Notre Dame Law Review

Volume 78 | Issue 5 Article 2

8-1-2003

The Reliability of the Administrative Office of the U.S. Courts Database: An Initial Empirical Analysis

Theodore Eisenberg

Margo Schlanger

THE RELIABILITY OF THE ADMINISTRATIVE OFFICE OF THE U.S. COURTS DATABASE: AN INITIAL EMPIRICAL ANALYSIS

*Theodore Eisenberg**
Margo Schlanger†

* Henry Allen Mark Professor of Law, Cornell Law School.
† Assistant Professor of Law, Harvard Law School.

The authors would like to thank Ann M. Eisenberg, Heather Hillman, Amanda Matey, Bradford P. Maxwell, Erica Miller, Douglas Schnell, and Sylvanie Wallington for their excellent research assistance. Thanks to the Harvard University Milton Fund for research support, and to Sam Bagenstos, Christine Jolls, and Louis Kaplow for helpful comments.

INTRODUCTION

Researchers have long used federal court data assembled by the Administrative Office of the U.S. Courts (AO) and the Federal Judicial Center (FJC). The data include information about every case filed in federal district court and every appeal filed in the twelve nonspecialized federal appellate courts.[1] Much research using the AO data spans subject matter areas, and includes articles on appeals,[2] caseloads and case-processing times,[3] case outcomes,[4] the relation be-

1 *See* INTER-UNIVERSITY CONSORTIUM FOR POL. & SOC. RES., FEDERAL COURT CASES: INTEGRATED DATA BASE, 2001, ICPSR Study No. 3415 (2002) [hereinafter ICPSR 3415]; INTER-UNIVERSITY CONSORTIUM FOR POL. & SOC. RES., FEDERAL COURT CASES: INTEGRATED DATA BASE, 1970–2000, ICPSR Study No. 8429 (2001) [hereinafter ICPSR 8429]. For additional information on the federal courts' recordkeeping, see TECH. TRAINING & SUPPORT DIV., ADMIN. OFF. OF THE U.S. CTS., CIVIL STATISTICAL REPORTING GUIDE (July 1999) [hereinafter CIV. STAT. REPORTING GUIDE] (on file with authors); 11 ADMIN. OFF. OF THE U.S. CTS., GUIDE TO JUDICIARY POLICIES AND PROCEDURES, at II-18 to -28 (1985) (district courts) (on file with authors); 11 ADMIN. OFF. OF THE U.S. CTS., STATISTICS MANUAL 7–43 (1989) (courts of appeals) (on file with authors).

2 *E.g.*, Paul D. Carrington, *Crowded Dockets and the Courts of Appeals: The Threat to the Function of Review and the National Law*, 82 HARV. L. REV. 542 (1969); Kevin M. Clermont & Theodore Eisenberg, *Anti-Plaintiff Bias in the Federal Appellate Courts*, 84 JUDICATURE 128 (2000); Kevin M. Clermont & Theodore Eisenberg, *Appeal from Jury or Judge Trial: Defendants' Advantage*, 3 AM. L. & ECON. REV. 125 (2001) [hereinafter Clermont & Eisenberg, *Defendants' Advantage*]; Kevin M. Clermont & Theodore Eisenberg, *Plaintiphobia in the Appellate Courts: Civil Rights Really Do Differ from Negotiable Instruments*, 2002 U. ILL. L. REV. 947 [hereinafter Clermont & Eisenberg, *Plaintiphobia*]; Richard A. Posner, *Will the Federal Courts of Appeals Survive Until 1984? An Essay on Delegation and Specialization of the Judicial Function*, 56 S. CAL. L. REV. 761 (1983); Todd E. Thompson, *Increasing Uniformity and Capacity in the Federal Appellate System*, 11 HASTINGS CONST. L.Q. 457, 459 (1984); Judah I. Labovitz, Note, *En Banc Procedure in the Federal Courts of Appeals*, 111 U. PA. L. REV. 220, 220 n.3 (1962).

3 *E.g.*, David S. Clark, *Adjudication to Administration: A Statistical Analysis of Federal District Courts in the Twentieth Century*, 55 S. CAL. L. REV. 65 (1981); Kuo-Chang Huang, *Mandatory Disclosure: A Controversial Device with No Effects*, 21 PACE L. REV. 203, 245–68 (2000); Judith Resnik, *Managerial Judges*, 96 HARV. L. REV. 374, 396 n.85 (1982); Hans Zeisel & Thomas Callahan, *Split Trials and Time Saving: A Statistical Analysis*, 76 HARV. L. REV. 1606 (1963).

4 *E.g.*, Jason Scott Johnston & Joel Waldfogel, *Does Repeat Play Elicit Cooperation? Evidence from Federal Civil Litigation*, 31 J. LEGAL STUD. 39 (2002); Daniel Kessler, Thomas Meites & Geoffrey Miller, *Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation*, 25 J. LEGAL STUD. 233, 248–57 (1996); Joel Waldfogel, *Reconciling Asymmetric Information and Divergent Expectations*

tween demographics and case outcomes,[5] class actions,[6] diversity jurisdiction,[7] and litigation generally.[8] Other research using the AO data covers particular subject matter areas, such as inmate cases,[9] contract cases,[10] corporate litigation,[11] antitrust litigation,[12] patent litigation,[13] employment litigation,[14] constitutional tort litigation,[15] and products

Theories of Litigation, 41 J.L. & ECON. 451 (1998); Joel Waldfogel, *The Selection Hypothesis and the Relationship Between Trial and Plaintiff Victory*, 103 J. POL. ECON. 229 (1995).

5 *See* Theodore Eisenberg & Martin T. Wells, *Trial Outcomes and Demographics: Is There a Bronx Effect?*, 80 TEX. L. REV. 1839 (2002); Eric Helland & Alexander Tabarrok, *Race, Poverty, and American Tort Awards: Evidence from Three Data Sets*, 32 J. LEGAL STUD. 27 (2003).

6 *See* Arthur R. Miller, Comment, *Of Frankenstein Monsters and Shining Knights: Myth, Reality, and the "Class Action Problem"*, 92 HARV. L. REV. 664, 691–92 (1979); Note, *Developments in the Law: Class Actions*, 89 HARV. L. REV. 1318, 1325 n.30 (1976).

7 *See* Eric Helland & Alexander Tabarrok, *The Effect of Electoral Institutions on Tort Awards*, 4 AM. L & ECON. REV. 341 (2002); David L. Shapiro, *Federal Diversity Jurisdiction: A Survey and a Proposal*, 91 HARV. L. REV. 317 (1977).

8 *E.g.*, Kevin M. Clermont & Theodore Eisenberg, *Litigation Realities*, 88 CORNELL L. REV. 119 (2002); Gary M. Fournier & Thomas W. Zuehlke, *Litigation and Settlement: An Empirical Approach*, 71 REV. ECON. & STAT. 189 (1989) [hereinafter Fournier & Zuehlke, *Litigation and Settlement*]; Gary M. Fournier & Thomas W. Zuehlke, *The Timing of Out-of-Court Settlements*, 27 RAND J. ECON. 310 (1996) [hereinafter Fournier & Zuehlke, *Out-of-Court Settlements*]; Marc Galanter, *Reading the Landscape of Disputes: What We Know and Don't Know (and Think We Know) About Our Allegedly Contentious and Litigious Society*, 31 UCLA L. REV. 4, 44 (1983); Marc Galanter, *The Life and Times of the Big Six; Or, the Federal Courts Since the Good Old Days*, 1988 WIS. L. REV. 921.

9 *E.g.*, Judith Resnik, *Tiers*, 57 S. CAL. L. REV. 837, 897, 940–65 (1984); Margo Schlanger, *Inmate Litigation*, 116 HARV. L. REV. 1555 (2003); David L. Shapiro, *Federal Habeas Corpus: A Study in Massachusetts*, 87 HARV. L. REV. 321, 332, 336 (1973); William Bennett Turner, *When Prisoners Sue: A Study of Prisoner Section 1983 Suits in the Federal Courts*, 92 HARV. L. REV. 610 (1979); Note, *State Court Withdrawal from Habeas Corpus*, 114 U. PA. L. REV. 1081, 1096 n.85 (1966).

10 *See* Marc Galanter, *Contract in Court; or Almost Everything You May or May Not Want To Know About Contract Litigation*, 2001 WIS. L. REV. 577.

11 *E.g.*, Terence Dunworth & Joel Rogers, *Corporations in Court: Big Business Litigation in U.S. Federal Courts, 1971–1991*, 21 LAW & SOC. INQUIRY 497 (1996).

12 *See* Note, *Nolo Pleas in Antitrust Cases*, 79 HARV. L. REV. 1475, 1478 & n.25 (1966).

13 *See* Gauri Prakash-Canjels, *Trends in Patent Cases: 1990–2000*, 41 IDEA 283 (2001).

14 *See* Gregory Todd Jones, Note, *Testing for Structural Change in Legal Doctrine: An Empirical Decision to Litigate Employment Disputes a Decade after the Civil Rights Act of 1991*, 18 GA. ST. U. L. REV. 997 (2002).

15 *E.g.*, Theodore Eisenberg & Stewart Schwab, *The Reality of Constitutional Tort Litigation*, 72 CORNELL L. REV. 641 (1987) [hereinafter Eisenberg & Schwab, *Reality*]; Stewart J. Schwab & Theodore Eisenberg, *Explaining Constitutional Tort Litigation: The Influence of the Attorney Fees Statute and the Government as Defendant*, 73 CORNELL L. REV. 719 (1988).

liability cases.[16] These varied uses of the AO database have led to it being called "by far the most prominent" database used by legal researchers for statistical analysis of case outcomes.[17]

For many years researchers relied on the data as published in the Annual Reports of the AO Director[18] or on specific inquiries answered by the AO staff. In recent years, the FJC has made the data available in electronic form through the Inter-university Consortium for Political and Social Research.[19] This easier access to the data, together with increasing use of computers and sophisticated statistical software programs, forecasts even greater future use of the AO data.

Like many large data sets,[20] the AO data are not completely accurate. Some reports exist relating to the AO data's reliability,[21] but no systematic study of the AO's non-bankruptcy data has been published. In the course of a substantive study of federal litigation brought by prison and jail inmates, one of us began to investigate the nature and

16 *See* Theodore Eisenberg & James A. Henderson, Jr., *Inside the Quiet Revolution in Products Liability*, 39 UCLA L. REV. 731 (1992); James A. Henderson, Jr. & Theodore Eisenberg, *The Quiet Revolution in Products Liability: An Empirical Study of Legal Change*, 37 UCLA L. REV. 479 (1990).

17 Frank B. Cross, *Comparative Judicial Databases*, 83 JUDICATURE 248, 248 (2000).

18 *See, e.g.*, ADMIN. OFF. OF THE U.S. CTS., 2000 JUDICIAL BUSINESS OF THE UNITED STATES COURTS (2001) (published annually).

19 *See, e.g.*, ICPSR 8429, *supra* note 1; ICPSR 3415, *supra* note 1. For a guide to merging the Internet-available data into one large database, see Margo Schlanger, Inmate Litigation Technical Appendix, at http://www.law.harvard.edu/faculty/schlanger/projects/.

20 *See* Utah v. Evans, 122 S. Ct. 2191, 2195 (2002) (noting the existence of gaps in the census data and of conflicts in the data); David Cantor & Lawrence E. Cohen, *Comparing Measures of Homicide Trends: Methodological and Substantive Differences in the Vital Statistics of Uniform Crime Report Time Series (1933–1975)*, 9 SOC. SCI. RES. 121, 143–44 (1980) (questioning the accuracy of homicide data collected and reported by the FBI and the National Center for Health Statistics); Michael G. Maxfield, *Circumstances in Supplementary Homicide Reports: Variety and Validity*, 27 CRIMINOLOGY 671, 675–81 (1989) (criticizing the data classification methods used in supplementary homicide reports data).

21 *See* THOMAS E. WILLGING ET AL., EMPIRICAL STUDY OF CLASS ACTIONS IN FOUR FEDERAL DISTRICT COURTS: FINAL REPORT TO THE ADVISORY COMMITTEE ON CIVIL RULES 197–200 (1996) (reporting inaccuracy of class action variable). *See also* Schlanger, *supra* note 9, at 1699–1704; sources cited *infra* notes 46, 47, 55. On the related (though separate and quite different) AO bankruptcy data, see DAVID T. STANLEY & MARJORIE GIRTH, BANKRUPTCY: PROBLEM, PROCESS, REFORM 170 (1971) (noting the difficulty the AO has in getting bankruptcy officials to submit accurate data); Jennifer Connors Frasier, *Caught in a Cycle of Neglect: The Accuracy of Bankruptcy Statistics*, 101 COM. L.J. 307 (1996) (reporting on systematic analysis of AO bankruptcy statistics); and Teresa A. Sullivan, Elizabeth Warren & Jay Lawrence Westbrook, *The Use of Empirical Data in Formulating Bankruptcy Policy*, LAW & CONTEMP. PROBS., Spring 1987, at 195, 222–24 (criticizing the accuracy and utility of AO bankruptcy data).

rate of errors, exploiting a technological innovation in federal court records: the availability of docket sheets over the Internet via the federal judiciary's Public Access to Court Electronic Records project (PACER).[22] This Article follows a similar method to begin more comprehensively the process of assessing the AO data's reliability. (Relatively little is known about the accuracy of other major law-related data sets although it is clear that another source of information about thousands of cases, jury verdict reporters, vary in their accuracy.)[23]

In the large majority of districts,[24] PACER allows public Internet-based access to docket sheets recorded since 1993; in some districts other case materials are also available. To test the AO data's reliability, we compare the characteristics of cases as coded in the AO data with what we believe to be the more accurate information recorded by clerks on individual case docket sheets, as obtained through the PACER system.[25] Even though the court personnel who update case dockets are frequently the very people responsible for the AO data collection (and indeed, such personnel may often fill in many, though not all, of the AO variables on the basis of the docket sheet itself),[26] the information on the docket sheets is likely to be more reliable because it is entered in narrative form and therefore without coding issues and as litigation events occur rather than retrospectively, and because maintenance of dockets (unlike data entry for AO statistical purposes) is a core function of court clerks' office personnel.

This study looks at two large categories of cases, torts and inmate civil rights, and separates two aspects of case outcomes: which party

22 Schlanger, *supra* note 9, at 1601.

23 For discussion of verdict reporters' reliability and relevant references, see Theodore Eisenberg et al., *Juries, Judges, and Punitive Damages: An Empirical Study*, 87 CORNELL L. REV. 743, 747–48, 748 n.17 (2002).

24 Of the ninety-four federal district courts, thirteen did not have Internet-accessible records at the time we gathered data for this study. They were the Southern District of New York, Eastern District of North Carolina, Western District of Kentucky, Southern District of Indiana, Western District of Arkansas, District of Alaska, District of Idaho, District of Montana, District of Nevada, District of New Mexico, Eastern District of Oklahoma, District of the Northern Mariana Islands, and the District for the Virgin Islands. These districts accounted for approximately 11% of the federal district court docket terminated in 2000. Because several of these districts have recently adopted the PACER system, the currently unavailable districts see only 6% of the federal district court docket (again, using 2000 terminations).

25 Except with respect to some pleadings prior to the start-date of the system (usually 1993), the PACER-available dockets are generally not summaries derived from some other, lower-tech docketing system, but rather are simply the case dockets, which are now maintained electronically.

26 Telephone Interview by Margo Schlanger with Virginia Hurley, Operations Manager, U.S. District Court for the District of Massachusetts (Jan. 14, 2003).

obtained judgment and the amount of the judgment when plaintiffs prevailed. With respect to the coding for the party obtaining judgment, we find that the AO data are very accurate when they report a judgment for plaintiff or defendant, except in cases in which judgment is reported for plaintiff but damages are reported as zero. As to this anomalous category (which is far more significant in the inmate sample than in the torts sample), defendants are frequently the actual victors in the inmate cases. In addition, when the data report a judgment for "both" parties (a characterization that is ambiguous even as a matter of theory), the actual victor is nearly always the plaintiff. Because such cases are quite infrequent, this conclusion is premised on relatively few observations and merits further testing.

With respect to award amounts, we find that the unmodified AO data are more error prone, but that the data remain usable for many research purposes. While they systematically overstate the mean award, the data apparently yield a more accurate estimate as to median awards. Moreover, researchers and policymakers interested in more precise estimates of mean and median awards have two reasonably efficient options available. First, as described below, they can exclude two easily-identified classes of awards with self-evidently suspect values entered in the AO data. Second, using PACER or courthouse records, they can ascertain the true award only in the suspect cases without having to research the mass of cases. Either technique seems to provide reasonable estimates of the median award. The second technique may provide a reasonable estimate of the mean award, at least for some case categories.

Concern about the remaining degree of error depends on the case category being studied and on the research question being asked. The second technique produces accurate mean and median estimates in our torts sample. For our inmate cases, however, it proves less helpful, probably because of the small size of awards in inmate cases. Even in inmate cases, however, the suggested techniques produce estimates of the median award that are within a few thousand dollars of the true award. In short, however, for researchers interested in understanding the central tendencies of award amounts by case category, the AO data can provide usable information. We offer no conclusion on whether the data can sustain more complex modeling techniques in which damages amounts are linked to other docket and district features.[27]

Our conclusions differ notably from those based on the only other published systematic inquiry into AO federal court data. The

27 *See infra* text accompanying notes 80–82.

AO gathers bankruptcy case data using a system quite like the district court database we discuss here. And there has been some sustained examination of the accuracy of that bankruptcy data system—examination that concluded that the data are so "error ridden"[28] as to "impoverish the bankruptcy debate."[29] Indeed, leading empiricist scholars in the bankruptcy field have concluded that the AO's bankruptcy "data are utterly inadequate for policy purposes."[30] Why, you may wonder, the difference? The most obvious answers lie both in the details of how the AO bankruptcy data differ from the AO district court data, and in the need for precision. On the first issue, the bankruptcy data about which the above scholars seem especially concerned relate to *filings* rather than outcomes—in particular, the "size and nature of filed cases."[31] The AO data on such matters is entered into the computerized data system by court personnel, but the source of the information is the "face sheet" filed by debtors. The debtors (or their lawyers), it turns out, very frequently misread the form or report their assets, liabilities, or the number of their creditors incorrectly for other reasons. These incorrect entries by individual debtors and their lawyers—non-court personnel—are reportedly the source of the bulk of the error in the bankruptcy statistics.[32] The AO district court outcome data do not suffer from a similar infirmity.[33] On the second issue—the need for precision—it may be that the kind of research for which many scholars (including us) use the AO's district court data is

28 Frasier, *supra* note 21, at 308.

29 *Id.*

30 Sullivan et al., *supra* note 21, at 210.

31 Frasier, *supra* note 21, at 309.

32 *Id.* at 340–41 ("*Filer carelessness is the single, most important cause of error*"; "bankrupcty clerk transcription errors do not significantly lower accuracy rates" though "local data entry practices" do exacerbate the error rate in the "nature of case data.").

33 In the district court data, the case categorization similarly depends on the choice of the filers (if they are not pro se, see *infra* notes 39–40 and accompanying text). For some reason, however, it appears to be extremely accurate. *See infra* note 41. We are, nonetheless, inclined to be quite suspicious of district court AO data that depend too heavily on filer accuracy. We would hesitate, for example, to trust the "demand" variable, which purports to record the amount of money in controversy in each case. First, the demand variable is intended to be recorded in thousands of dollars, like the "award" variable discussed below—but even more problematically, because plaintiffs rather than clerks fill in the amount of the demand. Our guess is that, as with bankruptcy filings, small-money cases are frequently coded as big-money cases as a result. *See infra* note 58. Second, there is no requirement that plaintiffs fill in this variable except in diversity cases, which makes its availability in non-diversity cases infrequent, and non-randomly so. Third, because the amount chosen has little further bearing on the case, there is correspondingly little reason to think it has much meaning.

simply less demanding of precise accuracy than the kind of research the bankruptcy scholars would like to do with the AO's bankruptcy data. The article examining the accuracy of bankruptcy data describes 75% to 83% accuracy as "unacceptably low."[34] From our perspective, even if the district court data had a similar error rate, that description would not necessarily hold. Seventy-five percent accuracy may be plenty accurate enough—or very far from it, depending on how errors are distributed and the research questions and design. We discuss these matters in some depth below.

Part I of this Article reviews some strengths and weaknesses of the AO data. Part II uses samples of tort and inmate cases to report on the AO data's accuracy in reporting the party obtaining judgment and award levels. It then uses the information revealed about award level accuracy to estimate award levels in employment discrimination cases. Part III discusses the implications of the findings and applies the techniques developed in Part II to estimate the median trial award in all large federal case categories and to suggest the magnitude of some miscoding problems across case categories.

I. THE ADMINISTRATIVE OFFICE DATA

The AO database was designed not for research into civil justice, litigation theory, or any substantive area of law but for court administration, a purpose that helps explain much of what is both good and bad about the data.[35] Court personnel who input the data are trained centrally by the AO; various quality assurance techniques are used to increase consistency and decrease certain kinds of errors.[36] Where a variable is useful to track court workload or assign resources, it is frequently used and, we believe, probably highly reliable.[37] Accordingly, one strength of the AO data set is its completeness. Unlike any other

34 Frasier, *supra* note 21, at 340.

35 Although it is not our topic here, we, along with all of the scholars we know who have worked with the AO data, could suggest a number of seemingly easy, even trivial, changes in the way variables are gathered or coded that would make the data set even more useful for substantive research. But even as they exist, the variables and allowed values allow a good deal of useful analysis.

36 The best guide to the AO system for researchers is actually a training document. *See* CIV. STAT. REPORTING GUIDE, *supra* note 1. It is quite comprehensive and explains a number of such techniques.

37 *See* Jay Lawrence Westbrook, *Empirical Research in Consumer Bankruptcy*, 80 TEX. L. REV. 2123, 2152 (2002) (noting that the AO gathers "data, it would seem, almost entirely with an eye to accountability, workload analysis, and management generally, but with little or no attention to what data would be useful to policymakers or scholars").

data set covering the federal courts, it purports to cover every case filed. And it seems more than likely that this is indeed its coverage. Cases get entered into the database on filing, and there is a built-in check because they get entered again, on termination.

Moreover, the most basic code for researchers' use of the AO data—the case category, which identifies cases as pertaining to a specified subject matter—appears, from the limited research already done, to be highly accurate. (This too is unsurprising, because the AO depends on the accuracy of reports on filings by case category code to allocate resources among courts.[38]) For cases with counseled plaintiffs, the case category in the data set is generally based on the JS-44 Civil Cover Sheet, which plaintiffs' lawyers are required to fill out simultaneously with filings.[39] The lawyers check off a simple description of the type of case (unlike in the bankruptcy face-sheet discussed above, which requires filers to complete the more complicated—and error prone—tasks of filling in amounts and summarizing various features of their cases). Pro se plaintiffs do not typically complete the civil cover sheet, and so in pro se cases usually the court clerks seem to fill in this variable based on their own understanding of a case's subject matter.[40] In any event, we are confident that the case codes used for tort and inmate cases are not terribly *over*inclusive, because the dockets we examined for this project would have evidenced any such errors (subject matter errors were indeed apparent, but in very small numbers).[41] Because we did not audit dockets that were *not* classified by the AO data as inmate cases or tort cases, we could not, however, detect *under*inclusiveness in those categories.[42] Nonetheless, for researchers seeking to identify all federal district court cases in a certain subject matter category, it is clear that the AO database is the easiest, and perhaps the most reliable, method of doing so, provided that the

38 *See* Federal Judicial Center, New Case Weights for Computing Each District's Weighted Filings per Judgeship (1994) (memorandum on file with the authors) (setting out results of comprehensive "district court time study" used to calculate workload measures for district courts based on substantive case categories).

39 *See* JS-44 Civil Cover Sheet, *available at* http://www.uscourts.gov/forms/JS044.pdf.

40 *E.g.*, Telephone Interview with Virginia Hurley, *supra* note 26.

41 Of the 176 cases in our inmate samples, two (1.1%) were not in fact inmate cases; we did not formally audit this aspect of the tort sample, but we did not notice any errors and believe that the error rate is extremely low.

42 Underinclusiveness was, however, a correspondingly small problem in one field study in which researchers read every filed complaint in one district court during the study's time period and found only a very few civil rights cases not so characterized in the AO data set. Theodore Eisenberg, *Section 1983: Doctrinal Foundations and an Empirical Study*, 67 CORNELL L. REV. 482, 524, 535 n.237 (1982).

subject matter of interest matches one or a group of the AO case categories.[43]

The AO data include only a fairly small number of other variables, each with a limited set of permitted values. They identify the case—district, office, docket number, parties. They specify the case's timing—filing date and termination date. They elaborate its procedural history, including its "source" (e.g., original filing, inter-district transfer, remand), jurisdictional basis (e.g., federal government defendant, federal question, or diversity), procedural progress (the point in the litigation life cycle at which the case was terminated). And they set out the outcome—the nature of the judgment (e.g., money, costs, injunction), the type of disposition (e.g., by settlement, dismissal, jury verdict); the victor (plaintiff, defendant, or both), and the amount of any damages awarded. In the past few years, new variables have addressed whether the parties have counsel and the use of magistrate judges and court-annexed arbitration. As in any large and longstanding database, a number of the variables have quirks; careful use of the available documentation is essential.[44]

Overall, both field studies and other data sets confirm the general picture of district court litigation suggested by the AO data, although as already described, bankruptcy scholars have questioned the AO's bankruptcy data's reliability,[45] and some aspects of the district court data have also been challenged.[46] For example, a field study comparing the characteristics of litigation as suggested by the AO data with the characteristics suggested by case-by-case inspection of records in courthouses confirmed findings based on AO data that constitutional

43 Searching for cases on a given subject-matter seems likely to be more rather than less error-prone than the AO database, which uses the expertise of litigants and court clerks to classify cases. A study of civil rights cases filed in one district court found that analysis of individual complaints by hand-searching for them in courthouse records missed approximately 20% of civil rights cases properly identified in the AO data as civil rights cases. *Id.*

44 The most comprehensive codebooks are available as Parts 94 and 57 of ICPSR 8429, *supra* note 1. *See also id.* pt. 117; Schlanger, *supra* note 9, at 1699–1704.

45 *See supra* notes 21, 28–34 and accompanying text.

46 In particular, the class action variable is authoritatively reported to have been quite unreliable, at least for a substantial period of time. *See* WILLGING ET AL., *supra* note 21, at 197–200. In addition, Kimberly Moore has questioned the usefulness and reliability of the AO data in patent cases. *See* Kimberly A. Moore, *Judges, Juries, and Patent Cases—An Empirical Peek Inside the Black Box*, 99 MICH. L. REV. 365, 381 (2000) [hereinafter Moore, *Judges*] (discussing limitations of the AO data for analysis of patent cases); Kimberly A. Moore, *Xenophobia in American Courts*, 97 NW. U. L. REV. (forthcoming 2003) (manuscript at 37, on file with authors) (questioning reliability of AO "judgment-for" data in patent cases).

tort plaintiffs fare relatively poorly at trial compared to other plaintiffs, and also obtain significantly fewer money judgments or settlements.[47] The field data also confirmed the AO data on amounts awarded in the sense that both sources suggested that perceptions about damages in constitutional tort litigation are overstated.[48] And a more recent study began the process of comparing AO data with Internet-accessible dockets, and confirmed that much of the AO data is consistent with dockets.[49]

Other data sets supply additional evidence relating to the AO data's reliability. For example, plaintiffs' rates of prevailing at trial appear to be quite consistent across data sets. The AO data suggest that plaintiffs in medical malpractice and products liability cases have low trial win rates relative to plaintiffs in most other classes of tort and contract litigation.[50] These low AO-data win rates are consistent with win rates in studies of products liability by the RAND Institute for Civil Justice studies of litigation, with studies of medical malpractice litigation,[51] with General Accounting Office data,[52] with the National Center for State Courts data obtained from state court clerks' offices,[53] and with jury verdict reporters.[54]

The AO data's reliability for award amounts is less secure.[55] It has been thought for years that the amounts are questionable, but the

47 Eisenberg & Schwab, *Reality*, *supra* note 15, at 680.

48 *Id.* at 684.

49 Schlanger, *supra* note 9, at 1699–1704.

50 Kevin M. Clermont & Theodore Eisenberg, *Trial by Jury or Judge: Transcending Empiricism*, 77 CORNELL L. REV. 1124, 1137 (1992).

51 NEIL VIDMAR, MEDICAL MALPRACTICE AND THE AMERICAN JURY: CONFRONTING THE MYTHS ABOUT JURY INCOMPETENCE, DEEP POCKETS, AND OUTRAGEOUS DAMAGE AWARDS 39 (1995) (noting the low win rates at trial for medical malpractice cases).

52 U.S. GEN. ACCOUNTING OFF., PRODUCT LIABILITY: VERDICTS AND CASE RESOLUTION IN FIVE STATES, H.R. DOC. NO. 89-99, at 24 (1989).

53 *E.g.*, CAROL J. DEFRANCES & MARIKA F.X. LITRAS, CIVIL TRIAL CASES AND VERDICTS IN LARGE COUNTIES, 1996, *in* BUREAU OF JUST. STAT., BULLETIN 1 (Sept. 1999), *available at* http://www.ojp.usdoj.gov/bjs/pub/pdf/ctcvlc96.pdf (last visited Mar. 22, 2003); CAROL J. DEFRANCES ET AL., CIVIL JURY CASES AND VERDICTS IN LARGE COUNTIES, *in* BUREAU OF JUST. STAT., SPECIAL REPORT 1 (July 1995), *available at* http://www.ojp.usdoj.gov/bjs/pub/pdf/cjcavilc.pdf (last visited Mar. 22, 2003).

54 STEPHEN DANIELS & JOANNE MARTIN, CIVIL JURIES AND THE POLITICS OF REFORM 82–83 (1995).

55 *See* Theodore Eisenberg, John Goerdt, Brian Ostrom & David Rottman, *Litigation Outcomes in State and Federal Courts: A Statistical Portrait*, 19 SEATTLE U. L. REV. 433, 439 n.13 (1996) ("[T]he federal method of recording awards may result in some awards being inflated."); Moore, *Judges*, *supra* note 46, at 381; Schlanger, *supra* note 9, at 1703; Stewart J. Schwab, *Studying Labor Law and Human Resources in Rhode Island*, 7 ROGER WILLIAMS U. L. REV. 384, 394–95 (2002) (discussing the inaccuracy of award data in the AO database).

precise nature and extent of the likely error has not been known. Several of the problems we now explore stem from the decision, made in an era of more expensive computer memory and storage space, to allow only four digits to record the amount recovered in a civil action.[56] This limitation means that the highest number that can be entered in the AO database is "9999," so award amounts are supposed to be recorded in thousands of dollars. A number of errors have resulted. First, even without any inputting mistakes, the AO's data design allows for award amounts of up to only $9,999,000. Logically, this suggests that AO reports of award amounts should be understated because award amounts in excess of $9,999,000 are deflated.[57] Cutting the other way, towards the problem of AO over-statement, are the systematic errors introduced by the system of recording award amounts in thousands of dollars. A $1000 award should be recorded as a "1" in the AO's amount field. But court personnel might easily instead record the $1000 as "1000," which is intended by the AO to be interpreted as an award of $1,000,000.[58] Moreover, the need to round actual award figures to thousands creates imprecision, and might even mean that small awards are omitted from the system.[59] Finally, and unrelated logically, the figure 9999 may also be used by court clerks in other ways, such as to indicate missing data. (Many other AO vari-

56 ICPSR 8429, *supra* note 1.

57 Eisenberg & Schwab, *Reality*, *supra* note 15, at 686 nn.187–88.

58 The AO itself warned in 1995: "Researchers should also be aware that the requirement that the Demand and Amount Received fields be reported in thousands of dollars is sometimes not followed correctly causing the information for those fields to be reported inaccurately. Although the problem is known the level of inaccuracy is undetermined." ICPSR 8429, *supra* note 1, pt. 94, at xxi.

59 In one place in the training manual currently used to instruct court personnel on data entry, the AO directs that any award under $500 be entered as zero. CIV. STAT. REPORTING GUIDE, *supra* note 1, at D:1. At the same time, however, the computer system is programmed to produce an error report whenever the "nature of judgment" in a plaintiffs' victory is a monetary award but the award entered equals zero. *Id.* at 4:4, 5:1. (Error reports can be overridden, but it seems likely that clerks avoid the error report by coding awards between $1 and $499 as 1; we have seen many such cases, and very few, if any, coded as the AO's manual suggests.) Prior to 1987, when the coding system was generally overhauled, the clerks apparently were instructed to code any award of less than $1000 as zero. *See* ICPSR 8429, *supra* note 1, pt. 94, at 62; *id.* pt. 57, at 49. We are not sure what the instruction was after 1987 but before 1999. In any event, interviews together with examination of the 1993 inmate data examined here along with a different inmate case sample, from 2000 terminations, demonstrate that court clerks have at least frequently and perhaps consistently used "1" to indicate any damages amount from $1 to $1499 since at least 1993. *See infra* Table 7; *see also* Telephone Interview with Virginia Hurley, *supra* note 26. To us, this makes the most substantive sense, because for low-damage cases, what is most important to capture is the distinction between some and no damages.

ables use repeated 9s as special codes.)[60] The possible confusion generated by the four-digit limitation, together with the differing uses of 9999 in the amount field, make it difficult to know precisely what to make of the amounts reported in the AO data.

II. Comparing the AO Data with Docket Sheet Data

To assess errors in the AO data set, we compare AO data with what we believe to be the more reliable docket sheets maintained in individual cases. We have not undertaken to travel to a variety of district courts and examine the actual case records (pleadings, orders, and so on) or to discuss the cases with the parties or lawyers. Rather, we have used PACER to gather electronic docket sheets, and our research assistants (checked by us) have entered data from the docket sheets into a new database.[61] The comparisons between the AO data and the docket sheet data provide a general sense of the magnitude and direction of the error in the AO data and, we hope, suggest reasonable approaches to correcting or interpreting the AO data.

A. The Data

The samples used here are a bit eclectic, reflecting the current interests of the co-authors, the availability of docket-sheet data via PACER, and limits on time and financial resources.

We used two different samples. To construct the first sample, we began with every tort case[62] terminated after trial in federal district court between January 1 and September 30, 2000.[63] According to the

60 *See, e.g.*, ICPSR 8429, *supra* note 1, pt. 94, at 12, 108, 182; *id.* pt. 57, at 8, 9, 33, 41–45, 47, 50–53.

61 The initial coding of the PACER data was done by research assistants without any access to the AO coding, to avoid biasing the results. For the inmate sample, one of us reviewed each entry against the PACER dockets; for the torts sample coding and results were reviewed in periodic meetings with our assistants.

62 The AO tort case categories (followed by their code values) are: Airplane Personal Injury (310); Airplane Product Liability (315); Assault, Libel, and Slander (320); Federal Employers Liability (330); Marine Personal Injury (340); Marine Product Liability (345); Motor Vehicle (350); Motor Vehicle Product Liability (355); Other Personal Injury (360); Workers' Comp./Industrial Accident Board (361); Personal Injury/Medical Malpractice (362); Personal Injury/Product Liability (365); Asbestos Personal Injury Product Liability (368); Other Fraud (370); Truth in Lending (371); Other Personal Property Damage (380); and Property Damage Product Liability (385). *See* ICPSR 8429, *supra* note 1, pt. 93.

63 To be precise, the first sample is every tort case with a "procedural progress" code indicating termination after a judge or jury trial, and with a specified victor, between January 1, 2000 and September 30, 2000. There are an additional eighty-five cases without information as to the victor in the AO data set. According to their

AO data, 786 such cases terminated in ninety of the ninety-four federal districts. We then excluded 105 cases in the districts that do not participate in PACER or in which PACER-based outcomes were otherwise unavailable.[64] The total number of cases included in panel A of Table 1, below, is thus 681.

The second sample has two parts. The first part, described in panel B-1 of Table 1 below, includes every available inmate civil rights case[65] terminated in federal district court in fiscal 1993[66] in which a positive plaintiff's award was recorded in the AO data.[67] The AO recorded 142 such cases, in fifty-eight district courts; we were able to obtain the relevant docket information for 126, from fifty-five courts.[68] The sample's second part, described in panel B-2 of Table 1 below, explores an oddity in the data: the AO data includes 330 inmate cases terminated in 1993 in which the amount of the judgment is coded as zero but the plaintiff is nonetheless coded as the victor.[69] We con-

"disposition" codes, these seem largely to be cases that settled or were otherwise disposed of without a verdict despite a trial having commenced or been completed. But it is not implausible that some of them were in fact tried to final judgment but for some reason the court clerk either did not know or failed to enter the victor. As discussed below, we also looked at a separate sample of cases terminated between 1996 and 1999; these too were tort cases tried to judgment, but limited to diversity cases. *See infra* note 78.

64 The number of cases omitted from districts not participating in PACER follows in parentheses after the name of the applicable district: Southern District of New York (39), Eastern District of North Carolina (3), Western District of Kentucky (9), Southern District of Indiana (3), Western District of Arkansas (6), District of Alaska (1), District of Idaho (3), District of Montana (3), District of Nevada (4), District of New Mexico (8), and Eastern District of Oklahoma (8). In addition, there were eighteen cases from scattered districts for which docket information was not available or in which we could not classify the outcome for some other reason.

65 The inmate case sample includes two AO inmate case categories—Prisoner Civil Rights (550) and Prison Conditions (555). See Schlanger, *supra* note 9, at 1699–1700, for a discussion of these two categories.

66 We follow the AO and use the federal fiscal year, October 1, 1992 to September 30, 1993.

67 For purposes of comparing the inmate sample with the tort sample, note that AO data indicates that of these 142 cases, only about half involved trials (eighty-one are coded with dispositions by jury or judge verdict; eighty-two are coded as resolved "during" or "after" jury or judge trials; and seventy-seven meet both criteria).

68 For the inmate sample, though not the torts sample, we made efforts to obtain photocopied docket sheets from court clerks' offices for cases in the districts that do not participate in PACER. In some cases, the clerks' offices were unable to identify the docket; in others, the records were unavailable for a variety of reasons. We were able to obtain 129 docket sheets; in three of them, the requisite information could not be gleaned from the docket sheet.

69 *See supra* note 59.

structed a 20% random sample of these anomalous case records, attempting to obtain sixty-seven of them from thirty-eight district courts. Of these, we were able to actually get docket sheets for forty-seven, from twenty-eight courts, and to glean the relevant information from all but six.[70]

Given the nature of our samples, a cautionary note is in order. We are reasonably confident that our results are valid for the case categories and times we study, at least for cases terminated after trial. Applying the findings to data sets covering different time periods, different case categories, and different procedural postures, as we ourselves do below, should be done with the samples' limitations in mind.

B. Win Rates

Because of differences in the tort and inmate samples, we explore accuracy in reporting judgments separately. Table 1 explores the rate of agreement between the AO coding of whom judgment was entered for and what inspection of individual docket sheets reveals.

Panel A describes the tort cases terminated in 2000 (again, cases with AO-reported judgments after trial). Its 313 AO-coded plaintiffs' judgments include 253 cases with AO-reported judgments for positive amounts and sixty cases in which the judgment was reported as zero, even though the plaintiff was reported as the victor. The seventeen cases coded with judgments coded for "both plaintiff and defendant" include a slightly higher proportion of awards reported as zero—six. We include all of these zero-award cases in Table 1, but will address them separately in analyses of award amounts.

Panel B-1 covers the 1993 inmate cases with AO-coded judgments for positive amounts; panel B-2 covers those with judgments entered by the AO as being equal to zero.[71] Because the inmate sample was constructed only from cases in which plaintiffs were listed as at least

70 The cases were selected using a random number generator, and we did not resample to make up for unavailable dockets. The poor retrieval rate is not surprising because the distribution of anomalous dockets across districts is extremely disproportionate; with a very large number (24%) from districts that did not participate in PACER, even though those districts accounted for a much smaller proportion (11%) of the inmate docket terminated in 1993. The Southern District of New York, in particular, reported forty-five of these cases in 1993 (about 14% of the total amount, though the district had less than 3% of inmate terminations that year) and is the source of much of the anomaly.

71 The last comprehensive codebook about the database, published in 1997, explains that a value of zero means "missing," ICPSR 8429, *supra* note 1, pt. 94, at 62, though this comment is not repeated in the more recent codebooks or in the training materials currently used by court clerks. *See supra* note 59.

partial victors, cases reporting judgments for defendants are not included. In addition, the panel B-2 data are drawn from samples. The column reporting judgments for plaintiffs is a sample of thirty (of 185) 1993 inmate trials with judgments for plaintiffs and zero-awards. The column reporting judgments for both is a sample of eleven of 145 cases from 1993 inmate trials with judgments for both and zero-awards.

Each panel shows all the sampled permutations of outcomes, where the AO records a victory for plaintiff, defendant, or both, and the PACER-obtained docket can be classified as for plaintiff or defendant. The shaded squares are those in which our two sources unambiguously agree.

TABLE 1. ACCURACY OF AO CODING OF PARTY OBTAINING JUDGMENT, TORT TRIALS AND INMATE CASES

PANEL A—All available tort trials terminated in 2000 (n = *681*) AO Judgment for—n (% of cases)			
PACER judgment for:	Plaintiff	Defendant	"Both"
Plaintiff	313 (46.0%)	10 (1.5%)	17 (2.5%)
Defendant	3 (0.4%)	337 (49.5%)	1 (0.1%)
PANEL B-1—All available inmate cases terminated in fiscal year 1993, AO Award > 0 (n = *126*) AO Judgment for—n (% of cases)			
PACER judgment for:	Plaintiff	Defendant	"Both"
Plaintiff	98 (77.8%)	—	24 (19.0%)
Defendant	4 (3.2%)	—	0
PANEL B-2—Sample of inmate cases terminated in fiscal year 1993, AO Award = 0 (n = *41*) AO Judgment for—n (% of cases)			
PACER judgment for:	Plaintiff	Defendant	"Both"
Plaintiff	5 (12.2%)	—	2 (4.9%)
Defendant (includes voluntary dismissals)	25 (61.0%)	—	9 (22.0%)

SOURCE: ICPSR 8429, *supra* note 1, supplemented by PACER docket research. The columns show who won according to the AO data; the rows show who won according to the more accurate Internet-available docket sheets. Shaded squares are unambiguously in agreement.

For the cases in panels A and B-1, the AO data prove extremely accurate. In panel A, there is agreement with the PACER-based data in about 95% of the cases. More than half of the errors, if they are

properly even considered errors, arise from the small portion of the docket in which the AO "judgment for" variable is coded "both," meaning that judgment was entered for both plaintiff and defendant. Such cases amount to 2.5% of tort trial verdicts in 2000 (or about 5% of the tort cases with a full or partial plaintiffs' victory coded). The portion of the sample in which victory is recorded for the plaintiff but the amount of damages is coded zero do not present a different error pattern. The chart does not separate such cases out—but of the sixty, just one (1.6%) has an incorrect "judgment-for" code. In panel B-1, judgments coded by the AO as simple plaintiffs' victories are similarly accurate. Nearly all the arguable errors are from the "judgment for" equals "both" category, which forms a far larger proportion of the inmate sample than of the torts sample[72]—19% of the inmate cases with full or partial plaintiffs' victory coded. As in panel A, our reading of case dockets in cases so coded in the AO data set cannot distinguish them from the plain vanilla plaintiffs' judgments.[73]

We consider the errors in the "judgment for both" category unsurprising, because the intended meaning of "both" is unclear. The AO apparently does not provide any guidance to court personnel on this point.[74] Judgment for "both" could mean simply that at least one defendant beat liability on at least one count of the complaint. But in that case, one would expect a far higher percentage of cases to be so coded; it simply cannot be the case that victorious plaintiffs win a victory on *all* counts against *all* defendants in all but 5% of their judgments. So if this is the intent, then "judgment for both" is being used far too little. Alternative interpretations of "judgment for both" are possible—for example, the category would make some sense if applied to the small group of cases in which defendants bring counterclaims and both the defendant and the plaintiff win on liability. Or the code might signal the presence of a pyrrhic plaintiff's victory—a case in which the plaintiff technically wins but is awarded only nominal damages, or some similar outcome. There is, however, little sign

72 The twenty-four such cases comprise about 2% of all inmate trial judgments; defendants won 900 of the trial judgments in inmate cases terminated in 1993.

73 Of the seventeen tort cases in 2000 coded as judgment for both, twelve were in the Fifth Circuit. Of the twenty-six similarly coded inmate cases, twelve were in the Eighth Circuit. We have no particular reason to think that this is anything other than random variation because when we looked at the entire AO data set of judgments from 1987 (when the AO began using the current coding system) to 2000, we found that the only notable outlier in the use of the "both" code was the Ninth Circuit, in which district courts disposed of 16% of the cases, but coded 32% of the "boths."

74 *See* CIV. STAT. REPORTING GUIDE, *supra* note 1, at 3:21, D:2 (indicating the code for "both plaintiff and defendant" with no further explanation).

in the dockets that any such guess actually matches how the code is being used. We are, in short, unable to come up with any consistent interpretation of its meaning. In the absence of a theory for what "both" *should* mean, it is hard to say that any use of it is erroneous. Still, researchers would be well advised to consider counting cases in which "judgment for" equals "both" as plaintiffs' victories, which is how nearly all appear to us.

Panel B-2 presents a far less favorable view of the accuracy of the AO data. Unlike the results in the torts sample, for inmate cases, in the anomalous category of purported plaintiffs' victories with zero (or missing)[75] damages, the AO's "judgment for" data seem to be too error-ridden to be of use. Further exploration is clearly required; we offer some preliminary thoughts here. First, in both tort and inmate cases, and over the federal docket taken as a whole, the problematic coding (that is, the conjunction of judgment for plaintiff and a zero-award) seems to be considerably more common in cases terminated without trials than in those terminated after trial.[76] Researchers looking at trial judgments have somewhat less to be worried about than those looking at overall, or just non-trial, outcomes. (Among the portion of the anomalous cases that had trials, however, the problem remains; of the forty-one cases in Panel B-1, eleven are coded by the AO as involving trials, of which seven have incorrect "judgment for" codes.) Second, the AO's coding in these anomalous cases seems to be erroneous in different ways in our two samples. In the torts sample, as already stated, there is no problem in the "judgment for" variable. Nonetheless, the anomaly does flag somewhat consistent error. As Table 2 demonstrates, that error lies in the "award" variable, which is correctly coded in only about half of the sixty cases we were able to check. (In half the cases, that is, the plaintiff really did win, and without any damages—these are declaratory judgment cases; in the other half, the plaintiff won damages incorrectly coded as not present.) In the inmate sample, the error lies in the "judgment for" variable. It may be, however, that inmate cases, with their extremely low rate of success for plaintiffs, are exceptional in this respect. Because of the varying relation between zero-award cases and error patterns, we report in Part III the percentage for each major case category of plaintiffs' awards for an amount of zero.[77]

---

75 *See supra* note 59.

76 In all fiscal year 2000 terminations, for example, the anomalous coding was present in 27% of cases ended after trial and 37% of other cases in which the AO data includes "judgment-for" values.

77 *See infra* Part III, tbl.9.

The particular lesson we draw is that the AO "judgment for" variable is reliably accurate except where something else looks suspicious—such as, for example, what looks like a large proportion of take-nothing plaintiffs' judgments.[78] The more general lesson is one of cautious optimism. The AO data contain their own error checks; different variables can be examined in relation to each other to assess the likelihood of error. So far as we have been able to determine, where the values make sense and seem consistent across variables, the data are very good indeed. But where there seems to be an anomaly, researchers would be foolhardy not to inquire further. And the availability of PACER dockets allows such inquiry with relative economy.

C. Awards

This section first discusses the accuracy of the AO reports of awards following trials in tort cases and in inmate cases. It then devotes separate attention to the import of these errors, looking at award means and medians within our samples. Finally, it applies the techniques developed in assessing tort and inmate case reliability to estimate awards in employment discrimination cases.

1. The Frequency and Nature of Errors in Award Amounts

Table 2 reports on error rates in the AO-reported awards in our 2000 tort sample, as checked against dockets available from PACER. Different error rates and types are associated with different awards, and (as will become evident in the discussion of inmate cases below) these associations may vary with the kind of case.

The table summarizes errors in columns, by error type. Its fourth column shows that, in our tort trial sample, a plurality of classifiable errors relates to rounding. These can be simple arithmetic mistakes; where an award is rounded up instead of down, for example. Other times rounding errors exist when clerk's office personnel seem to use less precision than the system allows; where, for example, a damages

78 We repeated the same analysis on a smaller sample of tort cases terminated from 1996 to 1999. Although these were a nonrandom set of cases (limited to districts in the First, Second, and Third Circuits and with different proportions of the trial sample for different years), the 1996–1999 data allow for a partial check on the results reported in the text. Our results for the second sample strongly confirm our finding of a very high level of accuracy for the AO's "judgment for" variable.

TABLE 2. ERRORS IN AO AWARD VALUES, YEAR 2000 TORT TRIALS[79]

AO award range (in 1000s)	Total n	Errors n (% of sample)	Type of error—n (% of errors)		
			Rounding	Digit	Other
0	61	31 (51%)	1 (3%)	0 (0%)	30 (97%)
1	2	0 (0%)	0 (0%)	0 (0%)	0 (0%)
2–199	130	38 (29%)	19 (50%)	1 (3%)	14 (37%)
200–9998	90	37 (41%)	13 (35%)	2 (5%)	31 (84%)
9999	24	19 (79%)	0 (0%)	0 (0%)	19 (100%)
TOTAL	307	125 (41%)	33 (26%)	3 (2%)	94 (76%)

SOURCE: ICPSR 8429, *supra* note 1, supplemented by PACER docket research. The rows group cases by the damages award recorded in the AO data set. The columns summarize the rate and accuracy of the AO coding.

award of $357,914 is coded as 360 (which should mean an award between $359,500 and $360,499). While the amounts subject to rounding error can be several thousand dollars or more, these are not errors that should greatly concern most analysts, because they are necessarily either small or small in relation to the actual award, and usually both. Digit errors, which we define to occur where an award is misstated in the AO data because of the need to input the amount in thousands of dollars, could pose a larger problem for research use of the AO data, but such errors are very rare in our tort trial sample. The final column aggregates a variety of other kinds of errors: typos, partial awards, and so on. Some may well not be errors at all, but rather disagreements between the two data sources about the proper way to categorize different kinds of awards (e.g., pre-judgment interest, costs).

As discussed above, awards coded by the AO as zero in which plaintiffs are simultaneously coded as victors merit special mention. The dockets set out the amount of the judgment in sixty-one of the seventy-seven such cases. Table 2 shows that half of those awards were correctly coded; they were nearly all declaratory judgments.[80] The

79 The cases included meet the following criteria: They are coded by the AO as terminating after a trial, with a judgment for plaintiff or for both, and we were able to obtain actual award information for them. Additional award errors, not presented in Table 2, stem from errors in the "judgment for" code—of the ten cases identified in Table 1's panel A as erroneously coded by the AO as defendants' victories, eight compounded the error with awards recorded as zero.

80 Twenty-nine of the thirty were in related cases in which an insurance company apparently successfully sought a declaratory judgment that the defendants did not have asbestosis. *See, e.g.*, Liberty Mut. Ins. Co. v. Carr, No. 97-125 (E.D. Tenn. filed Mar. 12, 1997); Liberty Mut. Ins. Co. v. Seabolt, No. 97-105 (E.D. Tenn. filed Mar. 12,

other half, incorrectly coded, consists of ordinary damage awards for plaintiffs, in varying (non-zero) amounts. (In addition, one of these cases was actually a defendant's victory.)

Finally, the cases in which the award is coded 9999 prove to be of two types. As intended by the AO, in a portion of our sample (five of twenty-four cases), 9999 indicates awards of $9.999 million or more. But the large majority are not cases in which the award is too high for proper coding in the AO system, but are rather errors.

Similar analysis can be applied to the sample of inmate cases from 1993, though the results are quite different. Table 3 groups the inmate cases by AO award range.

TABLE 3. ERRORS IN AO AWARD VALUES, FISCAL YEAR 1993 INMATE CASES[81]

AO award range (in 1000s)	Total n	Errors n (% of sample)	Type of error—n (% of errors)		
			Rounding	Digit	Other
1	52	1 (2%)[82]		0 (0%)	1 (100%)
2–999	48	13 (27%)	1 (8%)	8 (62%)	4 (31%)
1000–9998	17	16 (94%)[83]	1 (6%)	10 (63%)	5 (31%)
9999	5	5 (100%)		0 (0%)	5 (100%)
TOTAL	122	35 (29%)	2 (6%)	18 (51%)	15 (43%)

SOURCE: ICPSR 8429, *supra* note 1, supplemented by docket research. The rows group cases by the damages award recorded in the AO data set. The columns summarize the rate and accuracy of the AO coding.

As for Table 2's tort sample, the errors in the inmate case sample are summarized in columns by error type. One thing Table 3 demon-

1997). The other case was a take-nothing plaintiffs' judgment, in which the jury found fault but no damages.

81 The sample consists of all available inmate civil rights cases terminated in fiscal 1993, with AO coding for a positive award for plaintiff or "both." Five cases are omitted because the docket sheet did not include relevant information; thirteen because no docket sheet could be obtained, and two because they were not inmate cases at all.

82 According to their docket sheets, thirty-seven of the fifty-two cases in the first row have awards between $1 and $499, which, according to one of the directions the AO currently gives court personnel, should be coded with a zero award. For the reasons explained above, *supra* note 59, we think these cases are best considered as non-errors, but we report them in this note for the sake of complete transparency. In addition, awards on the breakpoint of rounding (for example, $1500) are not treated as erroneous whichever way they are rounded (for example, coded as either "1" or "2").

83 Some case entries reported in this row have errors of multiple types and are therefore listed more than once.

strates is a meta-point we want to emphasize: the AO data can vary a great deal across case categories. Compared to tort cases, the errors in inmate cases have a quite different feel. Rounding errors are rarely present here, perhaps because the awards are lower (as the awards are meant to be coded in thousands, there is less rounding to do). Instead, in this sample, a majority of errors are digit errors, which are common. These are likely to have a large impact on the accuracy of summary statistics from the AO data.

Table 2's total error rate of 41%, and Table 3's total error rate of 29%, each demonstrate that researchers' caution about errors in the AO's award data is merited. The absolute rate of error is high. But even a very high *rate* of error would not matter for most research purposes if errors are consistently small. And for some purposes, even large errors would not pose an obstacle to using the AO data if those errors were symmetrically distributed around zero (so that they would tend to cancel each other out). Thus we next consider the magnitude and distribution of errors.

One way to assess the size of an error is as a percentage of the actual damage award. Using this approach for the torts sample, it turns out that although errors are often small (6% or less of the actual value for about a quarter of the errors), they are as often equal to 100%, and nearly as often quite large (200% or more of the actual value for about a fifth of the errors, and more than 1000% for one-tenth of the errors). The median error amount is 81%. This may overstate error magnitude, however: leaving out the anomalous cases in which the award is coded as zero (for which the error amounts are, of course, 100%)[84] the median error among cases with errors is just 17%—quite small.

For assessment of error direction as well as magnitude, it's useful to consider a simpler error index—the true amount minus the AO-coded amount. Using this error figure, at least in our torts sample, errors again have a non-normal distribution, several aspects of which are worth noting. First, small errors are the most prevalent. Among these small errors there is a slight overrepresentation of *negative* errors—AO *under*statement of true awards. However, there are a fairly large number of very large errors, and these demonstrate substantial overrepresentation of *positive* errors—AO *over*statement of true awards. The 9999s and digit errors are the bulk but not all of these.

The inmate sample looks somewhat different. Using either error index, about half the errors are quite small, but about half (the digit

84 For the reasons discussed *infra* note 88 and accompanying text, omitting these cases seems unlikely to bias assessments of award amounts.

errors, mostly) are very large. And nearly all the errors are overstatement of true awards.

All this is conceptually simple. A significant portion of the awards seem to be erroneously coded; the 9999 code only rarely means what it purports to and there are, depending on the case category, more or fewer rounding, "digit," and other errors. In the torts sample, the magnitude of errors tends to be small, but is often quite large, with most large errors overstating actual damages. In the inmate sample, nearly all of the errors overstate actual damages—half by a small amount, half by a large amount. What is far harder is to assess how much all this error matters for actual research uses, and whether there are methods by which researchers might work around errors to obtain useful information from the AO database. The next section moves to this issue.

2. Research Implications of the AO Award Error Pattern

Researchers have tended to use the AO data on award amounts in two distinct ways. Some users of the data—especially recent users—have been interested in modeling quite complex litigation dynamics. Such researchers explore, for example, the prevalence of settlement, and its relationship with other docket features,[85] or the impact of demographic factors on award levels,[86] or the decision to appeal and the outcome of the appellate process.[87] For such uses, the devil may well be in the details. That is, whether the degree of error in award amounts undermines the AO data's ability to sustain this kind of research turns on the fine details of research design and model specification. All we can do here is offer a warning to such researchers to be aware of the issue and design their studies accordingly.

---

85 *See* Fournier & Zuehlke, *Out-of-Court Settlements*, *supra* note 8; Fournier & Zuehlke, *Litigation and Settlement*, *supra* note 8.

86 Helland & Tabarrok, *supra* note 5; Helland & Tabarrok, *supra* note 7.

87 *See, e.g.*, Clermont & Eisenberg, *Defendants' Advantage*, *supra* note 2; Clermont & Eisenberg, *Plaintiphobia*, *supra* note 2. These two studies find that appellate courts reviewing tried cases tend to be more favorable to defendants than to plaintiffs. Clermont & Eisenberg, *Plaintiphobia*, *supra* note 2, at 952. Appellate courts are more likely to reverse a trial victory for plaintiffs than for defendants. *Id.* But the studies use individual case data for award amounts and therefore could be affected by the error patterns reported here. Re-running the analysis used in those studies with (1) 9999 cases separately coded as such (award amounts in cases with awards of 9999 were treated by using a dummy variable), and (2) exclusion of 9999 cases, yields no material difference in results. In the models that treat 9999 cases separately, such cases are treated by using a dummy variable. The variable was not statistically significant in any of the models.

Other researchers use the AO data not for econometric modeling, but for the light shed on the political economy of particular flavors of litigation. These scholars and policymakers seek, that is, to understand the central tendencies of particular portions of the federal docket. The reliability problems reported above pose a more manageable challenge to this kind of work—one on which we may be able to make some progress in this Article. We next consider how such researchers might work around the AO award errors to obtain useful information from the AO database. To assess the overall impact of the errors, Table 4 reports on the mean tort awards in our 2000 tort sample, again as checked against dockets available from PACER; Table 5 looks at the distribution of awards, including the median; Tables 6 and 7 present the corresponding data from the inmate sample.

In comparing actual to AO-reported award levels, a preliminary decision must be made about what to compare with what. For our purposes, what seems most sensible is to compare the *apparent* universe of awards with the *true* universe of awards (rather than, as in Tables 3 and 4, comparing the AO values of some given set of cases with the PACER value of the same set of cases). More precisely, a researcher using just AO data to compute mean awards by case category or time period, for example, would, we believe, most reasonably proceed as follows: (1) limit the sample to cases in which a judgment was, according to the AO data, entered for plaintiff or for both plaintiff and defendant, and (2) further limit the sample to those cases in which the coded award for plaintiff exceeded zero. The second limitation is based on the reasonable assumption that awards of zero in damage actions won by plaintiffs are rare—so that the zeros are either erroneous, signify missing data, or mark the cases as injunctive or declaratory judgment cases rather than damage actions.[88] To understand how much the error in the AO data matters, then, the most

88 Prior to our work here, it might have been thought reasonable for some purposes to include zero awards in the mean computation on the ground that they are not known to be erroneous. Our work cautions strongly against this approach. Moreover, at least as far as our torts sample indicates, leaving out plaintiffs' victories with AO-coded zero awards is unlikely to bias the result because, statistically comparing the actual awards from the anomalous zero-award cases to those from the larger non-anomalous portion of the docket, one cannot reject the hypothesis that the cases actually present observations from the same distribution. More specifically, one cannot reject the hypothesis that the medians of the two distributions are equal (the p-value for a Mann-Whitney test is .55); one cannot reject the hypothesis that the overall distributions are equivalent (the p-value of a Kolmogorov-Smirnov test is .81); and one cannot reject the hypothesis that the means of log-transformed distributions are equal (the p-value for a T-test is .44 assuming equal variances and .48 not assuming equal variances).

useful comparison is between the results of the above reasonably constituted set of awards and the true set, comprising all the damage actions in which plaintiffs are recorded in docket sheets as the victors, regardless of the victor or the level of award noted in the AO database. This compares two slightly different sets of cases—but that is because the purpose is not to check the reliability of individual data points (an issue fully canvassed above) but rather to assess the impact of the errors on assessments of the distributional tendencies of awards. The tables following thus take this approach.

a. Tort Awards

TABLE 4. ACCURACY OF AO AWARD MEANS (AMOUNTS IN THOUSANDS), YEAR 2000 TORT TRIALS RESULTING IN PLAINTIFFS' JUDGMENTS

	Mean	95% confidence intervals	n
1. True mean	816	452–1180	286
2. AO mean, AO award > 0 (and true award found)	1387	1009–1765	246
3. True mean of AO zero-awards	618	129–1108	40
4. True mean of AO 9999 awards	4717	696–8738	24
5. AO mean, AO award > 0, excluding 9999 cases	456	317–595	222
6. Replace only 9999 awards with true data	872	453–1290	246

SOURCE: ICPSR 8429, *supra* note 1, supplemented by PACER docket research. The first row shows the true mean award as determined by inspecting PACER dockets. The second and fifth rows show mean AO award codes. The third and fourth rows show true mean awards for cases for which the AO award code is "0" and "9999." The final row combines true and AO awards, replacing AO codes only for AO awards coded 9999 and omitting AO awards coded zero.

Table 4's first row sets out what we refer to as the "true" mean award in tried tort cases in the sample. That mean is based on the 286 docket sheets in our tort trial sample whose dockets show judgments for plaintiffs, regardless of how the AO coded either the victor or the amount of the judgment. (We have omitted the non-damages plaintiffs' judgments, discussed above, because including non-monetary cases obscures the true award pattern in the damage actions.) The row shows that the mean award for the full tort sample, as determined by inspecting the docket sheets via PACER, is approximately $816,000, with a 95% confidence interval of $452,000 to $1,180,000. In contrast, the table's second row shows that the mean award that a researcher looking only at the AO data (excluding AO-reported zero-awards) would report for a similar case population would be $1,387,000. The AO-based mean award is thus far higher than the more accurate mea-

sure—the error is $571,000 on a base of $816,000, or 70%.[89] Table 4's second numerical row also shows that the 95% confidence intervals for the two means overlap only slightly.[90]

Because we are interested in the influence of both zero and 9999 awards on the AO-based mean, the table's next three rows explore these topics. We know from Table 2 that half of the zero and even more of the 9999 awards are erroneous. But how large is the resulting error in the estimate of the mean? Table 4's third row shows that the cases with zero judgments in which plaintiffs are coded as victors have a mean award, as reported on the docket sheets, of $618,000; the fourth row shows that the 9999 cases have a mean award, as reported on the docket sheets, of approximately $4.7 million.

So, for researchers seeking to use the AO data in future analyses, both the zero and the 9999 awards seem to pose significant problems. One possible solution is simply to discard such awards. Table 4's fifth row tries out this approach, and shows that if both the zero and 9999 awards are excluded, the mean award in the remaining part of the sample is $456,000. Thus, comparing an AO-based estimate to the true tort case mean of $816,000, the AO-based estimate shifts from substantially too high with the 9999 cases included to substantially too low if they (along with the zero cases) are omitted. The reason is clear: the AO-based mean is substantially too large because the 9999 cases are not in fact on average awards of nearly $10 million or higher. Yet the mean calculated by excluding the 9999 cases is too low because *some* of these cases' awards are correctly coded, and as a group they are therefore atypically high compared to the non-9999 cases.

A second possible adjustment that continues to economize on case-by-case research could employ detailed, docket-sheet-based, investigation only of the awards entered as 9999. Table 4's sixth row reports this calculation of the mean, based on replacing only the 9999 awards with the true award, as reported on docket sheets. That is, the sixth row is based on 222 non-zero trial awards as reported in the AO

89 The alternative approach, which (as just stated, see *supra* note 88) we believe is conceptually flawed, would be to compute the AO-based mean reported in the second row by including trials with zero awards. But this approach includes a fairly large number of cases known to be non-monetary declaratory judgments, *see supra* note 80, and an equal number of cases in which the zero-awards coded by the AO are known to be incorrect. In any event, this approach yields a sample of 308 awards to use in computing the AO-based mean, with a mean of $1.108 million and a 95% confidence interval of $800,000 to $1.415 million.

90 The difference between the means of the awards in the first two rows of Table 4 is highly statistically significantly ($p < 0.0001$). The difference persists even after a log transformation of award levels.

database, and twenty-four trial awards based on case-by-case inspection of cases whose awards are entered in the AO database as 9999. The mean award using this methodology is $872,000, reasonably close to the true tort-case mean of $816,000. On a percentage basis, the error is $56,000 out of $816,000, or 6.9%. Table 4 also shows that the 95% confidence intervals of the true mean and the 9999-replacement-based mean overlap nearly entirely. This second adjustment, then, is much more satisfactory than the first: the coding errors in the 222 non-9999 cases are not so substantial as to yield a distorted mean when the 9999 cases are corrected. Moreover, the basic analysis holds in a second, smaller sample as well (although with a somewhat less accurate estimate of the mean), so it seems to be quite robust, at least for tort cases.[91]

Table 5 continues the analysis but instead of mean awards reports percentiles, including the fiftieth percentile (the median). The initial distortion introduced by unquestioning use of AO data is substantially smaller than for the mean award. Table 5's first row shows the true median award to be $137,000. The second row shows the AO-based median to be $151,000, an error of $14,000 or 10.2% (of $137,000), compared to the 70% error in the AO-based mean. The error is in the expected direction—the AO data exceed the true median. Here, even without correction, aggregate statements about the AO data could be useful. The AO-based median is the right order of magnitude; $151,000 does not "feel" dramatically different from $137,000—and is, in fact, within the 95% confidence interval of the true amount. A policy maker who acted on the basis of the AO figure for a general sense of award levels would not be too far off for many purposes.[92]

The effect of excluding the twenty-four 9999 awards is helpful, though less so than in the case of the mean. Table 5's third row

91 Indeed, as to each point reported above, our results are similar in the second torts sample described previously. *See supra* note 78. The AO-based mean, $1.855 million (*n = 127*), was very high compared to the true mean, observed by inspecting the docket sheets, which was $799,000 (*n = 136*), with this smaller sample's seventeen 9999 cases left in. The AO-based mean was very low—$430,000 (*n = 106*) with the 9999 cases taken out—but approaching acceptable—$600,000 (*n = 126*) with docket-based corrections to the 9999 cases that could be found. The error rate of 24.5% from the true mean is substantially larger than the error achieved in the 2000 cases, but still supplies a more reasonable estimate of the mean award than do alternative methods. The higher error rate may be due to the ordinary random variation in the much smaller sample of cases, or to some nonrandom factor such as the small sample of districts.

92 The 1996–1999 tort data, *see supra* note 78, confirm this general analysis. In that second tort sample, the AO-based median award is $186,000; the true median award is $134,000.

shows that excluding the twenty-four cases yields a median estimate of $125,000. The error is now $12,000 out of $137,000, a bit smaller than the error resulting from using the AO-based median, and, again, well within the 95% confidence interval of the true median award.

TABLE 5. ACCURACY OF AO AWARD PERCENTILES (AMOUNTS IN THOUSANDS), YEAR 2000 TORT TRIALS RESULTING IN PLAINTIFFS' JUDGMENTS

	95% confidence intervals, by percentile (point estimates)				
	10th	25th	50th	75th	90th
1. True data (*n = 286*)	5–17 (9)	25–52 (36)	100–173 (137)	330–611 (412)	957–2052 (1346)
2. AO data, AO award > 0 (true award found) (*n = 246*)	7–17 (10)	25–59 (38)	123–217 (151)	400–936 (725)	2000–9999 (8600)
3. AO data, AO award > 0 excluding 9999 cases (*n = 222*)	5–15 (8)	23–45 (33)	95–166 (125)	276–574 (373)	832–1448 (963)
4. Replace only 9999 awards with true data (*n = 231*)	7–17 (9)	24–56 (35)	105–200 (144)	341–750 (426)	973–2000 (1324)

SOURCE: ICPSR 8429, *supra* note 1, supplemented by PACER docket research. Shaded squares best fit with row 1's true data.

Obviously, the AO-based median in row 2 is too high because only some of the 9999 cases are in fact awards in excess of $9 million. The twenty-four cases coded as 9999 in our sample actually have a median award of $998,000—substantially higher than the $125,000 median of the non-9999 cases, but not nearly so high as the coded 9999 figure suggests. The cases' relatively large awards also explain why Table 5's third row estimate—based on excluding only the 9999 cases—is too low. Excluding such cases eliminates a set of observations that are high relative to the mass of cases, thereby artificially depressing the median derived from the non-excluded cases.

Replacing only the 9999 awards with the true awards in such cases yields improvement for the median estimate. Table 5's fourth (and final) row shows that replacing only the 9999 awards produces a median of $144,000. This is $7000 above the true median of $137,000, an error of 5.1%. This is yet more accurate than the 8.8% error obtained by excluding the 9999 cases. Indeed, across all percentiles, replacing the 9999 awards with their actual values gets the closest to the true distribution of awards; each box in the row is therefore shaded grey.

Both the point estimates and the confidence intervals in this row, across all five percentiles, are reasonable.

But even though these *best* estimates depend on case-by-case inspection of a number of dockets, the third row, which simply leaves out the 9999 cases, yields confidence intervals and point estimates that are quite reasonable for the tenth, twenty-fifth, fiftieth, and seventy-fifth percentiles. They are low only for the ninetieth percentile. Indeed, even the second row, which includes the 9999 cases, is fairly reasonable up to the median point. This suggests that researchers may be able to obtain a reasonable estimate of the median award without any docket-research. If one is interested in an upper-limit estimate, one could simply use the AO-based median (as in row 2) and be reasonably confident that the estimate is conservative (in the sense that the true median is unlikely to be substantially higher than the number so reported). So, for example, if one wishes to report an upper limit on the median tort awards in federal court for a year or other time period, the AO-based median seems reasonable to use. Similarly, if one is interested in a lower-limit estimate, excluding the 9999 cases, as in row 3, gives a reasonable figure. The excluded 9999 awards tend to drive up the median, as they do the mean. So in our sample, the $125,000 figure is a reasonable lower-bound point estimate of the median award.

b. Inmate Awards

Table 6 reports on the mean awards in our 1993 inmate sample, again as checked against dockets available from PACER. Table 6's first row shows that the true mean for the inmate cases is $69,000. The AO data, used as published, yield a mean of $927,000.[93] While the $858,000 error in the mean is bigger in absolute amount than the mean error for our tort sample, the more relevant and more damning statistic is that the error in the inmate case sample amounts to well over 1000% (the analogous figure for the tort sample was 70%).[94]

93 There is one outlier award of over $6 million, correctly coded in the AO data set; if that award is taken out, the true mean is only $16,403; the AO mean also comes down to $882,000. So, if anything, the text understates the degree of error.

94 *See supra* note 88, and accompanying text.

TABLE 6. ACCURACY OF AO AWARD MEANS (AMOUNTS IN THOUSANDS) FISCAL YEAR 1993 INMATE CASES, AO AWARDS > 0

	Amount	95% confidence intervals	n
1. True mean	69	−35–174	122
2. AO mean	927	495–1360	122
3. True mean of AO 9999 awards	23	8–38	5
4. AO mean, excluding 9999 awards	540	257–822	117
5. Replace only AO 9999 awards with true mean	518	247–789	122
6. AO mean, excluding 9999 awards *and* adjust digit errors	199	35–364	117

SOURCE: ICPSR 8429, *supra* note 1, supplemented by PACER docket research. The first row shows the true mean award as determined by inspecting PACER dockets. The second and fourth rows show mean AO award codes. The third row shows true mean awards for cases for which the AO award code is "9999." The fourth row combines true and AO awards, replacing AO codes for AO awards coded 9999. The final row excludes all awards coded 9999 and adjusts all digit errors.

One hypothesis about the source of this large error is that inmate civil rights cases have an exceptionally high percentage of awards reported as 9999. This proves incorrect, however. As Table 6 shows, reported 9999 awards are just five of 122 cases (4.1%) with awards greater than zero in the 1993 inmate sample, compared to twenty-four of 246 (9.8%) in the analogous torts sample in Table 4. Moreover, in the full AO trial data set since 1991, which is described in Table 10 below, awards coded 9999 account for only 2.5% of the inmate awards.[95]

However, even though 9999 awards are not exceptionally frequent, the lower, true awards in inmate cases must be substantially more distorted by these erroneous large awards than are the larger awards in tort cases. Table 6's fifth row suggests that replacing the 9999 awards with their actual values from docket sheets does go some portion of the way towards estimating the true mean, though by no means far enough for most purposes. Replacing the five 9999 awards yields a mean estimate of $518,000, quite a bit closer to the real mean award of $69,000, though still dramatically higher.

Table 6's final row demonstrates that there is another, even larger source of error—digit mistakes, which typically overstate awards by a factor of 1000 (when, say, a judgment of $112 is entered as 112, which is supposed to mean $112,000). To obtain the statistics in the

95 The difference in 9999-award rates between 1993 terminations and the AO data set as a whole is not statistically significant.

last row, we took each case in the inmate sample that had a digit error, and substituted the correct code for the award; we also excluded the 9999 award cases.[96] The result is a much better estimate of the true mean though still substantially incorrect. In sum, it seems likely that the real prevalence of small awards in the inmate sample not only amplifies the effect of the erroneous 9999 award entries but, more importantly, has a strong tendency to promote digit errors. (It makes intuitive sense that awards under $1000 are the most easily miscoded, because they do not have more digits than there are spaces in the data system.)

The problem is that the correction in the final line of Table 6 relies on the cumbersome process of reading many dockets, which is quite impracticable in many circumstances. So we move on to Table 7, which examines the distribution of awards in the inmate sample, to see whether some other technique may be helpful in gleaning from the AO data a more accurate picture of the awards.

Table 7 demonstrates that for the inmate cases, the very issue that introduces error—the extremely modest awards—also makes that error matter less, if one is looking at and below the fiftieth percentile. The true median (in row 1) is just $950 (the tenth and twenty-fifth percentiles are smaller—so if they were rounded up, to match the AO data's capabilities, they would be accurate). Using the medians in the next three rows in the table certainly does not eliminate the error, which is $4000 or about 400% if the comparison is to the AO data in its entirety (row 2) and $2000 or about 200% if the comparison drops the five 9999 awards (row 3). On the other hand, the importance of the error's magnitude depends on the research question being asked. It seems likely that $5000 or $3000 could be used almost interchangeably with $950 in many discussions of inmate award issues. Indeed, the latter figure is not far off from the 95% confidence interval for the true number. A policymaker who acted on the basis of the AO figures for a general sense of award levels would not be far off. And the AO-based median would again provide a conservative upper bound estimate on the median inmate award.

Thus, as in the case of the tort data, researchers, without the need for case-by-case inspection, can obtain a reasonable estimate of the median award. If one is interested in an upper-limit estimate, one could simply use the AO-based median and be reasonably confident that the estimate is conservative in the sense that the true median is

96 If only the digit errors are adjusted, and the 9999 cases remain in the sample, the resulting mean is 493, with a confidence interval between 124 and 862. This alone is a more significant improvement than simply excluding the 9999 cases.

TABLE 7. ACCURACY OF AO AWARD PERCENTILES (AMOUNTS IN THOUSANDS), FISCAL YEAR 1993 INMATE CASES AO AWARD > 0

	95% confidence intervals, by percentile (point estimates)				
	10th	25th	50th	75th	90th
1. True data (*n* = *122*)	.001–.02 (0.001)	.025–.3 (0.10)	.5–2.5 (.95)	5–23 (10)	27–102 (47)
2. AO data: all (true award found) (*n* = *122*)	1–1 (1)	1–1 (1)	1–11 (5)	32–1000 (70)	1000–7172 (5000)
3. AO data, excluding 9999 cases (*n* = *117*)	1–1 (1)	1–1 (1)	1–9 (3)	25–305 (50)	382–5000 (1312)
4. Replace only 9999 awards with true data (*n* = *122*)	1–1 (1)	1–1 (1)	1–10 (5)	25–144 (46)	345–5000 (1098)
5. Adjust digit errors in AO data (*n* = *122*)	.21–1 (1)	1–1 (1)	1–4 (1)	6–50 (18)	52–5627 (813)
6. Exclude AO 9999 awards and adjust digit errors (*n* = *117*)	.14–1 (1)	1–1 (1)	1–4 (1)	5–35 (12)	38–1000 (101)

SOURCE: ICPSR 8429, *supra* note 1, supplemented by PACER docket research. Shaded squares best fit with row 1's true data.

unlikely to be higher than the $5000 so reported. Above the fiftieth percentile, however, the erroneous awards entirely dominate the sample, and the AO data cannot itself do much to inform an estimate. For that, once again, researchers would need to read dockets.

But how is one to know whether a given component of the AO data is more like our torts sample, or more like our inmate case sample, or unlike either? Again, the AO data itself may help to answer this question. We suggest above that the feature of the inmate cases that makes them error-prone is the low level of awards. Even though the AO data inflate the awards, they report a very large number of awards of "1": 52 of 122 (43%). This might be a potential tip-off in other case categories as well. We explore this issue briefly in Table 10, at the end of this Article; it turns out that there is no other category even close to inmate cases on this measure.

3. Estimating Employment Discrimination Awards

This section uses the information about award errors in the preceding sections to estimate the level of awards in federal employment

discrimination trials. The sample consists of all tried federal employment discrimination cases terminated from 1994 through 2000.[97] We wish to estimate the true award levels in such cases without using the labor-intensive technique, employed above, of inspecting individual case docket sheets. To do so, we use, without modification, the AO-reported awards in all cases other than those in which the entry in the "amount" field is suspect because it is "9999." For those cases reporting a 9999 award, we used PACER to inspect the actual docket sheets. Table 8 reports the results.

TABLE 8. APPLYING ESTIMATES TO EMPLOYMENT DISCRIMINATION TRIALS, MEAN TRIAL AWARDS, 1994–2000 (AMOUNTS IN YEAR 2000 THOUSANDS)[98]

	Estimated mean	Estimated median	n
1. AO data (no adjustment)	863	121	1298
2. AO data (excluding 9999 awards)	295	107	1220
3. Replace only AO 9999 awards	301	110	1292
4. True awards, 9999 cases only	410	170	70

SOURCE: ICPSR 8429, *supra* note 1 (supplemented by PACER docket research of cases with 9999 awards). Each row shows the estimated mean and median trial award for the indicated data set.

Table 8's first row shows the AO data's mean and median employment discrimination awards—$863,000 and $121,000, respectively. Our analysis of tort awards suggests that the mean is likely substantially too high because it includes many awards reported to be $9.999 million or higher that are in fact not so high. The second row of Table 8 reports what we expect to be low estimates of both mean and median, based on simply excluding the 9999 cases. Both figures turn out to be close to those computed in the third row, which is based on substituting the amount reported on the docket sheets for the seventy available cases with AO award codes of 9999. In this sample, simply excluding the 9999 cases yields mean and median estimates, $295,000 and $107,000, that are not too different from those we obtained by

97 More precisely, the sample is every case terminated between January 1, 1994 and September 30, 2000, with an AO case code of 442, in which the procedural progress is coded as after jury or judge trial and judgment is coded for plaintiff or for "both" plaintiff and defendant.

98 Adjustments for inflation are based on BUREAU OF LABOR STATISTICS, U.S. DEP'T. OF LABOR, CONSUMER PRICE INDEX, ALL URBAN CONSUMERS (2003), *available at* ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt (last visited Mar. 22, 2003). The true values of some 9999 cases are not available so the total number of AO cases in the table (1298) exceeds the sum of the number of 9999 cases and non-9999 cases.

the more laborious method of looking up the results in the seventy 9999 cases, $301,000 and $110,000 (Of course, we cannot ascertain the true mean or median without inspecting all of the non-9999 cases, which we have not done.) Thus, in estimating the employment discrimination award mean, the technique of excluding the 9999 awards may itself yield a reasonable estimate. No case-by-case inspection of docket sheets may be required at all.

Why are the 9999-excluded mean and the 9999-ascertained mean so much more similar in employment cases than in Table 4's tort cases or in Table 6's inmate cases? Two hypotheses are worth noting. First, the 9999 cases are a much smaller fraction of employment cases than of tort cases or inmate cases. Seventy of 1298 employment case awards (5.3%) have 9999 entered for the amount field. In Table 4's tort awards, the 9999 cases comprise twenty-four of 246 cases (9.8%). This difference is highly statistically significant ($p < .001$). Second, the absolute level of docket-verified awards in the seventy 9999 employment discrimination cases is noticeably smaller than the level of docket-verified awards in the twenty-four 9999 tort cases. In our torts sample, about a quarter of the 9999 cases are accurately coded (that is, had actual awards of $9.999 million or higher); taken together, the torts 9999 cases have a docket-verified mean of $4,717,000—more than ten times the award level in non-9999 cases, reported in Table 4 to be $456,000. In the employment sample, by contrast, just one (1.3%) of the 9999 cases was accurately coded, and taken together, the 9999 cases have a docket-verified mean of just $410,000 compared to the AO level of $295,000 for the non-9999 cases. Similarly, whereas the docket-verified median for our tort 9999 cases is nearly eight times greater than the docket-verified median for non-9999 cases, the true median of employment 9999 cases, $170,000, is just $63,000, or 59.9%, above the non-9999 case median, as reported in the AO data.

Thus the tort 9999 cases are a higher percentage of all cases, and they tend to have more extreme values than the employment discrimination 9999 cases. The result: our finding above that exclusion of the tort 9999 cases more substantially affects both the mean and the median than does exclusion of employment discrimination 9999 cases.

III. Implications and Further Applications

The implications of our findings depend in part on whether researchers are interested in assessing win rates or award levels. We briefly explore both below, and then apply the techniques developed here to estimate median trial awards for all large federal case categories.

A. *Implications for Win-Rate Studies*

Generally—that is, where no other anomaly exists to counsel caution—analyses using the AO's coding of which party obtained judgment are likely largely unaffected by errors in the AO data. Our evidence suggests that when the AO data show that judgment is entered for plaintiff or defendant (at least in cases coded with non-zero-awards) the reported victor is overwhelmingly accurate. In tried cases, moreover, relatively small fractions of the AO data report judgment codes other than for plaintiff or defendant. Still, for groups of cases that show substantial percentages of cases coded as judgment for both parties, such as our inmate case category, researchers should consider analyzing their data in the alternative: first without such cases and, second, with such cases treated as victories for plaintiffs. If the results of these alternative analyses are consistent with respect to the research question of interest, little basis for concern exists about possible inaccuracies in the judgment coding. If the results are not consistent, further consideration of how to deal with the ambiguous judgment code is necessary. In addition, as panel B-2 of Table 1 demonstrates, an anomaly, such as the miscoding of which party obtained judgment in cases coded with zero-awards, can render the "judgment for" data quite inaccurate, and needs to be accounted for with care.

B. *Implications for Studies of Amounts*

With respect to award levels, our findings suggest that relying on unmodified AO trial data substantially overstates mean awards. Tables 4 and 6 establish this in the areas of tort and inmate cases. And Table 8 suggests that this is the case in employment discrimination cases; checking just the cases with coded awards of 9999 establishes that the mean award derived from AO data is unreliable.

Tables 5 and 7, however, suggest that relying on the AO data to study median awards is often reasonable, depending on the research question being addressed. And Table 8's check on employment data does not falsify this hypothesis. For tort cases, and perhaps for employment discrimination cases, the error in using AO data seems to be within acceptable ranges for most purposes, and the error can be further reduced by the simple expedient of excluding awards coded as 9999. For inmate cases, and presumably for other classes of cases with typically small awards, the percentage error in the median is high. But the absolute difference in dollars between the AO-based median and the true median is small, precisely because most awards are small.

For studies that use award amounts in a more complex way — not looking at awards by case category, but rather performing more indi-

vidualized modeling or other analysis—we offer only a cautionary word. Whether the AO data are sufficiently reliable to support such research will depend on the precise details of the research design, and the issue requires close attention.

C. *Applications: Judgment Patterns and Awards Patterns for All Federal Case Categories*

We build on the results reported above to supply some possibly helpful information about the AO data for several case categories. We first report on the percentage of trials that the AO reports as ending in judgments for plaintiffs or both plaintiffs and defendants, but with zero-awards. This class of cases was especially troublesome for the inmate civil rights class of cases, but much less troublesome for tort cases. We then supply an estimate of the median trial award for all sizeable case categories.

1. Judgment Code Patterns

We have suggested that the successful use of AO data depends on close attention to anomalies. In both torts and inmate civil rights litigation, purported plaintiffs' judgments with zero damages are anomalous. Table 9 presents data on this cautionary signal in other case categories, as well as on the size of the category of judgment for "both." The table shows the total number of trial outcomes, and the percentage of those outcomes ending in judgments entered in the AO data as being for plaintiff or for both plaintiff and defendant. It is limited to those case categories with at least one hundred trials coded with judgments for plaintiff or "both" for years 1991–2000. The last column explores the percentage of the plaintiffs' judgments in which the damage award coded is zero. We exclude cases in which the AO data's "nature of judgment" code indicates that the judgment is an injunction, a forfeiture or condemnation, a costs-only judgment, and so on, in contrast to a monetary judgment.[99] That is, we intend the column to explore a possible data anomaly, not an ordinary non-monetary judgment.[100] (Of course, for some of these case categories,

99 ICPSR 8429, *supra* note 1.

100 Four "nature of judgment" codes remain: "-8," which codes missing information, "0," which codes "no monetary award," "1," which codes "monetary award only," and "2," which codes "monetary award and other." In every case category, all or nearly all of cases that contribute to the potential anomaly we are highlighting—judgment for plaintiff or both combined with a zero award—have a nature of judgment code of zero ("no monetary award"). This is the code typically used in conjunction with a defendant's victory. Since we know from our torts sample that many of the

plaintiffs' judgments with zero damages might be expected, rather than anomalous.)

TABLE 9. AO JUDGMENT CODES IN CASES WITH TRIAL JUDGMENTS, FISCAL YEARS 1991–2000

Case category	n	Judgment for plaintiffs	Judgment for "both"	AO award = 0, as % of victories for plaintiff or "both"
Insurance (110)	2265	49%	4%	25%
Marine (120)	533	63%	5%	15%
Miller Act (130)	217	70%	10%	10%
Negotiable Instruments (140)	307	63%	5%	10%
General Contract (190)	4661	58%	8%	14%
Contract Product Liability (195)	107	54%	7%	5%
Land Condemnation (210)	323	39%	11%	23%
Foreclosure (220)	126	72%	7%	24%
Torts to Land (240)	153	46%	5%	19%
Other Real Property (290)	253	53%	6%	41%
Airplane (310)	215	57%	1%	7%
Assault, Libel, Slander (320)	275	40%	5%	12%
Federal Employers' Liab. (330)	670	67%	1%	6%
Marine (340)	1124	53%	5%	9%
Motor Vehicle (350)	2678	62%	2%	7%
Motor Vehicle Product Liab. (355)	279	28%	1%	12%
Other Personal Injury (360)	4675	42%	2%	10%
Medical Malpractice (362)	1153	33%	1%	10%
Product Liability (365)	1713	29%	2%	12%
Asbestos (368)	307	82%	1%	6%
Fraud (370)	379	56%	6%	12%
Other Personal Prop. Damage (380)	477	55%	7%	10%
Property Damage Prod. Liab. (385)	198	37%	4%	12%
Antitrust (410)	180	47%	3%	16%
Bankruptcy Appeals (422)	178	28%	10%	81%
Bankruptcy Withdrawal (423)	142	39%	6%	16%
Other Civil Rights (440)	6179	28%	4%	17%
Voting (441)	112	40%	7%	84%
Employment Discrim. (442)	8200	32%	3%	14%
Accommodations (443)	272	44%	7%	22%
RICO (470)	185	57%	8%	11%
Habeas Corpus (530)	338	18%	2%	90%
Inmate Civil Rights (550, 555)	7261	9%	3%	24%
Drug-Related Prop. Forfeiture (625)	211	78%	4%	63%
Other Forfeiture & Penalty (690)	217	72%	4%	67%

cases with the anomalous "judgment for plaintiff" (or both) and zero-award combination are, nonetheless, actually plaintiffs' judgments with erroneous award codes, we conclude that, unfortunately, the nature of judgment code is unhelpful to our analysis here.

TABLE 9, CONTINUED

Case category	n	Judgment for plaintiffs	Judgment for "both"	AO award = 0, as % of victories for plaintiff or "both"
Fair Labor Standards Act (710)	588	53%	5%	13%
Labor/Mgt Relations (720)	235	49%	3%	29%
Other Labor Litigation (790)	378	39%	3%	19%
ERISA (791)	1077	45%	5%	21%
Copyright (820)	342	68%	7%	17%
Patent (830)	700	54%	9%	34%
Trademark (840)	434	61%	9%	33%
Securities, Commodities, Exch. (850)	323	53%	8%	17%
Tax Suits (870)	581	51%	6%	27%
Other Statutory Actions (890)	1135	49%	6%	30%
Environmental Matters (893)	200	56%	11%	28%

SOURCE: ICPSR 8429, *supra* note 1. The table includes all cases coded in the AO data as terminating after trial with a judgment for plaintiff or defendant or both plaintiff and defendant. It excludes cases with judgments coded as missing; and cases coded as injunctions, costs-only awards, and the like.

Table 9 demonstrates that cases with plaintiffs' victories combined with awards coded as zero are most prominent in a few categories, many of them cases of a type that rarely result in damages (for example, land condemnation, foreclosure, and habeas corpus). In these categories, the combination is not anomalous at all. In other categories, however, a high portion of such cases may well be a signal of erroneous coding. Unless our inmate case sample turns out (against our current belief) to be nonrepresentative of inmate cases, we know that a researcher who accepts the "judgment for" code at face value would overstate plaintiffs' success rate in that category. The same may be true for other case categories in which damage actions are prevalent, and nontrivial percentages of zero awards exist. Based, however, on the evidence from our torts sample, which seems likely to be more typical of the dataset as a whole, we suspect that in case categories in which plaintiffs are more frequently successful, errors signaled by an anomalous zero-award will be found more often in the award coding than in the "judgment for" code.

Table 9 also suggests that the influence of judgments entered as for "both" plaintiff and defendant varies by case category. Table 1 shows such judgments as nearly always for plaintiffs. But Tables 1 and 9 both indicate that the "both" code is a much higher fraction of possible pro-plaintiff judgment codes in inmate civil rights cases than it is in other case categories—indeed, using Table 9's figures, "both" judgments constitute nearly a third of the total pro-plaintiff judgments for

the inmate case category, a rate that is nearly the highest in any sizeable case category. So, the systematic coding of some plaintiff wins as wins for both has a larger effect on accurately stating plaintiff win rates in inmate civil rights cases than in most classes of cases.

2. Median Award Estimates and Rates of Suspicious Award Codes

Given the general reasonableness of the median estimates in our two sampled case categories, we apply the foregoing analysis to a larger set of AO cases to provide interested researchers and policymakers with likely-improved estimates of median awards across many case categories. Table 10's rows represent each AO case category (and its respective code value) for which at least fifty trials with judgments for plaintiffs were concluded with positive awards from fiscal years 1991 through 2000.[101] The first numerical column in each row reports the median dollar award, in inflation-adjusted year 2000 dollars, as computed from the unaltered AO data. The second numerical column reports the number of verdicts used to compute that median award. The third numerical row adjusts the AO median by recomputing the median after excluding awards of 9999. The fourth numerical column shows the number of verdicts used in computing this adjusted median award. The fifth numerical column, computed from the second and fourth columns, shows the percent of verdicts for each case category that report an award of 9999. And the sixth numerical column shows the percent of verdicts for each case category in which the award is coded "1." These low-award cases could be of special interest as a source of error because where awards in the hundreds and low

101 More precisely, the sample consists of terminations in which the judgment was after a jury or judge trial that resulted in a judgment for the plaintiff with a positive award noted. Dollar amounts are adjusted using Bureau of Labor Statistics inflation data. *See supra* note 98. Table 1's finding that cases coded by the AO as judgment for "both" plaintiff and defendant are actually plaintiff victories suggests checking Table 10's median results by including "judgment for both" cases. We have done so and, in the large majority of case categories, including the "judgment for both" cases does not materially change the median. Table 10 excludes such cases because we do not assume that Table 1's pattern holds for every category. Moreover, including the "judgment for both" cases seems to us most suspect in case categories in which including them generates a large change in the medians. In these categories, the "judgment for both" cases are most dissimilar in amounts from cases coded as plaintiff judgments, the proportion of "judgment for both" cases is especially high, or both the amounts and proportion are unusual. Either of these features may indicate, for a particular case code, that cases coded as "judgment for both" differ systematically from plaintiffs' victories: perhaps they err by more than simply incorrectly coding which party won, or perhaps they constitute a conceptually separate category of outcomes in some other way.

thousands are particularly prevalent, what we have called "digit error" is likely to abound. Researchers would do well to be particularly careful if their focus is on a case category with a high proportion of reported small awards.

For example, using the AO data without modification, the product liability case category (code 365) shows a median award of $486,000 based on 437 plaintiffs' verdicts. Excluding the 9999 awards yields a median products award of $368,000 based on 385 verdicts. We now hypothesize that the $368,000 figure is closer to the true median than is the $486,000 figure.[102]

The principal non-inmate civil rights categories, "Other Civil Rights" (code 440) and "Employment Discrimination" (code 442), have adjusted median awards of $78,000 and $116,000 respectively. Inmate civil rights cases, for which the AO data may be the least accurate (as a percentage of the true award), conform to the pattern of low awards suggested in Part II's detailed analysis of 1993 inmate cases. The $6000 median estimate in Table 10 is probably too high in light of that discussion.

Inmate civil rights cases also have by far the largest percentage of trials entered resulting in damages coded as "1" in the AO data. The 39% rate is more than triple the rate in most categories. This high rate of such awards is consistent with Table 3's report that, in our inmate case sample, fifty-two of 122 awards (42.6%) are coded as "1." The many low-award cases in this much larger sample further support the suggestion that the impact of the error pattern in inmate civil rights cases is likely not typical of the impact of the error pattern in other classes of cases.

One interesting implication of Table 10 is that even after deflation of awards by omission of the 9999 cases, the reported awards remain substantially higher than awards in state court litigation.[103]

---

102 Given interest in the size of awards, one noteworthy feature of Table 10 is that only three case categories, Asbestos, Antitrust, and Patent, have median awards greater than $1 million, even using the probably inflated medians based on all the AO awards, including the 9999 cases. The Antitrust and Patent categories have the highest percentage of 9999 awards; these 9999 awards may, indeed, be more-than-typically accurate in these large-award categories. The 9999 cases are not, however, contributing much to the high award level for asbestos cases, which are a world unto themselves. *See, e.g.*, Deborah R. Hensler, *As Time Goes By: Asbestos Litigation After* Amchem *and* Ortiz, 80 TEX. L. REV. 1899 (2002).

103 *E.g.*, Eisenberg et al., *supra* note 55, at 439.

TABLE 10. ESTIMATED MEDIAN AWARDS IN CASES WITH TRIAL JUDGMENTS, 1991–2000 (AMOUNTS IN YEAR 2000 THOUSANDS)

	Plaintiffs' verdicts with non-zero judgment		Excluding AO 9999 awards		% AO award =	
Case category	Median (1000s)	n	Median (1000s)	n	9999	1
Insurance (110)	173	769	149	715	7.0%	1.7%
Marine (120)	107	282	97	272	3.5%	1.8%
Miller Act (130)	48	135	42	128	5.2%	0.7%
Negotiable Instruments (140)	366	173	309	155	10.4%	1.7%
General Contract (190)	238	2261	201	2066	8.6%	1.5%
Contract Product Liability (195)	327	54	156	47	13.0%	0.02%
Torts to Land (240)	161	61	148	59	3.3%	4.9%
Other Real Property Actions (290)	98	68	76	63	7.4%	2.9%
Airplane (310)	681	114	483	103	9.6%	2.6%
Assault, Libel, Slander (320)	103	99	90	91	8.1%	4.0%
Federal Employers' Liab. (330)	233	410	198	387	5.6%	1.2%
Marine (340)	187	543	165	513	5.5%	1.5%
Motor Vehicle (350)	113	1509	95	1418	6.0%	2.1%
Motor Vehicle Prod. Liab. (355)	652	68	431	58	14.7%	1.5%
Other Personal Injury (360)	109	1734	90	1619	6.6%	2.5%
Medical Malpractice (362)	482	337	364	297	11.9%	0.9%
Product Liability (365)	486	437	368	385	11.9%	0.5%
Asbestos (368)	3799	238	3793	236	0.8%	0.0%
Fraud (370)	355	186	242	163	12.4%	2.2%
Other Pers. Prop. Damage (380)	169	230	144	208	9.6%	3.9%
Prop. Damage Prod. Liab. (385)	284	62	225	57	8.1%	3.2%
Antitrust (410)	2823	65	1190	44	32.3%	0.0%
Other Civil Rights (440)	99	1362	78	1262	7.3%	6.5%
Employment Discrim. (442)	129	2186	116	2064	5.6%	1.7%
Accommodations (443)	40	91	35	86	5.5%	4.4%
RICO (470)	631	94	422	79	16.0%	0.0%
Inmate Civil Rights (550, 555)	5	479	5	467	2.5%	39.0%
Fair Labor Standards Act (710)	47	260	46	255	1.9%	5.8%
Labor/Mgt Relations (720)	198	75	178	72	4.0%	2.7%
Other Labor Litigation (790)	111	114	91	104	8.8%	4.4%
ERISA (791)	60	364	53	350	3.8%	4.4%
Copyright (820)	62	187	59	180	3.7%	5.9%
Patent (830)	1694	250	625	194	22.4%	2.8%
Trademark (840)	172	139	134	129	7.2%	4.3%
Sec., Comm., Exchange (850)	547	130	357	118	9.2%	0.8%
Tax Suits (870)	133	198	108	183	7.6%	6.6%
Other Statutory Actions (890)	77	369	62	338	8.4%	5.1%
Environmental Matters (893)	607	76	524	70	7.9%	1.3%

SOURCE: ICPSR 8429, *supra* note 1. The table includes all cases coded in the AO data as terminating with a judgment for plaintiff and a positive award amount in following a trial.

CONCLUSION

Subject to the limitations of our samples, we tentatively conclude that AO data can provide reasonably accurate estimates of the proportion of cases in which plaintiffs win damages judgments. A possible systematic understatement of plaintiff win rates exists that is attributable to judgments recorded as judgments for "both" plaintiffs and defendants in fact tending to favor plaintiffs, but this outcome classification accounts for a small percentage of trial outcomes.

With respect to awards, it is necessary to distinguish between mean and median awards. The error resulting from using unmodified AO data to compute mean awards has a distinct direction in our two samples—the AO data systematically overestimate the mean award. Thus, studies that rely on AO data to address questions about the level of awards probably overstate amounts paid out in, for example, products liability litigation.[104] For case categories with fairly large awards, substantially improved mean-award estimates are likely obtainable by substituting awards recorded on docket sheets for awards coded by the AO data as 9999. Estimates of median awards based on the AO data without further investigation appear to be of reasonable size and to provide useful upper bounds of true median awards.

The AO database is likely to remain one of the major sources for civil justice research. We hope that this partial exploration of the accuracy of the data is helpful to other researchers, offering not only warnings but reasonably efficient solutions to identified accuracy problems.

104 Eisenberg & Henderson, *supra* note 16, at 739.