

GAO

Testimony

Before the Subcommittee on Courts, the
Internet, and Intellectual Property,
Committee on the Judiciary,
House of Representatives

For Release on Delivery
Expected at 2:00 p.m. EDT
Tuesday, June 24, 2003

FEDERAL JUDGESHIPS

General Accuracy of District and Appellate Judgeship Case-Related Workload Measures

Statement of William O. Jenkins, Jr., Director
Homeland Security and Justice Issues



Mr. Chairman and Members of the Subcommittee:

I am pleased to be here today to discuss the results of our review and assessment of case-related workload measures for district court and courts of appeals judges.¹ Biennially, the Judicial Conference of the United States, the federal judiciary's principal policymaking body, assesses the judiciary's needs for additional judgeships.² If the Conference determines that additional judgeships are needed, it transmits a request to Congress identifying the number, type (courts of appeals, district, or bankruptcy), and location of the judgeships it is requesting. In assessing the need for additional district and appellate court judgeships, the Judicial Conference considers a variety of information, including responses to its biennial survey of individual courts, temporary increases or decreases in case filings, and other factors specific to an individual court. However, the Conference's analysis begins with the quantitative case-related workload measures it has adopted for the district courts and courts of appeals—weighted case filings and adjusted case filings, respectively. These two measures recognize, to different degrees, that the time demands on judges are largely a function of both the number and complexity of the cases on their dockets. Some types of cases may demand relatively little time and others may require many hours of work.

My statement is based on our recent report, which you requested, on the relative accuracy of weighted case filings and adjusted case filings as a measure of the case-related workload of district and courts of appeals judges, respectively.³ Whether weighted case filings and adjusted case filings are reasonably accurate measures of case-related judge workload rests on the soundness of the methodology used to develop these

¹We recently testified on the methodology used to develop the case-related workload measure for bankruptcy judges. See U.S. General Accounting Office, *Federal Bankruptcy Judges: Weighted Case Filings as a Measure of Judges' Case-Related Workload*, [GAO-03-789T](#) (Washington, D.C.: May 22, 2003). This testimony is available on GAO's Web site at [www.gao.gov](#).

²The Chief Justice of the United States presides over the Conference, which consists of the chief judges of the 13 courts of appeals, a district judge from each of the 12 geographic circuits, and the chief judge of the Court of International Trade. The Conference meets twice a year.

³U.S. General Accounting Office, *Federal Judgeships: The General Accuracy of the Case-Related Workload Measures Used to Assess the Need for Additional District Court and Courts of Appeals Judgeships*, [GAO-03-788R](#) (Washington, D.C.: May 30, 2003). This report is available on GAO's Web site at [www.gao.gov](#).

measures. My statement and our report are based on the results of our review of documentation provided by the Federal Judicial Center (FJC) and the Administrative Office of the U.S. Courts (AOUSC) and interviews with officials in each organization. The scope of our work did not include how the Judicial Conference used these case-related workload measures to develop its current judgeship request for district court and courts of appeals judgeships. My statement includes the following major points:

- The district court weighted case filings, as approved in 1993, appear to be a reasonably accurate measure of the average time demands that a specific number and mix of cases filed in a district court could be expected to place on the district judges in that district. The methodology used to develop the case weights was based on a valid sampling procedure, developed weights based on actual case-related time recorded by judges from case filing to disposition, and included a measure (standard errors) of the statistical confidence in the final weight for each weighted case type.
- The case weights, however, are about 10 years old, and the data on which the weights are based are as much as 15 years old. Changes since 1993, such as the characteristics of cases filed in federal district courts and changes in case management practices, may have affected whether the 1993 case weights continue to be a reasonably accurate measure of the average time burden on district court judges resulting from a specific volume and mix of cases.
- The Judicial Conference's Subcommittee on Judicial Statistics has approved a research design for updating the current case weights, and we have some concerns about that design. The design would include limited data on the time judges actually spend on specific types of cases. The proposed design would not include collecting actual data on the noncourtroom time that judges spend on different types of cases. Estimates of the noncourtroom time required for specific types of cases would be based on estimates derived from the structured, guided discussions of about 100 experienced judges meeting in 12 separate groups (one for each geographic circuit). These noncourtroom time estimates are likely to represent the majority of judge time used to develop the new case weights. The accuracy of case weights developed on such consensus data cannot be assessed using standard statistical methods, such as the calculation of standard errors. Thus, it would not be possible to objectively, statistically assess how accurate the new case weights are—weights on whose reasonable accuracy the Judicial Conference will rely in assessing judgeship needs in the future.

-
- Adjusted case filings, the principal quantitative measure used to assess the case-related workload of courts of appeals judges, are based on available data from standard statistical reports from the courts of appeals. The measure is not based on any empirical data about the judge time required by different types of cases in the courts of appeals. The measure essentially assumes that all cases filed in the courts of appeals, with the exception of pro se cases—those in which one or both parties are not represented by an attorney—require the same amount of judge time. On the basis of the documentation we reviewed, there is no empirical basis on which to assess the accuracy of adjusted filings as a measure of case-related workload for courts of appeals judges.
 - Whether the district court case weights are a reasonably accurate measure of district judge case-related workload is dependent upon two variables: (1) the accuracy of the case weights themselves and (2) the accuracy of classifying cases filed in district courts by the case type used for the case weights. If case filings are inaccurately identified by case type, then the weights are inaccurately calculated. Because there are fewer categories used in the courts of appeals workload measure, there is greater margin for error. AOUSC said that its staff took a number of steps to ensure that individual cases were assigned to the appropriate caseweight category. These are described in appendix 1. We did not evaluate how effective these measures may be in ensuring data accuracy.

District Court Weighted Case Filings, as Approved, Are a Reasonably Accurate Measure of Case-Related Judge Workload

The demands upon judges' time are largely a function of both the number and complexity of the cases on their dockets. Some types of cases may demand relatively little time, and others may require many hours of work. To measure the case-related workload of district court judges, the Judicial Conference has adopted weighted case filings. The purpose of the district court case weights was to create a measure of the average judge time that a specific number and mix of cases filed in a district court would require. Importantly, the weights were designed to be descriptive not prescriptive—that is, the weights were designed to develop a measure of the national average amount of time that judges actually spent on specific types of cases, not to develop a measure of how much time judges should spend on specific types of cases. Moreover, the weights were designed to measure only case-related judge workload. Judges have noncase-related duties and responsibilities, such as administrative tasks, that are not reflected in the case weights.

With few exceptions, such as cases that are remanded to a district court from the courts of appeals, each civil and criminal case filed in a district

court is assigned a case weight. Each case filed in a district court is assigned a case weight based on the subject matter of the case. The weight of the overall average case is 1.0. All other weights were established relative to this national average case. Thus, a case with a weight of 0.5 would be expected to require on average about half as much judge time as the national average case, and a case with a value of 2.0 would be expected to require on average about twice as much judge time as the national average case. Case weights for criminal felony defendants are applied on a per defendant basis.⁴ For example, the case weight for heroin/cocaine distribution is 2.27. If such a case involved two defendants, the court would be credited with a weight of 4.54—two times the assigned case weight of 2.27. Of course, the actual amount of time a judge may spend on any specific case may be more or less than the national average for that type of case.

Total weighted filings for a district are determined by summing the case weights associated with all the cases filed in the district during the year. Weighted case filings per authorized judgeship—is the total annual weighted filings divided by the total number of authorized judgeships for the district. For example, if a district had total weighted filings of 4,600 and 10 authorized judgeships, its weighted filings per authorized judgeship would be 460. The Judicial Conference uses weighted filings of 430 or more per authorized judgeship as an indication that a district may need one or more additional judgeships. Thus, a district with 460 weighted filings per authorized judgeship could be considered for an additional judgeship.

The Judicial Conference approved the use of the current district court case weights in 1993. The weights are based on a “case-tracking time study,” conducted between 1987 and 1993, in which judges recorded the amount of time spent on each of their cases included in the study. The study included about 8,100 civil cases and about 4,200 criminal cases. Overall, the weighted case filings, as approved in 1993, are a reasonably accurate method of measuring the average judge time that a specific number and mix of cases filed in a district court would require. The methodology used to develop the case weights was reasonable. It used a valid sampling procedure, developed weights based on actual case-related time recorded by judges from case filing to disposition, and included a

⁴The weights do not include nonfelony criminal cases, which are generally the responsibility of magistrate, not district, judges.

measure (standard errors) of the statistical confidence in the final weight for each weighted case type.

Current Case Weights about 10 Years Old

The case weights are almost 10 years old, and the time data on which they were based are as much as 15 years old. Changes since the case weights were finalized in 1993, such as changes in the characteristics of cases filed in federal district courts and in case management practices, may affect how accurately the weights continue to reflect the time burden on district court judges today. For example, since 1993, new civil causes of action (such as telemarketing issues) and criminal offenses (such as new terrorism offenses) needed to be accommodated within the existing case-weight structure. According to FJC officials, where the new cause of action or criminal offense is similar to an existing case-weight type, the weight for the closest case type is assigned. Where the new cause of action or criminal offense is clearly different from any existing case-weight category, the weight assigned is that for either “all other” civil cases or “all other” criminal cases.

Concerns about the Research Design for Updating the District Court Case Weights

The Subcommittee on Judicial Statistics of the Judicial Conference’s Committee on Judicial Resources has approved the research design for revising the current case weights, with the goal of having new weights submitted to the Resources Committee for review in the summer of 2004. The design for the new case weights relies on three sources of data for specific types of cases: (1) data from automated databases identifying the docketed events associated with cases; (2) data from automated sources on the time associated with courtroom events for cases, such as trials or hearings; and (3) estimated time data from structured, guided discussion among experienced judges on the time associated with noncourtroom events for cases, such as reading briefs or writing opinions.

Although the proposed methodology appears to offer the benefit of reduced judicial burden (no time study data collection), potential cost savings, and reduced calendar time to develop the new weights, we have two principal concerns about the research design—the challenge of obtaining reliable, comparable data from two different automated data systems for the analysis and the limited collection of actual data on the time judges spend on cases.

The design assumes that judicial time spent on a given case can be accurately estimated by viewing the case as a set of individual tasks or events in the case. Information about event frequencies and, where available, time spent on the events would be extracted from existing

administrative data bases and reports and then used to develop estimates of the judge-time spent on different types of cases. For event data, the research design proposes using data from new technology (the Case Management/Electronic Case Filing System) that is currently being introduced into the court system for recording case management information. However, not all courts have implemented the new system, and data from the existing and new systems will have to be integrated to obtain and analyze the event data. FJC researchers, who would conduct the research, recognize the challenges this poses and have developed a strategy for addressing the issues, which includes forming a technical advisory group from FJC, the Administrative Office of the U.S. Courts, and individual courts to develop a method of reliably extracting and integrating data from the two case management systems for analysis.

Second, the research design does not require judges to record time spent on individual cases. Actual time data would be limited to that available from existing reports on the time associated with courtroom events and proceedings for different types of cases. However, a majority of district judges' time is spent on case-related work outside the courtroom. The time required for noncourtroom events would be derived from structured, guided discussions of groups of 8 to 13 experienced district court judges in each of the 12 geographic circuits (about 100 judges in all). The judges would develop estimates of the time required for different events in different types of cases within each circuit, using FJC-developed "default values" as the reference point for developing their estimates. These default values would be based in part on the existing case weights and in part on other types of analyses. Following the meetings of the judges in each circuit, a national group of 24 judges (2 from each circuit) would consider the data from the 12 circuit groups and develop the new weights.

The accuracy of judges' time estimates is dependent upon the experience and knowledge of the participating judges and the accuracy and reliability of the judges' recall about the time required for different events in different types of cases—about 150 if all the case types in the current case weights were used. These consensus data cannot be used to calculate statistical measures of the accuracy of the resulting case weights. Thus, it will not be possible to objectively, statistically assess how accurate the new case weights are—weights on whose accuracy the Judicial Conference will rely in assessing judgeship needs in the future.

A time study conducted concurrently with the proposed research methodology would be advisable to identify potential shortcomings of the event-based methodology and to assess the relative accuracy of the case

weights produced using that methodology. In the absence of a concurrent time study, there would be no objective statistical way to determine the accuracy of the case weights produced by the proposed event-based methodology.

Adjusted Case Filings: Accuracy of Courts of Appeals Case-Related Workload Measure Cannot Be Assessed

The principal workload measure that the Judicial Conference uses to assess the need for additional courts of appeals judges is adjusted case filings. We found that adjusted case filings are based on data available from standard statistical reports for the courts of appeals. The measure is not based on any empirical data about the judge time required by different types of cases in the courts of appeals.

The Judicial Conference's policy is that courts of appeals with adjusted case filings of 500 or more per three-judge panel may be considered for one or more additional judgeships. Courts of appeals generally decide cases using constantly rotating three-judge panels. Thus, if a court had 12 authorized judgeships, those judges could be assigned to four panels of three judges each. In assessing judgeship needs for the courts of appeals, the Conference may also consider factors other than adjusted case filings, such as the geography of the circuit or the median time from case filings to dispositions.

Adjusted case filings are used for 11 of the 12 courts of appeals. It is not used for the Court of Appeals for the D.C. Circuit. A FJC study of that court's workload determined that adjusted case filings were not an appropriate means of measuring the court's judgeship needs. The court had a high proportion of administrative agency appeals that occurred almost exclusively in the Court of Appeals for D.C. and were more burdensome than other types of cases in several respects—e.g., more independently represented participants per case, more briefs filed per case, and a higher rate of case consolidation.⁵

Essentially, the adjusted case filings workload measure counts all case filings equally, with two exceptions. First, cases refiled and approved for reinstatement are excluded from total case filings.⁶ Second, two-thirds of

⁵The Conference did not request any judgeships in 2003 for the D.C. Court of Appeals.

⁶Such cases were dismissed for procedural defaults when originally filed, but "reinstated" to the court's calendar when the case was later refiled. The number of such cases, as a proportion of total cases, is generally small.

pro se cases—defined by the Administrative Office as cases in which one or both of the parties are not represented by an attorney—are deducted from total case filings (that is, they are effectively weighted at 0.33). For example, a court with 600 total pro se filings in a fiscal year would be credited with 198 adjusted pro se case filings (600 x 0.33). The remaining nonpro se cases would be weighted at 1.0 each. Thus, a court of appeals with 1,600 case filings (excluding reinstatements)—600 pro se cases and 1,000 nonpro se cases—would be credited with 1,198 adjusted case filings (198 discounted pro se cases plus 1,000 nonpro se cases). If this court had 6 judges (allow two panels of 3 judges each), it would have 599 adjusted case filings per 3-judge panel, and, thus, under Judicial Conference policy, could be considered for an additional judgeship.

The current court of appeals workload measure represents an effort to improve the previous measure. In our 1993 report on judgeship needs assessment, we noted that the restraint of individual courts of appeals, not the workload standard, seemed to have determined the actual number of appellate judgeships the Judicial Conference requested.⁷ At the time the current measure was developed and approved, using the new benchmark of 500 adjusted case filings resulted in judgeship numbers that closely approximated the judgeship needs of the majority of the courts of appeals, as the judges of each court perceived them. The current courts of appeals case-related workload measure principally reflects a policy decision using historical data on filings and terminations. It is not based on empirical data regarding the judge time that different types of cases may require. On the basis of the documentation we reviewed, we determined that there is no empirical basis for assessing the potential accuracy of adjusted filings a measure of case-related judge workload.

Recommendations

In our report, we recommended that the Judicial Conference of the United States

- update the district court case weights using a methodology that supports an objective, statistically reliable means of calculating the accuracy of the resulting weights; and

⁷U.S. General Accounting Office, *Federal Judiciary: How the Judicial Conference Assesses the Need for More Judges*, GAO/GGD-93-31 (Washington, D.C.: Jan. 29, 1993).

-
- develop a methodology for measuring the case-related workload of courts of appeals judges that supports an objective, statistically reliable means of calculating the accuracy of the resulting workload measure(s) and that addresses the special case characteristics of the Court of Appeals for the D.C. Circuit.

In a May 27, 2003, letter to GAO, the Chair of the Committee on Judicial Resources said that the development of the new case weights will use substantial data already collected and that our report did not reflect the sophisticated methodology the FJC had designed for the study nor acknowledge the substantial increased costs and time involved in a time study that was likely to offer little or no added value for the investment. The letter also noted that the workloads of the courts of appeals entail important factors that have defied measurement, including the significant differences in the courts' case processing techniques. The Deputy Director of FJC, in a May 27, 2003, letter agreed that the estimated data on noncourtroom judge time in the new study would not permit the calculation of standard errors. However, the integrity of the resulting case-weight system could still be evaluated on the basis of adherence to the procedures that will be used to gather the data and promote their reliability.

We believe that our analysis and recommendations are sound and that the importance and costs of creating new Article III federal judgeships requires the best possible case-related workload data to support the assessment of the need for more judgeships.

That concludes my statement, Mr. Chairman, and I would be pleased to answer any questions you or other Members of the Subcommittee may have.

GAO Contacts and Acknowledgments

For further information regarding this testimony, please contact William O. Jenkins, Jr., at (202) 512-8777. Individuals making key contributions to this testimony included David Alexander, Kriti Bhandari, R. Rochelle Burns, and Chris Moriarity.

Appendix I: Quality Assurance Steps the Judiciary Takes to Ensure the Accuracy of Case Filing Data for Weighted Filings

Whether the district court case weights are a reasonably accurate measure of district judge case-related workload is dependent upon two variables: (1) the accuracy of the case weights themselves and (2) the accuracy of classifying cases filed in district courts by the case type used for the case weights. If case filings are inaccurately identified by case type, then the weights are inaccurately calculated. Because there are fewer categories used in the courts of appeals workload measure, there is greater margin for error. The database for the courts of appeals should accurately identify (1) pro se cases, (2) reinstated cases, and (3) all cases not in the first two categories.

All current records related to civil and criminal filings that are reported to the Administrative Office of the U.S. Courts (AOUSC) and used for the district court case weights are generated by the automated case management systems in the district courts. Filings records are generated monthly and transmitted to AOUSC for inclusion in its national database. On a quarterly basis, AOUSC summarizes and compiles the records into published tables, and for given periods these tables serve as the basis for the weighted caseload determinations.

In responses to written questions, AOUSC described numerous steps taken to ensure the accuracy and completeness of the filings data, including the following:

- Built-in, automated quality control edits are done when data are entered electronically at the court level. The edits are intended to ensure that obvious errors are not entered into a local court's database. Examples of the types of errors screened for are the district office in which the case was filed, the U.S. Code title and section of the filing, and the judge code. Most district courts have staff responsible for data quality control.
- A second set of automated quality control edits are used by AOUSC when transferring data from the court level to its national database. These edits screen for missing or invalid codes that are not screened for at the court level, such as dates of case events, the type of proceeding, and the type of case. Records that fail one or more checks are not added to the national database and are returned electronically to the originating court for correction and resubmission.
- Monthly listings of all records added to the national database are sent electronically to the involved courts for verification.

-
- Courts' monthly and quarterly case filings are monitored regularly to identify and verify significant increases or decreases from the normal monthly or annual totals.
 - Tables on case filings are published on the Judiciary's intranet for review by the courts.
 - Detailed and extensive statistical reporting guidance is provided to courts for reporting civil and criminal statistics. This guidance includes information on general reporting requirements, data entry procedures, and data processing and reporting programs.
 - Periodic training sessions are conducted for district court staff on measures and techniques associated with data quality control procedures.

AOUSC did not identify any audits to test the accuracy of district court case filings or any other efforts to verify the accuracy of its electronic data by comparing the electronic data to "hard copy" case records for district courts. Within the limited time for our review, AOUSC was unable to obtain information from individual courts to include in its responses. We have no information on how effective the procedures AOUSC described may be in ensuring that the data in the automated databases were accurate and reliable means of assigning weights to district court case filings.