International Review of Law & Economics 79 (2024) 106216

Contents lists available at ScienceDirect

International Review of Law & Economics

journal homepage: www.elsevier.com/locate/irle

Scale and scope economies in first-instance courts: Portuguese specialized vs non-specialized courts

Nuno Q.M. Lopes, Maria C.A. Silva *

Universidade Catolica Portuguesa, Catolica Porto Business School, CEGE, R. Diogo Botelho, Porto, Portugal

ARTICLE INFO

Keywords:
Data envelopment analysis
Courts efficiency
Economies of scale
Economies of scope

ABSTRACT

This paper analyses the technical efficiency of first-instance courts and investigates the existence of scale and scope economies. To assess the technical efficiency of specialized and non-specialized court benches, we use Data Envelopment Analysis (DEA). This study uses data from 2015 to 2021, encompassing every bench within the Portuguese first-instance courts, totalling 3249 observations.

Our findings reveal diseconomies of scale, with more than half of the benches experiencing increasing returns to scale, indicating that their performance would benefit from increased scale. The scale diseconomies varied by bench type: benches primarily handling civil cases and generic benches faced mostly increasing returns to scale. In contrast, those dealing predominantly with criminal cases experienced decreasing returns to scale.

Additionally, we observe diseconomies of scope, indicating that generic and non-specialized benches were less efficient than specialized ones. Overall, this paper provides empirical evidence supporting the notion that the specialization of benches enhances their efficiency.

1. Introduction

In recent years, the performance of judicial systems has become a critical concern for policymakers, legal scholars, and practitioners. As societies grapple with increasing caseloads, budget constraints, and demands for swift dispute resolution, understanding the factors contributing to court efficiency is paramount. The assessment of a court's performance is generally carried out through Key Performance Indicators (KPIs) that follow international guidelines. The European Commission for the Efficiency of Justice (CEPEJ) has put forward several KPIs for assessing courts (e.g. clearance rates, case turnover, backlog, case per judge, etc.) including quality indicators (e.g. distinction between processes finishing by merit decision and complete appreciation of the case, from the remaining cases). CEPEJ has a Dynamic database of European judicial systems (see www.coe.int/en/web/cepej/cepej-stat) that uses two main indicators to measure efficiency of the judiciary: The clearance rate (the number of cases solved divided by the number of incoming cases), and the disposition time (the pending cases divided by the resolved cases multiplied by 365) (see also Magalhães and Garoupa, 2020).

Despite the widespread use of KPIs, establishing a universally accepted tool for measuring a court's efficiency is challenging, and doubts may arise about the definition of efficiency itself. However, an efficient judicial system can be defined as one that delivers prompt, high-quality decisions without accumulating backlogs. For an economist the price at which this speed and quality are obtained matter, and therefore, the above can be regarded as effectiveness measures, since the efficiency should also consider a 'price dimension'. Under the efficiency paradigm, we have two concepts of efficiency: technical efficiency and cost efficiency as originally proposed by Farrell (1957). Both concepts take into account the resources consumed to produce a certain output, but the technical efficiency simply looks at the quantities of resources used to produce certain amounts of outputs, whereas economic efficiency will also look at the total cost at which the outputs are produced (therefore a mix of resources may be efficient from a technical perspective, but inefficient from a cost perspective).

Numerous researchers contributed to analysing court efficiency worldwide. Voigt (2016) offers a survey of studies quantifying judicial efficiency, highlighting that efficiency is only one aspect of judicial performance, which also encompasses accessibility, accountability, and effectiveness.

Court specialization has been a priority for many judicial reforms globally, especially in Europe (Arlota and Garoupa, 2016). Specialization can range from the appointment of special masters within general courts to the establishment of fully independent specialized courts (Casado Pérez, 2019). The rationale for specialization lies in its potential to expedite resolution and improve decision quality, while

* Corresponding author.
E-mail address: csilva@ucp.pt (M.C.A. Silva).

https://doi.org/10.1016/j.irle.2024.106216
Received 31 October 2023; Received in revised form 26 June 2024; Accepted 8 July 2024
Available online 22 July 2024

promoting uniformity and consistency in jurisprudence (Casado Pérez, 2019).

While the concept of specialization has gained traction, it remains a topic of debate given that empirical evidence regarding its benefits remains scarce. This study contributes to the empirical evidence base by confirming that specialized courts tend to be more technically efficient than non-specialized ones, as seen in the case of Portuguese first-instance courts. Portugal is a representative example of a country with a civil law system, which is characterized by codified statutes and comprehensive legal codes. This contrasts with common law systems, where case law and judicial precedents play a more significant role. Analysing Portugal helps to understand how specialized courts operate within the framework of a civil law system, where the judiciary's role is to apply the law rather than to interpret it as extensively as in common law systems. Although Portugal's size and number of courts are modest compared to other European countries, its demographic distribution is representative of Europe, with larger cities centralizing most of the population and caseloads. This makes Portugal an ideal starting point for a deeper understanding of European specialized jurisdictions and scope economies. The reform implemented in Portugal in 2014 (Law 62/2013 regulated by Decree Law 49/2014), which introduced specialized jurisdictions at the national level and led to the closure of most generic courts, centralizing their workload into specialized courts, was based on an implicit assumption, untested in the Portuguese context, that there were diseconomies of scope. This paper provides confirming evidence that indeed this was the case. This is especially pertinent as many civil law countries face similar reforms. It is worth noting that the 2014 reform was later amended by Law 40-A/2016 and Decree Law 86/2016 due to concerns that the initial reform had negatively impacted access to justice for many populations. The 2016 reform aimed to address these issues and improve access to justice for the population. Subsequent reforms, such as those outlined in Decree Law 38/2019, further reinforced the specialization of courts. Nonetheless, the primary reform remains the one implemented in 2014.

We measure court performance using a technical efficiency measure computed by means of Data Envelopment Analysis (DEA). DEA has been the prevalent technique used to measure the efficiency of courts, since the precursor study by Lewin et al. (1982) that analysed the efficiency of 100 criminal courts in the US state of North Carolina. To the best of our knowledge, the use of DEA in the context of assessing scale and scope economies in the judicial context is also a novelty of this paper; only Mattsson and Tidanå (2019) study focused on economies of scope applied to courts. The integration of scale and scope economies through Data Envelopment Analysis (DEA) is particularly significant from the authors' perspective. This approach facilitates the comparability of future studies with the present research and enhances the understanding of the true value of specialization within the European context.

As we apply the framework to Portuguese first-instance benches of courts of law, we also add to the literature concerning the previous analyses of court performance in this country. In concrete terms, Santos and Amado (2014) and Silva (2018) also applied DEA to measure the efficiency of courts, under different perspectives. While Santos and Amado (2014) focused on first-instance courts or judicial counties, Silva (2018) focused on benches of 'generic competence' that operate within courts. In this paper, we also focus on benches that are the most disaggregate unit of analysis of courts. Indeed, first-instance judicial courts are organized within judicial districts, which in themselves are organized in judicial counties (comarcas), usually referred to as 'courts'. Within courts (or judicial counties), judges are typically organized in benches (constituted in many cases by a single judge and some administrative staff). Benches can be of various types: generic benches or specialized (e.g. Criminal benches, civil benches, commerce benches, labour benches, etc.). In contrast to Silva (2018), we consider all types of benches in the same analysis, which raises issues of heterogeneity in the unit of analysis. However, it allows the computation of scope economies, which is the main objective of the present paper.

The present paper is organized as follows: Section 2 delves into the available literature concerning the analysis of court efficiency, as well as economies of scope. Section 3 presents the DEA model used to compute efficiency and the methods employed to assess scale and scope economies. Section 4 outlines the structure of the Portuguese judicial system and the data under analysis. In Section 5, we analyse the data and present our main findings, while Section 6 concludes.

2. Previous literature

In this section, we offer a concise overview of the literature, focusing specifically on studies addressing the efficiency of courts and those examining scale and scope economies as determinants of efficiency.

2.1. The notion of court efficiency and its measurement

As mentioned before, the performance of courts is measured many times by single KPIs, one of which is judges' productivity. Many studies have addressed judges' productivity (e.g. Choi et al., 2010; Ramseyer, 2012; Schneider, 2005; Pereira and Wemans, 2017), but the emphasis has been on the main determinants of this productivity (like the educational background of judges, incentives (promotions and monetary), and also reversals from courts of appeal Schneider, 2005).

Judges' productivity is closely intertwined with the concept of efficiency, which can essentially be viewed as a measure of relative productivity (e.g., Charnes et al., 1978). In simple terms, if Court A resolves 10 cases per judge while Court B resolves 20 cases per judge, Court B's efficiency would be 100%, and Court A's efficiency would be 50%. Therefore, comparing judges' productivity enables us to infer the efficiency of courts. However, it is important to note that courts handle various types of cases and involve more resources than mere judges, so judges' productivity may serve as an incomplete proxy for efficiency.

This reality has prompted the development and adoption of performance measures capable of simultaneously considering multiple dimensions of court performance. Frontier methods, such as DEA or Stochastic Frontier Analysis (SFA), facilitate the incorporation of multiple inputs and outputs in efficiency assessments. It is worth noting that SFA, as it is a parametric method, relies on models specifying the functional form of the frontier, where the output variable is singular. In the context of evaluating court efficiency, where accounting for case mix is crucial, reducing the primary output – "cases resolved" – to a single measure can be overly restrictive. It is essential to consider the various types of cases resolved to account for the diversity in caseload. This is particularly pertinent in our study as we evaluate different types of court benches. A single output measure aggregating cases resolved could potentially distort the analysis. Hence, we opted for the DEA method, which can easily accommodate multiple outputs and, therefore, case mix considerations.

Since the precursor study of Lewin et al. (1982), addressing court efficiency using multiple inputs and outputs through DEA, the literature on the topic has grown significantly. Voigt (2016) provides a survey of 34 studies that relied on parametric and non-parametric techniques to measure the efficiency of courts. Most of the studies in Voigt (2016) pertain to analysis of efficiency within country, and only six consider more than two countries. Since the 2016 review, there have been some further studies of which we would like to highlight the following: Peyrache and Zago (2016), Falavigna et al. (2018), Giacalone et al. (2020), Fusco et al. (2021), Falavigna and Ippoliti (2021) and Falavigna and Ippoliti (2022) on the Italian courts, Mattsson et al. (2018), Mattsson and Tidanå (2019), Agrell et al. (2020), and Månsson et al. (2022) on the Swedish courts, Bełdowski et al. (2020) on the Polish courts, Bajrić and Kadrić (2020) on the Bosnia–Herzgovina courts, Silva (2018) on the Portuguese courts, and Gupta and Bolia (2020) on the Indian

courts. Therefore, within Europe, the main countries studied include Italy and Sweden.

While traditional DEA models (see Charnes et al., 1978; Banker et al., 1984) are the most widely used models to measure court efficiency, other DEA models have also been applied to the context of courts. For example, Falavigna et al. (2015) and Falavigna and Ippoliti (2021) used Directional Distance Functions (DDF), while Tulkens (1993); De Sousa et al. (2005) used Free Disposable Hull (FDH). Besides these techniques, some authors performed additional analyses in relation to the computation of DEA efficiency scores through DEA. For example, regression analysis has been undertaken by Beenstock and Haitovsky (2004); Dimitrova-Grajzl et al. (2012); Bhattacharya and Smyth (2001), and Malmquist indexes have been employed by Mattsson et al. (2018), and both the techniques have been used by Falavigna et al. (2018); Giacalone et al. (2020).

Despite the many types of models used to assess efficiency, most studies assess technical efficiency (see e.g. Voigt, 2016). Cost efficiency in the court context has only been applied by Månsson et al. (2022) to Swedish courts. Results show that most cost inefficiency is indeed allocative and not technical, meaning that it stems from the use of the incorrect input mix or paying too much for inputs. Given regional differences in input prices (i.e., wages and rents), the study reveals that cost savings could be obtained by changing the mix of inputs and even from strategic measures like moving district courts out of city centres — which could improve cost efficiency, but deteriorate accessibility to justice.

Most existing studies also adopt an output-oriented perspective in measuring court efficiency — that is, the efficiency is given by the ratio between observed outputs and maximum outputs that could have been produced given the resources of the court. In many instances, an input perspective is not possible, as most of the time, inputs consist of the demand of the judicial services from the general public, and the number of personnel, variables over which the courts have no discretion.

2.2. Scale and scope economies and court specialization

Several studies have attempted to tackle the significance of court size in influencing its performance. In the realm of production economics, economies of scale refer to the phenomenon whereby cost advantages are achieved through the expansion of the size of operations. This concept is closely tied to strategic decisions observed in numerous public services, such as the consolidation and mergers of these services to yield efficiency improvements.

Some authors, analysed scale issues of courts and found scale inefficiencies mainly associated with small courts (e.g. Kittelsen, 1992, Santos and Amado, 2014 or De Sousa et al., 2005). Mattsson and Tidanå (2019), Agrell et al. (2020), Xiaoqing and Kerstens (2023), and Chen et al. (2024) are among the few studies that analysed mergers in courts, and all of them used the Swedish court case as an empirical application. Mattsson and Tidanå (2019) find that some mergers have little potential for efficiency gains, while others can generate significant merger gains. Agrell et al. (2020) report that the merged courts are more efficient than the non-merged ones. Xiaoqing and Kerstens (2023) found that horizontal mergers improve plant capacity utilization. Chen et al. (2024) found that horizontal mergers do not contribute to post-merger productivity gains.

Peyrache and Zago (2016) analysed these issues in the Italian case, and conclude that 35% of total inefficiency of the judiciary is due to size and together with reallocation of inputs, splitting large courts into smaller ones leads to higher efficiency scores. Tulkens (1993) says that most inefficient units are middle-sized courts as measured by clerical staff.

There is evidence suggesting that in Sweden, the consolidation of courts through mergers has enhanced the efficiency of most district courts (Hagstedt and Proos, 2008; Agrell et al., 2020). Conversely, in Italy, courts may be too large, and efficiency gains could potentially be achieved by dividing large courts into smaller ones. Both findings are plausible because scale economies are contingent upon the actual scale size and whether courts are experiencing increasing returns to scale (where enlarging the size is advantageous) or decreasing returns to scale (where increasing the size diminishes efficiency).

Many papers discussing the advantages of a specialized judiciary often lack empirical evidence on benefits (see e.g. Revesz, 1989 and Casado Pérez, 2019). Regarding the studies that do provide empirical evidence, it is often the case that they are not based on an analysis of court efficiency, as carried out in this paper, but rather on case-specific analyses.

For instance, Garoupa et al. (2010) examined whether specialized family courts in Spain were quicker in decision-making compared to non-specialized courts. They employed an ordered probit model to estimate the probability of a case being concluded within certain timeframes. While specialized courts appeared to handle litigation more swiftly than regular courts on average, the overall econometric evidence was not strong.

In another example, Arlota and Garoupa (2016) directly addressed the issue of specialization by analysing 630 cases and the differences observed in the outcomes of abstract review. Their analysis focused on whether specialization led to faster decisions (analysing duration), better quality of lawmaking (examining length and citations of superior courts), and more extensive legal discussions (considering dissent rates). Using regression analysis, they found “mixed results in terms of confirming the theoretical literature on the benefits of specialized constitutional review”, as the specialized courts did not consistently outperform others, although there were statistically significant differences.

Detotto et al. (2019) analyse the hypothesis that the establishment of specialized courts in Spain is associated with the increased use of the formal insolvency framework. The authors used an endogenous treatment regression model to predict the number of formal bankruptcies per 1000 firms in Spanish provinces. Conclusions point to the fact that specialized courts governed by competent judges, are necessary, but not sufficient *per se*, to make the reform effective.

Amaral-Garcia (2019) analysed 366 decisions in relation to medical malpractice cases and compared the decisions made by administrative and civil courts, concluding that there is no evidence that the state is treated differently regarding non-economic compensation amounts for cases (see also a previous study of Amaral-Garcia and Garoupa (2015) on the same type of cases).

More recently, Castelliano et al. (2021) analysed 149,455 social security related cases in Brazilian first-instance courts, where one half of the cases were adjudicated by judges in a titled judgeship position and the remainder by judges in a substitute judgeship position. The authors were interested in estimating the effect of the adjudication forum and court office specialization on case duration, and survival analysis was used in order to take into account pending cases. Castelliano et al. (2021) found evidence of systematic differences, but magnitude-wise overall small, between titled and substitute judgeships with regard to case duration, and also found that court office specialization is associated with reduced case duration.

In summary, many studies offering empirical evidence regarding the benefits of specialization suggest that it is likely to enhance judicial efficiency, leading to faster, more accurate, and less costly proceedings. However, the evidence is not entirely convincing, particularly when analysing similar cases handled by specialized and non-specialized courts.

Regarding studies that look at courts as the unit of analysis and at their efficiency, evidence is scarce. Elbialy and García-Rubio (2011) suggest that criminal district courts are more efficient (mean efficiency 68%) than their corresponding civil district courts (mean efficiency 64%), and that higher shares of criminal caseloads tend to reduce court inefficiency, while more civil and family caseloads do not influence court inefficiency. On the contrary, Pereira and Wemans (2017)

highlight the positive effect on judges' productivity when cases are judged in courts where the vast majority of cases are in the civil area. Mattsson and Tidanå (2019), although not specifically discriminating the specialized vs non-specialized courts, argue that quality may be enhanced with a higher degree of specialization, and concluded that an efficiency improvement between 10.1%–19% may be achieved for 17.2% of the Swedish courts if those courts were to be merged based on economies of scope.

The evidence on economies of scope is therefore still scarce in the judicial literature, and this is a gap that the current paper fills.

3. Methodology

3.1. Data envelopment analysis

First introduced by Charnes et al. (1978), DEA allows the construction of a deterministic, non-parametric production frontier. This frontier is then used to compute the efficiency of the Decision-Making Units (DMUs), based on the radial distance of each of those DMUs to the frontier. Cost or revenue efficiency can also be computed following Farrell (1957). As mentioned before, in the court setting, economic efficiency is rarely found, and furthermore, we also do not pursue this avenue in this paper.

Charnes et al. (1978) adopt a Constant Returns to Scale (CRS) model, which assumes that an increase in inputs will generate a proportional increase in the outputs. Banker et al. (1984) proposed a model that adopts Variable Returns to Scale (VRS), which allows for constant increasing and decreasing returns. Model (1) shows the VRS output-oriented model, where m and s are the number of inputs and outputs, respectively. The subscript n relates to the DMUs and y_{ro} and x_{io} are respectively the amount of output r generated by unit o and the amount of input i used by unit o under assessment. λ_j are intensity variables showing the weight attached to each DMU j to form the efficient benchmark for the DMU o under analysis.

$$\begin{aligned} \max \quad & \beta \\ s.t. \quad & \sum_{j=1}^{n} x_{ij}\lambda_j \le x_{io} \qquad (i,\dots,m) \\ & \sum_{j=1}^{n} y_{rj}\lambda_j \ge \beta y_{ro} \qquad (r,\dots,s) \\ & \sum_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \ge 0 \qquad (j,\dots,n) \end{aligned} \tag{1}$$

When the convexity constraint ($\sum_{j=1}^{n} \lambda_j = 1$) is dropped, (1) becomes a CRS model. If the convexity constraint is replaced by ($\sum_{j=1}^{n} \lambda_j \le 1$) then model (1) becomes a Non-Increasing Returns to Scale (NIRS).

The optimal solution from model (1) returns a value β^*, which is the expansion needed in outputs of DMU o that led this unit to the frontier without decreasing inputs. The efficiency score of DMU o is given by $\theta^* = 1/\beta^*$. When efficiency is equal to 1 the DMU is located at the technical frontier, while if θ^* is smaller than 1, the DMU is located on the interior of the frontier and therefore inefficient.

Model (1) can be solved in relation to pooled or meta-technology when $j = 1,\dots,n$ represent all units observed over time. If units are somehow grouped by specific criteria (e.g. in our empirical application, we have different types of benches that are also compared amongst themselves), we may consider a sub-sample of j, i.e. only those units belonging to a specific group. In the first case, we obtain a meta-efficiency score (θ^M); whereas in the second case, we obtain a group specific efficiency score (θ^G). Clearly, the ratio between the two scores ($\frac{\theta^M}{\theta^G}$) yields a gap between the two frontiers that we name frontier gap.

3.2. Economies of scale

There are various methods in the literature to assess economies of scale. In this paper we follow the method of Färe et al. (1985), which requires three efficiency estimates in relation to three technological Returns to Scale specifications: CRS, VRS, and NIRS. From the efficiency measures obtained from each of these models, conclusions can be reached concerning returns to scale: (1) If the CRS, VRS and NIRS models yield exactly the same efficiency measure, then the unit lies, or is projected, on a frontier region exhibiting local CRS; (2) If the CRS and NIRS efficiency measures are both equal and lower than the VRS efficiency measure, then the unit lies, or is projected, on an IRS region of the frontier; (3) If VRS and NIRS efficiency measures are both equal and higher than the CRS efficiency measure, then the unit lies, or is projected, on a DRS region of the frontier. The Färe et al. (1985) method has the advantage of being unaffected by the existence of multiple optimal solutions. Its main disadvantage seems to be the need to solve three DEA problems (Seiford and Zhu, 1999).

3.3. Economies of scope

Economies of scope were first described by Panzar and Willig (1981) as an intuitive production property based on cost savings resulting from the scope (instead of the scale) of a company. One can say there are economies of scope when producing two or more products in one firm is less costly than producing them separately. Panzar and Willig (1981) developed the theoretical concept behind economies of scope, and Morita (2003) was among the first author to operationalize the concept of economies of scope in the non-parametric setting of DEA where no cost information was required. Morita (2003) expressed economies of scope as $C(u_1,0)+C(0,u_2) > C(u_1,u_2)$ where $C(u_1,u_2)$ represents the cost of producing u_1 and u_2 units of product A and B, respectively. Since cost information may not exist, Morita (2003) proposes the use of technical efficiency improvements instead of cost savings to ascertain the existence of economies of scope. The main idea is to compare the DEA frontiers of joint production and separate production. If joint production increases the efficiency over an initial separate production, there are economies of scope. If the value remains the same, there are no economies of scope. Additionally, if the efficiency decreases, we face diseconomies of scope.

In the context of Portuguese courts, joint production is exemplified by the generic benches, which handle all types of cases. Conversely, separate production refers to courts that exclusively deal with two or fewer of the four types of cases considered (the structure of the Portuguese judicial system is outlined in the next section).

In our empirical application neither Morita (2003) nor Mattsson and Tidanå (2019) approaches to economies of scope could be employed. The reason is related to the fact that many inputs and outputs will have zero values, when we consider the full set of variables for all types of courts (for example criminal courts do not handle labour cases whose value is zero). This would result in a large number of undefined efficiencies that would lead to no conclusions. To tackle this situation, economies of scope are analysed, resorting to the concept of frontier gap. This means that, when the frontier gap is close to 1, the within-group frontier is closer to the meta-frontier. That is, gaps closer to 1 reveal the best performance of the group when compared to the other groups (specializations or generic). Depending on whether the best group performance occurs for specialized or generic benches, we may infer about economies of scope (when the frontier gap of the generic benches is the highest) or diseconomies of scope (when the frontier gap for the generic benches is the lowest).

4. The Portuguese system and data

4.1. Structure of the Portuguese judicial system

The Portuguese Judiciary system is administrated by the Ministry of Justice (government department), but unlike other countries, the head of that department does not exercise any authority over the Public

Ministry, nor do they head the public prosecutions.[1] Courts themselves are divided into different larger categories (orders), like the Constitutional, Judicial, Administrative, and Auditing and minor jurisdictions, like Peace Courts, Courts of Arbitration and Ecclesiastical Court. This paper only handles the Judicial courts, which are hierarchically divided into the Supreme Court of Justice (as its decisions are final in terms of Law, it is only possible to appeal to the Constitutional Court), Courts of Appeal (which handle the appeals from Courts of First Instance), and Courts of First Instance. The latter will be the sole focus of this analysis. There are 23 general jurisdiction constituencies *(comarcas)* that deal with generic or specific competence cases. Those constituencies are divided into municipalities, which subdivide into benches *(Juízos)*. The latter comprise the division between specialized competence, generic competence, and proximity. Specialized competence benches deal with cases according to their proximity, namely central civil and criminal, local civil and criminal, local minor crimes. The remaining benches are criminal prosecution, labour, family, commerce, and enforcement (civil law and administrative law enforcement of court decisions and not criminal law enforcement).

The organization of the Portuguese judiciary has been stable since 2014, irrespective of some of the changes introduced in 2016 to correct the errors made in the 2014 reform. It is worth mentioning that the objective of this paper is not to analyse the effects of the reforms since we do not have or use information prior to the main reform of 2014. In any case, by examining the existence of economies of scope, we are indirectly assessing if the reforms undertaken in the country were indeed in the right direction.

Cases are allocated to the right competence bench, according to their nature, by proximity to the plaintiff's address criteria. To avoid any conflict of interests, they are then randomly assigned to each judge within the bench.[2] Judges, however, are not randomly assigned to courts, since judges at the beginning of their career are generally assigned to generic benches and after a few years of experience they can ask to be assigned to specialized benches (see Law 21/85). This results in a possible selection effect biasing our results. That is, if one finds specialized benches to be more efficient, this may be a result of judges' experience rather than of specialization. Previous literature, however, does not seem to support the hypothesis of a selection effect, at least when we associate it with the experience of judges, as Pereira and Wemans (2017); Backes-Gellner et al. (2011) conclude that the experience of the workforce tends to negatively impact the number of cases resolved, although the quality of the decisions is allegedly improved.

4.2. Data

As previously mentioned, the data used for this study consists of all Portuguese first-instance benches from 2015–2021. We opted to leave military benches and criminal instruction out, as the former are only used on special occasions, and the latter are responsible for conducting the preliminary investigation to decide whether a case should go to trial or not (and do not perform trials *per se*).

Benches analysed are included in one of 11 types: generic, central civil, central civil and criminal, central criminal, civil, commerce, criminal, enforcement, family, minor crimes, and labour. The 11 types of benches were aggregated, for the purposes of this study, into a smaller number of groups, based on the type of case they handled: Agg_Civil (which contemplates central civil and civil), Agg_Criminal (which contemplate all criminal types of benches), generic, labour, enforcement, commerce, and family. This results in 7 different groups of benches, which deal with different mixes of the 4 types of cases that our dataset contains (civil, criminal, labour and family). Although enforcement and commerce deal with civil cases, just like the benches in Agg_Civil, it is clear that the procedures are different in each situation, therefore, this separation is maintained

After the data were cleaned of some errors such as absolute zeros in judges, staff and cases, 3249 observations remained (*R* software was used for both this purpose and data manipulation). The data were provided by the Portuguese Directorate-General for Justice Policy's statistics.[3] For the study at hand no benches were considered as outliers, therefore, there are benches with less than 100 cases incoming or solved, whereas bigger benches have over 200,000 cases.

The variables considered for measuring the efficiency of the Portuguese benches were separated into inputs or resources and outputs, as usual in efficiency assessments. Inputs considered are caseload defined as incoming plus pending cases of the four different types (civil, criminal, labour, and family) and personnel (judges and other staff). As outputs, cases resolved of civil, criminal, labour and family cases were considered.

The rationale for this choice of inputs and outputs encounters echo in the literature. The output considered in all court efficiency studies is resolved cases. Differences between studies occur in the way cases resolved are aggregated or disaggregated. While some authors aggregated the cases resolved as a single output (Hagstedt and Proos, 2008; Elbialy and García-Rubio, 2011), others specified the cases resolved taking into consideration their category (e.g., criminal, civil, labour) (De Sousa et al., 2005; Kittelsen, 1992; Silva, 2018; Gupta and Bolia, 2020; Agrell et al., 2020), and some went further by subdividing the outputs considering some type of weight that accounted for the complexity/heterogeneity of each case category (Santos and Amado, 2014).

Regarding inputs, there is more divergence in the literature, mainly as far as caseload is considered. Indeed, there is no agreement on whether it should be considered as aggregate caseload or as both new cases and pending cases, or merely one of the latter. Several authors, Deyneli (2011); Kittelsen (1992); Tulkens (1993); Finocchiaro Castro and Guccio (2014); Agrell et al. (2020), opted to use solely courts staff as input (judges or/and others). Schneider (2005) pointed to the importance of measuring workload in the analysis, since it is the measure of the demand for a court's services, and is directly linked to the output cases resolved (without caseloads, no cases could be resolved!). Therefore, several studies included pending and/or incoming cases as a measure of workload (Schneider, 2005; De Sousa et al., 2005; Peyrache and Zago, 2016; Silva, 2018; Nissi et al., 2019; Giacalone et al., 2020).

A descriptive summary of the variables used may be found in Table 1.[4]

On average benches are relatively small, with 2.6 judges and 8.71 other staff. The largest caseload is found in civil cases.

In Tables 2 and 3, we consider data disaggregated per year, and show the totals observed in each year.

On average, the total caseload over the period of analysis is 1,424,141 cases, which are mainly constituted of civil cases. There is a decreasing trend of caseload over time with the maximum value found in 2015 (1,915,924) and the minimum value found in 2021 (1,006,046). This decrease also occurs for the total number of judges and other staff, where the averages are 1205 and 4042, respectively. Regarding cases finished (outputs) in Table 3, the numbers differ. The average number of finished cases is 640,375, but 2015 is still the year with the most finished cases (745,777). The year with the least number is 2020 (466,901), probably as a result of the pandemic.

[1] Public prosecution is assured by a government-independent body of magistrates headed by the Attorney General's Office.

[2] An exception is made for the so called "mega-cases" that are still randomly assigned, but within a group of highly experienced judges. If those judges believe they are not able to accept new cases, they may request to be removed from the random assignment of "normal" cases.

[3] https://estatisticas.justica.gov.pt/sites/siej/pt-pt

[4] As this analysis is being carried out at the most disaggregated level, the zeros found in the medians show that the majority of benches do not deal with some types of cases.

Table 1
Descriptive summary of variables (Inputs/Outputs) from year 2015 until 2021.

		Years 2015–2021					
		Average per bench	Median	Standard deviation	Min	Max	Average per year
Inputs	Judges	2.6	2	2.65	1	33	1205
	Other Staff	8.71	6	7.53	1	80	4042
	Civil Case load	2460.65	509	9905.94	5	184,937	1,142,092
	Criminal Case load	256.40	50	594.27	0	8473	119,006
	Labour Case load	166.58	0	631.21	0	6929	77,315
	family Case load	184.70	0	644.94	0	7911	85,728
Outputs	Civil Finished	964.97	278	2819.96	0	46,383	447,885
	Criminal Finished	175.90	38	383.15	0	4461	81,644
	Labour Finished	108.33	0	425.21	0	5016	50,278
	family Finished	130.49	0	442.69	0	4462	60,566

Table 2
Year-to-year descriptive summary of inputs.

Year	Caseload civil	Caseload criminal	Caseload labour	Caseload family	Caseload total	Judges	Other staff
2015	1,575,684	152,083	77,623	110,534	1,915,924	1271	4164
2016	1,358,296	144,558	81,792	98,025	1,682,671	1230	4066
2017	1,201,026	123,990	85,436	87,576	1,498,028	1226	4111
2018	1,065,281	119,455	77,528	79,038	1,341,302	1188	4117
2019	1,113,232	116,931	77,640	81,986	1,389,795	1170	4042
2020	892,895	102,852	67,791	71,691	1,135,229	1176	3941
2021	788,230	73,175	73,395	71,246	1,006,046	1177	3854
Average	1,142,092	119,006	77,315	85,728	1,424,141	1205	4042

Table 3
Year-to-year descriptive summary of outputs.

Year	Finished civil	Finished criminal	Finished labour	Finished family	Finished total
2015	517,406	98,487	52,304	77,580	745,777
2016	517,036	96,143	53,112	72,149	738,440
2017	472,411	89,538	57,137	65,993	685,079
2018	435,557	79,140	54,981	56,422	626,100
2019	546,530	77,621	52,076	56,973	733,200
2020	327,878	58,523	34,243	46,257	466,901
2021	318,379	72,056	48,099	48,593	487,127
Average	447,885	81,644	50,279	60,567	640,375

Table 4
Technical efficiency over years.

Years	Count of benches	Average of VRS meta-efficiency
2015	453	72.31%
2016	452	72.62%
2017	459	71.99%
2018	459	68.64%
2019	473	71.02%
2020	476	60.68%
2021	477	74.19%

Table 5
RTS over years.

Years	Count of CRS	Count of IRS	Count of DRS
2015	33	220	200
2016	40	238	174
2017	33	236	190
2018	36	267	156
2019	37	278	158
2020	32	287	157
2021	56	273	148
Total	267	1799	1183

5. Technical efficiency results

Technical efficiency values for each bench were obtained from model (1), which was solved under various types of returns to scale and for the meta and group specific technologies. First, we report the results for the returns to scale analysis, and then we report the results for the scope economies analysis.

5.1. Returns to scale in Portuguese courts

The investigation of returns to scale used all data pooled together and model (1) was solved under constant, variable and non-increasing returns to scale. We call the technical efficiency scores computed in relation to the pooled frontier the 'Meta-efficiency' scores.

The use of pooled data overlooks potential changes in the efficient frontier over time, a consideration we deemed reasonable given our context. For instance, the VRS efficiency of courts showed minimal fluctuations over the analysis period, as indicated in Table 4. Although 2020 exhibited the lowest technical efficiency value, this can be attributed to the peak of the COVID-19 pandemic, which likely impacted efficiency without affecting technology. Conversely, in 2021, it is plausible that professionals made additional efforts to address pending cases from the previous year. Excluding 2020, Table 4 demonstrates that technical efficiency fluctuated within a narrow range of approximately 6%, from 68.64% (in 2018) to 74.19% (in 2021).

Regarding returns to scale, Table 5 shows the number of benches that were identified (through the procedure of Färe et al., 1985) in any types of returns to scale (increasing (IRS), decreasing (DRS) or constant (CRS)).

About 55% of the benches analysed experience increasing returns to scale, while about 36% experience decreasing returns to scale during the whole period of analysis. A minor percentage of around 9% experience constant returns to scale. It is worth noting that in Table 5 there are not many variations over time. Most variations in RTS are possibly by type of bench, and that analysis is shown in Table 6.

It should be highlighted that as evidenced in Table 6, the benches that deal predominantly with civil cases (Agg_Civil) mainly face IRS (around 75%), while the ones dealing primarily with criminal cases face DRS (58%). Generic benches also face mostly IRS (75%). This implies that in most cases the size of the benches is not optimal with some benches being too small (the ones experiencing increasing returns), whereas others are too large (the ones experiencing DRS).

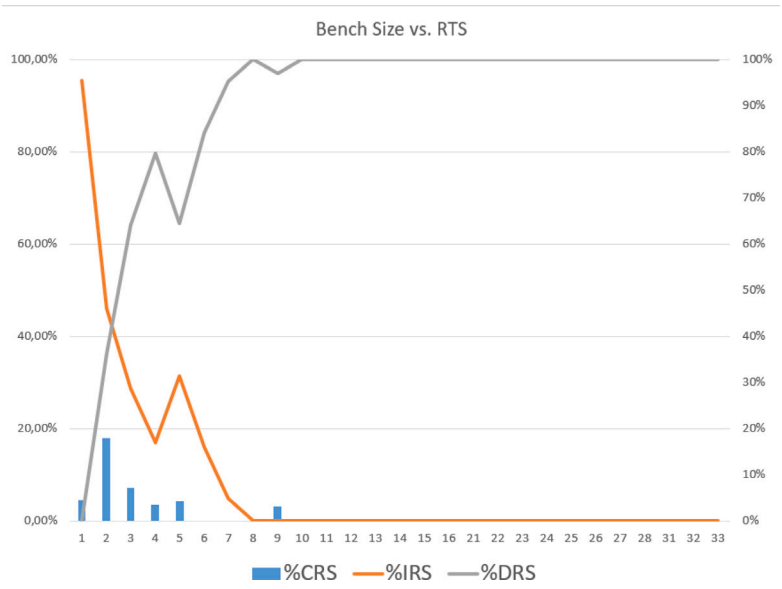

Fig. 1. Bench size vs. RTS.

Table 6
RTS by specialization.

Specialization	Count of CRS	Count of IRS	Count of DRS
Agg_Civil	38	551	145
Agg_Criminal	57	284	472
Generic	29	543	156
Commerce	7	51	91
Enforcement	52	52	67
family	46	157	141
Labour	38	161	111

Fig. 1 shows the percentage of benches identified under the three types of RTS per size of the bench (proxied by the number of judges) for the complete sample of benches observed in all years. From Fig. 1, we can conclude that all benches with more than nine judges experience DRS. The highest percentage of courts exhibiting CRS occurs for benches with two judges. This suggests that the ideal size of benches is around two judges, and with above nine judges marginal gains in efficiency are decreasing. The same picture can be produced per type of bench, as shown in Fig. 2.

Fig. 2 demonstrates that optimal scale differs depending on the type of bench. For example, generic benches seem to operate optimally with two judges, since with three judges, all generic benches experience decreasing returns to scale. Benches only dealing with civil cases should not be larger than eight judges, since after this value all benches experience DRS. Furthermore, benches only dealing with criminal cases should be no larger than six judges, since after that no courts experience IRS.

Given that most courts experience decreasing or increasing returns to scale, a VRS efficiency specification seems the most appropriate. Therefore, we report some details of this specification for the meta-efficiency analysis. The national map with the VRS meta-efficiency scores aggregated by district is shown in Fig. 3.

That is, the worst-performing districts in the country are the Algarve in the south, Leiria in the centre and Viana do Castelo in the north. Thus, there is not a geographical pattern in the efficiency of the Portuguese benches as the best and worst performances are spread all over the country. This is in opposition to what has been found in Italy where courts in the north seem to experience higher efficiency levels than courts in the south of the country (see e.g. Falavigna et al., 2015, Peyrache and Zago, 2016, Nissi et al., 2019).

In order to understand the general profile of the best-performing benches, we divided them into three categories of efficiency, namely P1 if they had efficiency above 66%, P2 if the efficiency was between 33% and 66%, and P3 if their efficiency was lower than 33%. The average of inputs and outputs for these three groups of benches can be seen in Fig. 4 (data normalized by the mean).

From Fig. 4, it is possible to see that the most inefficient benches (P3) are the ones that have the highest number of civil caseloads as well as the ones that resolve more civil cases. P3 courts also have a higher number of judges. It is also possible to see that courts in P2, although not the most efficient, tend to be very balanced in all aspects. The main difference between the three groups concerns the mix of cases, as P1 benches do receive/finish, on average, more cases than the remaining courts in family, labour and criminal classifications, while receiving the least number of civil cases. This shows the importance of the case mix, and the fact that part of the inefficiencies identified in the meta-analysis may be related to the type of bench and the mix of cases handled rather than the technical inefficiency.

5.2. Economies of scope in Portuguese courts

In the analysis of scope economies, we employed the VRS model to calculate efficiency scores within groups and against the meta-frontier. Table 7 presents the results obtained for each type of bench. Furthermore, Table 7 also provides information on the number of judges and other staff (OS) for each kind of bench. As we observed previously, bench size varies significantly across types, with generic benches (the non-specialized ones) markedly smaller than others.

Notably, enforcement benches exhibit the lowest meta-efficiency, whereas labour and family benches demonstrate higher meta-efficiency values. One primary explanation for this discrepancy is the urgency associated with concluding cases in labour and family matters. Labour cases are directly linked to the economic security of the population and the overall economic well-being of the country. Similarly, family cases, such as divorces involving children, necessitate swift resolution due to their sensitive nature. In contrast, enforcement cases often involve companies, and these proceedings are frequently prolonged by lawyers and stakeholders seeking to avoid potential fines or maximize gains, among other interests.

Table 7 also shows that generic benches (that were above the average in meta-efficiency, and had the second-highest within-group efficiency) have the worst gap to the meta-frontier, which is 83.23%.

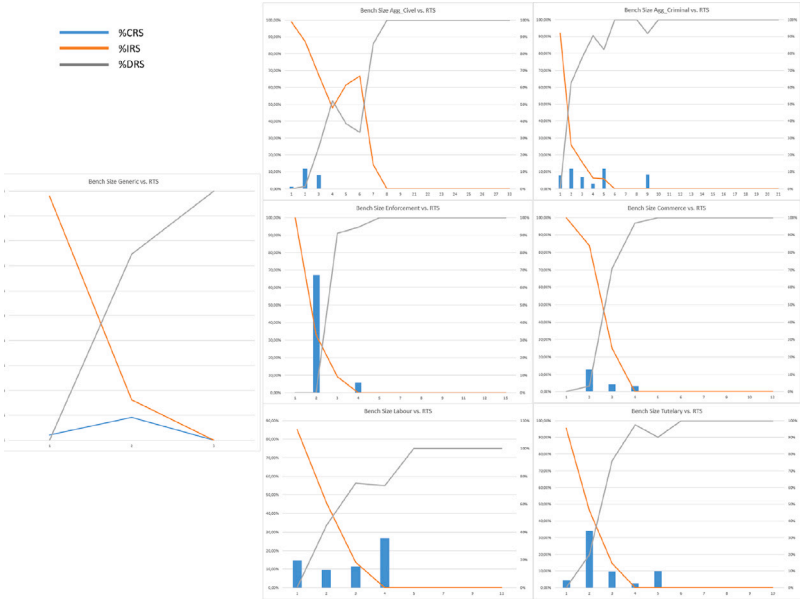

Fig. 2. Bench size for each type vs. RTS.

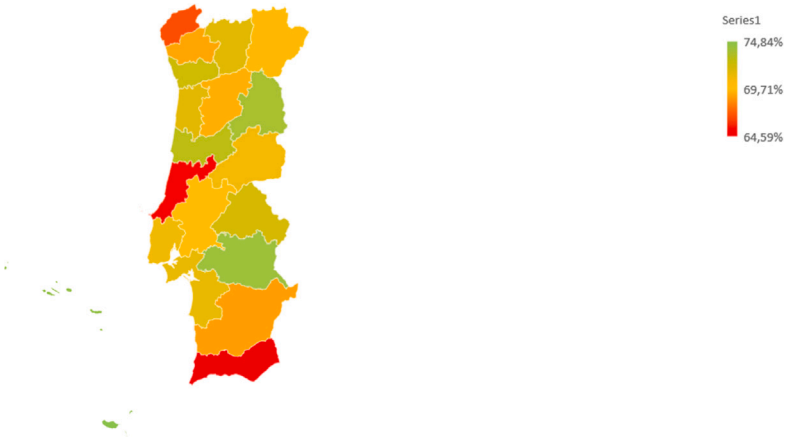

Fig. 3. Efficiency map by district.

Table 7
Meta-efficiency by specializations.

Specialization	N.	N. Eff	Avg judges	Avg OS	Avg OS/judge	Avg meta-eff	Avg with-grp eff	Avg ratio (gap)
Enforcement	171	12	3,06	12,04	3,93	58.44%	61.30%	95.72%
Agg_Civil	734	11	3.12	7.56	2.43	60.99%	66.73%	92.08%
Commerce	149	5	3.53	15.25	4.32	69.65%	78.07%	89.39%
Generic	728	26	1.30	5.14	3.97	70.32%	84.54%	83.23%
Agg_Criminal	813	63	3.29	10.68	3.25	72.66%	74.89%	97.24%
family	344	30	2.45	11.28	4.60	79.26%	81.18%	97.70%
Labour	310	50	2.07	6.82	3.29	81.70%	87.15%	93.80%

This means that the frontier of generic benches is the farthest from the meta-frontier composed by all other benches. Looking at this more closely, specialized benches, such as enforcement, labour, and family reach values above 90%. AggCivil, which, *per se*, is a mix of non-specialized civil cases, is also one of the worst. Therefore, it is possible to conclude that the more specialized, the closer the efficient frontier is to the overarching frontier of benches, and so, there are reasons to believe that diseconomies of scope apply in the Portuguese judiciary.

Using part of our data, we illustrate the concepts of Table 7 in Fig. 5. In this Figure, we represent two partial output frontiers considering two outputs and a single input (the true frontiers have many more inputs and outputs and therefore cannot be represented in a two-dimensional graph — that is why we call these partial frontiers as they use only part of the overall information). It is clear from this comparison that the generic benches' frontier is well below the frontier of the Agg_Criminal group, which in this case would be coincidental with the meta-frontier. The frontier gaps are the distance between the two frontiers, whereas the within-group efficiency is the distance of each observation to its own group frontier. The figure illustrates that in the generic benches, observations are closer to the within-group frontier, while in the criminal benches, we have many observations with very low efficiency scores, presenting a high potential for improvement.

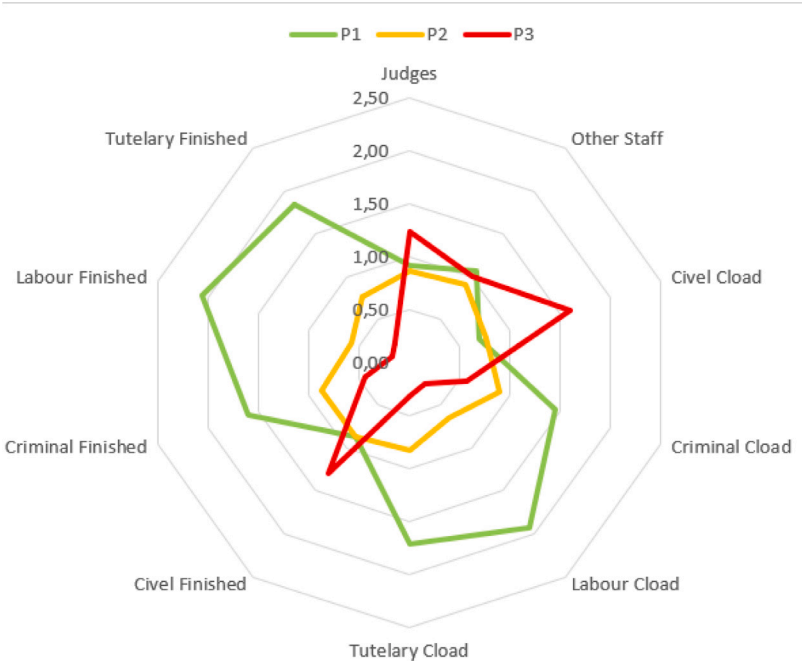

Fig. 4. Distribution of inputs/outputs per category.

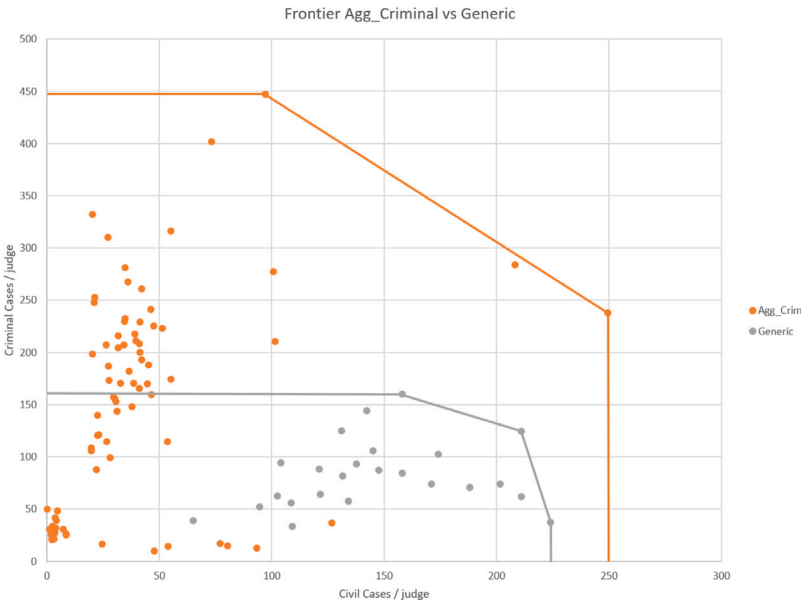

Fig. 5. Partial frontier Agg_Criminal vs Generic.

The above results may somehow be influenced by scale efficiencies, given that generic benches tend to have the smallest number of judges. As a result, we repeated the calculation of within-group and meta-efficiency scores for a subsample of benches — those with a dimension of between two and six judges in order to guarantee comparability in terms of size between the generic benches and the specialized ones (note that all generic benches with a single judge were excluded). The results from this analysis are shown in Table 8.

The analysis of the sub-sample of benches reveals the same conclusions. Generic benches have the lowest average frontier gap, implying that the frontier of generic benches is further apart from the best practices frontier of the remaining types. The specialized benches, dealing mostly with civil cases (and as we know also have criminal cases and therefore are not completely specialized), experience the second lowest frontier gap, implying that more specialization indeed generates a higher proximity of the frontier to the best practices frontier irrespective of the type. It is worth highlighting that parametric and non-parametric tests revealed that differences between groups of benches are statistically significant.

Average within-group efficiencies reveal the homogeneity or heterogeneity of benches within the same type. That is, if all benches of the same type are relatively close to the within-group frontier then the group is rather homogeneous and highly efficient (it should be noted that efficiency within a group does not imply good performance when compared with other types of benches). More heterogeneous groups, in our case, are enforcement benches and Agg_Civil, since within-group efficiency scores are the lowest. For the enforcement benches, for example, this implies that although the best practices frontier for this type of bench is very close to the best performance of the overall sample of benches, most enforcement benches are quite distant from

Table 8
Meta efficiency, within-group efficiency and Gap per specialization for benches with 2–6 judges.

Specialization	Count	Avg meta efficiency	Avg within-group efficiency	Avg ratio (gap)
Enforcement	73	0.524	0.541	0.969
Agg_Civil	278	0.540	0.640	0.849
Commerce	48	0.655	0.712	0.924
Generic	126	0.671	0.859	0.783
Agg_Criminal	220	0.724	0.752	0.961
family	114	0.749	0.776	0.966
Labour	104	0.828	0.897	0.922

that frontier. This implies that there is a lot of potential in terms of efficiency improvement for these types of benches. On the other hand, labour courts are the ones with the least potential for improvement, they are in general concentrated close to the within-group frontier, and this frontier is very close to the overall best practices frontier.

6. Conclusion

This paper investigated the presence of economies of scale and scope within the Portuguese judicial system through a DEA frontier analysis for the 2015 to 2021 period, focusing on the technical efficiency of courts. The findings reveal a small percentage of benches operating under constant returns to scale (8.2%), while 55% experience increasing returns to scale (IRS), implying that they are smaller than the optimal size. Generic benches dealing with civil cases are more prone to IRS, whereas diminishing returns to scale (DRS) are more common in criminal benches. Technical efficiency scores remained stable over time, with the exception of 2020, likely due to the impact of the pandemic. The results indicate that the mix of cases significantly affects technical efficiency, with benches handling more civil cases tending to have lower technical efficiencies. Regarding economies of scope within the Portuguese judicial system, the analysis revealed diseconomies of scope, suggesting that generic benches are further from the meta frontier compared to specialized benches. This finding is consistent across the entire sample of benches and a subsample of comparable-sized benches (between two and six judges). These results align with previous studies, such as those by Elbialy and García-Rubio (2011), who found that benches dealing with criminal cases are more efficient than those handling civil cases, and that generic benches exhibit lower efficiency. Similarly, Pereira and Wemans (2017) analysis of Portuguese data from 2007 to 2013 showed higher productivity in courts specializing in civil law, although the effect of specialization on productivity was less pronounced for other specific competences. Mattsson and Tidanå (2019)'s research on Swedish courts and Castelliano et al. (2021) study also support the conclusion that specialization reduces disposition time and enhances technical efficiency.

Declaration of competing interest

Authors have no interests to declare in this research.

Data availability

Data will be made available on request.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT in order to improve writing. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Acknowledgements

Financial support from Fundação para a Ciência e Tecnologia (through project UIDB/00731/2020) is gratefully acknowledged.

References

Agrell, P.J., Mattsson, P., Månsson, J., 2020. Impacts on efficiency of merging the Swedish district courts. Ann. Oper. Res. 288 (2), 653–679.
Amaral-Garcia, S., 2019. Medical malpractice appeals in a civil law system: do administrative and civil courts award noneconomic damages differently? Law Soc. Rev. 53 (2), 386–419.
Amaral-Garcia, S., Garoupa, N., 2015. Do administrative courts favour the government? Evidence from medical malpractice in Spain. J. Eur. Tort Law 6 (3), 241–259.
Arlota, C., Garoupa, N., 2016. Do specialized courts make a difference? evidence from brazilian state supreme courts. Eur. Bus. Law Rev. 27 (4).
Backes-Gellner, U., Schneider, M.R., Veen, S., 2011. Effect of workforce age on quantitative and qualitative organizational performance: Conceptual framework and case study evidence. Organ. Stud. 32 (8), 1103–1121.
Bajrić, H., Kadrić, E., 2020. Efficiency assessment of municipal courts in federation of Bosnia and Herzegovina using data envelopment analysis. Int. J. Qual. Res. 14 (2), 347.
Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manage. Sci. 30, 1078–1092.
Beenstock, M., Haitovsky, Y., 2004. Does the appointment of judges increase the output of the judiciary? Int. Rev. Law Econ. 24 (3), 351–369.
Bełdowski, J., Dąbroś, L., Wojciechowski, W., 2020. Judges and court performance: a case study of district commercial courts in Poland. Eur. J. Law Econ. 50, 171–201.
Bhattacharya, M., Smyth, R., 2001. Aging and productivity among judges: Some empirical evidence from the High Court of Australia. Aust. Econ. Pap. 40 (2), 199–212.
Casado Pérez, V., 2019. Specialization trend: Water courts. Environ. Law 49 (2), 587–629.
Castelliano, C., Grajzl, P., Alves, A., Watanabe, E., 2021. Adjudication forums, specialization, and case duration: Evidence from brazilian federal courts. Justice Syst. J. 42 (1), 50–77.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring efficiency of decision making units. European J. Oper. Res. 2, 429–444.
Chen, X., Kerstens, K., Tsionas, M., 2024. Does productivity change at all in Swedish district courts? Empirical analysis focusing on horizontal mergers. Soc.-Econ. Plan. Sci. 91, 101787.
Choi, S.J., Gulati, M., Posner, E.A., 2010. What do federal district judges want? An analysis of publications, citations, and reversals. University of Chicago Law & Economics, Olin Working Paper, (508), pp. 10–06.
De Sousa, M., Schwengber, S., et al., 2005. Efficiency estimates for judicial services in Brazil: Nonparametric FDH (Free Disposal Hull) and the expected order-M efficiency scores for Rio Grande Do Sul courts. Encontro da anpec, 33th.
Detotto, C., Serra, L., Vannini, M., 2019. Did specialised courts affect the frequency of business bankruptcy petitions in Spain? Eur. J. Law Econ. 47, 125–145.
Deyneli, F., 2011. Analysis of relationship between efficiency of justice services and salaries of judges with two stage DEA method. Eur. J. Law Econ. 34 (3), 477–493.
Dimitrova-Grajzl, V., Grajzl, P., Sustersic, J., Zajc, K., 2012. Court output, judicial staffing, and the demand for court services: Evidence from Slovenian courts of first instance. Int. Rev. Law Econ. 32 (1), 19–29.
Elbialy, N., García-Rubio, M.A., 2011. Assessing judicial efficiency of Egyptian first instance courts: A DEA analysis. Technical Report, MAGKS joint discussion paper series in economics.
Falavigna, G., Ippoliti, R., 2021. Reform policy to increase the judicial efficiency in Italy: The opportunity offered by EU post-covid funds. J. Policy Model. 43 (5), 923–943.
Falavigna, G., Ippoliti, R., 2022. Model definitions to identify appropriate benchmarks in judiciary. J. Appl. Econ. 25 (1), 339–360.
Falavigna, G., Ippoliti, R., Manello, A., Ramello., G., 2015. Judicial productivity, delay and efficiency: A directional distance function (DDF) approach. European J. Oper. Res. 240, 592–601.
Falavigna, G., Ippoliti, R., Ramello, G.B., 2018. DEA-based malmquist productivity indexes for understanding courts reform. Soc.-Econ. Plan. Sci. 62, 31–43.
Färe, R., Grosskopf, S., Lovell, C.A.K., 1985. The Measurement of Efficiency of Production. Kluwer-Nijhoff Publishing, Boston.
Farrell, M.J., 1957. The measurement of productive efficiency. J. R. Stat. Soc. Series A, general 120 (Part 3), 253–281.

Finocchiaro Castro, M., Guccio, C., 2014. Searching for the source of technical inefficiency in Italian judicial districts: an empirical investigation. Eur. J. Law Econ. 38 (3), 369–391.
Fusco, E., Laurenzi, M., Maggi, B., 2021. Length of trials in the Italian judicial system: An efficiency analysis by macro-area. Justice Syst. J. 42 (1), 78–105.
Garoupa, N., Jorgensen, N., Vazquez, P., 2010. Assessing the argument for specialized courts: Evidence from family courts in Spain. Int. J. Law Policy Family 24 (1), 54–66.
Giacalone, M., Nissi, E., Cusatelli, C., 2020. Dynamic efficiency evaluation of Italian judicial system using DEA based Malmquist productivity indexes. Soc.-Econ. Plan. Sci. 72, 100952.
Gupta, M., Bolia, N.B., 2020. Efficiency measurement of Indian high courts using DEA: a policy perspective. J. Policy Model. 42 (6), 1372–1393.
Hagstedt, K., Proos, J., 2008. Has the recent restructuring of the Swedish district courts improved efficiency?: A DEA analysis. Technical Report, Nationalekonomiska institutionen.
Kittelsen, F.R., 1992. Efficiency analysis of norwegian district courts. J. Prod. Anal. 3 (3), 277–306.
Lewin, A.Y., Morey, R.C., Cook, T.J., 1982. Evaluating the administrative efficiency of courts. Omega, Int. J. Manag. Sci. 10 (4), 401–411.
Magalhães, P.C., Garoupa, N., 2020. Judicial performance and trust in legal systems: Findings from a decade of surveys in over 20 European countries. Soc. Sci. Q. (Wiley-Blackwell) 101 (5).
Månsson, J., Andersson, C., Bonander, F., 2022. What lessons can be learned from cost efficiency? The case of Swedish district courts. Eur. J. Law Econ. 54 (3), 431–451.
Mattsson, P., Månsson, J., Andersson, C., Bonander, F., 2018. A bootstrapped Malmquist index applied to Swedish district courts. Eur. J. Law Econ. 46 (1), 109–139.
Mattsson, P., Tidanå, C., 2019. Potential efficiency effects of merging the Swedish district courts. Soc.-Econ. Plan. Sci. 67, 58–68.
Morita, H., 2003. Analysis of economies of scope by data envelopment analysis: comparison of efficient frontiers. Int. Trans. Oper. Res. 10 (4), 393–402.
Nissi, E., Giacalone, M., Cusatelli, C., 2019. The efficiency of the Italian judicial system: A two stage data envelopment analysis approach. Soc. Ind. Res. 146 (1), 395–407.
Panzar, J.C., Willig, R.D., 1981. Economies of scope. Am. Econ. Rev. 71 (2), 268–272.
Pereira, M.C., Wemans, L., 2017. Productivity in civil justice in Portugal: A crucial issue in a congested system. Banco de Portugal Econ. Stud. 3 (1), 1–29.
Peyrache, A., Zago, A., 2016. Large courts, small justice!: The inefficiency and the optimal structure of the Italian justice sector. Omega 64, 42–56.
Ramseyer, J.M., 2012. Talent matters: Judicial productivity and speed in Japan. Int. Rev. Law Econ. 32 (1), 38–48.
Revesz, R.L., 1989. Specialized courts and the administrative lawmaking system. U. Pa. L. Rev. 138, 1111.
Santos, S., Amado, C., 2014. On the need for reform of the portuguese judicial system - does data envelopment analysis assessment support it? Omega, Int. J. Manag. Sci. 47, 1–16.
Schneider, M., 2005. Judicial career incentives and court performance: An empirical study of the german labour courts of appeal. Eur. J. Law Econ. 20 (2), 127–144.
Seiford, L.M., Zhu, J., 1999. An investigation of returns to scale in data envelopment analysis. Omega, Int. J. Manag. Sci. 27, 1–11.
Silva, M.C.A., 2018. Output-specific inputs in DEA: An application to courts of justice in Portugal. Omega 79, 43–53.
Tulkens, H., 1993. On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts and urban transit. J. Prod. Anal. 4, 183–210.
Voigt, S., 2016. Determinants of judicial efficiency: a survey. Eur. J. Law Econ. 42, 183–208.
Xiaoqing, C., Kerstens, K., 2023. Evaluating horizontal mergers in Swedish district courts using plant capacity concepts: With a focus on nonconvexity. RAIRO-Oper. Res. 57, 219–236.