

STUDYING THE ICEBERG FROM ITS TIP: A COMPARISON OF PUBLISHED AND UNPUBLISHED EMPLOYMENT DISCRIMINATION CASES

PETER SIEGELMAN
JOHN J. DONOHUE III

Researchers often rely on published opinions to draw conclusions about cases decided by the courts, determinants of court decisions, and broader social phenomena. We demonstrate that 80 to 90 percent of employment discrimination cases filed in federal court do not produce a published opinion. There are good theoretical reasons to believe that the process generating a published opinion is not random and thus that samples of published cases will not be representative of all cases. Through a direct comparison of published and unpublished cases, we show that the two actually do differ in significant and predictable ways. Examining several studies that use cases with published opinions for a variety of purposes, we show how our understanding of the operation of employment discrimination law changes—in some instances, dramatically—when we look at all cases, whether or not they have published opinions.

I. INTRODUCTION

One of the few uncontested truths produced by the application of social science to law is that only a tiny part of the “action” in the legal system is revealed in appellate cases. Most potential disputes never get defined by the actors as such, most actual disputes don’t go to court, most court cases are settled rather than adjudicated, and most adjudicated cases are not appealed.¹ Moreover, at

We thank research assistants Dawn Jeglum Bartusch, Steven Martin, and Kirsten Alesch Muth for their superb assistance with all phases of this highly RA-intensive work. The staff at the Administrative Office of the United States Courts, especially Dave Cook, provided us with important data and was generous in answering our questions. The article benefited from helpful comments by Ted Eisenberg, James Hughes, Vicki Schultz, and an anonymous referee. Listing of the authors’ names in reverse alphabetical order is a random event in the course of their collaboration.

¹ Miller and Sarat (1981: 544) describe a dispute-processing pyramid: the broad, bottom layer is possible instances of conflict, while filing a legal claim is at the top of the pyramid. (The share of grievances that go on to become litigated cases with published opinions or appealed cases amounts in this analogy to a few grains of sand on top of the point of the pyramid.)

The pyramid seems particularly “fat bottomed” in the substantive legal area that will be the focus of this article—employment discrimination. That is,

LAW & SOCIETY REVIEW, Volume 24, Number 5 (1990)

least since the pathbreaking work of Priest and Klein (1984),² social scientists studying the law have come to appreciate that the process by which parties select disputes for litigation or appeal will not, in general, produce a random (representative) sample of all disputes or cases. If appellate cases constitute only a fraction of the population of disputes and potential disputes, and if that fraction is not only minuscule, but systematically unrepresentative of the larger population from which it is drawn, generalizing from such cases to larger conclusions about law and/or society is hazardous.³

The unrepresentative nature of appellate cases is by now widely accepted among social scientists. But a similar problem has commanded far less attention: district court cases with published opinions may be subject to the same kinds of sample selection problems that confront users of appellate cases.⁴

among those who say they have experienced employment discrimination, a smaller percentage go on to do anything about it—complain, see an attorney, file suit—than for almost any other type of dispute. See, e.g., Curran (1977: 108, 135, 161); Miller and Sarat (1981: 545); Bumiller (1988); Kuhn (1987).

Without wishing to start a battle of the metaphors, we suggest that an iceberg is also an appropriate analogy for the sociolegal system: what is visible above the metaphorical waterline are the readily accessible published cases, but these constitute only a small portion of the object being studied.

² A useful recent survey of the kinds of economic models of legal disputes (including litigation and settlement behavior) that motivate our argument is Cooter and Rubinfeld (1989).

³ Priest and Klein (1984: 1) note:

Virtually all systematic knowledge of the legal system derives from studies of appellate cases. Appellate cases, of course, provide the most direct view of doctrinal developments in the law. . . [and] may tell us which disputes courts find troublesome and which they find easy to decide. But this doctrinal information discloses very little about how legal rules affect the behavior of those subject to them or affect the generation of legal disputes themselves.

Of course, appellate cases are worthy of study in their own right, since they have an important influence on legal rules and on the construction of legal consciousness. Suppose, for example, that Willie Griggs had settled with the Duke Power Company rather than contesting its decision not to hire him. If so, the Supreme Court might never have been confronted with a case involving practices implemented on the effective date of the Civil Rights Act in 1965 by a firm that freely admitted its discriminatory practices before 1965 (*Griggs v. Duke Power*, 1971). Because the case *did* reach the Supreme Court, it presented both the general public and the legal profession with a new definition of what “discrimination in employment” meant—the use of ostensibly neutral job requirements that disqualify a disproportionate number of black applicants was held to be discriminatory. A careful reading of Chief Justice Burger’s decision suggests that had the question arisen in a different factual setting, the legal doctrine that emerged might have been far less momentous (Gold, 1985).

Cast in terms of the pyramid of disputes described in Miller and Sarat (1980–81), our theme is that while each level of the pyramid is worth studying on its own, caution must be exercised in trying to generalize about one level of the pyramid on the basis of information about another.

⁴ Some scholars make passing reference to the problem, although it seems more common to ignore it altogether. Two recent papers, however, have treated the broad set of issues we deal with here.

Olson (1989) does not discuss specific kinds of unrepresentativeness or present a theoretical treatment. She does demonstrate convincingly that the

The existence of specialized case reporters such as BNA's Fair Employment Practice Cases Digest, the growth of computerized legal research tools such as LEXIS and WESTLAW (see, e.g., Brenner, 1989), and the development of content analysis techniques have made it relatively easy for social scientists to tap the rich source of data contained in legal opinions. But if cases with published opinions are systematically different from those without, such sources and techniques need to be used with more caution than they have been.

Our article attempts to determine how district court cases with published opinions differ from those without them. We begin in section II by confirming what most readers probably know already—the potential unrepresentativeness of cases with published opinions is likely to be significant because only a few cases ever leave a published record. In section III, we briefly discuss the definition of unrepresentativeness and then analyze the process by

potential for unrepresentativeness is high, given that few cases produce published opinions. Surprisingly, she also finds that there is little relationship between attorneys' assessment of a case's importance and whether the case generates a published opinion.

Our article also echoes many of the themes in Eisenberg and Schwab (1989). Although our basic methods and insights are similar, Eisenberg and Schwab's purpose is slightly different from ours: to compare the perspective that law professors and students develop by looking at published opinions with the broader reality of all district court cases. As they put it (*ibid.*, p. 535),

[i]t is not surprising that published [appellate] opinions offer a different perspective . . . than do district court filings. Parties do not randomly decide which cases to settle and which to continue through to district court judgment. Nor do they randomly appeal these judgments. Nor do judges randomly decide which of these appeals to publish. This filtering process . . . makes appellate opinions a biased sample of district court filings.

But our primary concern is how published district court decisions—which as we will show are often used in social science research—differ from the broader class of all district court filings.

Although we both focus on civil rights cases, Eisenberg and Schwab examine "constitutional torts" while we look at employment discrimination. We both compare samples of cases with and without published opinions (although they use appellate opinions, while we use those from district courts). We, too, find that cases with written opinions are systematically different from those without. Because of the differing objectives of the two studies, we consider a broader set of variables in describing how published cases differ from unpublished, and present a more integrated theoretical treatment of why non-representativeness occurs. In addition, there are a number of significant differences in the precise method of sampling. We compare all employment discrimination cases from a single district over a fifteen-year period based on whether the case generated a published district court opinion. Eisenberg and Schwab focused on a slightly larger class of cases in three districts, and compared district court decisions for a single year (1980–81) with the encompassing circuit courts of appeals' decisions over a five-year period (1980–85). Thus, our study has the advantage of greater precision in terms of subject matter and in having a precise geographic and temporal overlap between the two classes of cases over a broader expanse of time, while their study has the advantage of greater breadth in terms of subject matter and region over a more sharply focused time period.

Cartwright (1975) also discussed problems in using published cases but concentrated on issues of ambiguous outcomes, missing data, and so on.

which a published opinion is generated; section IV suggests the ways in which cases with published opinions might differ from the much larger body of cases in which no opinion is published. Section V compares samples of district court cases with and without published opinions to assess how such cases actually differ. We conclude with a brief review of some of the ways in which social scientists studying employment discrimination have used published opinions, showing how a shift from published cases to *all* cases can at times dramatically alter our understanding of employment discrimination law.

We focus for the most part on a single substantive area of federal law—employment discrimination—in order to make use of what we already know about the factors that affect the volume and composition of cases that get brought, the settlement rate, and so on.⁵ We have already examined a rich source of data containing the complete listing of *all* federal employment discrimination cases filed over a fifteen-year period⁶ and have begun to analyze the data from a random sample of some twelve hundred employment discrimination case files.⁷ Thus, we have data that allow us to examine whether and how cases with district court published opinions differ from other federal court cases.⁸

We will at times be critical of some of the uses to which (samples drawn from) employment discrimination cases with published opinions have been put. We should therefore make it clear at the outset that we are not opposed in principle to using cases with published opinions (or appellate cases, or Supreme Court cases) to conduct social research.⁹ Our main theme, however, is that the

⁵ For example, Donohue and Siegelman (1991) and Siegelman (1991) demonstrate that the volume of employment discrimination cases filed, the settlement rate in such litigation, and the plaintiff win rate are all sensitive to the overall condition of the economy (as measured by the unemployment rate) at the time a case is filed.

⁶ These data come from a tape compiled by the Administrative Office of the U.S. Courts (AO) containing all 92,090 cases filed in federal district courts under Code 442 (Civil Rights, Employment) between 1 July 1972 and 31 March 1987. For further details about this data source and an extensive analysis, see Donohue and Siegelman (1991) or Siegelman (1991); Eisenberg (1989: 1574) also contains useful information on the Administrative Office data.

⁷ A research project of the American Bar Foundation (ABF) is in the process of compiling data coded from the original case files in these twelve hundred cases, drawn from seven cities. In what follows, we will refer to this data set as the ABF Survey.

⁸ Most of the analysis that follows will be limited to samples of cases with and without published decisions in the Northern District of Illinois. We focus on this single district largely for convenience, although it is the judicial district with the largest volume of employment discrimination cases in the country.

⁹ We agree with Karl Llewellyn, who wrote that “the judicial opinion is a human document and a fascinating record, there, for anybody’s use. From the standpoint of behavioral study these are data in which so many factors are held equal as to outrun the results of an ordinary ten- or even hundred-thousand dollar grant. All there. All waiting. Already gathered. Merely neglected” (Cartwright, 1975: 369, citing Llewellyn, 1960: 514). Since Llewellyn was writing in 1960, his “ten- or . . . hundred-thousand dollar grant” would need to be

body of cases with published opinions needs to be viewed with greater sensitivity to the ways in which such cases are unrepresentative of all cases. Rather than ignoring the issue of representativeness or relegating it to a footnote, we want to urge scholars and/or consumers of research to think more carefully about the kinds of biases introduced by concentrating on published opinions. If we know that there are systematic divergences between cases with published opinions and those without, we can interpret results accordingly.¹⁰ Our aim is to provide a sense of these differences.

II. ESTIMATES OF THE PUBLICATION RATE FOR EMPLOYMENT DISCRIMINATION CASES

This section compares the number of employment discrimination cases with and without published opinions. If most cases generated published opinions, the ability of published cases accurately to represent *all* cases would be of relatively little importance. To summarize our conclusion, however, roughly 80 percent of the 4,310 cases in the population we studied did not produce a published district court opinion.¹¹

Before we can say anything about the publication rate for employment discrimination cases, however, we need to explain the nature of our data.

A. *The Population of Employment Discrimination Cases*

Every case filed in federal courts is classified into one of eighty-five "Nature of Suit" categories established by the Administrative Office of the US Courts (AO).¹² "Civil Rights, Employment" (Code 442) contains cases brought under Title VII of the 1964 Civil Rights Act,¹³ as well as under sections 1981 and 1983 (the Reconstruction Era Civil Rights Acts), the Equal Pay Act, the Rehabilitation Act, and the Age Discrimination in Employment

multiplied by a factor of more than three to be expressed in current dollars. We discuss the costs and benefits of using published opinions in the conclusion.

¹⁰ Econometricians have developed sophisticated techniques to deal with what has become known as the sample selection bias problem. The pioneering reference is Heckman (1979). Greene (1990b: 262) describes the sample selection correction this way: "The general solution to the selectivity problem relies upon an auxiliary model of the process generating [the sample]. Information about this process is incorporated in the estimation of [the variables of interest]." However, alternative techniques—including nonrigorous means of compensating for selection bias—may be just as effective (Stolzenberg and Relles, 1990: 408).

¹¹ As explained in the Appendix, we started with the complete listing of federal employment discrimination cases for the Northern District of Illinois between 1 July 1972 and 31 March 1987.

¹² The civil cover sheet in which this and other information is coded is filled out by the plaintiff's attorney and checked by a clerk.

¹³ Title VII is the basic federal statute prohibiting discrimination in employment on the basis of race, sex, religion, or national origin.

Act; the category also contains some suits by government employees alleging violation of their Fifth and Fourteenth Amendment rights to due process (typically, through discharge without proper civil service procedures).¹⁴ In the Northern District of Illinois, there were 4,310 unique Employment Civil Rights cases filed between 1 July 1972 and 31 December 1986.¹⁵ Further details about the Administrative Office data may be found in the Appendix.

B. The Population of Employment Discrimination Cases with Published Opinions

“Employment discrimination cases with published district court opinions” is a less precise concept than “Employment discrimination cases” per se, because the words “published opinions” are subject to multiple interpretations.¹⁶ We take “published opinion” to mean “included in LEXIS,” rather than restricting the definition to those cases published in a reporter such West’s *Federal Supplement* or BNA’s *Fair Employment Practice Cases Digest*.¹⁷ There are two justifications for this decision. First, since social science researchers are now taking advantage of the ability to search computer libraries such as LEXIS and WESTLAW to elicit data about litigation, it is interesting to explore how accurate these uses are, given the selective nature of publication. Second, to the extent that we use a more expansive definition, we adopt an approach that is least favorable to our hypothesis. The broader the definition of

¹⁴ Although the statute under which the case is brought was not listed on the AO tape before 1989, the ABF’s survey—based on a sample of 1,200 cases from the tape—demonstrates that most (more than 80 percent) of the cases raise a Title VII claim, either alone or in conjunction with one or more of the other statutes. For further details about the AO’s coding procedures, see Administrative Office of the United States Courts, Statistical Analysis and Reports Division (June 1989).

¹⁵ As discussed in note 6, we obtained from the AO a tape containing the information for all of the “Civil Rights, Employment” cases filed between 1 July 1972 and 31 March 1987. This tape was used to generate the information about “all employment discrimination cases” that forms the baseline for our study.

Each case is supposed to be identified by a unique docket number, whose first two digits designate the year in which the case was filed, but we found that about 3 percent of all docket numbers were duplicates. On inspection, most of these cases had identical or virtually identical plaintiff and defendant names (e.g., “*Schmoo v. Weinberger*” and “*Schmoo v. Sec’y of Defense*”) and filing dates. We therefore deleted the second occurrence of all duplicate docket numbers.

¹⁶ Brenner (1989) citing Reynolds and Richman (1978: 1169–71):

In the 1970’s the federal circuits adopted rules which distinguish between “published” opinions, which are “precedent,” and “unpublished” opinions, which are *not* “precedent.” These rules equate “publication” with appearing in print in a West reporter. . . . An opinion is “published” when its full text appears in that reporter and “unpublished” when only a notation of the disposition of the case appears.

¹⁷ LEXIS now includes all cases published in hard copy by BNA’s *Fair Employment Practice Cases*, dating back through the start of the digest in July 1965. Conversation with Michelle Love, Mead Data Central, 27 April 1990.

“published cases,” the more such cases we will find and the higher the ratio of published cases to unpublished. Other things equal, published cases are more likely to be representative of unpublished cases if the ratio of published to unpublished is 1:2 than if it is 1:10.

The technique we used to determine the set of cases with published opinions involved matching the docket numbers of the cases found in an extremely broad LEXIS search with those listed on the Administrative Office tape. Interested readers will find a fuller description in the Appendix.

C. Defining the Publication Rate

Conceptually, the definition of the publication rate is fairly clear—it is simply the ratio of the number of published cases to the number of cases that are eligible for publication. In practice, however, defining the publication rate is not completely straightforward, with most of the difficulty stemming from the measurement of the denominator in the publication rate calculation.¹⁸

We chose to take the broadest possible definition of “cases eligible for publication,” assuming that any filed case is at least potentially capable of generating a published opinion. One might instead choose to limit the denominator—for example, to cases that required some judicial action. Such a limitation might be appropriate for research about the legal system—in computing judges’ workloads, for example. But our goal was to get as far down the pyramid of disputes as possible. Cases that required judicial action are likely to be subject to the same kinds of sample selection problems as published cases. To determine whether published cases yield an accurate picture of the *lowest* level of the pyramid, the appropriate comparison is therefore between published cases and all filed cases.

D. A Comparison of the Two Populations

We begin with two populations of cases: the first is the list of the 4,310 employment discrimination cases in the Northern District of Illinois filed over a fifteen-year period, and the second is a list of cases generated by LEXIS for the same district and time period using the very expansive search (described in the Appendix) to capture all employment discrimination cases that have generated published opinions. For the period as a whole, the volume of LEXIS cases averaged 71 percent of the volume of AO cases (although with considerable year-to-year variation).¹⁹

In fact, however, our expansive LEXIS search captured many

¹⁸ There may be some complications arising from our definition of “published cases,” as well, as noted earlier.

¹⁹ That is, we found 3,058 unique cases in our LEXIS search, and 4,310 unique cases on the AO tape, for a ratio of 0.709.

nonemployment discrimination cases, so that the true percentage of AO cases that turned up on LEXIS was far smaller than 71 percent. Figure 1 plots the number of cases on the AO tape that were also found in the LEXIS search as a percentage of all cases on the AO tape. The figure makes it clear that most of the LEXIS cases are not the same as those listed on the AO tape. This should not be surprising: given the deliberately broad search technique used, we were certain to pick up a large number of cases that were not about employment discrimination.²⁰

The percentage of AO cases that can be matched with a LEXIS case clearly has been rising over time: only 6 percent of the employment discrimination cases that were filed in 1974 generated an opinion that subsequently appeared on LEXIS, while by 1986 this percentage had risen to 27 percent. The reason for this strong upward trend is unclear.²¹ One possibility is that the trend arises from a behavioral change on the part of LEXIS: the company might now include cases that previously it ignored. The competition between WESTLAW and LEXIS to be the most comprehensive computer library of legal material might explain such a trend, although LEXIS officials contend that there has been no such change in policy. Another possibility is that the trend is caused by judges writing opinions in cases that they would not have written on in the past, or submitting for publication opinions that an earlier generation of judges would have consigned to the files.²² We suspect that

²⁰ For example, any opinion referring to price *discrimination* or citing *Title VII* of the Bankruptcy Code would be detected by our search. So would an opinion in which the judge wrote "Unlike Title VII, the Fair Housing Act provides for" Of course, many cases brought under §§ 1981 and 1983 do not concern discrimination in employment at all. It is especially likely that many of the 1404 additional citations found under § 1983 (i.e., those that did not mention Title VII) fell into this category (see Appendix Table A1).

²¹ Indeed, the estimated trend is actually biased downward, although the extent is probably minor. The bias results from the fact that not all cases filed by 1986 had been terminated by the time of our LEXIS search in 1989, and therefore some cases that will later produce a published opinion are not yet included. Since the duration of the average case is slightly more than one year (with a standard deviation of about 1.5 years), however, the bias should be small.

²² The official federal reporters of district court decisions, which appear on LEXIS, have been termed the vanity press of the federal judiciary, since any opinion that a federal judge submits to West Publishing Company will be published. As federal judges get younger, perhaps the average degree of vanity is rising.

One other possibility is that increasing numbers of law clerks are now able to write decisions that the judge would in the past have decided orally or in a briefer and less noteworthy written opinion. This is unlikely to be a factor, however, because the number of law clerks per federal district court judge has not changed since it rose to two in 1965 (Posner, 1985: 102). Nor is expanding use of word processors likely to be the explanation since the trend precedes the greater use of personal computing technology in the 1980s. Finally, we can reject the notion that the workload of district court judges has been shrinking over time, enabling them to write more publishable opinions: there were 365 filings per district court judge in 1972 and this number rose quite steadily to 572 in 1983 (*ibid.*, pp. 358–60, Table B.4).

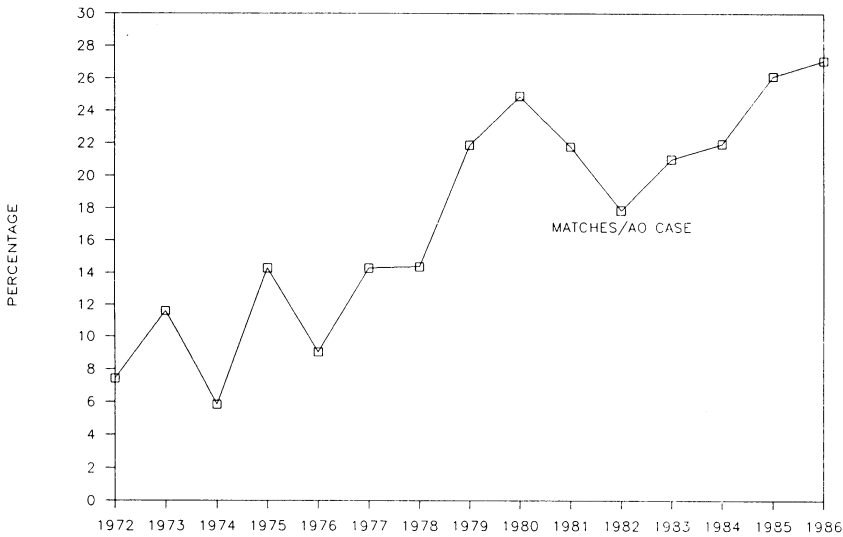


Figure 1. Matching docket numbers between tape from Administrative Office of the United States Courts (AO Tape) and LEXIS search (described in text) as a percentage of the number of cases on the AO tape, by year, Northern District of Illinois, 1 July 1972 to 31 December 1986.

the primary reason is that federal law in general and employment discrimination law in particular are becoming more complicated.²³

More central to our concerns in this article than the upward trend in published opinions per filed cases is that the match rate (publication rate) for the period as a whole averages only 20.1 percent:²⁴ only 867 of the 4,310 cases on the AO tape generated any sort of published record as indexed by LEXIS. Excluding all cases without published opinions thus eliminates 4 in 5 of all cases filed. For some purposes, using published cases as a method of sample selection may be innocuous; as we argue below, however, it will often produce biased samples.

²³ Because judges are less likely to write opinions that only address clear or settled legal issues, complexity and changes in law should generate more written decisions. The fact that the average length of federal circuit court opinions and the number of footnotes and citations per opinion has been increasing over time provides some evidence of the increasing complexity of the opinions themselves (Posner, 1985: 358–60, Table B.4).

²⁴ This is higher than the 12.2 percent (11 cases out of 90) publication rate for civil rights cases of all kinds found by Olson (1989). There are at least four possible explanations for this difference. First, Olson's data all come from the federal district court for Minnesota, and publication rates seem to differ by district. Second, her category includes civil rights cases of all kinds, while we focus on employment civil rights cases. Third, her data come from a two-year period (1982–84), and cover only a small number of cases (90 civil rights cases of all kinds, of which 11 were published). Our sample covers 4,310 cases over fifteen years. Fourth, the difference in publication rates across the two studies might merely reflect random sampling differences. Indeed, the χ^2 test (1 d.f.) reveals no significant difference in the two proportions ($\chi^2=3.44$, $p>0.05$).

E. Comparing Publication Rates Across Districts

1. Methodology. To provide some perspective on the finding that 20 percent of the employment discrimination cases filed in the Northern District of Illinois over a fifteen-year period generated a published district court opinion, we found it useful to compare this figure with publication rates in other districts, both for employment discrimination and other civil cases. The comparison was facilitated by the fact that we had already generated random samples of roughly 150 cases in each of seven districts using the AO tape of all "employment civil rights" cases.²⁵

To obtain publication rate estimates, we first determined which cases were *not* about employment discrimination,²⁶ and eliminated them from consideration. We then searched on LEXIS (by docket number) for each of the remaining employment discrimination cases.²⁷ The publication rate was computed by dividing the number of cases with published opinions by the total number of employment discrimination cases in our sample (col. (3) of Table 1).

Column (4) of Table 1 also gives the "pseudo-publication rate," defined as the total number of *opinions* found divided by the total number of docket numbers searched for. Because multiple opinions were not eliminated, these estimates are not comparable with those derived earlier from the Northern District of Illinois or with the estimated publication rates presented in column (3). Nevertheless, the calculation of the "pseudo publication rate" for the employment discrimination cases is useful because it provides a method of comparing the relative likelihood that employment discrimination and nonemployment discrimination cases will generate published opinions. Because it is impossible to eliminate multiple opinions in estimating the publication rate for all civil cases, we were forced to use the "pseudo publication rates" in making this comparison.

To estimate the pseudo-publication rate for all civil cases, we first conducted a LEXIS search for all civil docket numbers by district.²⁸ For each of the seven districts, we then took the total

²⁵ The seven districts were chosen because we had already coded data from random samples of employment civil rights cases from these districts as part of the ABF study referred to earlier.

²⁶ That is, we deleted those cases in which the complaint did not allege a cause of action based on the plaintiff's age, sex, race, handicap status, religion, national origin, or pregnancy. The cases eliminated were virtually all due process claims brought by government employees (plus a smattering of miscoded cases that were not about employment civil rights at all).

²⁷ The format for docket numbers varies across districts and, in some districts, also changes over time. We took care to include all possible variants of docket number formats in our searches.

²⁸ Searching LEXIS for all civil cases is not straightforward. Our basic procedure was to use a search request such as

Lib = GENFED, File = DIST Court (Northern or N. or N.D. or ND

Table 1. A Comparison of Pseudo-Publication Rates for Employment Civil Rights and All Civil Cases in Seven Judicial Districts, Fiscal Years 1973-87

Federal Judicial District	No. of Employment Civil Rights Cases (1)	No. of Civil Cases (2)	Publication Rate, Employment Discrimination (3)	Pseudo-Publication Rate, Employment Discrimination (4)	Pseudo-Publication Rate, All Civil Cases (5)
Southern New York (New York, <i>n</i> = 176)	2,450	119,756	30.1(1)	43.2(1)	14.3(2)
Northern Illinois (Chicago, <i>n</i> = 148)	4,435	97,589	21.8(3)	41.2(2)	14.4(1)
Eastern Pennsylvania (Philadelphia, <i>n</i> = 135)	1,580	78,530	24.4(2)	35.5(3)	7.4(3)
Northern California (San Francisco, <i>n</i> = 148)	2,490	74,536	8.8(4)	12.8(4)	2.5(6)
Northern Texas (Dallas, <i>n</i> = 158)	2,077	43,835	5.7(5)	8.8(5)	2.6(5)
Northern Georgia (Atlanta, <i>n</i> = 182)	2,168	44,744	5.5(6)	7.1(6)	4.1(4)
Eastern Louisiana (New Orleans, <i>n</i> = 162)	1,521	75,070	5.5(6)	6.2(7)	1.6(7)
Weighted Average			14.3	21.7	8.1

SOURCE: Administrative Office of the United States Courts Annual Reports and LEXIS searches described in text.

NOTE: *n* = number of employment discrimination cases whose docket numbers were searched for in cols. (3) and (4); ranks are given in parentheses.

Col. (1): From AO tape—number of cases filed.

Col. (2): From AO Annual Report—number of cases filed.

Col. (3): Number of unique cases found in LEXIS search (by docket number) divided by number of docket numbers searched for (*n*).

Col. (4): Calculated as total number of opinions found in a LEXIS search for a sample of roughly 150 employment discrimination cases per district. Not comparable with earlier data, since multiple opinions are included here.

Col. (5): Calculated from a LEXIS search for opinions in civil cases, divided by the total number of civil cases closing in the period (col. (2)). Multiple opinions per case are included more than once. An alternative definition of the denominator (all civil cases that required some court action) substantially raises the publication rate (by a factor of 1.25-3 times) but perfectly maintains the ranking among districts.

Correlations:
 Between cols. (3) and (4) = 0.977
 Between cols. (4) and (5) = 0.938
 Between cols. (3) and (5) = 0.886

number of opinions found in our LEXIS search and divided by the total number of closed civil cases in fiscal years 1973 through 1987 (obtained from the Annual Reports of the Administrative Office of the U.S. Courts) to obtain an estimate of the pseudo-publication rate.

2. Results. Table 1 presents the results of these searches. Three strong conclusions emerge. First, publication rates for employment discrimination cases vary dramatically across districts. In the three northern, urban districts with large caseloads (New York, Chicago, and Philadelphia) more than one-fifth of all employment discrimination cases filed generate a published opinion. In the four other districts (three in the South, plus San Francisco), only about one in twenty employment discrimination cases yields a published opinion. Somewhat surprisingly, the higher publication rates are found in the three districts with the largest caseloads. Second, the overall publication rate for employment discrimination cases is quite low—less than 15 percent. Finally, the correlations among all the publication rate estimates are extremely high: districts that publish a relatively high proportion of their employment discrimination cases also seem to publish an unusually large share of *all* civil cases.²⁹

These results suggest that relying solely on cases with published opinions will generally produce a geographically skewed sample of all cases filed in the United States. Suppose, for example, that one had a list of all the closed civil cases from each of the seven judicial districts listed in Table 1. The probability that a randomly selected closed case would come from the Southern District of New York is only 22.4 percent, while the probability that a randomly selected *published opinion* would come from the same district is 38.5 percent, with the difference being attributable to New York's higher-than-average publication rate.

In sum, differences in publication rates by judicial district appear to be substantial, and to be highly correlated across employment discrimination cases and all civil cases. These differences may pose problems for researchers who use published opinions, as suggested below.

and Illinois) and date aft 6/30/72 and date bef 7/1/87 and number (C!) and not number (CR!)

This was designed to capture civil docket numbers such as "74 Civ 192" and "74 C 1972," while excluding criminal docket numbers such as "74 CRIM 189" or "74 CR 189." Again, docket number formats varied substantially across districts, and sometimes within districts over time as well, so that modifications of this strategy had to be used. More precise details are available on request from the authors.

²⁹ This is also true of the pseudo-publication rate estimates.

III. WHY CASES WITH PUBLISHED OPINIONS MIGHT BE DIFFERENT FROM THOSE WITHOUT

A. A Definition of "Unrepresentative Sample"

We all have a common-sense understanding of what it means to say that one group is unrepresentative of a larger group. This section presents an extremely simple formal definition of unrepresentativeness, based on the concept of statistical independence, which will be useful in analyzing the problems we discuss below.

Intuitively, a sample will be biased with respect to some variable X if that variable is independently distributed with respect to the variable Z that is used to select the sample. For example, if the sample selection variable Z is whether or not the case is published, then the value of any variable X that is correlated with Z will be biased as measured in the sample.³⁰

Note that this definition raises the important possibility that the sample is unbiased for some variables but not for others. For example, suppose that cases with published opinions, on average, last longer than those without. Sampling only those cases with published opinions will bias the researcher's estimate of a case's duration, since the expected duration of cases in the sample will be longer than the average duration of the cases in the entire population. Suppose, however, that the variable of interest is not the duration of the case but the plaintiff's gender. If the gender of the plaintiff is independent of whether a case generates a written opinion, then selecting only those cases with published opinions will *not* influence any estimate of the proportion of plaintiffs who are women. This discussion demonstrates that, above all, the sample selection problem requires the researcher to have a clear conception of the process that generates the sample being used, and which, if any, of the variables of interest are likely to be related to this process.

B. The Process That Generates an Opinion

Using the existence of a published case as our selection (Z) variable, we can see that selection (publication) is unlikely to be independent of many of the variables in which social scientists are interested. To understand how the subset of cases with published opinions is likely to differ from the set of all cases, therefore, we need to analyze the process that generates a published opinion. Olson (1989: 3) provides a convenient summary of this process:

³⁰ Formally, we would say that bias exists whenever the expected value of X conditional on Z is not equal to its unconditional expectation, or $E(X|Z) \neq E(X)$.

In another context (Good, 1968: 296), "as Sir Arthur Eddington once pointed out, if you catch fish with a net having a 6-inch mesh, you are liable to formulate the hypothesis that all fish are more than 6 inches in length."

Table 2. A Hypothetical Depiction of the Process Generating a Published Opinion in an Employment Discrimination Case

Stage of Process	% of All Cases Filed That Reach This Stage	% of Previous Stage
1. Initial number of cases filed (settled, dropped, etc., without producing an opinion)	100 (60–40)	
2. Adjudicated (Kritzer's broad definition)	40–60	40–60
3. Judge writes decision	35	58–88
4. Judge submits decision for publication	12–22	34–63
5. LEXIS publishes decision ^a	11.75–21.5	98

^a Mead Data Central claims that it publishes everything it receives. Phone conversation with Michelle Love, 27 April 1990.

Four steps occur from the termination of a case to its being reported in one of the national . . . computerized sources. First, the case must be disposed of by a judicial decision. Second, the judge must write an opinion rather than merely issuing the decision orally. Third, . . . Mead Data Central [the company that runs LEXIS] must acquire the published opinion. Fourth, . . . [it] must decide to report the opinion.

We should offer one correction to this categorization. A case may well generate a published legal decision on a preliminary matter that is not dispositive. For example, a motion for class certification is often an important question in employment discrimination matters, and a case will often settle in the wake of the judge's ruling on this motion. But if significant issues of law and/or fact are raised, the judge might well make her decision in this preliminary matter in writing, and this decision could then be published. Nonetheless, it is almost certainly the case that rulings that dispose of a case are more likely to be written and more likely to be published. As long as this condition is true, then settlement (or the dropping of the case by the plaintiff) will tend to reduce the likelihood that a case will generate a published legal decision. Since so many cases are settled or dropped, this factor will have an important impact in reducing the likelihood of a case appearing in LEXIS: the best available evidence seems to suggest that settlement of one sort or another eliminates at least 40–60 percent of all filed cases.³¹ Table 2 provides more detail.

Borrowing from the language of demography, we can think of the various possible case outcomes as different “causes of death”

³¹ Kritzer (1986: 165) found that judges made an adjudicatory decision in roughly 43 percent of his sample of 809 federal civil cases of all kinds. The comparable figure for civil rights cases was 57 percent, but this included both state and federal cases, and was based on a small *N* of only 109; it is therefore not strictly comparable. Moreover, Kritzer's definition of adjudication is a rather broad one and probably includes situations in which judges are unlikely to produce a written opinion.

for a case. If settlement “kills off” 40–60 percent of all cases filed, and if 10–20 percent of all cases end in the publication of a published opinion, all other causes (rows 3 and 4 in Table 2) combined must account for only 20–50 percent of all “deaths.”³²

IV. CONSEQUENCES OF THE SELECTION PROCESS

A. *Settlement Effects*

The key insight of section III.A is that sample selection bias in the measurement of variable X will occur whenever the criterion for inclusion in the sample and the variable X are not distributed independent of each other. In other words, if anything that causes a case to “die” without a published opinion is correlated with any variable of interest to social scientists, that variable will be subject to sample selection bias when the sample consists only of published cases. This suggests that we should begin our analysis of selection bias with the first (and most important) cause of “premature death”—settlement.³³

Settled cases are less likely to produce published opinions than those that terminate with an adjudicated result for two reasons. First, intermediate rulings (for example, on motions to dismiss, motions for summary judgment, and so on) are substantially less likely to be published than are final adjudications, since judges do not typically produce lengthy opinions on such issues, and are less inclined to submit what they do write for publication. Moreover, many settled cases settle before the judge has a chance to produce any intermediate rulings.

As long as settled cases are less likely than adjudicated cases to produce published opinions, any variable that influences the probability of settlement will not be independent of the sample se-

³² One caution about applying these percentages to our sample should be noted. Olson and Kritzer both used samples of terminated cases only; we use *all* cases filed over a certain period, whether terminated or not. Given our previous claim that a judicial decision that terminates a case elevates the likelihood that a published decision will emerge, the proportion of cases in our sample that generates published decisions will be biased downward, although as argued earlier, this bias should be small given that the sample period ended more than two years before the LEXIS search was carried out.

In theory, one could correct for this problem by computing hazard rates showing the likelihood that cases in our sample will generate a published decision in the next time period. The benefit from estimating hazard functions is that a researcher can simply use the information from any case to indicate whether the event in question—the publication of a decision—occurred before the case ended or the period of observation came to a close.

³³ Employment discrimination cases are, of course, “disputes” before they are filed cases, and as such they are subject to prior screening by the EEOC, to negotiation by the parties before any complaint is filed, and so on and on. The utility of looking at cases with written opinions, or legal cases at all, depends of course on how far down the pyramid of disputes one wishes to look. For some limited evidence that EEOC screening serves as an essentially neutral filter that does not dramatically alter the composition of cases, see Donohue and Siegelman (1991).

lection criterion. The key question then becomes, what determines whether a case settles or not? One can draw upon the Priest/Klein model of the selection of disputes for litigation in order to derive theoretical predictions of how cases generating published opinions might differ systematically from cases that yield no published decision. For example, in other work we have shown that the overall win rate at trial for plaintiffs in employment discrimination cases (including some of those with published opinions and some without) is quite low, perhaps as low as 20 percent, and certainly less than 50 percent. The Priest/Klein model predicts that this pattern will be observed when the defendant's stakes in the litigation are greater than the plaintiff's, which might well be true in certain employment discrimination contexts. For example, a defendant might well believe that a finding that he unlawfully refused to hire blacks would subject his firm to a possible loss that was far greater—either from bad will in the community or from demands from other rejected black applicants claiming similar discrimination—than what the plaintiff would gain from trial. Thus, the defendant would be more willing to settle out cases that might impose this loss than he would if the stakes between the plaintiff and defendant were equal.

Suppose that our sample were evenly divided into fifty cases of discrimination against black job applicants, in which the stakes were higher for the defendant than for the plaintiff, and fifty cases of refusal to grant paid leave to Jewish workers on Yom Kippur, in which the stakes were even. We might expect *ceteris paribus* that the settlement rate would be higher in the first category. Now since we have claimed that cases that end in settlement are less likely to generate published opinions, this fact combined with the above differential settlement rate yields the following predictions of systematic differences between cases with and without published decisions: (1) published decisions will be more prevalent among the cases in which the stakes between the parties are equal; and (2) readers of published opinions will think that cases of discriminatory refusal to hire blacks are less common than they in fact are, both absolutely and relative to other types of discrimination cases.

A substantial literature also indicates that uncertainty about the likely outcome of a trial will diminish the chance that the case will be settled. The reason is that if both parties know what will happen, they are better off settling for this amount and pocketing the litigation costs they have avoided. But if settled cases will tend to be those with lower uncertainty (other things being equal), then more uncertain cases will tend to settle less and therefore will have higher rates of published opinions. Consequently, more complicated disparate impact cases (or perhaps age discrimination cases) should be more numerous among those with written opinions than among those without.

B. Other Effects Based on Judicial Behavior

The other three stages in the production of a published opinion (described above in Table 2) seem to be less useful in terms of providing predictions of systematic differences between cases with and without published decisions. For example, lacking a good theory about what makes a judge decide to write or submit an opinion, we are not well equipped to explain what sorts of biases are introduced by these decisions.

Nevertheless, it seems reasonable to suppose that other things being equal, judges will be more likely to write and submit an opinion in a case the more "important" that case is.³⁴ "Important" is obviously a highly subjective word.³⁵ One measure of importance might be the size of the stakes involved—larger monetary demands or complex efforts to reshape hiring or promotion policy. But considerations of *legal* importance are probably more relevant than factual importance. That is, judges will be more likely to write and submit an opinion if the case breaks novel legal ground than if it merely presents a new (or interesting) fact situation.³⁶

One might think of district court decisions that break new legal grounds as analogous to genetic mutations. There are more than five hundred federal district court judges, and when one of them breaks new ground it likely means that he or she is doing something differently than others have done. As with genetic mutations, some of which are highly favorable to the survival of the species but most of which are not, the preponderance of highly unorthodox district court decisions—Judge W. Brevard Hand's ruling that secular humanism is a religion that cannot be presented in the public schools³⁷ comes to mind—are aberrant and soon to be

³⁴ Olson's data, however, seem to refute this hypothesis. She finds that when lawyers are interviewed after the conclusion of a case, their assessment of whether the case raises broader legal issues (a measure of its importance) bears no relationship to the probability of publication (Olson, 1989: 9–10). She also cites political science research suggesting that "judges' egos and years of experience on the bench influence how many of their opinions they consider important enough for publication" (*ibid.*, p. 5, citing Carp and Rowland, 1983). Although such variables as age, experience, and self-esteem may influence the probability that a given judge produces a published opinion in a given case, since cases are assigned randomly to judges, such effects presumably wash out in larger samples and thus can be ignored. However, a reviewer pointed out to us that a major statute creating new judges "could certainly make the effects [of age and experience] non-random . . . for a study comparing publication rates over time."

³⁵ Eisenberg and Schwab (1989: 508) cite two studies by Franklin (1981: 799 n. 11; 1980: 464) demonstrating that "trial judges are more likely to publish opinions when making decisions that conclude cases, such as granting motions to dismiss or granting summary judgment, . . . than when denying such motions."

³⁶ Eisenberg and Schwab (1989: 535) write: "Judges themselves decide whether to publish an opinion. The general criterion for publication is that the case be noteworthy rather than routine or obvious, and thus will contribute to the development of the law."

³⁷ *Smith v. Board of School Commissioners of Mobile County* (1987).

reversed on appeal. Thus, one might expect to find (1) more unusual district court decisions among reported decisions and (2) higher rates of reversal for appeals from published decisions than for appeals from among nonpublished decisions.³⁸

V. AN EMPIRICAL ASSESSMENT OF PUBLISHED OPINION BIAS

The analysis of the previous section suggested several ways in which cases with published opinions are likely to differ from those without them. This section attempts to test these hypotheses by comparing a random sample of published employment discrimination cases from the Northern District of Illinois³⁹ with a random sample of such cases that did not produce published decisions.⁴⁰ Once we had selected our two samples, we collected data from the original district court case files at the Federal Records Center in Chicago.⁴¹

Tables 3–7 present some simple comparisons of the two samples. As we predicted above, the cases with published decisions tend to be more complex as evidenced by a number of different factors. First, the file is more than two and one-half times as thick in published cases as in nonpublished cases. Second, the average number of plaintiffs is considerably higher in the published cases, and the proportion of cases that were certified as class actions in published cases (13.1 percent) is more than six times the proportion in unpublished cases (1.9 percent). Third, in two-thirds of the cases from the published sample, violations were alleged to be continuous—for example, harassment occurring over a period of time, as opposed to a single refusal to hire—while only half the cases

³⁸ A reviewer pointed out to us that this proposition has been tested. Songer (1988: 211) compares reversal rates for cases not listed in the *Federal Supplement* and those that were listed (published, in his definition). He finds both surprisingly high reversal rates in unpublished opinions and only slightly higher reversal rates in published opinions.

³⁹ As discussed in sec. II.B, there were 867 cases (unique docket numbers) that were (1) listed on LEXIS and (2) matched one of the 4,310 docket numbers on the AO tape. We randomly selected 114 cases from this population of 867.

⁴⁰ To obtain this sample, we began by selecting 179 cases at random from the population of all employment discrimination cases filed in the Northern District of Illinois over the period 1972–87. (See Appendix for further detail.) We then searched on LEXIS for these 179 docket numbers and removed the 39 cases with published opinions. We were thus left with a sample of 140 cases. We also removed all cases that were not about employment discrimination per se, leaving a total of 105 employment discrimination cases without published opinions.

Note that this procedure also serves as a check on the estimate of the publication rate produced earlier. It yields an estimated publication rate of $39/179=21.8$ percent. This is not statistically significantly different from the 20.1 percent publication rate estimated above on the entire sample of 4,310 cases (χ^2 (1 d.f.) = 0.49, $p > 0.05$).

⁴¹ Copies of the code sheet, coding instructions, and the raw data used in the following analysis are available by request from the authors.

Table 3. Samples of Published and Unpublished Employment Discrimination Cases, Northern District of Illinois, 1972–86 (Standard Deviations in Parentheses)

A. Continuously Measurable Characteristics of Cases

Variable	Average in Unpublished Cases	N	Average in Published Cases	N	Z-value ^a
File size (in inches)	2.0 (4.0)	97	5.2 (11.5)	114	2.78 [0.01]
Total dollars awarded ^b	\$12,545 (\$12,979)	18	\$606,424 (\$1,687,853)	19	1.53 [0.12]
Dollars/plaintiff	\$12,375 (\$13,088)	18	\$47,907 (\$47,867)	15	2.79 [0.01]
Fees awarded	\$6,400 (\$6,505)	2	\$29,765 (\$34,700)	9	1.88 [0.06]
Lag between violation and filing, in quarters ^c	6.24 (6.07)	94	4.77 (6.85)	105	1.61 [0.10]
Average no. of plaintiffs	1.5 (3.1)	99	18.2 (136.2)	103	1.24 [0.21]
Plaintiff's job tenure (in years)	8.4 (5.9)	14	7.8 (9.8)	99	0.32 [0.74]
Plaintiff's age	48.3 (10.8)	6	49.1 (10.8)	45	0.17 [0.86]
Total cases in sample		105		114	

^a Z is the number of standard deviations from the mean of an $N(0,1)$ variable. The number in brackets is the significance level for a two-tailed test. The normal rather than the T test is used because we do not assume equal variances in the two samples. See, e.g., Hogg and Tanis (1977: 252).

^b Includes fees, costs, and damages; excludes cases in which plaintiffs had to pay defendants' fees or costs.

^c The median lag was 4.5 quarters in the unpublished cases and 3.0 quarters in the published cases.

B. Discrete (0,1) Characteristics of Cases

Cases in which:	Unpub-lished Cases (%)	N	Published Cases (%)	N	$\chi^2_{(1)}$	Significance Level
	There was more than 1 plaintiff	13.3	105	22.8	114	3.28
At least 1 plaintiff was black	71.7	99	63.0	73	1.13	0.31
A class action was certified	1.9	105	13.1	114	9.67	0.01
Violation was continuous ^d	49.5	103	66.9	112	6.73	0.01
Plaintiff alleged a disparate treatment theory	85.6	104	91.1	112	1.59	0.21
Plaintiff alleged animus-based discrimination	72.1	104	74.7	111	0.19	0.66
Outcome was appealed	1.9	105	20.1	114	17.9	0.01

SOURCE: ABF survey of case files.

^d A continuous violation occurs over some interval of time (e.g., harassment), as opposed to a discrete, one-point-in-time violation such as failure to hire or firing.

from the unpublished sample made this allegation. In general, the volume of evidence that is required to demonstrate a continuous violation will be greater, thus contributing to the legal and factual complexity of the case. Fourth, Table 6 reveals that plaintiffs in cases that generated published opinions seek more forms of relief

Table 4. Basis of Suit, Basis of Discrimination, and Type of Discrimination in Published and Unpublished Employment Discrimination Cases

	% of Cases Without Published Opinion	% of Cases With Published Opinion	$\chi^2_{(1)}$	Significance Level
A. Cases raising a claim under:				
Title VII	84.8	78.1	1.61	0.20
Age Discrimination in Employment Act	5.7	24.6	14.80	0.01
Equal Pay Act	6.7	0.9	5.21	0.02
§ 1981	32.4	27.2	0.77	0.40
§ 1983	14.3	10.5	0.71	0.40
Rehabilitation Act	1.0	1.8	0.26	0.62
Constitutional amendment	16.2	18.4	0.18	0.67
Pendant state claims	2.9	9.6	4.21	0.04
B. Cases in which basis of discrimination included:				
Age	6.7	25.4	14.02	0.01
Sex	34.3	31.6	0.18	0.67
Race	63.8	48.2	5.37	0.02
Handicap	2.9	6.1	1.35	0.26
Religion	1.0	2.6	0.86	0.38
National origin	9.5	12.3	0.43	0.51
Pregnancy	1.9	2.6	0.13	0.73
Other	1.0	0.0	1.09	0.28
C. Cases in which type of discrimination included:				
Hiring	13.3	17.5	0.74	0.39
Firing	75.2	62.3	4.25	0.02
Pay	25.7	19.3	1.30	0.26
Demotion/promotion	28.6	36.8	1.69	0.20
Seniority	5.7	6.1	0.02	0.89
Conditions of employment	36.2	38.6	0.14	0.62
Retaliation	11.4	20.2	3.11	0.08
Total no. of cases	105	114		

SOURCE: ABF survey of case files.

NOTE: Because plaintiffs could sue under more than one statute or allege more than one basis or type of discrimination, the percentages do not add to 100 percent and the χ^2 tests are therefore line-by-line comparisons.

than plaintiffs in cases that do not result in published opinions: note that in only one of the categories—punitive damages—is the proportion in the first column greater than the proportion in the second column, and even here there is no significant difference between the two.

The amount in controversy is similarly far greater for the published cases than the unpublished cases. The best evidence of this is the finding in Table 3 that the average number of dollars awarded per plaintiff is almost four times as great in published cases (\$47,907) as in unpublished cases (\$12,375).⁴² As one would

⁴² The figures for total dollars awarded in Table 3 show an even more dramatic difference—\$606,424 versus \$12,545—but one that is less statistically significant than the per plaintiff comparison. This apparent anomaly is explained by the unusually high award in one published case with numerous plaintiffs, which caused the average award in the published cases to rise,

Table 5. Plaintiffs' Occupations and Defendants' SIC Codes in Published and Unpublished Employment Discrimination Cases

	Cases Without Published Opinion	Cases With Published Opinion
No. of plaintiffs whose occupation was:		
Managerial and professional	17	31
Technical, sales, and other support	17	30
Service	16	15
Precision production	3	12
Operator, fabricator	26	18
Total	79	106
$\chi^2_{(4)} = 10.86$		
Significance level = 0.03		
No. of defendants whose SIC was:		
Mining	1	0
Construction	3	2
Manufacturing durables	13	14
Manufacturing nondurables	10	15
Transportation	11	12
Utilities and communication	3	8
Wholesale/retail trade	10	9
Finance, insurance, real estate	2	3
Services	16	8
Government	19	30
Total	88	101
$\chi^2_{(9)} = 9.09$		
Significance level = 0.43		

SOURCE: ABF survey of case files.

expect given the higher damage awards, fees awarded to prevailing plaintiffs were higher in the published cases, although not commensurately higher as evidenced by the fact that the ratio of award per plaintiff/fee award was 1.9 for unpublished opinions but only 1.6 for published opinions. Perhaps related to the issue of bigger stakes is the finding that cases that will ultimately generate a published opinion get to court significantly faster than those that do not generate a published opinion: indeed, only 44.8 percent of all unpublished cases were filed within one year of the violation, while 60.1 percent of all published cases were. This is consistent with the notion that there is more to gain from litigation in published cases. Since pre-judgment interest is unlikely to be available, the cost of delay in filing is greater for higher stakes cases.

Table 4 demonstrates once again that complexity correlates with whether the case generates a published opinion. For example, far more published cases have pendant state claims attached to their federal charges, and the relatively complex, novel, and evolving Age Discrimination in Employment Act is dramatically over-represented in published cases. At the other extreme, the old,

thereby greatly increasing the standard error of the variable. Since the award was not unusually high when computed on a per plaintiff basis, the standard error for the dollars per plaintiff variable is far smaller in relation to the mean dollars per plaintiff.

Table 6. Cases in Which Plaintiffs Made Various Demands in Published and Unpublished Employment Discrimination Cases

	% of Cases Without Published Opinion	% of Cases With Published Opinion	$\chi^2_{(1)}$	Significance Level
Cases in which plaintiff asked for:				
Reinstatement, hiring, or promotion	54.1	69.8	6.01	0.01
Back pay	72.4	89.6	9.92	0.01
Injunction	58.2	59.4	0.03	0.87
Punitive damages	59.2	55.7	0.26	0.61
Declaratory judgment	29.6	31.1	0.05	0.82
Change employment practices ^a	8.2	22.6	7.99	0.01
Pension or other benefits	16.3	35.8	9.96	0.01
Front pay	0.0	7.5	7.65	0.01
Clear record	0.0	1.9	1.88	0.17
Liquidated damages	1.0	14.2	12.25	0.01
Training	0.0	1.9	0.29	0.59
Other	5.1	16.0	6.30	0.01
No. of cases	98	106		

SOURCE: ABF survey of case files.

NOTE: Because plaintiffs could make more than one kind of demand, the percentages do not add to 100 percent, and the χ^2 tests are therefore line-by-line comparisons.

^a Includes affirmative action plan, recruitment, change seniority system, etc.

somewhat outdated (because it has largely been superseded by Title VII), and static Equal Pay Act almost never appears in published cases but is actually more common than Age Discrimination Act cases in the unpublished cases. Moreover, as we suggested above, firing cases—in which the stakes would seem to be roughly symmetric for plaintiffs and defendants—appear significantly less often in the published sample while hiring cases are overrepresented in the published sample.

Table 5 reveals that plaintiffs' occupations (but not defendants' industrial classifications, at an aggregate level) are significantly different in the two samples. As one would expect, managerial, professional, and technical workers—which are the high-wage job classifications—are more common in the published sample, while low-wage operatives are overrepresented in the unpublished sample. Similarly, if one looked only at published cases, one would suspect that cases against utilities and government employers were more common and cases brought in the service sector were less frequent than they in fact are. The differences might be explained by a wage effect, in that service jobs pay less and utility and (federal) government jobs pay more.

If the degree of complexity and novelty tends to be greater in published cases than in unpublished cases, one would expect this factor to have an effect on the rate of settlement. Specifically, complexity and novelty increase the uncertainty of the case outcome and therefore decrease the likelihood of settlement (see, e.g.,

Table 7. Outcomes in Employment Discrimination Cases With and Without Published Opinions

Outcome	No. of Cases Without Published Opinion	No. of Cases With Published Opinion
Settled ^a	70	39
Plaintiff won at trial	1	9
Defendant won	29	61
Dismissed, procedural reasons	1	5
Plaintiff dropped case ^b	11	4
Dismissed, summary judgment for defendant	7	22
Dismissed other motion	6	6
Dismissed, failure to state a claim	1	7
Tried, judgment for defendant	3	17
Other	3	3
Total no. of cases	103	112
Plaintiff Win Rate ^c	0.033	0.129
$\chi^2_{(1)} = 2.12$		
Significance level = 0.15		

SOURCE: ABF survey of case files.

^a May include some cases that were simply dropped by plaintiff.

^b Includes cases dismissed for want of prosecution.

^c The win rate is defined as WINRATE = # of Cases won by Plaintiff/# of Cases won by either Party.

Priest and Klein, 1984). Accordingly, we should expect to see fewer settlements in the sample of cases with published opinions. Table 7 confirms this prediction. The level of settlement is far higher in cases without published opinions—68 percent, versus 35 percent in published cases. Another interesting finding is that more than 10 percent of unpublished cases are simply dropped (or dismissed for want of prosecution) by plaintiffs. Not surprisingly, a far lower percentage of cases producing published opinions fall into this category. At the same time that uncertainty tends to decrease the likelihood of settlement, it also probably tends to bias the likelihood of victory toward 50 percent. The reason for this effect is that as the outcome of the litigation becomes more uncertain, the event becomes closer to a coin flip in terms of the probabilistic outcomes. Therefore, one would imagine that the win rates would be closer to 50 percent in the more uncertain, published cases, and the data supports this supposition. *Looking only at cases that are tried and reach a final adjudication*, we see that the rate of victory for plaintiffs in unpublished cases is 25 percent (1 of 4) while the rate of victory in published cases is almost 35 percent (9 of 26).⁴³ In general, though, it is evident that win rates in employment discrimination cases are significantly below 50 percent. According to the Priest/Klein model this probably suggests that the stakes are higher for defendants than for plaintiffs. Again, this

⁴³ Of all *nonsettled* cases, plaintiffs appear to have won only 1 of 33 (3.3 percent) among those without published opinions, whereas they won 9 out of 73 (12.3 percent) of those with published opinions.

seems plausible as an adverse judgment in an employment discrimination case might inflict greater monetary loss on the defendant than the award received by the plaintiff.

Overall, the most important pattern that emerges from this welter of numbers is a simple one: cases with published opinions are indeed significantly different from those without them. The published cases tend to be longer, more complicated, more heavily concentrated on newer areas of the law. They also seem to include a different mix of plaintiff occupations, to proceed at a different pace through the legal system, and to end in different kinds of outcomes. For some purposes, differences between published and unpublished cases may not matter. The next section argues that in some instances, however, important conclusions turn out to be sensitive to the sample definition that is used.

VI. HOW SOCIAL SCIENTISTS HAVE USED EMPLOYMENT DISCRIMINATION CASES WITH PUBLISHED OPINIONS

Samples of employment discrimination cases with published opinions have been analyzed for a number of purposes, using a variety of techniques: content analysis, simple counts (to construct variables used for other purposes), and analysis of legal rules and outcomes. In this section, we briefly review some of these studies and suggest ways in which their findings are affected by their sample selection criterion.

A. The Effect of Title VII Suits on Black Employment Levels

Jonathan Leonard (1984) has studied the effects of affirmative action requirements and Title VII lawsuits on black and female employment. He developed his Title VII variable by conducting a LEXIS search for Title VII class actions.⁴⁴ He then sorted the suits by the defendant firm, and used the number of such suits per firm as an independent variable in a regression analysis to explain changes in black and female employment levels in each firm. On the basis of cross-sectional regressions, he reports that the number of Title VII class action lawsuits per corporation is significantly re-

⁴⁴ Leonard did not describe his search request, making it difficult to describe exactly what he did. Moreover, our attempts to replicate his results were unsuccessful. In part, this may be due to the inclusion in LEXIS of additional cases that were not covered when Leonard undertook his search. But without an adequate description of his search request, it is difficult to interpret his results.

This difficulty is amplified by the lack of internal consistency in describing his search. On p. 146, Leonard writes that "between 1965 and 1977 the federal district courts decided more than 1700 class action suits brought under Title VII of the Civil Rights Act of 1964." Four pages later, he notes that "between 1964 and 1981 [note that the period covers four additional years at the end, and one fewer at the beginning] more than 5000 cases of litigation under Title VII, many of which were private suits, were decided in the federal district courts. More than 1700 of these cases were class action suits."

lated to the rate of increase in the percentage of black workers over the period from 1966 to 1978.

In one sense, most of the biases inherent in looking at LEXIS cases would not be problematic for Leonard's research because he might well want to focus on the very high stakes, complex, and novel cases that tend to be overrepresented in published decisions. On the other hand, our own experience in trying to find employment discrimination cases of any type raises some concerns about the hope that LEXIS alone can pick out the right cases. Although Leonard's use of LEXIS was an imaginative attempt to evaluate the effect of Title VII litigation, the fact that he gives no mention to the LEXIS search strategy he employed suggests that he may not have been aware of how difficult it is to achieve accurate figures by simply relying on a LEXIS search.

First of all, any attempt to capture all Title VII class action lawsuits will undoubtedly be far too inclusive; yet, if the search is too narrowly defined, many cases will be missed. For example, many Title VII class action suits are brought by white women: thus, a simple search for Title VII risks bringing in all sorts of cases that are unlikely to be correlated with better *black* performance. (Similarly, a substantial portion of cases containing the words "Title VII" are actually age discrimination cases brought largely by white males.) Nor is it obvious that all class actions (whether or not the request for class certification was certified) should be treated identically, as Leonard appears to have done. Finally, it is not clear that cases charging discrimination in hiring should be treated symmetrically with those alleging discriminatory failure to promote or firing. The latter charges can only be brought by *current or former* employees. Thus, they are likely to be *caused by* the employer's hiring practices, rather than to be a cause of such practices.⁴⁵

To examine the possible problems inherent in Leonard's construction of his Title VII class action lawsuits variable, we performed a simple LEXIS search designed, as best we could, to approximate his original sample.⁴⁶ This search yielded 5,566 Title VII opinions and 1,815 Title VII class action opinions. We then read and coded all the class action opinions for the years 1966 through 1969 (inclusive) and 1972, and half the class actions in 1976 and 1981, for a total of 304 opinions. Our findings are summarized briefly in Table 8.

⁴⁵ In fact, the ability of current workers to contest promotions and discharge provides employers with an incentive *not* to hire women and minorities (Donohue and Siegelman, 1991). Whether this disincentive is strong enough to outweigh the positive incentive provided by hiring suits is an open question.

⁴⁶ The search request was:

Library = GENFED

File = DIST

Level 1: Text (Title VII) and date (aft 12/31/63 and bef 1/1/82)

Level 2: Text(class action)

Table 8. "Title VII Class Action" Opinions Found in Attempt to Replicate Leonard's Results

Total opinions found ^a	1,815	
Opinions read	304	
<i>Less:</i>		
Class definition uncertain ^b	47	
EQUALS	257	
<i>Less:</i>		
Not class actions	29	
Not employment discrimination	26	
Multiple opinions (same underlying case)	12	
EQUALS	190	"Valid" Opinions, of which:
Class definition includes minorities ^c	130	Of which
Hiring claims only	25	
Post-hiring claims ^d only	63	
Both hiring and post-hiring	42	
Other, unclassifiable, etc.	10	

SOURCE: LEXIS search described in text.

^a The search request used was Library = GENFED, File = DIST,

Level 1: Text(Title VII) and date (aft 12/31/63 and bef 1/1/82)

Level 2: Text(class action).

^b E.g., opinion deals with an award of attorney's fees and does not describe the nature of the class certified.

^c As long as the designation of the class seeking representation included blacks or other racial minorities (even if it also included whites), we included the opinion. Thus, class designations such as "all minority workers," "all black female job applicants," or "all black males and all white females" were included. Class designations such as "all job applicants" or "all female job-holders" were not included.

^d Post-hiring claims are those dealing with promotion, pay, conditions of employment, and firing.

The conclusion that emerges from Table 8 is that the variable Leonard appears to have constructed to measure Title VII litigation is extremely imprecise. Of the 257 opinions (nominally class action Title VII cases) that we could classify, 26 percent were either not class actions at all, were not about discrimination, or were duplicates of other opinions. Of the remaining 190 opinions, 60 were not brought by a class that specifically included blacks or other minorities. Finally, only 25 of the remaining opinions (19.2 percent) contained challenges to hiring practices alone, while 63 (48.4 percent) contained only challenges to *post*-hiring practices, and might thus be expected to have a negative effect on the employment of black workers. In sum, the opinions used by Leonard do not seem to be an accurate measure of the kind of litigation that might be expected to increase black employment.

B. Changing Regional Distribution of Litigation

Jerome Culp (1985) argues that the South was the traditional focus of early antidiscrimination litigation, but that by the early 1970s, employment discrimination litigation had spread to the rest of the country and, in so doing, changed its targets, methods, and results. Culp used a LEXIS search⁴⁷ to demonstrate that the re-

⁴⁷ Culp's (1985: 899) search term was "Title VII or 42 pre/5 2000!"

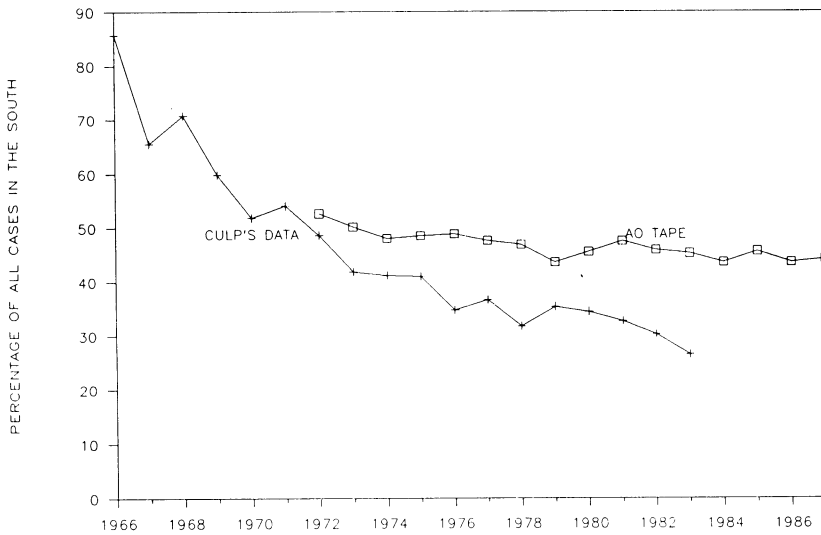


Figure 2. Two estimates of the share of employment discrimination cases filed in southern states, 1972–87, based on Title VII cases with published opinions and the AO Tape. “South” is defined to be the 11 states of the Confederacy plus Delaware, Kentucky, Maryland, Missouri, and Oklahoma. Source: Culp (1985) and AO tape.

gional distribution of cases brought under Title VII has changed over time. From Culp’s data, which are plotted in Figure 2, this story appears to be reasonable.

Suppose, however, that one takes a somewhat broader definition of employment discrimination and includes *all* cases filed, whether or not they generated a published opinion (and including non-Title VII cases as well). The top line in the figure (based on the data from the AO tape) reveals a somewhat different story. Many more of the employment discrimination cases that do not generate a written opinion are filed in the South; and their share of all employment discrimination cases has been dropping far more slowly than has their share of cases with written opinions. As demonstrated earlier, southern districts seem to have much lower publication rates than the rest of the country (for both employment discrimination and all civil cases). Thus, relying on published cases to measure the geographical distribution of *all* cases seems likely systematically to underrepresent certain regions of the country.⁴⁸

⁴⁸ Unless publication rates have been increasing faster outside the South, differences in publication rates by region do not explain Culp’s finding that the share of (published) Title VII cases in the South has been declining over time. Note, however, that age discrimination cases are more likely to generate published opinions than other employment discrimination cases and that most of the growth in age cases seems to have come in areas outside the South. It is thus possible that southern publication rates for employment discrimination cases have not been growing as fast as those in the rest of the country.

Moreover, the same concerns that we expressed above with respect to the

C. Content Analysis

In a series of articles, Paul Burstein has made extensive and imaginative use of a dataset consisting of federal appellate cases concerning discrimination. Burstein (1989: 662 n.2) acknowledges that “[d]isputes about allegedly discriminatory employment practices which reach the appellate courts are obviously not a random sample of such disputes, and little is known about how the many initial complaints about employment practices are winnowed to the few leading to appellate decisions.” (See also Burstein, 1988; Burstein and Monaghan, 1986.) He goes on to argue (1989: 662 n.2) that he is not attempting to generalize from appellate decisions to other kinds of cases. “The justification for analyzing appellate decisions is not that they are a random sample of disputes, however; rather it is that they constitute the universe of a type of dispute of special importance—an importance based on their consequences for [women and minorities] in the labor market.”

Nonetheless, in several places Burstein does claim that one can indeed learn about labor market discrimination and the disputes it generates from looking at appellate decisions.⁴⁹ Consider two of the conclusions in his 1989 article: (1) “women often act collectively against discrimination, frequently bringing class actions and getting organizations to act on their behalf” (*ibid.*, p. 660), and (2) “Women win more than half their cases, considerably more than prior work on confrontations between ‘haves’ and ‘have nots’ would lead us to expect” (*ibid.*, p. 660).⁵⁰

While our discussion has focused only on the differences between published and unpublished cases at the district court level, the same theoretical considerations would seem to apply a fortiori when looking at appellate decisions. Our theoretical analysis of sample selection bias suggests that both conclusions (1) and (2) are likely to be to some significant degree an artifact of the sample (appellate cases) that Burstein used. In other words, cases in which women act collectively and/or file class actions and/or get organizations to act on their behalf are much more likely to be litigated

Leonard study apply to Culp’s LEXIS search. Merely looking for “Title VII” can generate references to other statutes—Title VII of another act—or cases that are merely referring to Title VII but have nothing to do with discrimination. For example, a reference that the right to jury trial does not extend to Title VII litigants may be a case discussing the Seventh Amendment rather than employment discrimination.

⁴⁹ See, e.g., Burstein (1989: 660–61). He does note (1989: 661) that a “disadvantage . . . to focussing on E[qual]E[m]ploymentO[pp]ortunity cases” is that “they do not represent a random sample of discriminatory practices.”

⁵⁰ There is some ambiguity about the referent in both of these conclusions. For example, “often” might be interpreted to refer only to those cases that are appealed. But there is a dilemma here: either Burstein is referring only to appellate cases (in which case, as we show below, his findings are likely to be in large measure the result of sample selection bias), or he is extrapolating to *all* cases, in which case he appears to be overstating the case.

vigorously (and subsequently to be appealed) than are typical cases.

In the ABF's sample of 1,106 employment discrimination cases nationwide (including published and unpublished cases), 356 (32.2 percent) alleged sex discrimination.⁵¹ Of these, in only 58 (16.2) was there more than a single plaintiff suing; thus, it seems misleading to say that women "often" act collectively against discrimination.⁵² Although they shed some light on the way women have chosen to attack sex discrimination in the work place, Burstein's data are thus better understood instead as revealing something about which cases are selected for appeal.

A similar criticism can be leveled at Burstein's second conclusion. Overall, plaintiff win rates in district court employment discrimination cases (both published and unpublished) seem to be closer to 20 percent than to 50 percent (Siegelman, 1991; Eisenberg, 1989). Thus, Burstein's finding that plaintiffs appear to win about half of the appealed sex discrimination cases suggests not that women are unusually successful, but rather that something like the Priest/Klein selection effect is probably at work. Appellate cases are likely to be those in which the win rate is systematically closer to 50 percent than in all cases. Indeed, the win rate was nearly four times higher in our sample of *published* employment discrimination cases than in the unpublished cases in Table 7,⁵³ and it seems reasonable that the sample selection problems would be even stronger in appellate cases.

In a similar sort of analysis, Stidham, Carp and Rowland (1983) make no mention of the possibility of sample selection bias in their study of district court race and sex discrimination cases with written opinions. Their conclusion that 53 percent of the cases were "liberal" (i.e., favored women plaintiffs) (*ibid.*, p. 209) is subject to exactly the same criticism as Burstein's result discussed earlier. We would expect women or minorities to do better in cases with written opinions than in those without them, so this high win rate suggests more about how the sample of cases was selected

⁵¹ Note that this figure includes some cases in which the basis of discrimination was both sex and some other characteristic, such as age or race.

⁵² Plaintiffs in sex discrimination cases requested class certification for their suits in only 36 of the 320 cases (11.2 percent). In only 14 of these cases (4.3 percent of all sex cases) was this request granted. Clearly, most sex discrimination suits are individual rather than collective actions; conversely, it is relatively unusual for women to oppose discrimination collectively. We are not suggesting that all suits are equal in importance—one class action can have a hundred times the impact of a single-plaintiff suit. But the claim we wish to refute is not that women work collectively *in the important suits* (which will tend to be true almost by definition of "important"). The conclusion that women *often* act collectively, however, seems to be true only for the limited and unrepresentative sample on which it was based.

⁵³ The difference was only statistically significant at the 0.15 level, given the small sample.

than about the political or social forces at work in shaping litigation.

Finally, Mills (1981) seems to have been largely inattentive to the concerns that we have tried to articulate. She claims to have “all sex discrimination suits filed in U.S. District Court in the Second, Fourth, Fifth, Ninth and District of Columbia circuit under Title VII of the Civil Rights Act from its passage in 1964 through 1975” (*ibid.*, p. 199) apparently without realizing that she is really only working with those cases that generated published opinions in the BNA reporter.

D. The Effect of Performance Appraisal System Characteristics on Outcomes

Feild and Holley (1982) set out to examine the relationship between an employer’s personnel system (in particular, how the system evaluates job performance for determining promotion, pay, and retention) and the outcome of employment discrimination litigation against that employer. Their hypothesis is that unsound performance appraisal techniques (such as the absence of specific written instructions for evaluating performance) will increase the probability of a plaintiff victory in a discrimination suit. Using discriminant analysis, they conclude that five characteristics of performance appraisal system do have a significant influence on trial outcomes in their sample of sixty-six employment discrimination cases with written opinions.⁵⁴

The authors’ interesting analysis is somewhat tarnished by their complete failure to consider sample selection problems, however. In fact, as Hughes and Snyder (1989) demonstrate, the statistical assumptions underlying such an analysis using published cases will necessarily be violated if publication induces sample selection bias. One possible response might be to argue that estimates are unbiased *conditional on the sample selection criterion* (i.e., for the universe of published cases). However, Hughes and Snyder (*ibid.*, pp. 427–28) show clearly that

The phrase “the estimates are unbiased *conditional* on the stage of litigation” must be interpreted cautiously. The “conditional” estimates . . . are not conditional in the statistical sense of the term, since [they] place no restrictions on the value of [the error term]. . . . Estimates . . . are conditionally unbiased only to the extent that the set of claims settled or litigated [or published] is assumed to be *the population of interest in its own right*, complete with a nor-

⁵⁴ The factors they identified as influencing outcomes were (1) defendant’s organization type (industrial or nonindustrial); (2) whether written instructions were provided to the performance evaluator; (3) whether evaluation used “trait” or “behavioral” model in assessing performance; (4) whether employer used “job analysis” in the evaluation system; and (5) whether results were reviewed with employee (the first, or the negative, answer to each question was associated with lower employer win rates).

mally distributed error term, rather than the selected subset of a larger population. "Conditional" estimation requires assuming statistical independence between [publication and the win equation], an assumption that [is] . . . at odds with the self-selection features of the underlying theoretical model [such as Priest/Klein]. (Emphasis in original)

In other words, if publication is correlated with any of the variables of interest, the statistical properties of the error term necessary to estimate discriminant or logit analysis of outcomes are not met.

E. Doctrinal Explanations for Caseload Growth

As we and others have pointed out, the federal employment discrimination caseload has grown dramatically over the past twenty years, at a rate several times faster than the rest of the federal civil caseload (Donohue and Siegelman, 1991). Some authors have attributed this rapid increase to changes in legal doctrine that create new rights for plaintiffs. For example, the American Law Institute (1989: 1) suggests that "in recent years there has been an 'explosion' in the number of employment discrimination . . . cases filed in the federal courts. The tremendous growth in such cases has been due primarily to landmark decisions by the U.S. and state Supreme Courts, federal and state appellate and trial tribunals, and recently enacted statutes." Similarly, the Equal Employment Advisory Council (an employers' lobbying group) states (1990: 3) that "there has been a tremendous increase in federal employment discrimination litigation." In the Council's view, "this increased litigation has been brought about by two primary factors—additional statutory private rights of action created by Congress and the courts plus the 1972 Title VII amendments that gave litigation authority to the EEOC."⁵⁵

The gist of these legal or doctrinal explanations is that additional rights created by courts or by Congress are responsible for the increase in litigation. The more plaintiffs' rights are given protection under the law, the argument goes, the more plaintiffs will be in position to defend these rights through litigation, and hence the greater the volume of lawsuits.

There are many potential flaws in this argument; here we concentrate only on whether it is supported by the facts. Suppose that changes in legal doctrine *were* driving the increase in litigation we have observed, as the doctrinal theories suggest. This should have obvious implications for changes in the *composition* of cases. But these implications are almost universally falsified by the available data.

⁵⁵ As we demonstrate elsewhere (Donohue and Siegelman, 1991), the argument that the EEOC bears any direct responsibility for the growth in cases is almost certainly wrong.

Using our nationwide sample of employment discrimination cases described earlier, we can disaggregate the caseload by an (admittedly crude) measure of the “subject matter” or right being asserted.⁵⁶ Doctrinal theories require that areas of the law in which new rights have been created should be the source of most of the caseload increase, while areas that have not seen new rights created should experience no (or less) growth in litigation.

An obvious counter-example is provided by the disparate behavior of hiring and firing suits. The legal rules for proving discrimination in such cases are identical.⁵⁷ Yet, as we show elsewhere (Donohue and Siegelman, 1991), almost all the growth in cases filed has come from an increase in firing suits. This is clearly at odds with a doctrinal explanation.

Moreover, most of the areas in which new plaintiff rights have been created have, in fact, produced very little of the additional litigation that doctrinal theories would predict. For example, according to doctrinal theories, recognition of the disparate impact theory of discrimination in *Griggs v. Duke Power* (1971)—perhaps the most important innovation in employment discrimination doctrine since the passage of Title VII—should be a major source of the growth in litigation. In fact, however, the volume of disparate impact cases has grown more slowly than the overall employment discrimination caseload, and constitutes less than 5 percent of all employment discrimination cases filed.⁵⁸ Thus, *Griggs* could not possibly be responsible for much of the growth in litigation.⁵⁹

Our analysis of sample selection effects in publication provides a clue about why doctrinal explanations for the caseload growth

⁵⁶ As noted earlier, the sample was randomly selected from all the cases filed in a nonrandom group of seven cities that together encompass more than 25 percent of all employment discrimination cases filed. As we argue elsewhere (Donohue and Siegelman, 1991), however, there are good reasons to think that it contains enough internal variation to make it a relatively good proxy for the caseload as a whole.

⁵⁷ According to Schlei and Grossman (1979: 5), “numerous courts have continued to apply the order and allocation of proof set forth in *McDonnell Douglas* to discharge cases” (citing *McDonald v. Santa Fe Trail Transportation Co.*, 1976; *Flowers v. Crouch-Walker Corp.*, 1977). The rules defining discriminatory behavior are thus essentially identical in the two kinds of employment discrimination cases. Note that the Supreme Court has recently held that § 1981 protection extends only to contract formation (hiring) and not to on-the-job discrimination such as harassment or promotion (*Patterson v. McLean Credit Union*, 1989). Whether § 1981 covers firing has been left uncertain, but the implications of this decision for the present discussion are not important, since it occurred too late to have any effect on the volume of cases in the population under study.

⁵⁸ These data come from an analysis of our nationwide sample of 1,247 employment discrimination cases. In 1972–73, disparate impact cases made up 8.7 percent of all cases filed; in 1986–87, their share had fallen to only 3.7 percent.

⁵⁹ This is not to say that *Griggs* had no influence on employer behavior, which could well have changed substantially to prevent future lawsuits under a *Griggs*-type theory. The argument is only that *Griggs* is not responsible for much of the observed growth in litigation.

may miss the boat. Most legal academics work *only* with cases that have published opinions. These cases tend to be those that raise novel legal issues or fact situations. Thus, sample selection effects may make it appear that novel legal issues are the source of the growing caseload, when in fact the volume of *unpublished* cases (which constitute 80–90 percent of all cases filed) is growing for entirely different reasons.⁶⁰

VII. CONCLUSION: REPRESENTATIVE OF WHAT?

Both theory and empirical work suggest that employment discrimination cases with published opinions are unlike those without them in certain systematic and important ways. We suggest that the sample selection bias inherent in using published cases is generally likely to be important enough to warrant attention. In the context of employment discrimination litigation, the occupational distribution of plaintiffs, the kinds of discrimination being complained about, the laws allegedly being violated, and the outcome of litigation all differ significantly between published and unpublished cases.⁶¹

Social scientists and legal academics interested in evidentiary rules, precedent, judicial rhetoric and similar issues now have a large and important body of employment discrimination cases with published opinions on which to draw for evidence.⁶² We enthusiastically support such research. With care, legal cases with published opinions can also be used to generate some useful conclusions about the society from which they come. If we have any disagreement, it is with scholars who restrict their sample to published

⁶⁰ Eisenberg and Schwab (1989) develop a version of this argument in a different context.

⁶¹ Moreover, we suspect that the differences we found between published and unpublished cases would be even greater if we had used a stricter definition of “published” (inclusion in BNA, rather than in LEXIS, which publishes all opinions submitted to it). For example, consider the following slip opinion, reproduced in its entirety:

These actions came on to be heard on the motions of the defendants for summary judgment, and the court having considered the pleadings in the actions, and having found that there is no genuine issue of fact to be submitted to the trial court on Counts I and II, and having concluded that defendants are entitled to judgments as a matter of law, it is hereby ORDERED, ADJUDGED AND DECREED that defendants’ motions for summary judgment on counts I and II are in all respects, granted, and judgments are entered herein in the defendants’ favor and against the plaintiffs. And the court further having determined that these would be inappropriate cases for the exercise of our pendant jurisdiction, IT IS FURTHER ORDERED, ADJUDGED AND DECREED that judgments be entered herein in the defendants’ favor dismissing Count III.

Slip Opinion No. 77 C 3511, *Presley-Robinson v. Fredric* (1979). Clearly, this opinion would not be found in BNA.

⁶² For a superb example of research that uses social science to explain and critique the law, without overgeneralizing from its sample, see Schultz (1990).

cases and then attempt to generalize from their results to all cases, or to broader social phenomena, without considering the possibility of sample selection bias.

We are certainly not advocating the abandonment of research using published opinions, however. For one thing, the costs of using such data are obviously dramatically lower than attempting to survey *all* cases: In the federal district courts, this is likely to require extended visits to a Federal Record Center or courthouse to examine documents.⁶³ We do want to suggest, however, that researchers—including those doing qualitative work—need to make more of an effort to correct for the ways in which focusing on cases with published opinions yields unrepresentative results.⁶⁴

⁶³ The availability of the Administrative Office of the U.S. Courts' complete records (the Integrated Database of the Federal Judicial Center) through the Inter-University Consortium on Political and Social Research should at least make it easier for interested researchers to locate unpublished case files by subject matter. We take it as a promising sign that these data have already begun to be used.

⁶⁴ For researchers using quantitative methods, powerful econometric techniques are now available in many statistical packages to correct for sample selection bias. Such techniques are beyond the scope of this article, but two cogent and accessible applications in a legal context are Hughes and Snyder (1989) and Schwab and Eisenberg (1990). Both give clear explanations of methodology and interpretations of their results and cite other relevant work. An excellent modern introduction to the econometrics of sample selection is Greene (1990a). For a more critical assessment of sample selection correction methods, and a sense of some of the possible alternatives, see Stolzenberg and Relles (1990).

APPENDIX

We obtained a data tape from the Administrative Office of the U.S. Courts (AO) that included all “employment civil rights” cases filed between 1 July 1972 and 30 June 1987. The tape included all cases still pending (open) as of 30 June 1987, but did *not* include cases filed before 1 July 1972, even if they closed after this date.

In constructing the set of employment discrimination cases with published opinions, one alternative would be simply to search in LEXIS for each of the 4,310 cases on the AO tape that were filed in the Northern District of Illinois. Instead, however, we took advantage of some sophisticated UNIX operating system software that allows relatively easy comparisons of (matching between) files. (More precise details are available from the authors on request.) Our strategy, therefore, was to search the LEXIS database using as broad a definition as possible of what constitutes employment discrimination.⁶⁵ The resulting list of cases was then compared with the set of docket numbers on the AO tape, and all matching docket numbers were recorded.⁶⁶ Table A1 presents details of the precise search request used and the number of citations found.

The table shows that the search produced 3,817 citations during the period comparable to that covered by the AO Tape. The adjustment between lines 9 and 10 was necessary because our search term started picking up LEXIS cases in January 1972, while the AO tape started six months later. We also picked up some citations to cases filed after 31 March 1987. Line 10 corrects for this “slippage” at either end of the data by eliminating cases filed outside of our comparison period.

In addition, it is important to note that LEXIS uses the judicial *opinion* as its unit of analysis, not the case. The same case might produce more than one opinion, as the judge rules first on a motion to dismiss, then on a motion for summary judgment, then on evidentiary issues, and ultimately on the final outcome of the case. In fact, if one counts only unique docket numbers (thereby eliminating multiple counting of cases with more than one opinion), the number of citations falls from 3,817 to 3,058, a drop of 20 percent. Put another way, the average case in our LEXIS sample contained about 1.25 opinions. We therefore use the term “LEXIS cases” to refer to unique docket numbers, reserving “citations” for the

⁶⁵ An alternative strategy would be to frame a LEXIS search that would come closest to approximating the population of employment discrimination cases appearing on the AO tape. The advantages of this alternative strategy are that it is easier to do; it also more closely approximates the manner in which social scientists use LEXIS to find opinions. On the other hand, this strategy will not yield as accurate a listing of employment discrimination because of the obvious problems of underinclusiveness and overinclusiveness that occur whenever one is trying to formulate the search terms.

⁶⁶ As a check, we did undertake a LEXIS search for 179 of the cases on the AO tape. The results substantiate our methodology, as discussed in detail in note 40.

Table A1. Terms and Number of Additional Citations Found in LEXIS Search for Employment Discrimination Cases

Terms	Additional Citations Found
Lib. = GenFed; File = Courts;	
1. Court (Northern and Illinois) and Date Aft 1971 and Discriminat!	2,943
2. or Title VII	137
3. or Equal Pay Act	3
4. or Rehabilitation Act	13
5. or 42 U.S.C. § 1981	28
6. or 42 U.S.C. § 1983	1,404
7. or 42 U.S.C. § 206	1
8. or 42 U.S.C. § 2000(e)	0
9. or [42 w/3 (1981 or 1983 or 2000(e))]	187
Total citations (opinions) found	4,716
10. Of which were opinions in cases filed after 6/31/72 and before 4/1/87	3,817
11. Of which were <i>unique</i> cases	3,058

number of entries detected (including multiple opinions from the same underlying case). In this article, we concentrate on cases rather than citations.

REFERENCES

- ADMINISTRATIVE OFFICE OF THE UNITED STATES COURTS (AO), STATISTICAL ANALYSIS AND REPORTS DIVISION (June, 1989) "Instructions for Completing District Court Reporting Forms," ch. 5 in 11 "Statistics Manual." Washington, DC: unpublished.
- AMERICAN LAW INSTITUTE (ALI) (1989) *Employment Discrimination and Civil Rights Actions in Federal and State Courts* Philadelphia: ALI.
- BRENNER, S. (1989) "Of Publication and Precedent: An Inquiry into the Ethnomethodology of Case Reporting in the American Legal System." Unpublished.
- BUMILLER, Kristin (1988) *The Civil Rights Society*. Baltimore: Johns Hopkins University Press.
- BURSTEIN, Paul (1989) "Attacking Sex Discrimination in the Labor Market: A Study in Law and Politics," 67 *Social Forces* 641.
- (1988) "Labor Market Discrimination, Social Science, and the Law." Unpublished.
- BURSTEIN, Paul, and Kathleen MONAGHAN (1986) "Equal Employment Opportunity and the Mobilization of Law," 20 *Law & Society Review* 355.
- CARP, R. A., and C. K. ROWLAND, *Policymaking and Politics in the Federal District Courts* (1983).
- CARTWRIGHT, Bliss (1975) "Conclusion: Disputes and Reported Cases," 9 *Law & Society Review* 369.
- COOTER, Robert, and Daniel RUBINFELD (1989) "Economic Analysis of Legal Disputes and Their Resolution," 27 *Journal of Economic Literature* 1067.
- CULP, Jerome M. (1985) "A New Employment Policy for the 1980s: Learning from the Victories and Defeats of Twenty Years of Title VII," 37 *Rutgers Law Review* 895.
- CURRAN, Barbara A. (1977) *The Legal Needs of the Public*. Chicago: American Bar Foundation.
- DONOHUE, John J., and Peter SIEGELMAN (1991) "The Changing Nature of Employment Discrimination Litigation," Forthcoming, *Stanford Law Review*.

- EISENBERG, Theodore (1989) "Litigation Models and Trial Outcomes in Civil Rights and Prisoner Cases," 77 *Georgetown Law Journal* 1567.
- EISENBERG, Theodore and Stewart J. SCHWAB (1989) "What Shapes Perceptions of the Federal Court System?" 56 *University of Chicago Law Review* 501.
- EQUAL EMPLOYMENT ADVISORY COUNCIL (1990) Memorandum to William K. Slate, Director, Federal Courts Study Committee, RE: Tentative Recommendations of the Federal Courts Study Committee (December 22, 1989), Washington, DC.: Unpublished.
- FEILD, Hubert, and William HOLLEY (1982) "The Relationship of Performance Appraisal System Characteristics to Verdicts in Selected Employment Discrimination Cases," 25 *Academy of Management Journal* 392.
- FRANKLIN, Marc (1980) "Winners and Losers and Why: A Study of Defamation Litigation, 1980 *American Bar Foundation Research Journal* 455.
- (1981) "Suing the Media for Libel: A Litigation Study" 1981 *American Bar Foundation Research Journal* 797
- GOLD, Michael (1985) "Griggs' Folly: An Essay on the Theory, Problems, and Origin of the Adverse Impact Definition of Employment Discrimination and a Recommendation for Reform," 7 *Industrial Relations Law Journal* 429.
- GOOD, Irving John (1968) "Fallacies, Statistical," *International Encyclopedia of the Social Sciences*, Volume 5: 292. New York: Macmillan Company & Free Press.
- GREENE, William (1990a) *Econometric Analysis*. New York: Macmillan & Co.
- (1990b) *LIMDEP Manual*. New York: LIMDEP.
- HECKMAN, James (1979) "Sample Selection Bias as a Specification Error," 47 *Econometrica* 153.
- HOGG, Robert, and Elliot TANIS (1977) *Probability and Statistical Inference*. New York: Mcmillan & Co.
- HUGHES, James, and Edward SNYDER (1989) "Policy Analysis of Medical Malpractice Reforms: What Can We Learn from Claims Data?" 7 *Journal of Business and Economic Statistics* 423.
- KRITZER, Herbert (1986) "Adjudication to Settlement: Shading in the Gray," 70 *Judicature* 161.
- KUHN, Peter (1987) "Sex Discrimination in Labor Markets: The Role of Statistical Evidence," 77 *American Economic Review* 567.
- LEONARD, Jonathan (1984) "Antidiscrimination or Reverse Discrimination: The Impact of Changing Demographics, Title VII, and Affirmative Action on Productivity," 19 *Journal of Human Resources* 145.
- LLEWELLYN, Karl (1960) *The Common Law Tradition: Deciding Appeals*. Boston: Little, Brown.
- MILLER, Richard, and Austin SARAT (1980–81) "Grievances, Claims and Disputes: Assessing the Adversary Culture," 15 *Law & Society Review* 525.
- MILLS, Heidi (1981) "On the Use of Equal Employment Laws," 24 *Pacific Sociological Review* 196.
- OLSON, Susan (1989) "Studying Federal District Courts Through Published Cases." Unpublished.
- PRIEST, George, and Benjamin KLEIN (1984) "The Selection of Disputes for Litigation," 13 *Journal of Legal Studies* 1.
- POSNER, Richard (1985) *The Federal Courts: Crisis and Reform*. Cambridge, MA: Harvard University Press.
- REYNOLDS, W. L., and W. M. RICHMAN (1978) "The Non-precedential Precedent—Limited Publication and No-Citation Rules in the United States Court of Appeals," 78 *Colum. L. Rev.* 1167.
- SCHLEI, Barbara, and Paul Grossman (1979) *Employment Discrimination Law* (1979 Supplement). Washington, DC: Bureau of National Affairs.
- SCHWAB, Stewart J., and Theodore EISENBERG (1990) "The Influence of Judicial Background on Settling and Winning Cases and a Study of the Disputing Pyramid." Unpublished, Cornell University Law School.
- SCHULTZ, Vicki (1990) "Telling Stories About Women and Work: Judicial Interpretations of Sex Segregation in the Workplace in Title VII Cases Raising the 'Lack of Interest' Argument," 103 *Harvard Law Review* 1749.
- SIEGELMAN, Peter (1991) "An Economic Analysis of Employment Discrimination Litigation." Ph.D. Diss., Yale University.

- SONGER, Donald (1988) "Nonpublication in the U.S. District Courts: Official Criteria versus Inferences from Appellate Review," 50 *Journal of Politics* 206.
- STIDHAM, Ronald, R. A. CARP, and C. K. ROWLAND (1983) "Women's Rights Before the Federal District Courts, 1971-77," 11 *American Politics Quarterly* 205.
- STOLZENBERG, Ross, and Daniel RELLES (1990) "Theory Testing in a World of Constrained Research Design: The Significance of Heckman's Censored Sampling Bias Correction for Nonexperimental Research," 18 *Sociological Methods and Research* 595.

CASES CITED

- Flowers v. Crouch-Walker Corp.*, 552 F.2d 1277, (7th Cir. 1977).
- Griggs v. Duke Power*, 401 U.S. 424 (1971).
- McDonald v. Santa Fe Trail Transportation Co.*, 427 U.S. 273 (1976).
- McDonnell Douglas v. Green*, 411 U.S. 792 (1972).
- Patterson v. McLean Credit Union*, 109 S. Ct. 2363 (1989).
- Presley-Robinson v. Fredric* (July 27, 1979), 7th Circ. Slip Opinion No. 77 C 3511.
- Smith v. Board of School Commissioners of Mobile County*, 655 F. Supp. 939 (1987).

STATUTES CITED

- Civil Rights Act, 42 U.S.C. 2000 (1964).
- Equal Pay Act, 29 U.S.C. 201 (1963).
- Rehabilitation Act, 29 U.S.C. 791 et seq. (1973).
- Age Discrimination in Employment Act, 29 U.S.C. 621 et seq. (1967).