EKONOMIA I PRAWO. ECONOMICS AND LAW
Volume 24, Issue 4, 2025
p-ISSN 1898-2255, e-ISSN 2392-1625
www.apcz.umk.pl/EiP

Original article
received 07.04.2025; accepted 11.12.2025; published 31.12.2025
doi: 10.12775/EiP.2025.20

Limitations of Two-Dimensional Indicators and unconstrained radial DEA models in Evaluating Judicial Efficiency: Insights from Poland's Appellate Court System

WALDEMAR FLORCZAK

corresponding author
Institute of Economics, Finance and Management, Faculty of Management and Social Communication, Jagiellonian University in Kraków, ul. prof. St. Łojasiewicza 4, 30-348 Kraków
waldemar.florczak@uj.edu.pl
ORCID: https://orcid.org/0000-0003-3408-7784

Abstract

Motivation: Two-dimensional efficiency indicators provide only a narrow perspective on a system's functional efficiency, a limitation that is addressed by multidimensional methods such as the widely used Data Envelopment Analysis. However, the uncritical application of this class of methods—which appears to be a common practice in judicial efficiency research—can result in misleading conclusions regarding efficiencies of Decision-Making Units (DMUs).

Aim: The aim of this study is to evaluate the technical efficiency of Poland's appellate court system in the first two decades of the 21st century (years 2002-2021) using a constrained version of Data Envelopment Analysis (DEA) applied to time-series data. Unlike two-dimensional judicial performance indicators, such an approach facilitates a comprehensive assessment of the aggregated efficiency of the legal, administrative, procedural, organizational, and economic macro-frameworks—consistently implemented across all appellate courts—which collectively shape the system's performance during each year under review.

Results: The outcomes of this research show the fundamental shortcomings of classical applications of DEA, raising valid concerns about the reliability and interpretability of the outcomes derived from this approach. To mitigate these challenges, it presents a novel procedure specifically aimed at addressing these limitations in evaluating judicial efficiency, which is next successfully implemented.

Keywords: judicial efficiency, appellate court system, (un)constrained radial DEA models, disposition time

JEL: C44; C80; K19

1. Introduction

Ensuring careful management of public funds, reflected in the commitment to rational economic principles across all areas of the public sector, should be a genuine — not merely declarative — objective of every modern state. The well-being of contemporary societies increasingly depends on the quality and efficiency of public institutions. Although much of the impact of public initiatives cannot be directly monetized, this does not relieve governments from the responsibility to pursue objective valuation and assessment of public sector inputs and outcomes. Numerous methods are available to facilitate the objectivity of such efforts.

An efficient and independent general judiciary, as a critical component of the judicial system, significantly influences the quality of social, public, and economic life. It forms the backbone of democracy and a free-market economy, generating numerous positive externalities. Given the judiciary's pivotal role in the functioning of society and the state, it should operate as a particularly effective sector within the public sphere. Achieving high judicial efficiency, as well as the reforms needed to accomplish this, requires an accurate assessment of the prevailing trends and current state through objective quantitative analysis tools. However, in Poland, public, academic, and media discussions surrounding the state of the judiciary, evaluations of its past performance, and suggested directions for reform are often dominated by legalistic debates. These discussions frequently overlook or selectively reference basic, two-dimensional efficiency indicators published on the Ministry of Justice's website, particularly the widely referenced metric of average case processing time (Ostaszewski, ed., 2020).

The study in this paper is fundamentally quantitative, engaging with legal debates around judicial system reforms only insofar as these reforms have measurable effects on the efficiency of the appellate court system. The focus remains on how changes in the judiciary manifest in quantitative performance metrics, rather than on the qualitative discussions or policy arguments driving possible reforms. This approach ensures that the analysis remains grounded in empirical data, directly assessing the impact of reforms on the operational effectiveness of the appellate courts.

The aim of this study is to estimate the technical efficiency of Poland's appellate court system over the 21st century, specifically from 2002 to 2021, by employing an adequate variant of Data Envelopment Analysis (DEA), the primary analytical tool for evaluating the efficiency of public sector entities (e.g., Banker et al., 2014; Data Envelopment Analysis in the Public Sector, 2014; Avkiran, 2015), including the judiciary (e.g., Voigt, 2016). This study's macro-level perspective enables a holistic evaluation of the effectiveness of systemic legal, administrative, procedural, organizational, and economic frameworks—uniformly applied across all appellate courts—that synergistically influence the operations of judicial entities and thereby shape overall systemic efficiency. Only through the use of time series data aggregated at the macro level is it feasible to assess the impact of these systemic determinants on efficiency. Such an approach allows for an examination of the years in which these solutions were relatively more or less effective, ultimately helping to identify possible underlying causes for these variations.

The research presented in this article builds on international studies in this field by spanning a 20-year period focused on a single object: the last-instance court system. This approach facilitates a deeper understanding of systemic efficiency. Furthermore, by treating system load as a non-discretionary input and employing a DEA variant that maximizes outputs, the study addresses concerns related to discretionary influences. Analysing the weight values obtained from the classic DEA model and applying constructive adjustments enhances the substantive interpretation of the results, ultimately providing valuable insights into the efficiency of Poland's appellate court system over time. These features contribute to the study's originality and practical significance.

The structure of this paper is organized as follows. The next section contains a review of the literature on DEA applications for assessing efficiency within the judiciary, addressing various perspectives and stakeholders. Subsequently, a concise overview of the DEA methodology is provided, along with a detailed description of the database and some initial descriptive observations. The main empirical findings are then presented, and the paper concludes with a summary of key insights.

2. Literature review

Efficiency is a fundamental concept and one of the most significant topics of research in economics. This is underscored by the assertion of some scholars that the pursuit of effective utilization of scarce resources is at the core of this discipline. E.g., according to Krajewski and Milewski (2018, p. 11), economics: *...shows how people use scarce resources, how they use them to conduct economic activities, how they allocate them between different uses that compete with each other, and what guides them in making such choices. It also shows whether the use of scarce resources is efficient or not and analyses the factors on which this depends.*

Strategies aimed at enhancing the organizational efficiency of individual decision-making units can lead to low-cost or even cost-free improvements in the productivity of various public sector institutions, including the judiciary, once the extent of their inefficiency has been identified and comparable efficient models have been recognized. However, a critical prerequisite for the successful implementation of such measures is a thorough diagnosis of the current situation, which must take into account prevailing quantitative trends over the historical period.

Measuring efficiency in the public sector, particularly within the judiciary, presents a range of complex challenges that stem from the complexity of judicial functions, the diverse perspectives of stakeholders, and the qualitative dimensions of justice. To enhance effectiveness measurement, there is a need for a multifaceted approach that integrates both qualitative and quantitative indicators, respects judicial independence, and accounts for the diverse nature of judicial work. Developing such comprehensive frameworks is crucial for fostering a fair, efficient, and effective judicial system that meets the needs of society.

In the quest to measure effectiveness in the public sector, particularly within the judiciary, quantitative methods provide essential insights. Among these, traditional metrics such as case clearance rates, average processing times, and judicial workload analyses offer straightforward, yet very limited, perspectives on performance. However, advanced methodologies like Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) offer more sophisticated frameworks for assessing efficiency, presenting both advantages and limitations compared to conventional metrics.

Of the two afore-mentioned it is indisputably – as judged by literature review – DEA and it's modifications prevail in investigations into efficiency of public sector institutions. This is so due to its key features, such as the capacity to account for multiple inputs and outputs common to all decision-making units (DMUs) without requiring arbitrary assumptions about the production function. Furthermore, the optimization techniques employed in DEA favour each of the DMUs being analysed, which makes the methodology particularly advantageous from their standpoint. These characteristics have led to DEA's widespread use in evaluating the efficiency of various entities in the public sector (Banker et al., 2014; Data Envelopment Analysis in Public Sector, 2014; Avikran, 2015), including the judicial system.

Two-dimensional productivity and efficiency indicators, such as those available on the Ministry of Justice's website offer only a limited view of a system's functional efficiency since they provide only a partial representation of its performance. Even among these two-dimensional metrics, those that focus on the 'final effects' of a system's activity (for instance, the commonly discussed indicators on the average duration of court proceedings) fall short of delivering an accurate measure of systemic efficiency because they fail to incorporate the relationship between outputs and inputs. Consequently, fluctuations in these indicators could signal either a shift in efficiency levels or simply reflect changes in the system's workload or burden.

Unlike all two-dimensional efficiency measures DEA enables estimation of overall, technical efficiency of homogeneous public sector institutions because it (i) takes into account the diversity of functions/tasks performed by the institutions, whereby it is not possible to objectively (monetarily) assess the importance/value of the performance of individual functions/tasks, (ii) takes into account the main factor inputs that serve to fulfil the functions/tasks entrusted to such institutions, (iii) needs no knowledge regarding the analytical form of the production function according to which factor inputs are transformed into the realisation of individual functions/tasks by the institutions analysed, i.e. into outputs, (d) offers the possibility of objectively assessing the degree of empirical efficiency achieved by individual institutions in comparison with other homogeneous institutions included in the sample.

When investigating judicial efficiency through DEA, several fundamental issues dictate its application and the interpretation of results. These issues encompass the subject of research (DMU), the scope of the investigation, and the data utilized. The research subject, the extent of the analysis, and the data can vary significantly, from micro-survey data relating to factors affecting judicial efficiency (e.g., Ostrom, Hanson, 2003; Schneider, 2005) to evaluations of selected first-instance courts in specific countries (e.g., Garcia-Riubio, Rosales-Lopez 2010; El Bialy, Garcia-Rubio 2011; Falavigna et al., 2015), as well as assessments of selected higher courts (e.g., Yeung, Azevedo 2011; Yeung, 2018). Some studies even attempt to analyse all first-instance courts over a selected period (e.g., Buonanno Galizzi 2014; Peyrache, Zago 2016; Mattsson, Tidana 2019) or higher courts (Gupta, Bolia 2020), and others explore international cross-sectional environments (e.g., Deynelli, 2012; Malcarne, Ramello 2015; Voigt S, El Bialy, 2016).

The selection of variables defining inputs varies widely in empirical studies. While the number of judges is almost always included as a "duty variable" (exceptions include Lewin et al., 1982 and Tulkens, 1993), other variables are often selected based on data availability. Along with the number of judges, DEA studies frequently consider variables such as the number of pending cases (caseload), which is often paired with the number of judges (e.g., Pedraja, Salinas 1996; Schneider 2005; Peyrache, Zago 2016; Mattsson, Tidana 2019), as well as the size of non-judicial staff (e.g., Kittelsen, Forsund 1992; Pedraja, Salinas 1996; Agrell et al., 2019), office and IT equipment, work-space (e.g., El Bialy, Garcia-Rubio 2011; Yeung, Azevedo 2011; Agrell et al., 2019), and expenditures on court operations (e.g., Deynelli 2012). Notably, in the majority of DEA applications concerning the judiciary, the number of inputs considered simultaneously rarely exceeds three, with Peyrache and Zago (2016) serving as an exception.

When it comes to selecting outcome-defining variables, most empirical applications reviewed include an aggregate of cases resolved in a given year as their sole output factor. This presents a significant simplification, as it does not adequately account for the variation in litigation durations and labour inputs across different legal subjects. Such a homogenized approach raises concerns about the reliability of the efficiency scores derived from these broad aggregate results. Some exceptions to this generalization are found in studies like Kittelsen and Forsund (1992), which categorize cases by subject of law into seven categories, Tulkens (1993) with four categories, and Mattsson and Tidana (2019) and Agrell et al. (2019), which each utilize three categories, as well as Santos and Amado (2014), who provide a comprehensive breakdown into 43 categories.

The issue of non-discretionary variables in judicial efficiency analyses is often overlooked in DEA. While some researchers address the influence of such factors and external conditions in later stages of their analysis using two-stage DEA (2SDEA) (e.g., Schneider 2005; Falavigna et al., 2015), it is critical to recognize that the caseload variable—which aggregates backlog cases from previous periods with newly received cases—is inherently a non-discretionary variable. Consequently, DEA analyses should adopt either analytical variants that directly account for the presence of non-discretionary factors (Ruggiero, 1997; Muniz et al., 2006; Cordero-Ferrera et al., 2010) or use the output maximization version of DEA.

Regarding the choice of DEA version, the empirical literature predominantly employs the outcome-maximizing version of DEA, which aligns with the judiciary's operational specifics, particularly given the increasing demands placed on court systems in recent years. The only noted exception is Schneider (2005). In instances where a non-discretionary variable appears among the inputs (e.g., caseload), it can only be treated the same way as discretionary inputs if the chosen version of DEA is output-oriented (Dyson et al,. 2001; Cook et al., 2014).

In terms of scale returns, studies utilizing constant returns to scale (CRS) dominate the literature, while variable returns to scale (VRS) are employed when analysing the effect of court size on efficiency, as seen in the works of Garcia-Rubio, Rosales-Lopez (2010), El Bialy and Garcia-Rubio (2011), and Peyrache and Zago (2016). However, the selection between CRS and VRS tends to be somewhat arbitrary, resulting in ambiguous conclusions regarding scale effects within the judiciary.

The limited degree of discrimination in the performance indicators obtained—especially with a small number of DMUs and numerous inputs and outputs—often leads to a high percentage of efficient DMUs. To address this, authors frequently resort to techniques aimed at enhancing the differentiation of efficiency indicators. One common method is the super-efficiency concept, introduced by Andersen and Petersen in (1993), to amplify the variation in efficiency among DMUs already deemed efficient through classic DEA. Classical super-efficiency (SE-DEA) models are constructed as radial extensions of the CCR or BCC models. This means that their objective function maximizes or minimizes the radial efficiency measure (input- or output-oriented), while slack variables do not appear in the objective function and are not active decision variables in the construction of rankings above one.

Furthermore, the reflection—or lack thereof—on the assessment of weights assigned to inputs and outputs, particularly the zero and unit weights, is a significant issue. To the author's knowledge, with the exception of a single study by Santos and Amado (2014), no other research has thoroughly investigated the substantive acceptability of the weights attributed to individual inputs and outputs. This lack of scrutiny raises doubts about the substantive validity of the obtained results.

Finally, all of the studies referenced above employed cross-sectional data to examine variations in the efficiency of individual courts within a single year of analysis. However, as noted by researchers such as Maital and Vaninsky (1999) and Avkiran (2015), DEA can also be used for longitudinal or time series data, applying it to a single entity—such as the entire court system at a particular judicial level—to track changes in efficiency over time. The prevalent use of DEA in the judiciary with cross-sectional data is likely due to limited access to time-series data at the necessary level of detail. Nonetheless, DEA still lacks a universally accepted framework for clearly separating temporal and cross-sectional effects in panel-data settings.

Empirical research on the efficiency of the Polish judiciary using Data Envelopment Analysis (DEA) remains relatively limited but has developed along a clear trajectory. The first systematic applications concerned entities within the justice sector rather than courts sensu stricto: Guzowska and Strąk (2010) analysed 45 public prosecutor regions, estimating technical and cost efficiency under CRS and VRS assumptions and identifying sizeable potential savings, thereby demonstrating the suitability of DEA for Polish justice-sector units. A subsequent milestone was Major (2015), who applied CCR-DEA to 26 district civil courts in the Kraków region, treating human resources as inputs and resolved cases as outputs, and showing scope for backlog reduction without additional resources, thus setting a reference framework for later court-level studies.

More recent work has both deepened and diversified the approach. Kapelko (2024) investigates Polish district courts nationwide for 2017–2021 using input- and output-specific DEA-based models, decomposing inefficiency by factors (judges, other staff) and types of cases, and showing strong heterogeneity across courts and a marked deterioration in 2020 linked to the COVID-19 shock. At the regional-court level, Świtłyk, Sompolska-Rzechuła and Oesterreich (2025) estimate technical and scale efficiency of 43 regional courts (2015–2022) in separate civil, criminal and labour divisions, documenting generally high CCR scores but also persistent scale inefficiencies and size effects in civil divisions.

3. Methods and data

In economic terms, efficiency measurement typically involves assessing the relationship between multifarious inputs and outputs. When both all inputs and all outputs can be quantified in monetary terms, efficiency can be expressed using the following formula:

$$Efficiency = \frac{\sum_{o=1}^{O} p_o Y_o}{\sum_{i=1}^{I} p_i X_i} \tag{1}$$

Where:
Y_0 – volume of the *o-th* output,
X_i – volume of the *i-th* input,
p_0 – unit price of the *o-th* output,
p_i – unit price of the *i-th* input.

One such measure that is based directly on formula (1) and extensively used in economic and managerial analyses is the Benefit–Cost Ratio, which evaluates whether the aggregated monetary value of outputs exceeds the aggregated monetary value of inputs. Closely related is the well-known Return on Investment (ROI), which expresses the excess of benefits over costs relative to the total value of inputs and is routinely applied in corporate finance and performance evaluation. Both BCR and ROI have been widely employed in empirical research, for example in project appraisal in public infrastructure (Boardman et al. 2018) or in evaluating firm-level investment efficiency (Damodaran 2012).

This approach, however, is practically limited to economic agents operating within markets, where nearly all inputs and outputs are measurable in monetary terms. In public service provision, such as within the judiciary, this framework falls short, as most public sector activities cannot be valued monetarily. Nonetheless, the general formula holds if unit price parameters are appropriately replaced by weights that reflect the relative importance of each output produced by a public institution. The primary and most challenging issue lies in determining these weights accurately, leading directly to the relevance of Data Envelopment Analysis (Allen et al. 1997, Podinovski 2016).

The starting point of the classical DEA method (e.g., Panwar et al., 2022) is the following quotient of the weighted sum of outputs to the weighted sum of inputs, defining the technical efficiency of a given decision-making unit (for ease of reading, the equivalents of general terms used in this study are provided in parentheses):

$$E_z = \frac{\sum_{t=1}^{T} w_{tz} Y_{tz}}{\sum_{s=1}^{S} v_{sz} X_{sz}} \tag{2}$$

E_z – technical efficiency of the *z-th* decision-making unit (efficiency of the appellate court system in each year of the 2002–2021 period);

$z = 1, 2 \dots Z$ – index of the *z-th* decision-making unit (total Z =20, i.e. number of years);

X_{sz} – the *s-th* type of input in the *z-th* decision-making unit (the amount of the *s-th* input - out of the three considered in the study: the number of judges, outlays on the appellate court and the system load – in a given year of analysis);

$s = 1, 2 \dots S$ – index of the *s-th* input (total of S=3 types of inputs, listed above, common to all decision-making units);

Y_{tz} – *t-th* type of output in the *z-th* decision-making unit (amount of the *t-th* type of output – out of the six included in the study: number of resolved civil, criminal, family, labour law, social security and commerce cases – in the given year of analysis);

$t = 1, 2 \dots T$ – index of the *t-th* output (total of T=6 output types, listed above, common to all decision-making units;

W_{tz}, V_{sz} – weights assigned to the *t-th* outcome and *s-th* input in the *z-th* DMU, respectively.

A key property of the DEA method is that the weights appearing in formula (2) are determined through a linear programming procedure, rather than by arbitrarily assigning specific values to them. Thus, the most important—and at the same time, most 'inflammatory'—issue of determining the values of these weights, which implicitly conveys the gradation/importance/significance of individual inputs and outputs, is objectivized in a way that ensures the most favourable set from the perspective of individual decision-making units. This is achieved by maximizing outputs and minimizing inputs for all the DMU-s in the sample, represented by the following mathematical problem (e.g., Panwar et al., 2022):

$$max\ E_z = \frac{\sum_{t=1}^{T} w_{tz} Y_{tz}}{\sum_{s=1}^{S} v_{sz} X_{sz}} \tag{3}$$

with respect to the following constraints:

$$0 \leq \frac{\sum_{t=1}^{T} w_{tz} Y_{ti}}{\sum_{s=1}^{S} v_{sz} X_{si}} \leq 1 \quad (4)$$

where „*i*” denotes all DMU-s in the sample $0 \leq W_{tz;}\ 0 \leq V_{sz}$ (5)

The above problem can be reduced to an equivalent linear programming problem, which determines the optimal weights. The optimization is conducted separately for each DMU included in the sample:

$$max\ E_z = \sum_{t=1}^{T} w_{tz} Y_{tz} \quad (6)$$

with respect to the following constraints:

$$\sum_{t=1}^{T} w_{tz} Y_{ti} - \sum_{s=1}^{S} v_{sz} X_{si} \leq 0 \quad (7)$$

$$\sum_{s=1}^{S} v_{sz} X_{sz} = 1 \quad (8)$$

$$W_{tz}, V_{sz} \geq 0 \quad (9)$$

The classic version of Data Envelopment Analysis (DEA) is hampered by several significant assumptions and undesirable properties that can obscure the clarity of its results. This is especially true when these issues are overlooked, as is common in studies concerning the judicial sector. One notable concern is the low discrimination power of the results, which becomes problematic in analyses involving a limited number of Decision-Making Units (DMUs). To enhance result discrimination, researchers have begun to favour a super-efficiency variant of DEA known as the SE-DEA model (Andersen, Petersen 1993). A concise representation of this extension of the DEA-CRS model is shown below:

$$max\ E_z = \sum_{t=1}^{T} w_{tz} Y_{tz} \quad (10)$$

with respect to the following constraints:

$$\sum_{t=1}^{T} w_{tz} Y_{ti} - \sum_{s=1}^{S} v_{sz} X_{si} \leq 0, \forall i \neq z \quad (11)$$

$$\sum_{s=1}^{S} v_{sz} X_{sz} = 1 \quad (12)$$

$$W_{tz}, V_{sz} \geq 0 \quad (13)$$

The difference between the classical DEA-CRS and its super-efficiency extension consists in excluding the *z-th* DMU from constraint (7) and using constraint (11) instead (see formulae (7) and (11)). As a result, the efficiency scores of DMUs that were previously considered efficient in the classical DEA-CRS model (with scores of 1) are no longer capped at 1, allowing for a more meaningful discrimination among these DMUs.

Additionally, traditional DEA produces efficiency ratios that range from 0 to 1, potentially leading to situations where certain efficient Decision-Making Units —those achieving a score of 1—attribute their efficiency solely to a single input or output. This aspect raises significant concerns regarding the interpretation of findings obtained from conventional DEA analyses. Although the DEA algorithm's ability to optimally choose weights to enhance each DMU's interests is a key feature, the choice of inputs and outputs remains vital for evaluating efficiency and cannot be overlooked. Granting full credit for efficiency to one factor implies that all other factors receive zero weights, complicating the interpretation. Moreover, it is important to acknowledge that DEA measures relative or empirical efficiency, not potential efficiency, which raises questions about the appropriateness of presenting this efficiency in what may be perceived as the "most favourable light" — one that has never been observed.

An effective approach to addressing the above issues is the use of the constrained SE-DEA model, which ensures that all inputs and outputs meaningfully contribute to the analysis—a feature intrinsic to the structure of any DEA model. Although this version of DEA has not yet been applied to judicial efficiency contexts, it compels optimal weights to align with specific, logically based assumptions, ensuring that no weights are zero and thereby addressing previously raised concerns.

Generally, three types of weights are distinguished in constrained SE-DEA analysis (Allen et al. 1997): absolute, price, and virtual weights. In this research, the latter two were used, resulting in the following presentation of the constrained SE-DEA model:

$$max\ E_z = \sum_{t=1}^{T} w_{tz} Y_{tz} \quad (14)$$

with respect to the following constraints:

$$\sum_{t=1,z\neq a}^{T} w_{ti} Y_{ti} - \sum_{s=1}^{S} v_{si} X_{si} \leq 0, \forall i \quad (15)$$

$$\sum_{s=1}^{S} v_{sz} X_{sz} = 1 \quad (16)$$

$$W_{tz}, V_{sz} \geq 0 \quad (17)$$
$$W_{sd} \leq V_{sz} \leq V_{se} \quad (18)$$
$$W_{tb} \leq V_{tz} \geq W_{tc} \quad (19)$$
$$f \leq W_{wz} y_{tz} g_t \quad (20)$$

The additional constraints (18)-(20) introduce the following modifications to the unconstrained weights in the classical SE-DEA:

(i) Subscripts *b* and *c* denote DMU-s whose input weights cannot be greater or smaller than those of the referenced *z-th* DMU, respectively (these are referred to as input *price weights*).

(ii) Subscripts *d* and *e* denote DMU-s whose input weights cannot be greater or smaller than those of the referenced *z-th* DMU, respectively (also referred to as output *price weights*).

(iii) Constants *f* and *g* put the lower and the upper limits upon the contribution/share of the *t-th* output in the production of total outputs (referred to as *virtual weights*).

To complete the outlined analysis, data collection at the macroeconomic level was essential, requiring a combination of multiple sources and the subsequent processing of gathered information. Data concerning the activities of the appellate court system, categorized by type of case (including criminal, civil, family, labour law, social security, and commercial cases) for the years 2011–2021, were obtained from publicly accessible databases provided by the Ministry of Justice. The primary resource was the .xls-format dataset titled *Register of Cases in Common Courts in Poland* (available at https://isws.ms.gov.pl/pl/baza-statystyczna/opracowania-wieloletnie). To extend the scope of the analysis to cover 2002–2010, the document *Statistical Analysis of the Activity of the Judiciary in the Years 2002–2011* (https://isws.ms.gov.pl/pl/baza-statystyczna/opracowania-jednoroczne-w-tym-pliki-dostepne-cyfrowo/rok-2011/) was also utilized.

In preparing the data for application to DEA efficiency analyses, several identity-based adjustments were necessary. These adjustments entailed defining fundamental relationships within the statistical court records to ensure the accuracy and consistency required for proper DEA computation. Specifically, the following relationships were essential:

$$Caseload_t = Pending\ cases_{t-1} + Incoming\ cases_t \tag{21}$$

$$Pending\ cases_t = Caseload_t - Resolved\ cases_t \tag{22}$$

$$Case\ processing\ time_t = \frac{Pending\ cases_t}{Resolved\ cases_t} \tag{23}$$

where the subscript t stands for *t-th* year.

For analysing the performance of appellate courts, the number of cases processed by area of law was selected as a central measure. This choice of variables ensures both homogeneity and completeness for the decision-making units under review—here, each year of appellate court activity—meeting a critical requirement for the DEA method's validity.

Data for input variables were gathered from various sources. The non-discretionary caseload variable, representing system load, was calculated based on equation (21). Information on the average number of judges in appellate courts from 2006 to 2021 was obtained through public information access requests to the Ministry of Justice. While corresponding data for 2002-2005 were not available, the number of judges could be accurately estimated for these earlier years. This estimation relied on judicial appointment statistics for the full 2002-2021 period, drawn from the statistical yearbooks published by the Central Statistical Office (CSO), and applied the perpetual inventory method to bridge the data gap.

The starting point was the following identity linking stocks and streams of relevant variables:

$$K_t = K_{t-1} + I_t - D_t \tag{24}$$

K_t – stock/capital in period t (in our case – the number of judges in a given year),
K_{t-1} – stock/capital in the preceding period (the number of judges in the preceding year),
I_t – stream of gross investment in stock K in period t (number of newly appointed judges in a given year),
D_t – the stream of depreciated stock/capital in period t (the number of judges who stopped – for whatever reason – working in the courts in year t).

Formula (24) is the same as the following relation:

$$K_t = (1 - \sigma_t)K_t - I_t \tag{25}$$

σ – is the stock/capital depreciation rate (the fraction of judges employed in year *t-1*, but ceasing – for whatever reason – to work in year *t*).

Assuming a constant depreciation rate from 2002 to 2005, equivalent to the average rate observed from 2006 to 2021, the number of judges in the appellate courts during 2002–2005 was estimated using a recursive form of formula (25):

$$K_{t-1} = (K_t - I_t)/(1 - \sigma) \tag{26}$$

σ – the average depreciation rate for 2006-2021, which is 0.04 (the average depreciation rate proved relatively stable, remaining close to 0.04 for variously selected time ranges, not necessarily based on the full sample for 2006–2021).

Data on state budget expenditures for the appellate judiciary, adjusted to constant prices, required several recalculations and sources. First, a time series of current expenditures on the broader judiciary (encompassing the appellate judiciary) was compiled from Ministry of Finance data sources (https://www.gov.pl/web/finanse/sprawozdania-roczne and https://mf-arch2.mf.gov.pl/web/bip/ministerstwo-finansow/dzialalnosc/finanse-publiczne/budzet-panstwa / wykonanie-budzetu-panstwa/sprawozdanie-z-wykonania-budzetu-panstwa-roczne). This approach was necessary because the Central Statistical Office's (CSO) statistical yearbooks do not provide budget expenditure data at the required level of disaggregation.

Second, in calculating the share of total judicial budget directed to the appellate judiciary, the fraction determined by Siemaszko and Ostaszewski (2013) was applied. Third, the budget allocation for the appellate judiciary in current prices was obtained by multiplying the total judiciary budget by the appellate judiciary fraction. Finally, this time series was adjusted to real terms by dividing it by the 2015 Consumer Price Index (CPI) deflator, yielding a constant-price series. The processed data used in this research is presented in Table 1.

Both the graphical analyses of the analysed categories and the linear correlation coefficients[1] for the first three variables suggest a low degree of collinearity among the input variables. This finding implicitly indicates that each variable can independently influence output levels, a conclusion supported by both theo-

1 For want of space, all figures elaborated on the basis of the data reported in Table 1—and supporting the descriptive conclusions in the following sections—as well as the Pearson correlation coefficients, have been omitted.

retical and empirical evidence (as detailed in the literature review). It should be noted that there is no need to examine the dataset for potential outliers, as such information is not relevant in the present study. This is because the dataset employed here contains macro-aggregated and complete information on the functioning of the Polish appellate court system over a twenty-year period and therefore represents a full population rather than a statistically defined sample. In contrast, detecting potential outliers is highly advisable—although not always undertaken—when efficiency analyses rely on data for individual courts, as is the case in all empirical studies cited in the "Literature Review" section.

Throughout the analysed period, the number of judges increased moderately, with an average annual growth rate of 1.17%. In contrast, real expenditures on the appellate judiciary experienced substantial growth, exceeding 4% per year over the same timeframe, resulting in total budget allocations for the appellate judiciary more than doubling by the final year of the analysis compared to the initial year. Consequently, the per-judge budget outlay also saw a significant increase.

Conversely, the number of cases pending resolution remained relatively stable. This stability led to consistently variable yet overall stable volumes of cases per judge, alongside a regular rise in unit costs associated with case resolution. As a result, judges are faced with a fairly constant caseload while their salaries have been progressively increasing. Notably, the average growth rates of pending cases outpaced those of resolved cases—1.4% compared to 1%, respectively—resulting in a continually rising number of unresolved cases, particularly evident since 2014.

Significant disparities in the dynamics of the output variables are evident. The most pronounced increase is observed in criminal cases, which nearly doubled during the analysis period. Additionally, civil cases represent another category that exhibited a growth trend, with levels at the conclusion of the sample period surpassing those at its commencement. In contrast, commercial and insurance cases displayed high variability, yet maintained relatively stable expected values. Meanwhile, volumes of labour cases have been in a consistent decline.

It is essential to note that the fluctuations in the number of resolved cases correspond to parallel changes in the volume of incoming cases for each category of law. Furthermore, the Pearson linear correlation coefficients among the various output variables are relatively low. This indicates that none of the aforementioned variables replicate the informational content of the others, as the dynamics of all variables differ significantly.

Consequently, beyond the substantive rationale—characterized by a set of mutually exclusive and collectively exhaustive elements, namely the number of cases resolved by category as outcome variables—the statistical rationale further underscores the importance of examining all these variables simultaneously. This approach stands in contrast to the predominant tendency in most studies on judicial efficiency utilizing the Data Envelopment Analysis (DEA) method, which often relies on a singular aggregate figure.

A similar conclusion applies to the input variables incorporated into the analysis. While the role of judges in adjudicating cases within appellate courts is undeniably critical and irreplaceable, other input variables—specifically, real budgetary expenditures allocated to the judiciary and the total number of unresolved cases—also significantly influence the volume of cases processed by the courts, largely independent of the number of judges. Therefore, when such data is available, incorporating these additional input variables alongside the number of judges in DEA analyses enhances the credibility of the conclusions drawn from such assessments. This inclusion not only enriches the analysis but also affirms its legitimacy.

Table 2 illustrate the fluctuations in the average duration of case processing across various areas of law, derived using formula (23). Notably, there is a consistent annual increase in processing times, despite the relatively stable number of cases. A significant observation pertinent to the analyses presented in this article is the systematic gradation in average case resolution times across different legal categories. Regardless of the year analysed, criminal cases consistently exhibited the shortest processing times, whereas insurance cases demonstrated the longest. In relation to the other legal categories, no discernible pattern emerges, characterized by a regular minority/majority gradation. This means that the actual average processing times for these categories fall between the disposition times recorded for criminal and insurance cases.

If the estimates of disposition times by legal category presented above are accepted as definitive, it becomes feasible to ascertain the proportion of the total time dedicated to resolving all cases within a specific area of law throughout the analysed period as shown in formula (27):

$$s_i = \frac{d_i \cdot n_i}{\sum_{j=1}^{6} d_i \cdot n_i} \tag{27}$$

S_i – fraction of the courts' time devoted to resolving the cases of the *i*-th subject of law (i=1, 2, ...6). By definition we have: ,

d_i – average processing time of cases for *i-th* subject of law, determined according to formula (23),

n_i – number of cases per *i-th* subject of law.

The significant variability in the proportions of total time allocated to resolving cases by legal category underscores temporal fluctuations in both the disposition times and the volume of cases. Table 3 displays the values of all quantities, alongside their respective minimum and maximum values.

Before proceeding to the primary focus of this article, it is essential to highlight that both the estimates of average case processing durations by area of law and the corresponding proportions of these durations in relation to the total time dedicated to court operations represent the actual, historical, and empirical ranges of variation for these metrics over the analysed period, as opposed to hypothetical ranges. This observation is pivotal and will be further developed and empirically operationalized to provide a clear interpretation of the results obtained through DEA in assessing the efficiency of the Polish appellate court system.

Let us conclude this section of the discussion with a brief comparative analysis of several two-dimensional judicial performance indicators that are frequently utilized in studies concerning the efficiency of judicial systems (e.g., CEPEJ 2018). These indicators are listed in Table 4:

a) Judges' work productivity: ratio of the total number of cases resolved to the total number of judges in a given year (i.e. average number of cases resolved by one judge, in thousands).
b) Unit cost of resolving an average judicial case: ratio of total expenditure to total number of cases resolved in a given year (thousands of PLN, real prices of 2015).
c) Average case handling time (calculated according to formula (23) for the total number of cases (in months)).
d) Clearance rate: ratio of the total number of resolved cases to the total number of incoming cases in a given year; in percentage.

Percentage normalization (relative to *Labour productivity* and *Clearance rate*) and reverse percentage normalization (relative to *Overall disposition time* and *Unit Costs*) enable direct comparisons among these measures. First, it is noteworthy that all indicators exhibit high variability, and consequently, a significant degree of unpredictability concerning projections of their future values. Second, the best (and thus, after normalization, the highest) values concerning systemic efficiency for these measures are observed in different years: specifically, for *Labour productivity*, *Unit costs*, *Overall disposition time*, and *Clearance rate*, these years are 2019, 2002, 2009, and 2007, respectively. Third, from the perspective of the entire two-decade period, a concerning trend of deterioration in all indicators can be observed.

Fourth, depending on which indicator is considered, various sub-periods of relatively high and relatively low "efficiency" can be distinguished. If conclusions are based on the indications of the *Clearance rate* and *Disposition time*, a relatively high "efficiency" is noted in the years following Poland's accession to the European Union (2006-2012). In contrast, using *Labour productivity* as a reference leads to entirely opposing conclusions. Similarly, drawing inferences about the degree of "efficiency" based on *Unit costs* favours the beginning of the sample period. However, regardless of the specific indicator chosen, all of them clearly deteriorate in the last years of the analysed period.

Fifth, the most stringent evaluation of the effects of the appellate judiciary's activities is reflected in the *Disposition time*, which has more than tripled in duration in the last years of analysis compared to 2009. In contrast, the least severe evaluation is indicated by the *Clearance rate*, with the lowest recorded value at 90% in 2013 and the highest nearly 111% in 2007. This situation regarding the Clearance Rate is entirely natural, as a sustained ratio of resolved to incoming cases — that is, the definition of the Clearance Rate — remaining below 100% (referring to the non-normalized value) would indicate a steadily worsening crisis in the judiciary, akin to a "slippery slope".

Unfortunately, the situation appears quite alarming in light of the indications from the second measure mentioned above. Although one should refrain from hastily making an overall assessment of the functioning of the judicial system using two-dimensional indicators it is nonetheless evident that from the perspective of the "court client/petitioner", *Disposition time* is undoubtedly a key factor influencing the public perception of the judiciary and, more importantly, directly affecting the functioning of society and the economy.

The variability of conclusions that can be drawn based on two-dimensional indicators highlights the complexity and multifaceted nature of measuring the effectiveness of the judiciary's functioning. There is a risk that the selective choice of one such indicator may be motivated by a desire to prove pre-established theses. Therefore, any attempts to evaluate the functioning of judicial systems should consider a range of metrics. Even then, inferring the effectiveness of the judiciary—whether in its current state or from a historical perspective—will be methodologically burdened, as such metrics do not account for the simultaneous occurrence of numerous factors that undoubtedly influence the system's functioning and its efficiency, whether on the input side (e.g., the number of judges, budget allocations for the judiciary, or system overload due to the volume of incoming cases) or on the output side (aggregating all judicial cases into a single category instead of disaggregating them).

In summary, methodologically sound research on the effectiveness of public institutions, including the judiciary, must employ analytical tools that take these conditions into account. This observation leads us directly to the next section of the article.

4. Results

In the analyses presented in this section, variants of DEA that employ constant returns to scale and maximized outputs were selected. Firstly, investigations that explore the potential influence of the size of court units on their observed efficiency typically focus on cross-sectional data, where differentsized court units are treated as DMUs. Conversely, this study, as highlighted in previous sections, examines the appellate court system as a whole across an extensive 20-year period, rather than individual courts. Secondly, it is important to note that the total number of operating appellate court units exhibit very low variability, with 10 units present from 2002 to 2004 and increasing to 11 thereafter. Finally, the incorporation of a non-discretionary systemic burden variable, specifically caseload, serves to mitigate any potential 'economies of scale' that might arise due to the changes in the volume of cases processed by the appellate courts, should the individual years of analysis be treated as standard DMUs.

The same theoretical and empirical rationale that underpinned the previous analyses employing DEA methods on judicial efficiency, as already discussed in this paper, was also applied to the chosen variant focused on performance maximization. The primary argument for selecting this DEA variant stems from the significant rigidity observed in both the supply of judges and the budgetary provisions governing the financing of the judiciary. Additionally, in light of the rising average duration of court case processing noted in many EU countries, including Poland (e.g., Study on the Functioning of Judicial Systems in EU Member States, 2022), it is imperative that any potential improvement strategies emphasize the maximization of outcomes rather than simply maintaining existing outputs while minimizing incurred expenditures. Moreover, as already noted, the output-oriented DEA framework makes it possible to treat non-discretionary inputs in the same manner as discretionary ones. Furthermore, given that the vast majority of empirical studies on judicial efficiency employ an output-oriented approach, the present analysis may be regarded as a methodological continuation of the current mainstream research.

This paper presents the findings of a study examining the efficiency of appellate courts through the classic CCR method of performance-maximizing DEA, alongside its modification that incorporates the super-efficiency concept and variants that impose restrictions on model parameters and weights. The sequential analysis aims to highlight the inherent limitations of the classic DEA method, which remain unaddressed by the super-efficiency variant. Instead, a viable and substantively sound approach to resolving these issues is to employ constrained DEA. These features, combined with the utilization of time-series data encompassing the entire judicial system and the treatment of this system as a DMU, underscore the scientific originality of this study.

The second column of Table 6 presents the efficiency scores derived from the classic DEA approach. As anticipated, the scores produced by this method demonstrate limited discriminatory capability, a shortcoming primarily attributable to technical rather than substantive factors[2]. To enhance the dis-

2 In practice, the minimum number of effective DMUs obtained using the classic DEA

criminative power of the DEA model, one can employ the concept of super-efficiency. In essence, this approach involves calculating a "new" efficiency score for a specific DMU deemed efficient in the classic DEA model by excluding that DMU from the analysis. Consequently, the efficiency scores may surpass the value of 1 to varying extents, thereby allowing for differentiation among DMUs that are classified as "classically" efficient but are capped at the upper end of the scale. By rescaling these scores to a maximum of "1" for the highest score, a "traditional" interpretation of the results can be achieved.

Table 5 presents the efficiency scores derived from the DEA super-efficiency procedure, including both the original and rescaled values (refer to columns 3 and 4 of Table 5, respectively). At first glance, the results appear quite satisfactory. Despite the limited number of Decision-Making Units, it seems that a commendable level of discrimination among their efficiency scores has been achieved, resulting in a notably diverse empirical set of scores that may misleadingly suggest a more realistic representation of efficiency. In fact, this represents the pinnacle of what many researchers examining judicial efficiency through DEA aspire to attain, with the majority stopping at the classic DEA stage. Why then not interpret the results right now, what would be wrong about it or what is wrong about such interpretation, especially if that's what everyone does?

The adage, "The less people know about how sausages and laws are made, the better they'll sleep at night," attributed to Otto von Bismarck, can be rephrased to reflect the concerns of DEA practitioners: "The less DEA practitioners—and even more so, DEA stakeholders—understand about the parameters and weights derived from their DEA analyses, the more comfortably they rest at night." This observation reveals a troubling reality: it appears that very few individuals regard the actual parameter estimates obtained in DEA models as significant. However, these are not trivial matters; they are essential to establishing trust in the resulting efficiency scores. The stark fact that only one publication referenced in this paper[3] provides explicit, albeit not exhaustive, consideration of parameter estimates underscores the extent

model is equal to the product of the inputs and outputs included in the analysis, as noted by Dyson et al. (2001). Thus, if one were to limit the study to the classical DEA model, given the number of DMUs (20), inputs (3), and outputs (6), it would be devoid of any deeper meaning.

[3] Santos S.P., Amado C.A.F. (2014), *On the need for reform of the Portugese Judicial System – Does Data Envelopment Analysis assessment support it?*, Omega – The International Journal of Management Science, vol. 47, pp. 1–16. However, even this publication refrains from disclosing the actual weights generated in its DEA analysis, suggesting that reliance on relative weights alone may be inadequate to fully address the problem of 0–1 weights highlighted in this paper. This omission indicates that a more rigorous approach may be necessary to ensure the robustness and interpretability of DEA results, particularly regarding the consistency and significance of the derived weights.

of both intentional and unintentional ignorance surrounding this critical issue.

Examining the weights assigned to outputs in the super-efficiency DEA model (Table 6), both in absolute and virtual terms (for want of space the latter not reported), reveals a concerning trend: many weights take on a value of zero and, at times, all but one weight are zero, indicating that only one factor, whether output or input, appears to contribute to the efficiency score of a DMU. Economically interpreting these weights suggests an improbable scenario where factors with zero weights contribute nothing to efficiency, or that a single factor exclusively drives overall efficiency—a conclusion that, though internally consistent within the model, is practically unrealistic. Such outcomes challenge the foundational assumptions of DEA as a reliable tool for assessing efficiency.

DEA requires that every input and output included in the model exert a measurable influence on efficiency—outputs should reflect a comprehensive representation of the DMU's activities. Otherwise, their inclusion is methodologically unjustifiable. This issue becomes even more pronounced when the DMUs are successive years within a given time frame (2002–2021). It defies logic to suggest that in one year, an output fully absorbs the inputs, while in another year, it has no impact at all (as shown in the weights for *LABOUR* in 2002-2004 in Table 6). Similarly, it is implausible to assume that in one period an input accounts for all outputs, yet in another, it contributes nothing (e.g., see weights for *OUTLAYS* in DMU04 and DMU05 in Table 6). These inconsistencies underscore the need for a robust and balanced approach to parameter weighting within DEA models.

This prompts the question of whether such outcomes have any explanatory value. The logical answer appears to be negative, leading to the subsequent question: what can be done to improve interpretability? The primary goal should be to ensure that the estimated weights are meaningful. This objective can be achieved through constrained DEA, wherein the optimization of efficiency scores includes logical and empirically justified constraints on weights, established on substantive grounds, as part of the optimization process.

A straightforward approach might involve assigning a meaningful value to each resolved case based on the time required for its resolution, aligning with the maxim "time is money"—especially pertinent when the precise costs associated with resolving an average case are difficult to ascertain. Additionally, by employing the super-efficiency model alongside constraints on both absolute and virtual weights (on the inputs and outputs side), one can derive consistent and robust efficiency scores for individual decision-making units without the anomalous zero or unitary weight values that undermine interpretability. This structured methodology is explained and implemented step by step in the following sections.

We begin with the approach proposed by Santos and Amado (2014), who utilized empirical data to establish appropriate minority-to-majority ratios for output weights assigned to different case types, according to their average time requirements derived using formula (23) and summarized in Table 2. Following this approach results in a set of relations regarding the value weights assigned to cases, differentiated by subject matter in accordance with their time absorption rates (refer to Table 2)[4]:

ACIVIL / *ACRIME* >= 1
ALABOUR / *ACRIME* >= 1
ACOMMERCE / *ACRIME* >= 1
AINSURANCE / *ACRIME* >= 1
AINSURANCE / *ACIVIL* >= 1
AINSURANCE / *ALABOUR* >= 1
AINSURANCE / *ACOMMERCE* >= 1

While innovative compared to prior studies in this area, this approach fails to address the significant issue of unacceptable output weights, as many weights still default to values of 0 or 1[5]. Notably, Santos and Amado, like most others in the field, omit discussion on this critical matter in their analyses of judicial efficiency. Only by introducing additional constraints—this time on what are known as virtual weights, which clearly indicate the share of total time allocated to cases of specific types relative to the total time spent on all cases—do the resulting output weights begin to hold meaningful interpretive value. These constraints are precisely specified with a help of formula (27) and are shown in Table 7.

Similarly to the minority/majority relationships established for absolute weights, realistic variability ranges for virtual weights assigned to respective case types by legal subject area must also be defined in accordance with logical criteria. These empirically derived values, which provide the most practical and realistic representation of each category, were derived on the basis of Table 3 and are presented in Table 7.

Applying a DEA variant with simultaneous restrictions on both price and virtual weights does, ultimately, lead to appropriate and substantively accurate weights on the output side (though the specifics are not presented here). However, the previously highlighted issues regarding input weights persist. This prompts a reconsideration of the admissibility of these results. To frame the discussion, consider two rhetorical yet insightful questions that underscore the essential role of judges in case resolution:

4 This notation aligns with the conventions established in the *pyDEA* software, where the weight names are intuitively labelled, with the prefix "A(?)" indicating absolute weights. This notation allows for the direct application of the previously discussed constraints on the respective terms in formula (2).

5 Due to space limitations, the relevant outcome tables have not been included.

1) Would judges, tasked with adjudicating all pending cases, undertake their responsibilities without compensation?
2) Would the judicial institution exist at all if there were no cases to be adjudicated?

The answer, self-evidently negative, reinforces that any challenges to the relevance of unconstrained DEA—especially regarding the appropriateness of output weights and the practicality of rankings it produces—are equally pertinent for input weights. Thus, empirical realism in input weights is essential, as this is a prerequisite for the reliability and validity of efficiency scores generated by DEA analyses.

In contrast to outputs, for which both price weights and virtual weights could feasibly be estimated based on actual time absorption, such an approach proves impractical for inputs. Instead, a viable alternative centres on the evidently indispensable role of judges in the case disposition process. As underscored by the earlier questions, judicial functions are contingent on the presence of two other key inputs: caseload and financial outlays. This insight leads to a heuristic approach that leverages the foundational role of judges as the critical input. By integrating this observation, a method can be devised to establish realistic, empirically grounded price weights among the inputs used, thereby positioning judges as the central input in the efficiency analysis (Allen et al., 1997).

The initial step involves establishing the empirical variability range of the ratios of outlays to judges and caseloads to judges. The respective extreme values observed during the analysed period from 2002 to 2021 are as follows:

min (*CASELOAD/JUDGES*) = 0.205
max (*CASELOAD/JUDGES*) = 0.377
min (*OUTLAYS/JUDGES*) = 0.660
max (*OUTLAYS/JUDGES*) = 1.213

Consequently, based on the assumption regarding the primary role of judges in the overall case disposition process, it follows that the price weights attributed to the other two inputs must fall within the intervals defined by the aforementioned extreme points. The absolute weights assigned to the "judges" input will serve as the reference category. An excessively high weight would be detrimental, indicating that the relative price (or significance) of the other inputs exceeds that of judges. Conversely, an excessively low weight would undervalue the empirical significance of these inputs. Thus, the constraints imposed on the price weights for all inputs are as follows:

ACASELOAD / AJUDGES > = 0.205
ACASELOAD / AJUDGE < = 0.377
AOUTLAYS / AJUDGES > = 0.660
AOUTLAYS / AJUDGES < = 1.213

It is only after imposing all the necessary constraints on both price and virtual weights for outputs and inputs that we arrive at the final results, which are devoid of the critical adverse properties that typically remain unacknowledged yet significantly undermine the credibility and interpretability of efficiency scores derived from classic DEA. At this stage, no weight—whether associated with outputs or inputs—assumes a value of "0" or "1" (refer to Table 8). Furthermore, due to the manner in which these constraints were applied, the estimated range of weight variability aligns with the historically observed minimum and maximum intervals throughout the analysed period. Consequently, we assert that the final outcomes of this empirical study are fully reliable and interpretable, facilitating a robust assessment of the actual performance of the Polish regional court system over a 20-year timeframe.

The portrayal of the Polish appellate court system that emerges from the analysis conducted over the past 20 years is rather pessimistic (see the last column of Table 5). The peak of performance has long since passed, and there are no signs of a return to relatively high efficiency levels (above 90%) that were recorded just before Poland's accession to the European Union and shortly thereafter. The technical maximum was precisely in 2004, and only the efficiency levels noted between 2002 and 2005 can be considered satisfactory. In subsequent years, the efficiency of the appellate courts began to deteriorate noticeably, reaching a local minimum in 2010 (with a 35 percentage points decline in efficiency compared to the most efficient year, 2004).

In the years following 2010 and up to 2019, there was a certain strengthening of the system, manifesting as an increase in efficiency by 5 to 15 percentage points compared to 2010. Consequently, the efficiency during these years fluctuated around 80%, still significantly lower than the levels that characterized the early years of the 21st century. Finally, towards the end of the sample period, a significant decline in the functional efficiency of the system occurred once again, with 2020 recording the lowest efficiency throughout the entire analysed period, at a mere 60%.

The year 2004, when Poland became a formal member of the European Union on May 1st, not only marks the most efficient period in the functioning of the Polish appellate court system but also reveals that *DMU04* serves as the sole peer DMU for all years included in the study. In other words, according to the measures used in this research to evaluate the efficiency of the Polish appellate court system, the standards and legislative-organizational-administrative frameworks that were in place in 2004 most effectively promoted the efficiency of this system. Therefore, in seeking answers to potential questions regarding the possibilities for reforming the current system, it would be advisable to conduct comprehensive, multidimensional comparisons between the current state and the operational status of the Polish appellate judiciary at the time of Poland's accession to the European Union. Clearly, not all changes—whether discretionary or external (though this suggestion primarily concerns the former)—that occurred in the Polish judiciary over the years following Poland's accession to the European Union have been effective from the standpoint of systemic efficiency.

Accurate estimation of the efficiencies of a set of decision-making units (DMUs) is one aspect, while the effective identification—and, ideally, quantification—of the factors influencing those efficiencies at various levels is quite another. However, if the first step is executed poorly, it inevitably leads to erroneous conclusions regarding the second stage of analysis. To illustrate the potential pitfalls in this regard, let us consider the outcomes of the simple, unconstrained super-efficiency DEA (SE-DEA) as reference estimates for the efficiencies of the Polish appellate court system presented in this research.

Figure 1 addresses this issue through the use of bar charts and simple linear trend regressions applied to the rescaled efficiency scores. A cursory visual examination of the two charts and their respective trend lines clearly reveals markedly different overall patterns in the dynamics of efficiency scores derived from unconstrained versus constrained SE-DEA models. There is little to no numerical or logical correlation between these outcomes, with the unconstrained SE-DEA exhibiting a more lenient assessment of the historical performance of the Polish appellate court system.

Notably, the scores for individual years not only vary in an erratic manner, but the overarching conclusions drawn about the long-term dynamics of the system also differ significantly. This discrepancy can lead to a fundamentally altered evaluation of the performance of the Polish appellate court in the 21st century. Thus, relying on the unconstrained SE-DEA results may obscure critical insights and misrepresent the actual state of judicial efficiency over time.

The results from the unconstrained super-efficiency DEA (SE-DEA) indicate that, although the system is not consistently efficient, it maintains relatively high efficiency scores exceeding 80% across individual years. This finding is further reinforced by the absence of a discernible trend line suggesting any significant decline in performance. In contrast, the outcomes from the constrained variant of SE-DEA, which serves as our methodologically rigorous reference, present a markedly less favourable assessment of the Polish appellate court system.

Specifically, the efficiency levels recorded in the constrained model are 10-20 percentage points lower than those derived from the unconstrained approach. Moreover, unlike the unconstrained variant, the constrained SE-DEA demonstrates a consistent downward trend in the system's functional performance over time.

In summary, the methodological framework adopted significantly influences the interpretation of judicial efficiency. Investigators who limit their analysis to a few decision-making units using classical SE-DEA—a prevalent practice—may overlook critical insights. Conversely, extending the analysis to incorporate constrained SE-DEA can yield qualitatively different conclusions and, consequently, inform distinct political recommendations. Such differences underscore the importance of rigorous methodological choices in evaluating the efficiency of judicial systems.

To conclude, we will conduct a concise descriptive statistical analysis, supplemented by graphical representation, to explore potential relationships between simple, two-dimensional judicial performance indicators and the efficiency scores derived from the constrained super-efficiency DEA (SE-DEA). For this analysis, all measures will be expressed in normalized percentage terms, utilizing data from the last column of Table 5 (constrained SE-DEA scores) and the last four columns of Table 4, which include labour productivity, unit costs, disposition time, and clearance rate.

Figure 2 illustrates the normalized efficiency scores from the constrained SE-DEA alongside the two-dimensional judicial performance indicators for the years 2002 to 2021. Additionally, Table 9 presents the results of the correlation analysis, including both Pearson and Spearman correlation coefficients. This analysis aims to provide insights into the interrelations between the performance indicators and the derived efficiency scores, contributing to a deeper understanding of judicial efficiency dynamics over the specified period.

Both the bare-eye “analysis” and the formal correlation analysis, utilizing both Pearson’s linear correlation coefficient and Spearman’s rank correlation coefficient, indicate a lack of any relationship between the efficiency indicators obtained through constrained SE-DEA and the clearance rate and disposition time. Regarding labour productivity and unit costs, the conclusions are less straightforward. Although, from a formal-statistical perspective, the null hypothesis of no correlation between these performance indicators and the efficiency coefficients should be rejected at the 5% significance level, an examination of the trajectories of these three categories (see Figure 2) allows for the conclusion that a rather loose relationship existed between the examined measures only in the first half of the analysed sample period—an observation that can explain the significance of the correlation analysis. In contrast, in the second half of the period, there is a complete divergence of these trajectories. Thus, asserting that there is no need to conduct complex, systemic DEA analyses to determine the scale of efficiency/inefficiency in judicial systems—since this objective could equally well be achieved using two-dimensional judicial performance indicators—would not be fully justified.

In summary, it can be definitively stated that two-dimensional judicial performance indicators, whether considered in isolation or in any configuration, cannot replace well-conducted DEA analyses. However, they can facilitate such analyses and significantly complement/enrich them, particularly concerning the category of disposition time.

First, this specific measure has allowed for the imposition of substantively correct constraints on the parameters/weights associated with the outcome variables, thereby addressing a key issue that effectively undermines the credibility and interpretability of results obtained through classical DEA: the presence of zero or unit values in the input variables.

Second, from the perspective of the "court user/client," disposition time is undoubtedly a key factor influencing the social perception of the judiciary and, more importantly, directly affecting the functioning of society and the economy. Thus, even in those years when the system's efficiency is relatively high—when measured by appropriate, multidimensional methods that take into account various inputs and outputs—if there is an accompanying increase in the average duration of court proceedings, it is difficult to assess the functional state of such a system in a unambiguously positive and uncritical manner. This situation is not hypothetical; the issue highlighted emerges from the current analysis (see Figure 2).

5. Conclusion

Disregarding the inherent limitations of classical or super-efficiency unconstrained radial DEA models can result in substantial inaccuracies in the efficiency ratings attributed to individual DMUs. As a consequence, policy recommendations based on these methods, aimed at improving the actual performance of specific DMUs, may prove to be unreliable and potentially misleading. The outcomes of this research show the fundamental shortcomings of classical applications of DEA, raising valid concerns about the reliability and interpretability of the outcomes derived from this approach. To mitigate these challenges, it presents a novel procedure specifically aimed at addressing these limitations in evaluating judicial efficiency, which is next successfully implemented.

Identifying the most historically efficient benchmark years through empirical analysis is merely an initial step toward a more qualitatively oriented investigation aimed at uncovering the factors that contributed to the superior efficiency of certain decision-making units (DMUs) during those years. This endeavour necessitates interdisciplinary expertise and the involvement of various stakeholders, as Data Envelopment Analysis (DEA) is primarily a descriptive method used to estimate relative efficiencies among homogeneous DMUs. While DEA effectively reveals variations in the effectiveness of the individual DMUs analysed over a specified period, it does not provide insights into the underlying reasons for such discrepancies.

A promising starting point for this exploration could be the application of two-stage DEA (2SDEA), as seen in studies by Schneider (2005), Deynelli (2011), Malcarne and Ramello (2015), and Nissi et al. (2019), utilizing the efficiency scores obtained from the current research. However, a more robust research agenda would require panel data analyses that encompass individual courts across different years, integrating both spatial (across various locations) and temporal dimensions. This approach would necessitate the incorporation of micro-level formal and administrative data, as well as survey-based data. Once the results of such comprehensive research are available, it would be possible to differentiate between inefficiency factors stemming from unique characteristics inherent to specific courthouses and those related to systemic or procedural issues that are prevalent across all operational courthouses.

Another potentially valuable extension of the analyses presented in this article would involve applying window analysis or the Malmquist index to assess the dynamics of productivity within the judicial system under consideration. Moreover, by employing panel data and applying the same classification of court cases by domain of law as in the present study, it would be possible—at the level of individual courts of appeal—to attempt to identify and disentangle three distinct classes of determinants of efficiency. First, this would concern system-wide macro-effects that uniformly affect all judicial units and arise from, inter alia, legislative changes, binding procedural rules, guidelines issued by central authorities, and other institutional conditions. Second, one could isolate meso-level effects reflecting the socio-economic heterogeneity of the regions under the jurisdiction of the respective courts of appeal, which may have implications for their caseload and operational performance. Third, further analyses could focus on micro-level factors associated with idiosyncrasies in the internal organisation and operational practices of individual courts. Such analyses would, however, require access to survey-based data, which in turn would necessitate the cooperation and approval of the competent judicial authorities for their collection.

In light of the above, a wide range of potential avenues exists for further developing the research initiated in this article. These may include an in-depth, quantitative assessment of the state of the Polish judiciary from the perspective of its efficiency, as well as the identification and empirical evaluation of factors influencing that efficiency at the macro, meso, and micro levels. The purpose of such research would be not only to estimate the strength and statistical significance of the effects of particular determinants, but also to indicate areas in which institutional intervention could yield the greatest benefits. Given the fundamental importance of an effective judiciary—for the functioning of a modern state, economy, and society—as well as its well-documented performance shortcomings, further empirical studies in this field should be regarded as both justified and socially desirable.

References

Agrell, P.J., Mattson, P., Mansson J. (2019). Impacts on efficiency of merging the Swedish district courts. *Annals of Operations Research*, *288*, 653-679. https://doi.org/10.1007/s10479-019-03304-0.

Allen, R., Athanassopoulos, A., Dyson, R.G., Thanassoulis, E. (1997). Weight restrictions and value judgements in Data Envelopment Analysis: Evolution, development and future directions. *Annals of Operations Research*, *73*, 13-34. https://doi.org/10.1023/A:1018968909638.

Andersen, P., Petersen, N.Ch. (1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science*, vol. 39(10), 1261-1264. http://dx.doi.org/10.1287/mnsc.39.10.1261.

Avikran, N.K. (2015). Productivity in the Service Sector with Data Envelopment Analysis. *SSRN Working Papers.* DOI:10.2139/SSRN.2627576.

Banker, R., Emrouznejad, A., Lopes, A.L.M., Rodrigues de Almeida, M., eds. (2014). Data Envelopment Analysis in the public sector. *Socio-Economic Planning Sciences*, *48*(1), 1–104. https://doi.org/10.1016/j.seps.2013.12.005.

Boardman, A.E., Greenberg, D.H., Vining, A.R., & Weimer, D.L. (2018). Cost–Benefit Analysis: Concepts and Practice (5th ed.). Cambridge University Press.

Buonanno, P., Galizzi, M. (2014). Advocatus, et non Latro? Testing the excess of litigation in the Italian courts of justice. *Review of Law and Economics*, vol. 10(3), pp. 285-322. https://doi.org/10.1515/rle-2014-0022.

Revised Saturn Guidelines for Judicial Time Management (2018). CEPEJ, European Commission for the Efficiency of Justice. https://rm.coe.int/cepej-2018-20-e-cepej-saturn-guidelines-time-management-3rd-revision/16808ff3ee.

Cook, W.D., Tone, K., Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega*, *44*, 1–4. https://doi.org/10.1016/j.omega.2013.09.004.

Cordero-Ferrera, J.M., Pedraja-Chapparo, F., Santin-Gonzalez, D., (2010). Enhancing the inclusion of non-discretionary inputs in DEA. *Journal of the Operational Research Society*, *61*(4), 578–584. https://doi.org/10.1057/jors.2008.189.

Data Envelopment Analysis in Public Sector (2014). *Socio-Economic Planning Sciences*, special issue, *48* (1–2).

Deynelli, F. (2012). Analysis of relationship between efficiency of justice services and salaries of judges with two-stage DEA method. *European Journal of Law and Economics*, *34*, 477–493. https://doi.org/10.1007/s10657-011-9258-3.

Damodaran, A. (2012). Investment Valuation: Tools and Techniques for Determining the Value of Any Asset (3rd ed.). Wiley Finance.

Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S., (2001). Pitfalls and Protocols in DEA. *European Journal of Operational Research*, *132*, 245–259. https://doi.org/10.1016/S0377-2217(00)00149-1.

El Bialy, N., Garcia-Rubio, M.A. (2011). Assessing judicial efficiency of Egyptian first instance courts: A DEA analysis. *MAGKS Joint Discussion Paper Series in Economics*. https://www.econstor.eu/bitstream/10419/56541/1/657907855.pdf.

Falavigna, G., Ippoliti, R., Manello, A., Ramello, G. (2015). Judicial productivity, delay and efficiency: A direction distance function (DFF) approach. *European Journal of Operational Research*, *240*, 592–601. https://doi.org/10.1016/j.ejor.2014.07.014.

Garcia-Rubio, M.A., Rosales-Lopez, V. (2010). Justice & Economy: Evaluating judicial efficiency in Andalusia. *SSRN Working Papers*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1738227.

Gupta, M., Bolia, N.B. (2020). Efficiency measurement of Indian high courts using DEA: A policy perspective. *Journal of Policy Modeling*, *42*, 1372––1393. https://doi.org/10.1016/j.jpolmod.2020.06.002.

Guzowska, M., & Strąk, T. (2010). An examination of the efficiency of Polish public sector entities based on public prosecutor offices. *Operations Research and Decisions*, 20(2), 41–57. IDEAS/RePEc.

Ippoliti, R., Tria, G., (2020). Efficiency of judicial systems: model definitions and output estimation. *Journal of Applied Economics*, *23*(1), 385–408. https://doi.org/10.1080/15140326.2020.1776977.

Kapelko, M. (2024). Evaluating input- and output-specific inefficiency in courts of justice: An empirical study of Polish district courts. *International Transactions in Operational Research*, 32(5), 2767–2797. https://doi.org/10.1111/itor.13503

Krajewski, E., Milewski, R., (201.8). Podstawy ekonomii [Fundamentals of economics], *PWN.*

Kittelsen, S.A.C., Forsund, F.R. (1992). Efficiency analysis of Norwegian district courts. *Journal of Productivity Analysis*, *3*(3), 277–306. https://doi.org/10.1007/BF00158357.

Lewin, A.Y., Morey, R.C., Cook, T.J. (1982). Evaluating the administrative efficiency of courts. *Omega*, *10*(4), 401–411. https://doi.org/10.1016/0305-0483(82)90019-6.

Maital, S., Vaninsky, A. (1999). Data envelopment analysis with a single DMU: A graphic projected-gradient approach. *European Journal of Operational Research*, *115*(3), 518-528. https://doi.org/10.1016/S0377-2217(98)00178-7.

Major, W. (2015). Data Envelopment Analysis as an instrument for measuring the efficiency of courts. *Operations Research and Decisions*, 25(4), 19–34. https://doi.org/10.5277/ord150402.

Malcarne, A., Ramello, G.B. (2015). Judicial Independence, Judges' Incentives and Efficiency. *Review of Law and Economics, 11*(2) 149–169. https://doi.org/10.1515/rle-2015-0024.
Mattsson, P., Tidana, C. (2019). Potential efficiency effects of merging the Swedish district courts. *Socio-Economic Planning Sciences, 67*, 58–68. https://doi.org/10.1016/j.seps.2018.09.002.
Muniz, M., Paradi, J., Ruggiero, J., Yang Zijiang (2006). Evaluating Alternative DEA models used to control for non-discretionary inputs. *Computers and Operations Research, 33,* 1173–1183. https://doi.org/10.1016/j.cor.2004.09.007.
Nissi, E., Giacalone, M., Cusatelli, C. (2019). The Efficiency of Italian Judicial System: A Two Stage Data Envelopment Analysis Approach. *Social Indicators Research, 146*, 395–407. https://doi.org/10.1007/s11205-018-1892-5.
Ostaszewski, P. (ed.) (2020). Efektywność sądownictwa powszechnego – oceny i analizy [Efficiency of the common courts – assessments and analyses], *Wydawnictwo Instytutu Wymiaru Sprawiedliwości*, Warszawa.
Ostrom, B.J., Hanson, R.A. (2003). Efficiency, timelines and quality: A new perspective from nine state criminal trial courts. *National Institute of Justice and the State Justice Institute.* https://www.ojp.gov/pdffiles1/nij/178403-1.pdf.
Panwar, A., Olfati, M., Pant, M., Snasel, V. (2022). A Review on the 40 Years of Existence of Data Envelopment Analysis Models: Historic Development and Current Trends. *Archives of Computational Methods in Engineering, 29,* 5397–5426. https://doi.org/10.1007/s11831-022-09770-3.
Pedraja, F., Salinas, J. (1996). An assessment of the efficiency of Spanish courts using DEA. *Applied Economics, 28*, 1391–1403. https://doi.org/10.1080/000368496327651.
Peyrache, A., Zago, A. (2016). Large courts, small justice! The inefficiency of the optimal structure of the Italian justice sector. *Omega, 64*, 42–56. https://doi.org/10.1016/j.omega.2015.11.002.
Podinovski, V.V. (2016). Optimal weights in DEA models with weight restrictions. *European Journal of Operational Research, 254,* 916–924. https://doi.org/10.1016/j.ejor.2016.04.035.
Ruggiero, J. (1997). Non-discretionary inputs in data envelopment analysis. *European Journal of Operational Research, 16*, 461–469. https://doi.org/10.1016/S0377-2217(97)00306-8.
Santos, S.P., Amado, C.A.F. (2014). On the need for reform of the Portuguese Judicial System – Does Data Envelopment Analysis assessment support it? *Omega – The International Journal of Management Science, 47,* 1–16. https://doi.org/10.1016/j.omega.2014.02.007.
Schneider, M.R. (2005). Judicial career incentives and court performance: An empirical study of the German labour courts of appeal. *European Journal*

of Law and Economics, *20*(2), 127–144. https://doi.org/10.1007/s10657-005-1733-2.

Siemaszko, A., Ostaszewski, P. (2013). Efektywność kosztowa sądownictwa powszechnego [Cost efficiency of the common courts]. *Instytut Wymiary Sprawiedliwości*, 1–57, https://iws.gov.pl/wp-content/uploads/ 2018/08/ Efektywno% C5%9B%C4%87-s%C4%85downictwa-powszechnego--oprac-12-1.pdf

Study on the function.ing of judicial systems in the EU Member States (2022). European Commission for Efficiency of Justice (CEPEJ). https://commission.europa.eu/system/files/2022-05/part_1_-_eu_scoreboard_-_indicators_-_deliverable_0.pdf.

Świtłyk, M., Sompolska-Rzechuła, A., & Oesterreich, M. (2025). Efektywność techniczna sądów okręgowych [Technical efficiency of regional courts]. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 69(3), 133––145. https://doi.org/10.15611/pn.2025.3.10.

Tulkens, H. (1993). On FDH efficiency analysis: Some methodological issues and application to retail banking, courts and urban transit. *Journal of Productivity Analysis*, *4*(1/2), 183–210. https://doi.org/10.1007/978-0-387-25534-7_17.

Voigt, S. (2016). Determinants of judicial efficiency: a survey. *European Journal of Law and Economics*, 42, 183–208. https://doi.org/10.1007/s10657-016-9531-6.

Voigt, S., El Bialy, N. (2016). Identifying the determinants of judicial performance: Taxpayers' money well spent? *European Journal of Law and Economics*, 41, 283–319. https://doi.org/10.1007/s10657-014-9474-8.

Yeung, L., Azevedo, P. (2011). Measuring efficiency of Brazilian courts with data envelopment analysis (DEA). *IMA Journal of Management Mathematics*, *22*(4), 343–356. https://doi.org/10.1093/imaman/dpr002.

Yeung, L. (2018). Measuring Efficiency of Brazilian Courts: One Decade Later. *SSRN Working Papers*. https://ssrn.com/abstract=3200588.

Acknowledgements

Author contributions: author has given an approval to the final version of the article. Author's total contribution to the manuscript: Waldemar Florczak (100%)

Appendix

Table 1. Input and output values used in DEA of appellate court efficiency

DMU/YEAR	INDICATORS				NORMALIZED INDICATORS			
	Labour productivity	Unit costs	Overall disposition time	Clearance rate	Labour productivity	Unit costs	Overall disposition time	Clearance rate
DMU02	0.244	2.709	4.294	93.438	88.478	100.000	39.963	84.379
DMU03	0.246	2.804	4.185	97.638	89.351	96.634	41.003	88.172
DMU04	0.247	2.834	4.109	98.330	89.689	95.612	41.761	88.798
DMU05	0.235	2.996	4.120	99.534	85.282	90.422	41.657	89.885
DMU06	0.221	3.436	3.339	108.117	80.227	78.841	51.392	97.636
DMU07	0.210	3.809	2.352	110.735	76.380	71.122	72.967	100.000
DMU08	0.186	4.109	2.144	103.129	67.585	65.940	80.051	93.131
DMU09	0.185	3.981	1.716	103.269	67.035	68.049	100.000	93.258
DMU10	0.177	4.378	1.915	99.090	64.260	61.886	89.615	89.484
DMU11	0.188	4.076	2.013	98.231	68.394	66.463	85.266	88.708
DMU12	0.201	3.713	2.464	95.202	72.849	72.963	69.653	85.973
DMU13	0.217	3.512	3.588	90.210	78.854	77.138	47.835	81.465
DMU14	0.233	3.396	3.469	99.052	84.524	79.782	49.477	89.449
DMU15	0.228	3.576	3.803	97.825	82.736	75.750	45.128	88.341
DMU16	0.238	3.711	3.761	100.348	86.396	72.997	45.627	90.620
DMU17	0.255	3.615	3.419	101.838	92.490	74.945	50.188	91.965
DMU18	0.265	3.893	4.055	95.840	96.222	69.584	42.326	86.549
DMU19	0.275	3.872	4.407	94.825	100.000	69.962	38.944	85.632
DMU20	0.228	5.319	6.302	94.879	82.787	50.933	27.230	85.681
DMU21	0.238	4.765	5.546	99.506	86.389	56.851	30.945	89.860

Inputs: *Jugdes* – number of judges in a given year; *Outlays* – expenditure on appellate courts, million PLN, constant 2015 prices; *Caseload* – number of cases to be handled in appellate courts in a given year, in thousands. **Outputs:** *Crime*: number of criminal law cases handled, in thousands; *Civil*: number of civil law cases handled, in thousands; *Family*: number of family law cases handled, in thousands; *Labour* - number of labour law cases handled, in thousands; *Insurance* - number of insurance law cases handled, in thousands; *Commerce* - number of business law cases handled, in thousands; **Aggregate measures**: *Resolved* – number of all cases handled in appellate courts, in thousands; *Incoming* – number of all new cases incoming to the court in a given year, in thousands; *Pending* – number of all unsolved cases in a given year, in thousands

Source: Own elaboration based on the data described in section *Methods and data*

Table 2. Average disposition time in Polish appellate courts system by subject of law in the years 2002-2021

Subject of law	YEAR/DMU																				MIN	MAX
	DMU02	*DMU03*	*DMU04*	*DMU05*	*DMU06*	*DMU07*	*DMU08*	*DMU09*	*DMU10*	*DMU11*	*DMU12*	*DMU13*	*DMU14*	*DMU15*	*DMU16*	*DMU17*	*DMU18*	*DMU19*	*DMU20*	*DMU21*		
CRIME	1.079	1.059	0.927	0.952	0.785	0.777	0.833	0.787	0.973	0.915	0.853	0.892	0.980	0.942	0.903	0.947	1.176	1.296	1.825	1.648	0.777	1.825
CIVIL	2.373	2.276	2.063	2.135	1.643	1.389	1.338	1.396	1.474	1.444	1.812	2.264	2.893	3.426	3.915	4.179	4.049	4.528	5.843	6.742	1.338	6.742
LABOUR	1.714	1.800	1.800	2.182	2.824	2.571	1.000	2.182	2.400	1.795	2.513	2.582	3.532	3.611	4.572	4.608	3.050	3.618	5.502	5.702	1.000	5.702
INSURANCE	10.737	9.993	9.337	9.095	7.758	4.895	4.489	3.160	3.904	4.729	6.056	9.549	6.277	7.903	7.258	8.051	8.529	7.714	15.643	10.204	3.160	15.643
COMMERCE	2.545	2.051	2.146	2.164	1.404	1.412	1.600	1.524	1.714	1.592	2.127	3.515	5.776	4.800	3.592	4.321	5.051	6.293	8.493	7.123	1.404	8.493

Source: Own elaboration based on the data in table 1 and formula (23)

Table 3. Shares of total time spent on resolving appellate courts cases by subject of law

Subject of law	YEAR/DMU																				MIN	MAX
	DMU02	*DMU03*	*DMU04*	*DMU05*	*DMU06*	*DMU07*	*DMU08*	*DMU09*	*DMU10*	*DMU11*	*DMU12*	*DMU13*	*DMU14*	*DMU15*	*DMU16*	*DMU17*	*DMU18*	*DMU19*	*DMU20*	*DMU21*		
CRIME	0.075	0.069	0.055	0.054	0.061	0.097	0.126	0.153	0.173	0.155	0.118	0.079	0.085	0.075	0.069	0.070	0.094	0.093	0.119	0.120	0.054	0.173
CIVIL	0.162	0.159	0.152	0.158	0.155	0.177	0.182	0.244	0.259	0.250	0.258	0.223	0.283	0.330	0.389	0.400	0.388	0.366	0.288	0.330	0.152	0.400
LABOUR	0.009	0.009	0.008	0.011	0.014	0.016	0.006	0.015	0.014	0.010	0.010	0.007	0.009	0.008	0.010	0.009	0.008	0.007	0.006	0.007	0.006	0.016
INSURANCE	0.688	0.705	0.725	0.717	0.730	0.667	0.635	0.527	0.489	0.525	0.550	0.600	0.457	0.457	0.427	0.379	0.408	0.428	0.486	0.446	0.379	0.730
COMMERCE	0.065	0.058	0.061	0.060	0.040	0.043	0.050	0.061	0.065	0.060	0.064	0.090	0.167	0.130	0.105	0.141	0.103	0.105	0.100	0.097	0.040	0.167

Source: Own elaboration based on the data in table 1 and formula (27)

Table 4. Two-dimensional "efficiency" measures for Polish appellate court system in the years 2002-2021 along with their percentage (reversed percentage) normalization regarding the "most effective" values

DMU/YEAR	INDICATORS				NORMALIZED INDICATORS			
	Labour productivity	Unit costs	Overall disposition time	Clearance rate	Labour productivity	Unit costs	Overall disposition time	Clearance rate
DMU02	0.244	2.709	4.294	93.438	88.478	100.000	39.963	84.379
DMU03	0.246	2.804	4.185	97.638	89.351	96.634	41.003	88.172
DMU04	0.247	2.834	4.109	98.330	89.689	95.612	41.761	88.798
DMU05	0.235	2.996	4.120	99.534	85.282	90.422	41.657	89.885
DMU06	0.221	3.436	3.339	108.117	80.227	78.841	51.392	97.636
DMU07	0.210	3.809	2.352	110.735	76.380	71.122	72.967	100.000
DMU08	0.186	4.109	2.144	103.129	67.585	65.940	80.051	93.131
DMU09	0.185	3.981	1.716	103.269	67.035	68.049	100.000	93.258
DMU10	0.177	4.378	1.915	99.090	64.260	61.886	89.615	89.484
DMU11	0.188	4.076	2.013	98.231	68.394	66.463	85.266	88.708
DMU12	0.201	3.713	2.464	95.202	72.849	72.963	69.653	85.973
DMU13	0.217	3.512	3.588	90.210	78.854	77.138	47.835	81.465
DMU14	0.233	3.396	3.469	99.052	84.524	79.782	49.477	89.449
DMU15	0.228	3.576	3.803	97.825	82.736	75.750	45.128	88.341
DMU16	0.238	3.711	3.761	100.348	86.396	72.997	45.627	90.620
DMU17	0.255	3.615	3.419	101.838	92.490	74.945	50.188	91.965
DMU18	0.265	3.893	4.055	95.840	96.222	69.584	42.326	86.549
DMU19	0.275	3.872	4.407	94.825	100.000	69.962	38.944	85.632
DMU20	0.228	5.319	6.302	94.879	82.787	50.933	27.230	85.681
DMU21	0.238	4.765	5.546	99.506	86.389	56.851	30.945	89.860

Source: Own elaboration on the basis of data reported in table 1

Table 5. Efficiency scores obtained by classical DEA, super-efficiency DEA, and by constrained super-efficiency DEA

DMU/YEAR	Classical DEA	Super-efficiency DEA	Super-efficiency DEA (rescaled)	Constrained super-efficiency DEA	Constrained super-efficiency DEA (rescaled)
DMU02	1.0000	1.2017	1.0000	0.9732	0.9335
DMU03	1.0000	1.0332	0.8598	0.9858	0.9457
DMU04	1.0000	1.0463	0.8707	1.0425	1.0000
DMU05	1.0000	1.0415	0.8667	0.9657	0.9264
DMU06	1.0000	1.0009	0.8329	0.8878	0.8516
DMU07	1.0000	1.0683	0.8890	0.8341	0.8001
DMU08	1.0000	1.0071	0.8381	0.7368	0.7068
DMU09	1.0000	1.0445	0.8692	0.7354	0.7054
DMU10	1.0000	1.0042	0.8357	0.6729	0.6455
DMU11	1.0000	1.0394	0.8650	0.7012	0.6726
DMU12	1.0000	1.0467	0.8711	0.7503	0.7197
DMU13	0.9950	0.9950	0.8280	0.8047	0.7719
DMU14	1.0000	1.0035	0.8351	0.8785	0.8427
DMU15	0.9902	0.9902	0.8240	0.8183	0.7850
DMU16	0.9981	0.9981	0.8306	0.8336	0.7996
DMU17	1.0000	1.1963	0.9956	0.8147	0.7815
DMU18	1.0000	1.0758	0.8952	0.8261	0.7924
DMU19	1.0000	1.1036	0.9184	0.9171	0.8797
DMU20	0.9976	0.9976	0.8302	0.6376	0.6116
DMU21	1.0000	1.0789	0.8978	0.7201	0.6908

Source: Own computations using data from table 1 and *pydea* software

Table 6. Efficiency scores and absolute weights obtained in the super-efficiency DEA

DMU	INPUTS			OUTPUTS					
	Efficiency	*JUDGES*	*OUTLAYS*	*CASELOAD*	*CRIME*	*CIVIL*	*LABOUR*	*INSURANCE*	*COMMERCE*
DMU02	1.20165445	0	0.00342443	0	0	0	0.47619048	0	0
DMU03	1.03317391	0.0011711	0.00178322	0	0.0049284	0	0	0	0.07401261
DMU04	1.04631632	0	0.00318202	0	0	0	0	0.0232382	0.01744301
DMU05	1.04151627	0	0	0.00668155	0	0	0.17660635	0.01539828	0.00632342
DMU06	1.0008754	0.00013026	0.00067279	0.00549412	0	0.02106154	0.01056189	0.01021377	0
DMU07	1.06831453	0.00037647	0	0.00675124	0	0	0	0.03289474	0
DMU08	1.00710013	0	0	0.00946568	0.01360623	0	0.2327803	0.01217793	0
DMU09	1.04451548	0	0.00049195	0.00743765	0.02676973	0	0	0.0070047	0
DMU10	1.00422348	0	0	0.00985935	0.03365019	0	0	0.0001892	0
DMU11	1.03938282	0	0.00044616	0.00737758	0.00040259	0.02556763	0.15420645	0	0
DMU12	1.04674801	0	0.00147585	0.00339972	0.01964336	0.00943961	0	0	0
DMU13	0.99495884	0	0.00147215	0.00321508	0.00490988	0.01834992	0	0.00549203	0
DMU14	1.00349273	0.00050405	0.00050358	0.00371466	0.00952239	0.00956086	0	0.00884951	0.00363945
DMU15	0.99022274	0	0.00138715	0.00302945	0.0046264	0.01729044	0	0.00517493	0
DMU16	0.99805907	0.0001132	0.00060831	0.00468057	0	0.01817272	0	0.00935118	0
DMU17	1.19631575	0.00182511	0	0	0	0.00747862	0	0	0.04419966
DMU18	1.07576812	0.00177397	0	0.00115073	0	0.01929521	0.13596099	0	0
DMU19	1.1035576	0.00204551	0	0	0.00962054	0.00479668	0	0.01404903	0
DMU20	0.99761241	0.00064286	0	0.00506232	0.02081856	0.00623842	0	0	0
DMU21	1.07887073	0.00081228	0	0.00346868	0.01699333	0	0	0.00944341	0

Source: Own computations using data from table 1 and *pydea* software

Table 7. Constraints imposed on virtual weights

Virtual weights (minima)	Virtual weights (maxima)
VCRIME >= 0.054	*VCRIME* <= 0.173
VCIVIL >= 0.152	*VCRIME* <= 0.400
VLABOUR >= 0.006	*VLABOUR* <= 0.016
VINSURANCE >= 0.379	*VSOCIVIL* <= 0.730
VCOMMERCE >= 0.040	*VCOMMERCE* <= 0.167

Source: Own computations using data from table 3

Table 8. Efficiency scores and virtual weights obtained in the super-efficiency DEA variant with minority/majority relationships imposed on output and input weights as well as with constraints put on virtual output weights

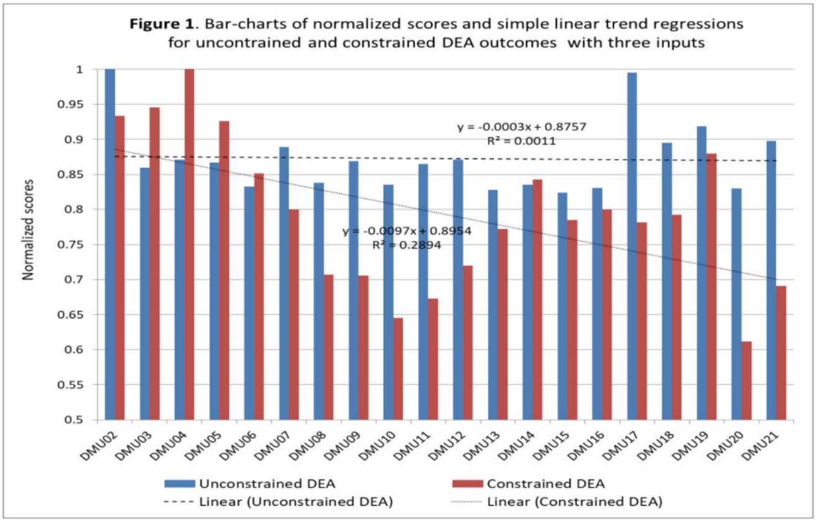
DMU	INPUTS			OUTPUTS					
	Efficiency	*OUTLAYS*	*JUDGES*	*CASELOAD*	*INSURANCE*	*LABOUR*	*CRIME*	*CIVIL*	*COMMERCE*
DMU02	0.97315407	0.44043645	0.54984293	0.03730714	0.37900001	0.016	0.173	0.36785394	0.06414607
DMU03	0.98582205	0.44571647	0.53245267	0.03621273	0.37900001	0.016	0.173	0.2806587	0.1513413
DMU04	1.0424726	0.42489613	0.50031683	0.03404487	0.64760519	0.006	0.054	0.152	0.14039481
DMU05	0.96574548	0.30377426	0.65384474	0.07785049	0.73	0.016	0.054	0.16	0.04
DMU06	0.88780861	0.3509147	0.70007152	0.07538268	0.44212767	0.00723404	0.1106383	0.40000002	0.04
DMU07	0.83406952	0.39057988	0.73831217	0.0700488	0.62058824	0.00823529	0.16352941	0.16764706	0.04
DMU08	0.73684436	0.43158952	0.85482523	0.07072391	0.39725235	0.00720833	0.173	0.3825393	0.04
DMU09	0.73538037	0.42170927	0.86903882	0.06909244	0.38092979	0.00623934	0.173	0.39983089	0.04
DMU10	0.6728777	0.4784261	0.93533997	0.07238809	0.381	0.006	0.173	0.4	0.04
DMU11	0.70117447	0.45470195	0.89699867	0.07447798	0.381	0.006	0.173	0.40000001	0.04
DMU12	0.75030632	0.41408532	0.84192237	0.07678129	0.38099999	0.006	0.173	0.40000001	0.04
DMU13	0.80467773	0.38869952	0.77191458	0.08211941	0.379	0.006	0.16682664	0.4	0.04817336
DMU14	0.87849223	0.36333064	0.69620008	0.07878327	0.37900001	0.006	0.16119038	0.39999998	0.05380962
DMU15	0.81832028	0.39818342	0.74008094	0.08375107	0.37900001	0.006	0.16030042	0.4	0.05469958
DMU16	0.83355978	0.41121896	0.70532627	0.08312879	0.37900001	0.006	0.15525322	0.4	0.05974678
DMU17	0.81467076	0.43099709	0.70898002	0.08751265	0.379	0.006	0.14942641	0.4	0.06557359
DMU18	0.82606523	0.45435504	0.66701989	0.08918324	0.37899999	0.006	0.17158721	0.4	0.04341279
DMU19	0.91705654	0.41589083	0.59067795	0.08387656	0.38070466	0.006	0.173	0.4	0.04029534
DMU20	0.63761167	0.65003158	0.8118533	0.10646787	0.379	0.006	0.173	0.39999999	0.042
DMU21	0.72012572	0.55298118	0.73874485	0.09692037	0.381	0.006	0.173	0.40000001	0.04

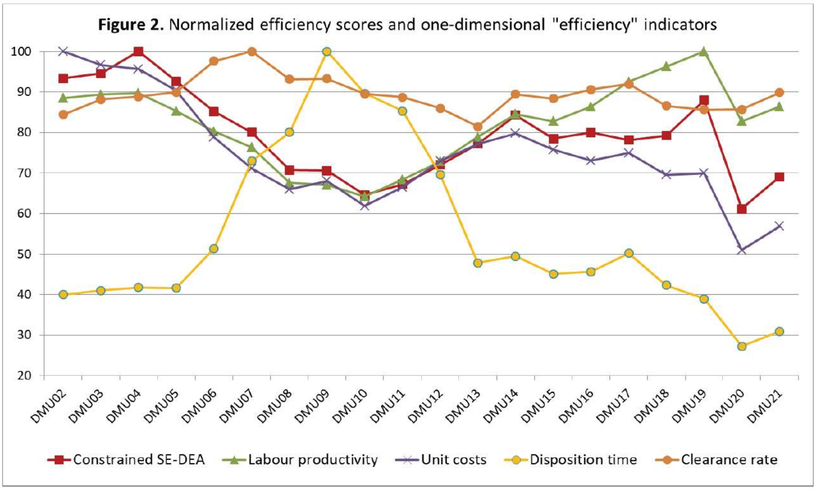
Source: Own computations using data from table 1 and *pydea* software

Table 9. Pearson's and Spearman's rank correlation coefficients between normalized constrained SE-DEA efficiency scores and normalized two-dimensional judicial performance indicators

Correlation test	Productivity	Unit costs	Disposition time	Clearance rate
Pearson's correlation coefficient (5% critical value = 0.4438)	0.6106	0.9070	-0.4442	-0.0309
Spearman's rank correlation (in parentheses *p*-values are given)	0.6150 (0.0073)	0.8615 (0.0002)	0.3774 (0.0999)	-0.0586 (0.7982)

Source: own elaboration based on research outcomes





Source: own elaboration based on research outcomes