

Performance indicators and evaluation for judges and courts

*Dr. Pim Albers
Special advisor of the CEPEJ¹
Council of Europe*

The evaluation of the performance of courts and judges can be carried out at different levels: at a national level, the level of a court, a department within a court and at the level of the individual judges. It is also dependent on who is responsible for the evaluation. In certain countries this may be the task of a ministry of Justice or an inspection organ (of a ministry), whilst in other countries the evaluation is conducted by a High Council for the judiciary, a Supreme Court or another Higher instance court. Mostly the evaluation at the national level might be limited to the monitoring of the performances of courts. Due to (legal) measures related to protection of the independent position of a judge, national executive authorities often do not intervene at the level of the individual judge (however this can be the case for a High Council for the Judiciary or a High Court). The evaluation of the performance of individual judges is primarily the responsibility of a superior judge at a level of a department within a court or the court president.

The debate about court performance is always a delicate one; since judges do not like to compare their judicial work with for example administrative work that is carried out in a department or another governmental agency. An 'overestimation' of the attention to the subject of 'performance' may lead to heated discussions on the 'I'-word (the independent position of a judge). Is the protection of the independent position of a judge a real valid argument for not looking at the subject? No, since the independence of a judge is connected with freedom of

¹ For further information on the European Commission for the Efficiency of Justice (CEPEJ) see the website: www.coe.int/cepej

decision making and non interference of the executive (and legislative) power in the work of a judge. This means not that a judge is not *accountable* for the work he or she is delivering. Someway or somehow, the performance of a judge and a court must be evaluated, because courts are financed by public means and play an important role in the protection of the rule of law in countries and the day-to-day life of citizens and companies.

Court as a system

When you take a closer look at a court, the court can be described by using a simple 'system-model' which differentiates between: input, throughput and output. The input part of a court can be distinguished between resources and cases. The resources of a court are: personnel (judges and court staff), material (court buildings, office equipment, etc.) and financial resources (the budget of a court). Influencing the level of these three types of resources can have an impact on the productivity of a court. I.e. a lack of resources (in terms of judges, staff, equipment, and budget) can lead to an increase of the length of proceedings and a growing backlog of cases.

As already has been said before incoming 'cases' belongs to the input-part as well. A high influx of cases (for example as a result of a increased level of crimes in a country, a growing economic progress or a decline of the economic climate) – with a same level of court resources – can lead to an increase of the length of proceedings and more court files on the 'bookshelf' too.

The throughput of a court is the process where the incoming cases are treated by judges and court staff, resulting at a decision of a judge (the output). One of the indicators to measure the throughput of courts is – logically – the length of proceedings and the backlog of cases.

Why is explaining the court in terms of a system model important? This is necessary just to draw attention to the fact that the performance of individual

judges can be influenced by external factors. Changes in: society, the budget of the State, legislation, etc. can lead to a fluctuation of cases received by the courts and thereby also may lead to a fluctuation in the workload of cases that can be handled by judges. For example an introduction of a new civil code, can have a high impact on the productivity of judges if this legal code reduces the number of legal steps in a judicial proceeding.

It must be underlined though that some factors are out of reach of the individual judge. An integral approach of looking at the subject of the performance of judges and courts is necessary.

How can this integral approach be realized and which examples of performance indicators can be used to evaluate judges and courts? In this article I will draw the attention to six 'efficiency' performance indicators:

1. the caseload per judge;
2. (labour) productivity;
3. the duration of proceedings;
4. cost per case;
5. clearance rate;
6. the budget of courts.

One of the main critics of the judiciary is that – mostly – the executive power is focusing on the 'productivity' and 'efficiency' of judges and courts. Legal and judicial quality is 'under-estimated' and under-valued. At each evaluation of judges and courts a balance should be realized between 'productivity/efficiency' and 'quality'. That is also one of the reasons that I will introduce at the end of the article 'quality models'. The history of these models will be described as well as some examples of the application of quality models in the judiciary and the role of the client as one of the assessors of the quality and services delivered by the courts and the judges.

The caseload per judge

The most basic methodology to calculate the caseload per judge is the division of the total number of incoming cases (and pending cases) to the total number of judges (at a level of a court or a department of a court). In a situation of more incoming cases (or a growing number of pending cases) the caseload per judge will increase. To prevent a future increase of backlog of cases and a longer duration of court proceedings, this information can be useful for a court manager to ask for a higher court budget.

In a more sophisticated model an additional precise distinction is made between civil, commercial, criminal and administrative law cases in terms of defining specific case-categories (for example in the Netherlands a workload model is used with 48 different types of case categories). For each case-category an estimation of the time that is needed for a judge or the court staff to prepare and finalize case (in minutes) is given. The level of complexity of a case is indicated too. This is necessary because a simple description of the caseload per judge only at a general level (for example at the total level of civil cases) is too broad and cannot be used to identify judges with a high caseload or a low caseload. Some judges with a high caseload might be working on simple cases, whilst judges with a low average caseload might be working on many complex court cases. The use of a more detailed workload model is necessary.

The caseload per judge must be seen as a 'raw' indicator and it can be used by the courts to estimate how much work they will receive given a certain number of incoming and pending cases and how many personnel resources are necessary to treat the cases in due time.

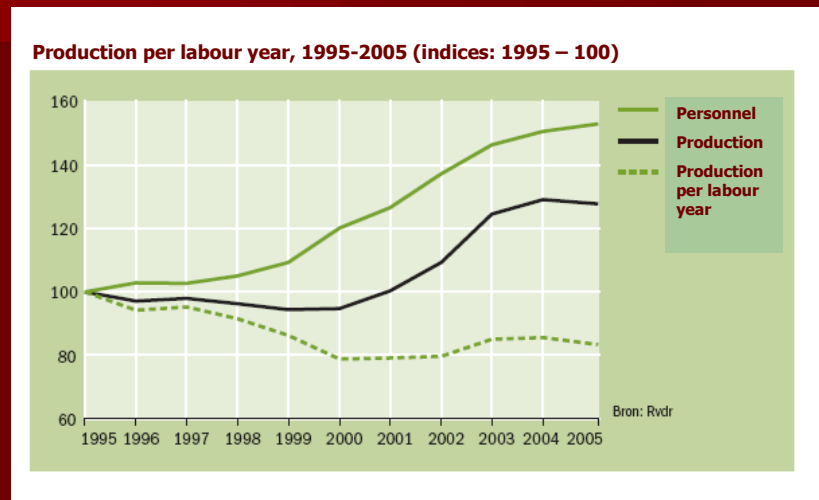
(Labour) productivity

Labour productivity is the most widely used measure and is usually calculated by dividing the total output (of cases) by the number of personnel or the number of hours worked. The total factor productivity attempts to measure the overall productivity of the inputs used by an organization.

Labour productivity is a performance indicator to receive information about the production (in terms of judicial decisions for example) delivered by the judges (and or court staff). In case for example 'efficiency' measures have been introduced in a court, this can have a positive effect on the labour productivity (more output with the same personnel resources). In the following graph the labour productivity of the period 1995 – 2005 is presented of the Dutch judiciary (at a national level). The dotted line expressed the variation in labour productivity over the years. Over the years 2000 – 2005 extra judges (and court staff) were recruited. However, in the period 2000 – 2002 no increase of the productivity was visible. After the year 2002 the labour productivity of the Dutch courts was increased (see figure 1).

Figure 1

Labour productivity at national level: an example



In the study conducted by the Dutch Council² for the Judiciary a similar analysis has been carried out at the level of the district courts. In the period 2002 – 2004 the labour productivity of the courts has been increased and after that period there was a decrease of productivity.

The level of labour productivity – in addition to inefficiencies related to for example complicated court proceedings – may have a correlation with the size of a court. The Study in the Netherlands showed that the labour productivity in 2002 was the lowest at the largest district courts. The mid-sized courts (in terms of number of judges) were the most productive ones. The authors of the study suggested that there is a relation between the size of courts and productivity. Too

² Social Cultural Plan Office and the Council for the Judiciary (April 2007), Administration of Justice: productivity in perspective, The Hague.

large sized courts can have a negative scale effect in terms of efficiency and productivity.

Instead of focusing on the labour productivity courts might use another performance indicator, namely the total productivity of a court. As has been said earlier the total productivity is useful to analyze the ratio between the output of a organization and all the allocated means (personnel *and* non-personnel resources, for example: ICT, office equipment and court buildings). More investments (for example in computers) may lead at the beginning to a decrease in the total productivity, however at the long run it can have a positive effect on the total productivity of a court.

One of the disadvantages of the use of labour productivity as a performance indicator in the courts is that it does not take into account the 'quality' aspects that are related to the work in the courts. The data concerning labour productivity can be improved by making use of a detailed definition model of the (weighed) *output* of courts. In the Netherlands for this purpose the Lamicie workload model is used (which differentiates between 48 case categories). Some cases receives a lower weight in the total contribution of the output of courts, because lesser time of a judge or a court clerk is spend on the preparation and the finalization of a judicial decision (for example a landlord and tenant case costs' lesser time compared to the preparation of a employer dismissal case). By making use of this weighed workload model, combined with the information of the available personnel resources a more reliable view can be given on the issue of labour productivity. Despite this refinement, the level of complexity of a case is not included in the quantitative analysis of the labour productivity.

Length of proceedings (and productivity)

The Italian research centre IRSIG-CNR has carried out a study where the indicators of the judges' performance were defined as: the duration of a trial, the probability of a disposition in a given time and the average unexpected delay

between the actual and the announced date of a hearing³. One of the main arguments for the excessive duration of trials in Italy is the lack of resources. However the Italian researchers found that the average length of proceedings significantly varies between judges. They have indicated also that if all the Italian judges worked as hard as the first quartile of the selected judges in their research project the median duration of the trials would decrease by 17 to 42 percent.

One of the conclusions that can be derived from this exercise is that length of proceedings is an important indicator for measurement of court performance. However, at the interpretation of the figures it is necessary not only to look at the resources available, but also at the working methods of judges, their expertise and methods of positive incentives to reward productive judges. I.e. productive judges are judges that have produced a high number of judicial decision in a relative short time period.

The importance of measurement of length of proceedings (and backlog of cases) for evaluating court performance is acknowledged by the CEPEJ of the Council of Europe. A special task-force has been introduced for producing information on the issue of length of proceedings and backlog of cases. One of the products of this taskforce was the drafting of a time-management checklist. It is a practical tool for courts, for evaluating their level of attention on the issue of reducing backlog of cases and length of proceedings⁴.

Cost per case

Another orientation on evaluating the performance of the courts in terms of 'efficiency' is the 'cost per case' approach. The US National Center for State Court stated that "efficiency within a context of cases resolution means that resources are used in their most productive fashion" (see NCSC Courtool on

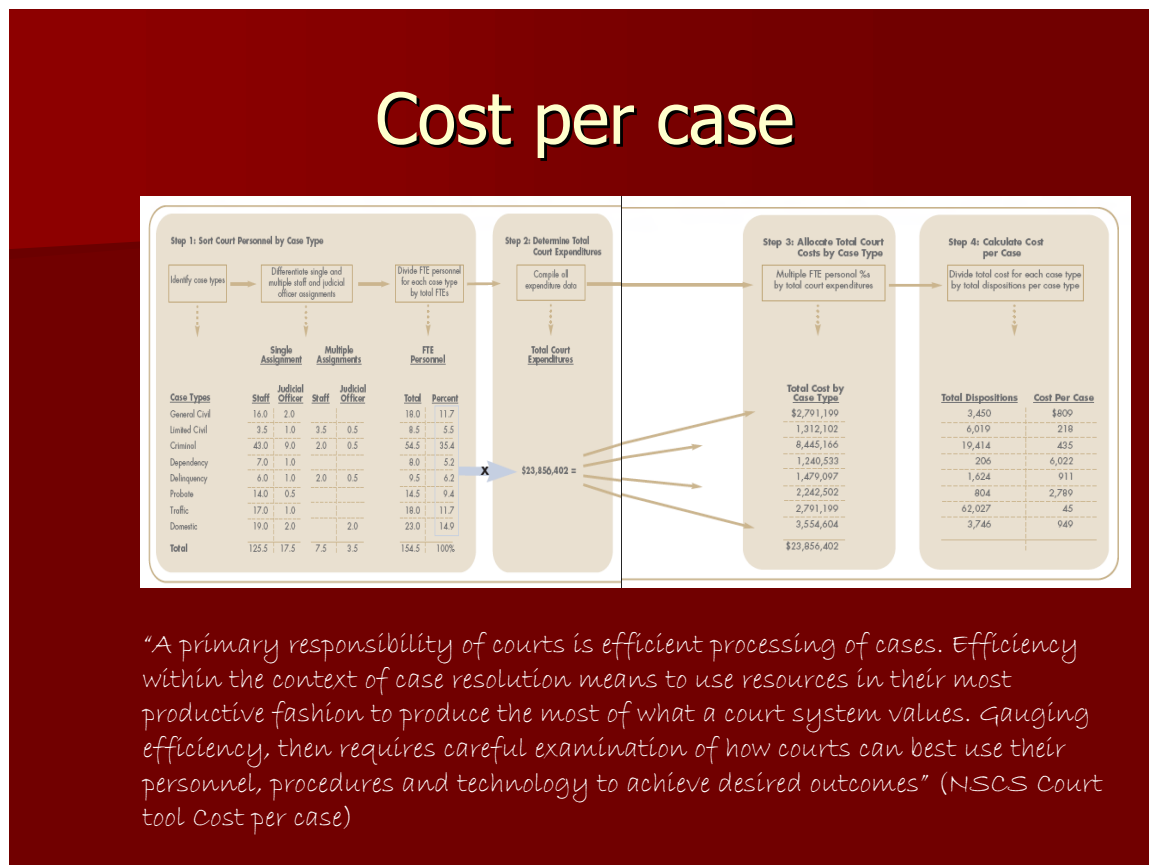
³ IRSIG CNR (February 2007), the duration of trials and the individual productivity of judges, Bologna.

⁴ See CEPEJ (2006), Time management checklist, Strasbourg.

Cost per case)⁵. The lower the costs' per case the more efficient is a court (or a department within a court).

The costs per case is derived by taking the aggregate costs by case type and dividing this figure by the total number of cases disposed in the year (see figure 2: steps to be taken to calculate the costs per case (source NCSC).

Figure 2



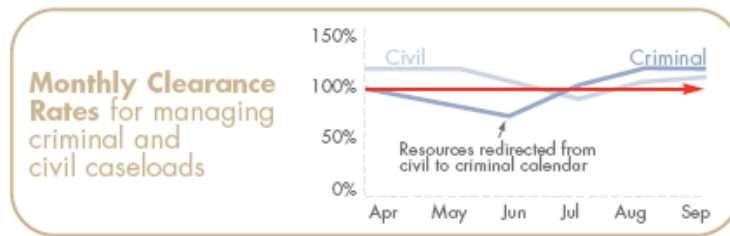
As is the case for the productivity performance indicator, one major disadvantage by making use of the cost per case approach is that it does not take into account the quality of the work delivered by the judges and their court staff.

⁵ See www.ncsconline.org.

Clearance rate

The clearance rate is defined as the number of outgoing case as a percentage of the incoming cases. It measures whether the courts are keeping up with their incoming caseload. If this is not the case, the backlog of cases will increase. In figure 3 an example has been given for the use of the clearance rate.

Figure 3 Source: NCSC Courtool Clearance rate



As can be seen from figure 3 the civil section of the court was functioning above the 100 percent norm, whilst the criminal section has a lower clearance rate. In this example the management of the court decides to reallocate financial resources from the civil section to the criminal section in the month June. After this reallocation the clearance rate in the criminal section of the court has been increased.

The budget

For determining the (labour) productivity in courts and the costs' per case it is necessary that courts systematically collect financial information. A planning and control cycle can be introduced as a method for rationalization of the financing of courts and the monitoring of their performances. In this approach courts are invited by superior authorities (for example a ministry of Justice or a High Judicial Council) to prepare an annual budget proposal. In this proposal information can be found of an estimation of the number of expected incoming cases for a year (per type of cases), the available personnel and material resources, other court performance information (length of proceedings, cases in stock, expected output of a court in terms of number of judicial decisions) and a proposal for the budget that is necessary to realize the expected output. At the

end of a budget year the management of the courts must prepare an annual report (spend budget and the court performance). This report, combined with a new budget proposal can be used in the negotiations between the court and the financing organization (Ministry, High Judicial Council, Superior Court, etc.). Examples of important budget items can be found in the section of the Judicial Reform Index of the American Bar Association.

Quality

One of the many disadvantages of the 'productivity' and 'efficiency' approach towards courts and judges is that it doesn't look at the quality of the products and services delivered by the courts. Mostly quality is defined by judges in terms of legal quality (the quality of judicial decisions), the system of quality protection by making use of courts of appeal and the role of the Supreme Court as a guard for judicial quality. Critics' say that the financiers of the judiciary are only interested in 'efficiency' and not in quality.

For a long period this was true, however a couple of years ago in several countries 'quality systems' has been introduced in courts. Most of these systems are derived from models that have been introduced in business companies. Two examples of well known Quality models are: the balance score card methodology and the EFQM-model (the European Foundation on Quality Management). The general principle of the balance score card model is that the quality of the organization can be measured by looking at four areas: the financial area, the working processes area, the learning and growth area ('the knowledge and the personnel of the organization') and the customer area. Similar principles can be found in the EFQM-model. One of the major aspects of both models is that the information from the client is important to assess the quality of an organization and that an organization must not only focus on 'efficiency' and 'productivity'. A balance should be struck between at the one hand 'quality' and at the other hand 'efficiency or productivity'.

The idea's of the general quality models has been a source of inspiration for developing specific models for the judiciary. One of the oldest models is the Trial Court Performance Standards (TCPS). It is one of the most important product of the Trial Court Performance Standards Project initiated and developed by the National Center for State Courts (NCSC), the US Bureau of Justice Assistance (BJU) and the US Department of Justice (started in 1987). In 1990 – after three years of development – the standards were published. During the following years trial courts in several US States has applied the standards and tested the utility and the feasibility of the system.

The Standards are developed around five performance areas: (1) access to justice, (2) expedition and timeliness, (3) equality fairness and integrity, (4) independence and accountability and (5) public trust and confidence. In total there are 68 measures included in this model. For each of the measures specific tools for data collection has been developed. For example the area of public trust and confidence a court client survey is drafted and can be used to assess the level of satisfaction of the clients regarding the services delivered by the judges and the court staff.

Evaluation of the Standards shows that it is possible to measure the quality in the courts. However, one of the drawbacks of the model is the level of complexity. It takes too much resources and time to evaluate the court quality by making use of the TCPS methodology.

As a part of the experiences of using TCPS, the National Center for State Courts has recently developed ten practical court tools. The court tools are defined by the Center as a set of trial court performances measures that offers court managers a balanced perspective on court operations. In the 'courtools' the major performance areas defined by the Trial Court Performance Standards has been integrated and extended with relevant concepts from successful performance measurement systems used in the public and private sector. Examples of concrete performance indicators that form a part of the set of court

tools which has been described in this paper are: the cost per case and the clearance rate.

Other countries that have developed a quality system for the judiciary are the Netherlands and Finland. The Dutch Rechtspraak model is based on the framework of the European Foundation for Quality Management. One of the elements in this model is the so-called 'measurement system'. It's a set of indicators and measurement tools that can be used to assess the quality in the Courts. The measurement system is based on five areas of measurement:

- Independence and impartiality
- Timeliness of proceedings
- Expertise of the judges
- Treatment of the parties at court sessions
- Judicial quality

Tools that are used to measure the quality are: the use of court statistics, an audit and surveys. The quality system of the Rovamieni court in Finland is based on similar principles.

One of the important aspects of the quality systems developed in the US, the Netherlands and in Finland is that it uses information of clients in addition to the 'efficiency' data. In figure 4 a sample is given of how a client survey can look like and shows some examples of items that are included in the survey.

Comparable surveys has been used in the Netherlands and what is interesting to notice is that when you are asking clients to rank the aspects that are important for a high quality of services delivered by the courts, the timeliness of the proceedings is not on the top. The expertise of the judge, the quality of the motivation of the decision and the expertise of the court staff are from the viewpoint of the clients the most important quality aspects for the courts.

Figure 4

Access and Fairness Survey- Adapted from NCSC Court Tools

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree	Not Applicable
Section I: Access to the Court						
1. Finding the courthouse was easy.	1	2	3	4	5	N/A
2. The forms I needed were clear and easy to understand.	1	2	3	4	5	N/A
3. I felt safe in the courthouse.	1	2	3	4	5	N/A
4. The court makes reasonable efforts to remove physical and language barriers to service.	1	2	3	4	5	N/A
5. I was able to get my court business done in a reasonable amount of time.	1	2	3	4	5	N/A
6. Court staff paid attention to my needs.	1	2	3	4	5	N/A
7. I was treated with courtesy and respect.	1	2	3	4	5	N/A
8. I easily found the courtroom or office I needed.	1	2	3	4	5	N/A
9. The court's Web site was useful.	1	2	3	4	5	N/A
10. The court's hours of operation made it easy for me to do my business.	1	2	3	4	5	N/A

Conclusion

In this paper some examples has been given regarding performance indicators for courts and judges. It is important to notice that when evaluating a judge or the court performance not to limit this evaluation to 'efficiency' and 'productivity' aspects but also to take a look at 'quality' aspects. Examples of an integral approach can be found in American, Dutch and Finnish courts which may apply quality system models.

For both the collection of data regarding 'efficiency' and (court) performance, as well as for (other) quality aspects it is important that a proper court management information system has been implemented. Without proper and reliable information it will not be possible to evaluate the quality and the performance of judges and court staff.