model:

$$max\ _{\lambda,\theta}\ \delta\ subject\ to\colon Y\lambda \geq \delta y_i\,; X\lambda \leq x_i\,; e\lambda = 1; \lambda_i \geq 0 \quad (2)$$

Where e is a row vector with all its elements equal to 1. When VRS are assumed, only efficiency between courts of similar scale is compared. As a result, some courts that were inefficient under the assumption of CRS, can be efficient when VRS are assumed. Efficiency scores obtained under VRS should be greater than or equal to those obtained under the assumption of CRS.

Banker and Natarajan (2004) develop a test of returns to scale using efficiency levels estimated under the assumptions of CRS (EC^C) and VRS (EC^V). By definition, $E^V \geq E^C$.

The non-parametric Kolmogorov-Smirnov Test, can be used if there is no assumption of a probability distribution for E^V, given by the maximum vertical distance between $F^V[\ln(1/E^V)]$ and $F^C[\ln(1/E^C)]$.The Test determines whether two samples differ significantly or whether they come from the same distribution, without making assumptions about the distribution itself. The Kolmogorov-Smirnov Test uses the maximum vertical distance between the two distributions, calling that distance D. The value of statistic D only considers the relative distribution of the data.

This test evaluates the null hypothesis that there is no scale efficiency (no VRS) or, alternatively, the null hypothesis indicates the presence of CRS. The D statistic varies in the range [0, 1]. A high value indicates the existence of significant scale efficiency in the sample. Thus, the VRS model would better represent the phenomenon under study.

One form to address the progress in efficiency from one period to the next is using Malmquist Index. This analyzes the total factor productivity (TFP) change of a decision unit over time. The TFP change has two main components: the shift in the production frontier over time ("technical change"), and the shift in the decision unit situation with respect to the former over time ("technical efficiency change").

3.2 Database

The precedent literature review highlighted the variables which are worth considering in explaining output, when analyzing a non-manufacturing, labor-intensive sector: staff (differentiating quality if possible), caseload, backlog and some indication of the length of trials. In a second stage, the DEA scores can be regressed against the characteristic of workers as age, seniority, their squares, education, the condition of temporary or tenured, genre, and the variables which can give some information of the heterogeneity of the judges.

In this case, the database provides the following possibilities for the inputs in the production function of courts: personnel discrimination between judges and the rest of the court staff; and two categories of the latter by qualification, those who hold a college degree and those who do not. For the second stage, we have information of age and seniority of the staff, genre; average time to promotion; and the condition of temporary or tenured personnel (a temporary judge is known as surrogant. A surrogant judge is normally a tenured one of another court). The squares of age and seniority of both personnel and judges, yield some clues on the increasing, decreasing or linear effect of those variables on output.

Concerning the output, the database contains information about Existent Cases, Incoming Cases, Re-Incoming Cases, Out-Coming Cases, Stopped Cases, Sentences and Caseload (Stock). Every period, a court starts with a stock of cases, which can be understood as the raw material of the productive process. Those cases can end as Out-Coming Cases (not ever by sentences). Incoming and Re-Incoming Cases arrive during the period. Of the out-coming cases, those resolved are the ones with sentences, and that will be our output.

Two qualitative variables were built: one establishes the average age, in days, of the cases in a court from its arriving to its sentence, the other one measures the proportion of sentenced cases which are appealed. The implicit assumption is that most technically solid sentences have less likelihood of being appealed.

In the Table 2 are listed and described each database variable and its unit of measurement. There are variables to design output, inputs (personnel and its characteristics of type, qualifications, age, seniority, promotion, genre, and its condition of temporary or tenured, and the unfinished cases).

Table 2: Definition of the Database Variables

Variable	Meaning	Type of Variable	Unit
OUTCOMING	Total Solved Cases in the year	Output	Number/Year
TRAMITING	Total Stopped Cases in the year	Output	Number/Year
SENTENCED	Total Sentences in the year	Output	Number/Year
EXISTENT	Initial Stock of Cases in the year	Input (raw material)	Number/Year
INCOMING	Total Incoming Cases in the Year	Input (raw material)	Number/Year
RE-INCOMING	Total Re-Incoming Cases in the Year	Input (raw material)	Number/Year
CASELOAD	Total Cases (existent + incoming + re-incoming) in the year	Input (raw material)	Number/Year
PROF	Total Staff with College Degree	Input (qualified labor)	Number/Year
NOPROF	Total of Personnel with no College Degree	Input (not qualified labor)	Number/Year
SENSTAFF	Seniority of all Personnel in Months	Input quality (seniority)	Months
SENJUDGE	Seniority of Judges in Months	Input quality (seniority)	Months
AGESTAFF	Age of All Personnel in Months	Input quality (age)	Months
AGEJUDGE	Age of Judges in Months	Input quality (age)	Months
SURROGANT	Surrogant Judge (1 yes, 0 no)	Environmental	Dummy
TENURED	Proportion of Tenured Personnel on Total	Environmental	Proportion
PROMOTION	Average Time to Promotion in Months	Environmental	Months
FEMALE	Proportion of Female Personnel on Total	Environmental	Proportion
APPEALED	1/Re-Incoming	Quality	Proportion
TURNOVER	Average time from incoming to sentenced case	Quality	Days
PRODUCTIVITYSTAFF	Ratio Sentenced/Staff	Quality	Proportion

Source: Own Elaboration based on data provided by the PJN.

In the Table 3 the descriptive statistics of the database are presented. There are 80 courts in our sample, with data for 7 years (2006-2012), totaling 560 for most of the variables (there are some missing observations in some variables).

The average Resolved Cases (Sentenced) is 200, with a standard deviation of 49. The smaller court produced 82 and the larger yields 541 sentences. On average, each court has a staff of 12 (7 professionals and 5 not professionals, with mean seniorities of 241 and 188 months respectively, that is 20 and 15 years). Judges are on average older and with more seniority than the personnel (704 and 349 months respectively, or 58 and 29 years). The 27 percent of the judges are Surrogant (not tenured), while on average 81 percent of the personnel are tenured. The mean time to promotion is 108 months (11 years) and 65 percent of the personnel are female.

The average sentence demands 511 days, with a standard deviation of 167 days, a minimum value of 169 and a maximum one of 1029. A simple measure of average productivity of personnel is 17.56 sentences per employee, with a standard deviation of 8, a minimum of 7.5, and a maximum of 64.6

Table 3: Descriptive Statistics of the Database (2006-2012)

Variable (Short Denomination)	Observations	Average	Standard Deviation	Minimum	Maximum
OUT-COMING	560	400.59	85.62	232.00	778.00
TRAMITING	560	712.06	329.74	155.00	1746.00
SENTENCED	560	200.43	49.16	82.00	541.00
EXISTENT	560	589.48	265.92	136.00	1406.00
INCOMING	560	509.56	135.87	288.00	839.00
RE-INCOMING	560	12.20	16.86	0.00	134.00
CASELOAD	560	1111.26	373.39	471.00	2133
PROF	560	7.38	2.14	2.00	13.00
NOPROF	560	5.07	1.99	0.00	10.00
SENSTAFF	560	220.56	29.29	122.26	309.58
SENJUDGE	560	348.90	134.25	0.00	744.12
AGESTAFF	560	536.94	35.29	438.06	684.59
AGEJUDGE	560	705.10	105.93	425.60	1003.23
SURROGANT	560	0.27	0.44	0.40	1.00
TENURED	560	0.81	0.15	0.40	1.00
PROMOTION	560	108.68	25.37	49.76	206.89
DEGREE	560	0.60	0.13	0.22	1.00
FEMALE	560	0.65	0.14	0.20	1.00
APPEALED	560	0.29	0.30	0.01	1.00
TURNOVER	560	511.41	167.51	169.40	1029.14
PRODUCTIVITYSTAFF	560	17.56	8.06	7.56	64.67

Source: Own Elaboration based on data provided by the PJN.

3.3. The Models

The productive process estimates sentences (resolved cases or output), the average length of trials in days and the appealed sentences as qualitative indicators of the outcomes. The inputs are professional and non-professional staff (proxies for skilled and non-skilled labor respectively), the raw material (caseload). We estimate output oriented models, exploring CRS and VRS versions and testing RS by a Kolmogorov-Smirnoff Test.

The results are tentatively sensitive to some characteristics of the labor force: its age, genre, seniority, time since last promotion, and temporary or tenured condition, and age, seniority and temporary or tenured condition for the judge in charge of the court, and we test those characteristic in the second stage, after the estimation of the DEA models. It is important to highlight some features of the judiciary career in the country under study. First, temporary personnel are by definition not tenured and subject to be ceased by the not renewal of their contracts, while permanent personnel are tenured and cannot be fired (except by a felony). Second, judges can be surrogant (temporary) or tenured (lifelong if appointed before 1994, when the federal constitution was amended and established a mandatory age of retirement of 75, intended for judges appointed after the amendment). Third, the age of retirement for the personnel is 65 for male and 60 for female.

We estimate a pair of models, named A1 and B1, where CRS and VRS versions are estimated. The output variables are SENTENCED (cases), the TURNOVER (or average time length to settle) and APPEALED (proportion of sentences). The inputs are PROF, NOPROF and CASELOAD.

4. Discussion of Results

4.1 Efficiency

The estimates show a consistent pattern. The CRS model has lower average efficiency scores than the VRS one, with higher deviation, and less number of efficient courts (Table 5). The evolution of efficiency scores across the periods does not show stark variations.

Table 4: Results: Efficiency Scores and Efficient Units

Model A1 CRS	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Average	0.866	0.854	0.879	0.875	0.910	0.852	0.899
Standard Deviation	0.071	0.072	0.069	0.069	0.066	0.078	0.073
Mean Deviation	0.053	0.053	0.052	0.050	0.051	0.061	0.055
Coefficient of Variation	0.082	0.085	0.079	0.079	0.073	0.092	0.082
# Efficient Courts	5	4	6	6	6	6	6
Model B1 VRS	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Average	0.893	0.926	0.912	0.918	0.935	0.887	0.934
Standard Deviation	0.070	0.054	0.055	0.057	0.052	0.070	0.058
Mean Deviation	0.055	0.045	0.043	0.043	0.043	0.058	0.047
Coefficient of Variation	0.078	0.058	0.061	0.063	0.056	0.079	0.062
# Efficient Courts	12	13	9	14	15	9	18

Source: Own Elaboration

The Table 6 groups the results for the whole period, which permits corroborating the impressions given for the yearly comparisons. We also add the statistics of the variable PRODUCTIVITYSTAFF which is simply a sentenced on staff ratio. The Table 6 also shows the correlation between the efficiency scores of the different models. The correlation between A1 and B1 models (basic ones) is 0.67. Also the correlation among A2 and B2 is strong (0.83). The correlation between both CRS models (A1 and A2) is 0.61, and between both VRS models (B1 and B2) is 0.72. The correlation between the simple productivity measure with the four models is low, from 0.2 with model A1 to 0.4 with model B1 (the highest).

Table 5: Descriptive Statistics and Correlation among Efficiency Scores and Labor Productivity

Variable	**Observations**	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
A1	560	0.8764571	0.0739181	0.5780	1.00
B1	560	0.9151518	0.0620277	0.6810	1.00
PRODUCTIVITYSTAFF	560	17.558340	8.0643470	7.5600	64.67
Correlation	**A1**	**B1**	**Productivity**		
A1	1.0000				
B1	0.6794	1.0000			
PRODUCTIVITYSTAFF	0.1973	0.3958			

Source: Own Elaboration

The Kolmogorov-Smirnov Test was employed to test RS. At a significance level of 5%, the null hypothesis of CRS was rejected in some years of the sample (2007, 2008, 2009 and 2012). Therefore, it is not possible to conclude that the VRS model better represents the phenomenon under study in every year.

4.2 Second Stage DEA: Determinants of Efficiency Scores

In the second stage, we take the DEA models and run an OLS regression of the efficiency scores against the variables which permit characterizing the staff and the judges.

In this subsection we first estimate an encompassing model, including all possible drivers, on pooled data, and afterwards, we proceed using the method of stepwise regression, eliminating successively all non-significant variables. The definitive model includes all significant variables at 95 percent level.

The model A1 scores depend positively of AGESTAFF, SENJUDGESQ, and AGEJUDGE, while they depend negatively of AGESTAFFSQ, FEMALE, SENJUDGE and AGEJUDGESQ (Table 7). That means that efficiency scores increase with the age of the staff and judge seniority and the age of the judge. Experience matters and the age of the judge seem to influence efficiency too. On the other hand, the square of the age of the staff, the seniority of the judge and the square of the age of the judge decrease linearly the efficiency scores as well as the proportion of female personnel. Nevertheless, the R-square of the regression is low (in the vicinity of 0.1). The linear effect of age and seniority of the staff, the proportion of tenured personnel or the condition of surrogant of the judge are not significantly different from zero.

The model B1 scores depend positively of SENSTAFFSQ, AGESTAFF, and AGEJUDGESQ, while the efficiency depends negatively of SENSTAFF, AGESTAFFSQ, AGEJUDGE and FEMALE (Table 8). That means that efficiency scores increase with the square of the staff seniority, the age of the staff, and the square of the judge age, while decrease linearly with the seniority of the staff, the age of the judge and the proportion of female personnel, as well as with the square of the age of staff. The R-square of the regression is low (in the neighborhood of 0.19). The linear and square effect of the seniority of judges, the proportion of tenured personnel and the condition of surrogant for judges are not significantly different from zero.

Table 6: Explanatory Variables of DEA Efficiency Scores (Model A1)

A1 Scores (dependent)	**Initial Model**			**Final Model**		
Independent Variables	**Coefficient**	**Std. Error**	**P>|t|**	**Coefficient**	**Std. Error**	**P>|t|**
Senstaff	0.0012445	0.001144	0.277			
Senstaffsq	-2.41e-06	2.59e-06	0.352			
Agestaff	0.0031656	0.0017934	0.078	0.0041978	0.0016769	0.013
Agestaffsq	-2.45e-06	1.63e-06	0.134	-3.30e-06	1.53e-06	0.032
Tenured	0.0116855	0.0204876	0.569			
Female	-0.0494806	0.0224095	0.028	-0.0448219	0.0219698	0.042
Senjudge	-0.0003396	0.0000945	0.000	-0.0002694	0.0000839	0.001
Senjudgesq	3.94e-07	1.31e-07	0.003	3.19e-07	1.23e-07	0.010
Agejudge	0.0009411	0.0003332	0.005	0.0008355	0.0003283	0.011
Agejudgesq	-6.66e-07	2.35e-07	0.005	-6.00e-07	2.32e-07	0.010
Surrogant	0.0109431	0.006893	0.113			
Constant	-0.5108267	0.4747443	0.282	-0.6274459	0.470042	0.182
N	560			560		
F-statistic	6.72			9.86		
Rsq	0.1189			0.1111		
Adj Rsq	0.1012			0.0998		

Source: Own Elaboration

Table 7: Explanatory Variables of DEA Efficiency Scores (Model B1)

B1 Scores (dependent)	**Initial Model**			**Final Model**		
Independent Variables	**Coefficient**	**Std. Error**	**P>|t|**	**Coefficient**	**Std. Error**	**P>|t|**
Senstaff	-0.0036649	0.0009535	0.000	-0.0037466	0.008787	0.000
Senstaffsq	8.45e-06	2.16e-06	0.000	8.48e-06	2.02e-06	0.000
Agestaff	0.0045878	0.0014948	0.002	0.0050955	0.0014749	0.001
Agestaffsq	-3.74e-06	1.36e-06	0.006	-4.14e-06	1.35e-06	0.002
Tenured	0.0220751	0.0170763	0.197			
Female	-0.033901	0.0186782	0.070	-0.035276	0.0186316	0.059
Senjudge	0.0000226	0.0000787	0.775			
Senjudgesq	-6.89e-08	1.10e-07	0.530			

B1 Scores (dependent)	Initial Model			Final Model		
Independent Variables	Coefficient	Std. Error	P>|t|	Coefficient	Std. Error	P>|t|
Agejudge	-0.0007099	0.0002777	0.011	-0.0006857	0.0002517	0.007
Agejudgesq	5.40e-07	1.96e-07	0.006	5.03e-07	1.75e-07	0.004
Surrogant	0.0088188	0.0057453	0.125			
Constant	0.1710515	0.3956965	0.666	0.0517801	0.3858615	0.893
N	560			560		
F-statistic	12.67			18.98		
Rsq	0.2027			0.1940		
Adj Rsq	0.1867			0.1838		

Source: Own Elaboration

5. Conclusion

We aimed to respond two questions on the technical efficiency of labor courts in Argentina: What are the average efficiency levels of the Argentine labor courts and how do they behave individually with respect to the average? Which are the determinants of the relative efficiency levels of those courts?

To respond the questions, we estimated by means of DEA models (CRS and VRS versions) the relative efficiency of a sample of 80 Labor Courts for the period 2006-2012. The estimates measure efficiency scores using as outputs the sentences, the average time to finish a case and the proportion of sentences that are appealed, against the inputs personnel with college degree, personnel without college degree and caseload (the raw material of the process). The results yield a mean efficiency of 0.87 and 0.91 in the CRS and VRS models respectively, with standard deviations of 0.07 and 0.06. On average, 5.6 and 12.8 decision units were efficient according to each model. No measure of capital input is used, thus the analysis is for the short run.

In a second stage, we regressed the efficiency scores against their possible drivers. Some explanatory variables are robust to both models but others change significance or sign. The variables to distinguish between tenured and non-tenured staff and tenure and non-tenured judges are not significant in any of the models. The proportion of the variance of the model explained in each case is low (0.1 and 0.19 in the CRS and VRS versions). The analysis of the efficiency scores against its determinants shows that more variables than those present in our database play a role.

References

Banker, R.D., A. Charnes and W. W. Cooper (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 30: 92-178.

Banker, R. and R. Natarajan (2004). Statistic Tests Based on DEA Efficiency Scores, in W. Cooper, L. Seiford & J. Zhu, Handbook on Data Envelopment Analysis, Norwell: Kluwer.

Beenstock, M. and Y. Haitovsky (2004). Does the appointment of judges increase the output of the judiciary? International Review of Law and Economics 24: 351-369.

Bhattacharya, M., and R. Smyth (2001). Aging and Productivity among Judges: Some Empirical Evidence from the High Court of Australia. Australian Economic Papers 40(2): 199–212.

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

Deyneli, F. (2012). Analysis of the 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.

Dimitrova-Grajzl, V., P. Grajzl, J. Sustersic, and K. Zajc (2012). Judicial Incentives and Performance at Lower Courts: Evidence from Slovenian Judge-Level Data. Review of Law and Economics 8(1): 215-252.

Djafari, S. (2011). Court Efficiency Gets Major Boost? Ph.D. Thesis, Utrecht University.

Elbialy, N. and M. A. García Rubio (2011). Assessing Judicial Efficiency of Egyptian First Instance Courts. A DEA Analysis. Joint Discussion Paper Series in Economics by the Universities of Aachen-Giessen-Götingen-Kassel-Marburg-Siegen Nº 19.

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

Falavigna, G., R. Ippoliti, A. Manello and G. B. Ramello (2015). Judicial Productivity, delay and Efficiency: a Directional Distance Function (DDF) Approach. European Journal of Operational Research 240(2), 592-601.

Färe, R., S. Grosskopf and C. A. K. Lovell (1985). The Measurement of Efficiency of Production. Kluwer-Nijhoff Publishers.

Ferrandino, J. (2014) Testing the Packer Theorem: The Efficiency of Florida's Criminal Circuit Courts. American Journal of Criminal Justice 39(2), 375-393.

Ferro G,, C. Romero and E. Romero-Gómez (2018). Efficient Courts? A Frontier Performance Assessment. Benchmarking: An International Journal: 25(9), 3443-3458.

Finocchiaro Castro, M. and C. Guccio (2018). Measuring Potential Efficiency Gains From Mergers of Italian First Instance Courts Through Nonparametric Model. Public Finance Review 46(1): 83-116.

Fusco, E. M. Laurenzi and B. Maggi (2018). A Data Envelopment Analysis of the Italian Judicial Efficiency. Dipartimento di Scienze Statistiche, "Sapienza" Universitá di Roma. DSS-E3 WP 2018/2.

Gorman, M. and J. Ruggiero (2009). Evaluating U.S. judicial district prosecutor performance using DEA: are disadvantaged counties more inefficient? European Journal of Law and Economics 27(3): 275-283.

Guzowska, M. and T. Strack (2010) An examination of the Efficiency of Polish Public Sector Entities Based on Public Prosecutors Offices. Operations Research and Decisions 2: 41-57.

Hagstedt, K. and J. Proos (2008). Has the recent restructuring of the Swedish district courts improved efficiency? A DEA analysis. Uppsala University, Department of Economics, Bachelor's Thesis.

Ippoliti, R. and G. B. Ramello (2016). Efficiency and Productivity of Tax Courts. Institute For Research in Economic and Fiscal Issues (IREF) Working Paper 201604.

Kittelsen, S. and F. Forsund (1992). Efficiency Analysis of Norwegian District Courts. The Journal of Productivity Analysis 3: 277-306.

Lewin, A., R. Morey, and T. Cook (1982). Evaluating the Administrative Efficiency of Courts. Omega 10: 401-411.

Major, W. (2015). Data Envelopment Analysis as an Instrument For Measuring the Efficiency of Courts. Operation Research and Decisions 4: 19-34.

Mattson, P., J. Mansson, C. Andersson and F. Bonader (2018). A Bootstrapped Malmquist Index Applied to Swedish District Courts. European Journal of Law and Economics 46: 109-139.

Nissi, E. and Agnese Rapposelli (2010). A Data Envelopment Analysis of Italian Courts Efficiency. Statistica Applicata-Italian Journal of Applied Statistics 22(2): 199-209.

Odhiambo, O. J. (2014). Technical Efficiency of the Kenyan Judiciary: A Case of the Magistrates' Courts. Masters Degree Thesis Paper. University of Nairobi, School of Economics.

Pedraja-Chaparro, F. and J. Salinas-Jimenez (1996). An Assessment of the Efficiency of Spanish Courts Using DEA. Journal of Applied Economics 28: 1391-1403.

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

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

Rushid, A. R. (2018). Technical Efficiency of Swedish District Courts. A Stochastic Distance Function Analysis. Economics, Master's Degree Project.

Santos, S. and C. A. F. Amado (2014). On the Need For Reform of the Portuguese Judicial System. Does Data Envelopment Analysis Assessment Support It? Omega. The International Journal od Management Science 47: 1-16.

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 20: 127-144.

St. Aubyn, Miguel (2008). Law and Order Efficiency Measurement. A Literature Review. School of Economics and Management, Uniersidade Técnica de Lisboa WP 19/2008/DE/UECE.

Tulkens, H. (1993). On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit. The Journal of Productivity Analysis 4: 183-210.

Voigt, S. (2016). Determinants of Judicial Efficiency: a Survey. European Journal of Law and Economics 42(2): 183-208.

Yeung, L. (2018). Measuring Efficiency of Brazilian Courts: One Decade Later. Encontro de Administraçao de Justiça ENAJUS.

Yeung, L. and P. Azevedo (2011). Measuring Efficiency of Brazilian Courts With Data Envelopment Analysis (DEA). IMA Journal of Management Mathematics, dpr002.