Research on the caseload management of courts: methodological questions

Andreas Lienhard
Daniel Kettiger

Introduction

Caseload management – as part of court management – is of key importance for guaranteeing the adjudication of cases, especially with regard to preventing both legal delays and the denial of legal rights as well as with regard to the quality of judgments. Caseload management can also serve as a method for the allocation of judicial resources and for the allocation of cases to judges. Weighted caseload systems are well established in the USA, but not to such an extent in continental Europe. This article deals with the methodology of caseload studies and shows that there is a considerable variety in methods of evaluating caseloads and in the use of weighted caseloads for assessing the needs of judicial officials. A review of the literature makes it apparent that in the near future the emergence of a unitary approach to caseload measurement as the preferred method or even as a standard of good practice is unlikely to happen.

1. Caseload management as an element of court management

1.1. Increasing need for court management

In most countries of the world the judiciary is under increasing pressure to reform: on the one hand, workloads, the complex issues being dealt with and procedural requirements are all increasing, while, on the other, scarcely any additional resources are available to cope with the problem. The result is that judicial authorities (conciliation and arbitration authorities, justices of the peace, public prosecutors’ offices, the courts) are forced to increase their efficiency.
However, this can ultimately be achieved only through a truly effective system of court management. Court management means the administration of the courts, i.e. the ‘administrative activity that creates and maintains the resources and personnel required for arriving at court judgments and rulings’. The essential matter is the administrative and financial management of the courts. The former President of the Cantonal Supreme Court of the Canton of Zurich (Switzerland), Rainer Klopfer, described the importance of court management as follows: ‘A court, as a major institution providing services, and as the most important supervisory body, needs a professional, efficient administration. This does not happen without management, but this in no way means that the independence of judges is compromised, just the opposite. It produces better working conditions for the judges and means that they can better fulfil their core duty, namely to adjudicate.’

1.2. Caseload management as a key element in court management

In recent times, endeavours to improve court management in theory and practice have led to the identification of a number of factors that are crucial to a system of good court management. The most important resources of the judiciary are human resources. To manage a court properly, one needs to know what capacity the judges and clerks have in properly dealing with cases brought before the court. Hence, caseload management is of crucial importance for guaranteeing adjudication, both with regard to preventing legal delays or the denial of legal rights (settling cases in good time) as well as with regard to the quality of judgments (settling cases in the correct way in procedural and material terms). To implement caseload management one first needs to measure the caseload (also referred to as workload). The main focus of this article is the methodology of caseload studies (cf. 3 and 4). Nevertheless, for a better understanding of the methodological questions, there is a need to focus first on caseload management as an instrument of allocation (cf. 2).

2. Caseload management as an instrument of allocation

2.1. Allocation of judicial resources

When discussing the allocation of judicial resources, one of the main questions concerns the number of judges, quasi-judicial officers, and court support staff which a court needs to serve the public. A clear measure of the court workload is central in determining how many judicial officers (judges, judicial assistants and clerks) are needed to resolve all the cases coming before the court. Weighted caseload systems are the most common and perhaps the best method used by court management systems to assess the judicial workload and resource requirements.

---

6 Not to be confused with case management: caseload management is focusing on the workload per case, while case management additionally also focuses on other elements like caseload.
7 Cf. Lienhard, supra note 2, no. 33 & 61; Henley & Suhr, supra note 2.
These systems are well established in the USA; their origins date back to the late 1970s. The National Center for State Courts, for example, has carried out this type of study in at least 11 states. In 2000, it was estimated that weighted caseload systems are used in at least 15 states. There are also long-lasting experiences with caseloads in federal district courts. One can find a considerable amount of literature on the subject, reflecting the development of caseload studies in the USA over almost 40 years.

In Continental Europe, however, caseload studies are not very well known and therefore caseload systems are seldom used for the allocation of judicial resources. Caseload analyses were carried out in German civil courts of the first and second instance at the beginning of the 1970s, and one of the pioneers in caseload management has been the Netherlands. However, it is only recently that caseloads have again become a major subject of research. The Belgian Ministry of Justice, for example, has commissioned a series of studies, whereas in Switzerland several caseload studies have recently been commenced as part of different initiatives towards good court management.

2.2. Allocation of cases

Another purpose of caseload management is the allocation of cases within the courts, to chambers or judges. It goes without saying that not all cases generate the same amount of work. In order to provide for proper case assignment and the allocation of the required resources on a reliable basis, it is necessary to have some idea of the staff resources, and especially of their working hours, that a case in any legal field (or case category) requires on average. A good climate in the court can thereby also be maintained and the presence of ‘legal harmony’ between the court divisions can be guaranteed. Caseload management is thus closely connected with controlling; it may also be a component of an integrated controlling system. The former Chief Justice of the Swiss Federal Supreme Court, Arthur Aeschlimann, summed this up as follows: ‘Each division president will have to work with objectives for his or her division members and staff and keep up to date on individual caseloads.’

10 Cf. Wisconsin Director of State Courts Office, supra note 8, p. 7; St. Caylor, Measuring the Need for Judges, 2000, p. 35.
11 Cf. B.J. Ostrom et al., Florida Delphi-based Weighted Caseload project, History of the Project, 2000, pp. 3 et seq.
12 Cf. Caylor, supra note 10, p. 35.
16 R. Depré et al., Etude de faisabilité de la mise en œuvres d’un instrument de mesure de la charge de travail destiné au siège, p. 3.
17 Ibid.
18 Cf. A. Lienhard & D. Kettiger, ‘Geschäftslastbewirtschaftung bei den Gerichten’, 2009 ZBl, pp. 413 et seq.; the authors are currently carrying out a caseload study on the Swiss Federal Administrative Court.
3. Methodological approaches to weighted caseloads

3.1. Weighted caseloads: from caseloads to workloads

Simply stated, a weighted caseload system is used to transform the court caseload into the workload of the judicial officials.21 Cases vary in complexity, and different types of cases require different amounts of time and attention from judges and court support staff.22 Raw case counts offer little help in distributing the workload equitably among judges, quasi-judicial staff, and court support staff. Merely adding up the total number of cases filed is not a good indicator of the amount of time it will take to dispose of that caseload. In the absence of explicit case weighting, all cases are counted equally, i.e. given a weighting of one unit. Focusing on case numbers without assessing the differences in workload means that one uncontested traffic case is equivalent to one contested case of intellectual property rights, although it is well known that some types of cases are more burdensome than others. As unweighted cases are not directly tied to workload, they offer only minimal guidance for estimating the need for judges and court support staff. Therefore, an estimate of the amount of work/time needed is a precondition for appraising the resources needed.23

The benefits of weighted caseloads are less important for courts whose jurisdiction is limited to the extent that they deal with similar cases in the sense of their variability of complexity. In those courts, the case mix remains constant and the raw, unweighted assessments of caseload allow the workload to be distributed fairly equitably.24

Workload in this context means the amount of a particular type of work which a qualified person can handle within a determined time. Yet only when working according to a defined method and quality standards, following a predetermined process within a specific organisational framework.25 It is important to evaluate the case workload in terms of the number of full time staff required to deal with it (the judicial full-time equivalent, FTE).26 This can be done, for example, by expressing the workload as a percentage of the annual working hours of a full-time member of judicial staff per case or by using the ‘judge year value’27 as a reference. Important factors that influence the workload are the type of case (the matter being judged), the type of procedure and the organisational framework. Therefore, it makes a difference whether the judicial officials simply deal with a case in their office, doing their work alone and exchanging opinions by email, or whether the judges hold court hearings with the parties and their lawyers.28 The size of the court also seems to be of some importance.29

---

21 Cf. Flango et al., supra note 9, p. 19; Lienhard, supra note 2, no. 61.
22 Cf. Flango et al., supra note 9, p. 19.
24 Cf. Flango et al., supra note 9, pp. 19 et seq.; Caylor, supra note 10, p. 35.
26 Cf. Lienhard & Kettiger, supra note 18, p. 434.
27 The ‘judge year value’ means the average amount of time that one full-time judge can expect to spend annually on case-related matters, cf. Stenz, supra note 9, p. 380; Caylor, supra note 10, p. 38.
Judicial officials cannot devote their entire working day to case-related matters. Therefore weighted caseload systems also need to know the time spent on non-case-related activities. Four main forms of non-case-related activity can be distinguished:\textsuperscript{30}

- General (non-case-related) court administration: This includes all non-bench or non-case-related working time spent on activities such as routine office matters, court committee meetings, staff meetings, personnel matters, work on court projects or chief judge duties.
- Judicial education and training: This includes all non-bench or non-case-related working time spent on basic judicial training and advanced training for qualified staff, attending judicial conferences and the reading of reports of decisions, law review articles and new legal literature.
- Community activities and public outreach: This includes all non-bench or non-case-related working time spent on public outreach, such as parliamentary or government boards and commissions, community education (including lectures at universities), meetings with the boards of other judicial authorities and bar association meetings.
- Private matters: This includes all non-bench or non-case-related working time spent travelling away from court\textsuperscript{31} or due to vacations, personal commitments or illness.

When measuring the workload or caseload of judicial officials, there is no need to measure the non-case-related working time in different categories. From the point of view of court management, knowing how much time is spent on different categories of non-case-related working time is more interesting in the context of measuring court overheads than in the context of caseload management.

3.2. Measuring the time taken for specific events

Probably the best known and most commonly used method\textsuperscript{32} for obtaining a weighted caseload is to measure the amount of judicial work by means of time studies for all major events within the case according to case types.\textsuperscript{33} In a first step, the events that occur in a particular type of case must be identified. These range from preliminary matters, meaning all of the routine matters that occur in cases before a disposal is made or before a case comes to trial, up to post-judgment work.\textsuperscript{34} In a second step, the frequency with which the event occurs must be determined. The third step is to measure the average time spent by judicial officials on each particular event of a case type. The total average time spent on all these events corresponds to the weighted caseload.

3.3. Measuring the total average time per case

Another approach is to measure the total time that is needed for a given case – this means the time required by all judicial officials involved in the case over the whole runtime of the case. This is the method favoured by the authors and is used in an ongoing caseload study in the Swiss

\textsuperscript{30} One finds an overview of non-case-related activities in Wisconsin Director of State Courts Office, supra note 8, Appendix 1, p. A-1; the authors use an almost identical list in their recent project with the Swiss Federal Administrative Court; see also B. Ostrom & N. Kauder, \textit{Examining the Work of State Courts}, 1997, p. 97.

\textsuperscript{31} In the USA travelling time is often considered a category of its own, cf. Wisconsin Director of State Courts Office, supra note 8, Appendix 1, p. A-1.

\textsuperscript{32} At least in the USA; the method was also used in Germany, cf. Bundesrechtsanwaltskammer (ed.), Vol. I, supra note 15, pp. 184 et seq.


\textsuperscript{34} Cf. Wisconsin Director of State Courts Office, supra note 8, p. 10; Bundesrechtsanwaltskammer (ed.), Vol. I, supra note 15, pp. 184 et seq.
Research on the caseload management of courts: methodological questions

Over a longer period (e.g. six months) all judicial officials record the time worked on each particular case. After the determined period, it is possible to calculate the average time needed for cases of a particular case type. This results in a weighted caseload per case type.

The disadvantage of this method is that the whole procedure has to be looked at as a ‘black box’. No information can be gained about the workload for specific events or working phases over the time taken to deal with the case. Therefore, in the recent study at the Swiss Federal Administrative Court, two different categories of time working on a case were distinguished: the time used for work in the pre-judgment phase, on the one hand, and the time used for work directly related to the judgment, on the other. Self-reports are considered the best way to gather information on judges and court support staff because they require direct participation (and buy-in) by the court and because self-reports cover all activities, whether or not they occur in a courtroom. But there can be compliance problems if judges refuse to participate.

3.4. Estimated values

To obtain weighted caseloads, methods based on estimated working time for each case type can also be used. In 2008, the University of Bern conducted a blanket survey of all the highest level cantonal administrative and social insurance courts in Switzerland. One part of the survey consisted of an assessment of the average time spent per case in 19 specified categories of cases. The persons involved had to indicate the estimated time spent on every case type as a percentage of FTE.

Another method based on estimated working time is the Delphi Method. This approach to building case weightings can also be done in two ways: focusing on case types or on case events. The weighted caseload is estimated by experts (the judges themselves or external experts appointed by the judges) in a first round. Then the experts review the opinion of their peers after each iteration (normally there are two or three iterations) and they may then modify their individual estimates based upon the results of the group estimates. The Delphi Method has been used in several states in the USA (e.g. Alabama, Arkansas, Georgia, Maryland, Michigan, Pennsylvania, and South Dakota). The method has been applied both to supplement actual working time studies (whether for judges, quasi-judicial officers or court staff) and to create weighted caseloads without an underlying empirical survey of processing times for case events.

The weaknesses of the Delphi Method are mainly that the responses to specific questions are subject to the question design, it relies on snap opinions, it forces a consensus and it creates the illusion of precision despite being based on personal estimates (often of ‘experts’ directly involved).

Attempts at a weighted caseload within the court – at least in Switzerland – are often primarily based on estimated values: the values they survey in some cases represent individual assessments by persons from the respective statistical population (court in plenary session,
division), made on the basis of a predefined scale (a scaled response to a questionnaire). The scale employed is abstract, i.e. it represents a weighting system sui generis, which cannot necessarily be linked to any other system of measurement or assessment, certainly not to FTE.\textsuperscript{43} In addition, the scale is limited by an upper and lower value; the lowest possible value (e.g. 1) and the highest possible value (e.g. 5) are prescribed beforehand. It was furthermore decided in advance that the middle value of the workload scale (in the case of a scale from 1 to 5, the value 3) signifies an average load. With such a procedure, it is not clear from the outset whether the scaling is linear (i.e. whether the distance or value difference between the values is always the same size) or whether it takes a different course.\textsuperscript{44}

On the one hand, the disadvantage of a weighted caseload system, based on estimated values, is the lack of reliability concerning the values of working hours or working days estimated per case. It is therefore difficult to reconcile the results with FTE or ‘judge days’. On the other hand, the research carried out by the authors shows that if one compares the estimated workload per case for particular case types resulting from the previously mentioned blanket survey\textsuperscript{45} with the results of several internal caseload approaches carried out by courts at the cantonal and federal\textsuperscript{46} level,\textsuperscript{47} one finds a constant relation. There is, for example, a constant factor of 3:5 when the workload of invalidity insurance cases is compared with the workload of planning, construction and environmental law cases. The relationship between the categories of workloads established by an estimation by judicial officials is relatively constant and may be regarded as a reasonably secure value.\textsuperscript{48}

4. The problem of benchmarking in federal states

There is one thing that caseload studies do not reveal: they do not answer the question of whether or not a court is working efficiently and effectively. To answer this question, benchmarks which enable comparisons between courts are needed. In centralised states, where the judiciary has a unified structure and there are unified procedural rules, it should be possible by implementing a standard method for weighting caseloads to devise benchmarks that allow information to be obtained about the efficacy and effectiveness of courts of the same type. In states with strongly federal structures, this is virtually impossible, as can be seen from the examples of the USA and Switzerland.

In the USA, the variation in the number of general jurisdiction case types that are weighted means that judicial workload assessment studies must be individually tailored to the needs of each state.\textsuperscript{49} As a consequence, it is very difficult to compare case weights across the states. Even

\textsuperscript{43} Cf. Lienhard & Kettiger, supra note 18, p. 422.
\textsuperscript{44} The question of whether a scale with a scaled response is linear depends on the precise, verbal definition of the values. For example, the scale of values ‘completely exceeded/partly exceeded/satisfactory/partly unsatisfactory/completely unsatisfactory’ is not linear, because in our experience the interviewees tick ‘partly exceeded’ for work that is slightly better than satisfactory, but only tick the highest grade in the case of absolutely perfect work. In practice, the same scales are also worded and used differently in different cultures: the traditional (Swiss) school marking system with grades from 1 to 6 in state schools is mainly used as a linear system of assessment (often in a test there are 12 questions and each question carries half a point, with the result that 8 points are enough to pass), at universities the tendency is more exponential in the satisfactory range.
\textsuperscript{45} Cf. Lienhard & Kettiger, supra note 18, p. 431.
\textsuperscript{46} The Swiss Federal Court made a survey in 2008 by allowing judges and clerks to estimate the weight of cases per case type within a scale from 1 to 5 (3 being the average caseload).
\textsuperscript{47} The results of those studies have not been made accessible to the public.
\textsuperscript{48} Cf. Lienhard & Kettiger, supra note 18, p. 435.
\textsuperscript{49} Cf. Ostrom & Kauder, supra note 30, p. 95; it seems that also a benchmark between federal district courts, using the same statistical system, is not that easy, cf. Meierhoefer et al., supra note 13, pp. 13-20.
if states appear to regard a case in the same way, case weights may vary because of differences in how the case is defined or processed.\textsuperscript{50} There are also considerable differences in how to calculate ‘judge time’.\textsuperscript{51} Finally, there is a huge variety in the details of the methods used.\textsuperscript{52}

One of the main findings of the study carried out by the authors in the field of Swiss administrative courts is that the landscape of the administrative justice system in Switzerland, including the highest administrative courts in the cantons, is far more diverse than it was assumed at the beginning of the research project.\textsuperscript{53} This diversity has led to the assumption that in a comparison with the Swiss-wide average, a distortion may arise, in particular due to the values returned by the small, rural cantons. The small, rural cantons (specifically Appenzell-Innerrhoden, Appenzell-Ausserrhoden, Nidwalden, Obwalden and Uri) have administrative courts that make extensive use of part-time judges. In addition, the number of cases in each of the categories is essentially different from those in the more urban areas or in cantons with both urban and rural or suburban regions. In various cantons (specifically Basel-Stadt, Schaffhausen, Solothurn), all the administrative and social insurance judgments at a higher level are issued by the Cantonal Supreme Court, as well as the judgments in civil and criminal cases, so that the judges basically all have experience in cases from the entire spectrum of court judgments. Given the small workloads per category of case, it becomes scarcely possible to properly estimate the workloads involved.

\section*{5. Final remarks}

This brief overview of caseload management, and in particular of weighted caseload systems, shows, firstly, that the weighted caseload has regularly been the subject of studies and research for a considerable period of time both on the European continent and in the USA. The USA, in particular, also has long-term experience with weighted caseloads. Secondly, the overview shows that there is a considerable variety in the methods used for caseload studies and in the use of weighted caseloads for assessing the need of judicial officials. A review of the literature makes it apparent that no unitary approach to caseload measurement is likely to emerge in the near future as the preferred method or will even become some sort of standard of good practice.\textsuperscript{54}

For the purpose of the allocation of judicial resources and for the allocation of cases within a court, this diversity in methods does not preclude the use of weighted caseloads. The methods in most cases match the needs of the courts and of the legislature and, for most purposes, it is normally sufficient to work within a closed system that is restricted to one court or one state. For the purpose of the allocation of resources, detailed accuracy is not necessary: it is sufficient to use a rough scale of approximate values, subdivided into a total of three to four general case types.\textsuperscript{55}

When it comes to the issue of efficiency and the efficacy of the courts and the need for benchmarks, however, there is a need for the standardisation of caseload methods – at least at a national level. To achieve standardisation, a great deal of research needs to be done and it will be necessary to pool existing knowledge in an international research project.

\begin{flushright}
\begin{footnotesize}
\textsuperscript{50} Cf. Ostrom & Kauder, supra note 30, p. 95.
\textsuperscript{51} Cf. Ostrom & Kauder, supra note 30, pp. 95 et seq.
\textsuperscript{52} Cf. Hurst, supra note 33, p. 5.
\textsuperscript{53} Cf. Lienhard & Kettiger, supra note 18, pp. 424 et seq.
\textsuperscript{54} Cf. the same findings by Hurst, supra note 33, p. 5.
\textsuperscript{55} Cf. Lienhard & Kettiger, supra note 18, pp. 434 et seq.
\end{footnotesize}
\end{flushright}