6-1-1990

Testing the Selection Effect: A New Theoretical Framework with Empirical Tests

Theodore Eisenberg

Cornell Law School, ted-eisenberg@lawschool.cornell.edu

Follow this and additional works at: <http://scholarship.law.cornell.edu/facpub>

Part of the [Applied Statistics Commons](#), [Civil Procedure Commons](#), and the [Litigation Commons](#)

Recommended Citation

Eisenberg, Theodore, "Testing the Selection Effect: A New Theoretical Framework with Empirical Tests" (1990). *Cornell Law Faculty Publications*. Paper 370.

<http://scholarship.law.cornell.edu/facpub/370>

This Article is brought to you for free and open access by the Faculty Scholarship at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Faculty Publications by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact jmp8@cornell.edu.

TESTING THE SELECTION EFFECT: A NEW THEORETICAL FRAMEWORK WITH EMPIRICAL TESTS

THEODORE EISENBERG*

RECENT law and economics scholarship has produced much theoretical and empirical work on how and why legal disputes are settled and litigated. One of the most significant developments in this literature, attributable to the work of William Baxter and the combined efforts of George Priest and Benjamin Klein,¹ has been the formation of a theory about both the selection of disputes for trial and the rates of success that plaintiffs enjoy for those cases that are resolved at trial. The basic theory contains two components. The selection effect refers to the proposition that the selection of tried cases is not a random sample of the mass of underlying cases. Rather, those cases that tend to be clear for either the plaintiff or the defendant under the applicable legal rules settle relatively quickly, leaving only the difficult cases for trial.

Since these tried cases are not representative of the larger class of disputed cases, it is risky to draw any inference from the outcome of tried disputes to the soundness of the legal rules by which these disputes are decided. The sample of tried cases may contain many victories for the plaintiff and many for the defendant. One cannot infer from that fact alone, however, the fairness or desirability of the underlying legal rules. If

* Professor of Law, Cornell Law School. I wish to thank Gregory Alexander, Kevin Clermont, Robert Hillman, David Kaye, and Martin T. Wells for their comments on a draft of this article. The data used in this study were received from the Administrative Office of the United States Courts under the auspices of National Science Foundation grant SES-8510284.

¹ William Baxter, *The Political Economy of Antitrust: Principal Paper* by William Baxter (Robert D. Tollison ed. 1980); George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. Legal Stud. 1 (1984). For references to more technical models complementing and building on these works, see Theodore Eisenberg, *Litigation Models and Trial Outcomes in Civil Rights and Prisoner Cases*, 77 Geo. L. J. 1567, 1570 n.19 (1989).

[*Journal of Legal Studies*, vol. XIX (June 1990)]

© 1990 by The University of Chicago. All rights reserved. 0047-2530/90/1902-0004\$01.50

the rules are heavily weighted for the plaintiff, then the closeness of the tried cases is consistent with there being many cases in which plaintiffs recover handsomely without litigation. Similarly, if the rules are weighted heavily for the defendant, then the total mass of cases brought will, by the time of trial, be reduced, as defendants will be able to settle many cases on favorable terms. The observation of hotly contested trials is therefore wholly consistent with the underlying rules that are skewed toward the plaintiff, toward the defendant, or toward neither side. The power of the selection effect is generally recognized in the academic literature. It has been the basis of complex litigation models² and has been subjected to empirical testing and debate.³

Closely akin to but clearly distinguishable from, the selection effect, is the so-called 50 percent hypothesis. This hypothesis is a more specific prediction than the selection effect. The 50 percent hypothesis posits that the set of tried cases culled from the mass of underlying disputes will result in 50 percent victories for plaintiffs and 50 percent victories for defendants.⁴ As generally developed, this hypothesis is subject to two important limitations. First, it is inapplicable where the issue in litigation is the extent of damages instead of the existence of liability. Second, it

² Authorities cited in Eisenberg, *supra* note 1, at 1570 n.19; John J. Donohue, Law and Economics: The Road Not Taken, 22 Law & Soc'y Rev. 903, 911 (1988) criticizes law and society scholars for not accounting for the Priest-Klein theory; Harold W. Elder, Trials and Settlements in the Criminal Courts: An Empirical Analysis of Dispositions and Sentencing, 18 J. Legal Stud. 191 (1989); David Trubek, Joel Grossman, William Felstiner, Herbert Kritzer, & Austin Sarat, Civil Litigation Research Project: Final Report (1983).

³ Elder, *supra* note 2; Donald Wittman, Is the Selection of Cases for Trial Biased? 14 J. Legal Stud. 185 (1985); George L. Priest, Reexamining the Selection Hypothesis: Learning from Wittman's Mistakes, 14 J. Legal Stud. 215 (1985). Studies of federal civil rights and prisoner cases confirm that the picture one sees of litigation varies substantially depending on whether one examines trials, filings, or appeals. Theodore Eisenberg & Stewart J. Schwab, What Shapes Perceptions of the Federal Court System? 56 U. Chi. L. Rev. 501 (1989); Eisenberg, *supra* note 1; Theodore Eisenberg & Stewart J. Schwab, The Reality of Constitutional Tort Litigation, 72 Cornell L. Rev. 641 (1987); Theodore Eisenberg, Section 1983: Doctrinal Foundations and an Empirical Study, 67 Cornell L. Rev. 482 (1982); Stewart J. Schwab & Theodore Eisenberg, Explaining Constitutional Tort Litigation: The Influence of the Attorney Fees Statute and the Government as Defendant, 73 Cornell L. Rev. 719 (1988).

⁴ The emphasis on tried cases is partly a consequence of available data. More systematic data are available on tried cases than on the many more disputes not tried, and the leading selection effect theoreticians' most substantial data set comes from jury verdict reporters, which are, of course, limited to tried cases (Priest & Klein, *supra* note 1). The emphasis on tried cases in 50 percent hypothesis literature, however, is not merely one of social scientific convenience. Trial is the end of the trial court level legal process. If ever the parties are going to settle based on mutually agreed predictions of case outcomes, it must be before or at trial. It is just before trial that the parties know the most about the case without actually knowing the outcome and, therefore, should be in the best position to arrive at common expectations about the outcome.

applies only where the stakes for the two sides are the same, a condition that is not satisfied where one party (such as an institutional defendant) worries about the effect that the outcome of the case will have on other litigation, whereas the other party (such as the ordinary tort plaintiff) does not.

The 50 percent hypothesis can conveniently be regarded as the limiting case of the selection effect theory. It follows, therefore, that any empirical corroboration of it generates powerful support for the more general selection hypothesis. But while the two theories are linked inextricably in the literature, the converse proposition does not hold: it is possible to reject the 50 percent hypothesis for all cases and still maintain the selection hypothesis. Given the theoretical arguments made on behalf of the stronger 50 percent hypothesis, many scholars have taken it seriously enough to seek empirical verification and to count in favor of the selection effect groups of cases in which trial success rates are near 50 percent.⁵ Baxter sketched reasoning leading to a prediction that plaintiffs should win approximately 50 percent of tried cases.⁶ He characterizes his anti-trust data, which departs sharply from 50 percent, as not supporting his hypothesis.⁷ Priest and Klein subjected the hypothesis to broader empirical testing through, *inter alia*, the study of tort trials in Cook County, Illinois.⁸ They counted trial win rates near 50 percent as supporting the selection effect and explained departures from 50 percent win rates by factors such as differential stakes to the parties.⁹ Wittman challenged Priest and Klein's tests with California rear-end automobile accident case data that departed from 50 percent success.¹⁰ Priest responded relying in part on Cook County rear-end collision cases.¹¹

This article suggests the incompleteness of existing methods of statistically testing the 50 percent hypothesis and reformulates the criteria for accepting or rejecting the hypothesis. It tests the hypothesis and concludes that it can be rejected as a description of all civil litigation. The

⁵ Priest & Klein, *supra* note 1. I term a case successful if the Administrative Office records a judgment for the plaintiff. A case is unsuccessful if the Administrative Office records a judgment for the defendant. There is, of course, room for debate about when a case is successful, even if the plaintiff obtains a settlement or a jury verdict. For discussions of defining success in litigation, see Eisenberg & Schwab, *What Shapes Perceptions?* *supra* note 3, at 516–17, 523–24, 532 n.87; Priest & Klein, *supra* note 1, at 29; Schwab & Eisenberg, *supra* note 3, at 726–28.

⁶ Baxter, *supra* note 1, at 13–16.

⁷ *Id.* at 16.

⁸ Priest & Klein, *supra* note 1.

⁹ *Id.*

¹⁰ Wittman, *supra* note 3.

¹¹ Priest, *supra* note 3.

hypothesis fares better as a description of tort litigation and, with downward adjustment of the predicted success rate, may accurately describe products liability cases.

More important, the methods presented for assessing the 50 percent hypothesis can be applied to other hypotheses about litigation outcomes. The 50 percent hypothesis may be rejected while the basic selection effect is retained.¹² Tried cases clearly are a skewed subset of filed disputes.¹³ They happen not to show a 50 percent trial success rate, but each category or subcategory of case displays some mean success rate at trial. For example, medical malpractice cases in federal court have an overall trial success rate of .38.¹⁴ Tests of whether the distribution of observed malpractice success rates across districts comport with the selection effect theory can employ or build on methods similar to those presented here.

I. THEORETICAL REFINEMENT OF TESTS OF THE 50 PERCENT HYPOTHESIS

Refining tests of the 50 percent hypothesis requires isolating the causal mechanism leading to the prediction. There are at least two separate reasons why one might advance this hypothesis. The first reason is that the legal system regards it as desirable for 50 percent of the disputes that reach trial to be decided for the plaintiff. Commitment to this principle of symmetry would create nonsensical social effects, providing plaintiffs with unsound incentives for bringing suit in the hope that they will fall in the half of tried cases that win.

The 50 percent hypothesis here does not derive, however, from any such quixotic view of the ends of the legal system. Quite the opposite, a more stochastic approach drives the theory. By the time the case is well into the litigation process and reaches trial, cases in which parties agree about the outcome will settle. The parties' different expectations about prevailing in light of the decision standard generates trials. In a simple model, "[i]ntuitively, the distribution of plaintiffs' errors to overestimate success is the same as the distribution of defendants' errors to underestimate plaintiffs' success (that is, approximately half of those disputes litigated will be attributable to plaintiffs' overestimates and half attributable to defendants' underestimates)."¹⁵ Similarly distributed error terms lead, in the limiting case, to 50 percent victories for plaintiffs.

¹² For example, Elder's evidence supports the selection effect theory without addressing the 50 percent hypothesis. See Elder, *supra* note 2.

¹³ Eisenberg, *supra* note 1, at 1584.

¹⁴ See the Appendix.

¹⁵ Priest & Klein, *supra* note 1, at 19 n.42.

Thus, for purposes of testing the 50 percent hypothesis, a trial can only have one of two outcomes, which Baxter, Priest, and Klein treat in the limiting case as having equally likely chances of occurring. The outcome of tried cases under Baxter-Priest-Klein selection theory is a binomial random variable with a probability of plaintiff success equal to one-half, or $P = 0.5$. In other words, one can model the outcome of litigated cases before a particular court by analogy to flips of an unbiased coin. There are only two possible outcomes, and theory suggests a tendency toward 50 percent over many cases or flips. In no specific series of coin flips or cases would one predict exactly 50 percent success.¹⁶

Pushing the analogy one step further, one can treat each court's success rate as the result of flips of a different coin (or as a separable set of flips of the same coin). If one were testing a group of 100 coins for bias or for a fair tendency toward 50 percent heads and flipped each coin 100 times, a test of the group's tendency toward 50 percent would compare the distribution of the 100 observed success rates with the predicted distribution of success rates. Any particular success rate (corresponding to one set of coin flips) would be of relatively little interest in testing the bias of the set of coins. It is predictable that some of the coins would show substantial departures from 50 percent heads. Similarly, the single overall success rate for all coins would not be of prime interest. That all coins taken together yield a 50 percent success rate does not confirm that the coins are unbiased. The overall 50 percent result could be the consequence of many coins being biased in the opposite direction from one another.

The coin flip analogy has important implications for empirical tests of the 50 percent hypothesis. It counsels analyzing trial outcome data by emphasizing the distribution of plaintiff trial victory rates. The distribution of plaintiff success rates, rather than any particular plaintiff success rate in a given year or court, tests whether a binomial random selection process is a useful analogy for the outcome of litigated cases. For example, Priest and Klein note where a year of trials yields a success rate that is statistically significantly different from the hypothesis. If the coin flip analogy is an apt one, such departures from the 0.5 mean do not necessarily count as evidence against the hypothesis. Some such departures are expected even when an underlying 50 percent process is at work. Put another way, Priest and Klein might have viewed the theory as fully vindicated if more of the sets of trial data had yielded 0.5 probabilities of success for plaintiffs. Yet, had all the data points shown success rates very close to $P = 0.5$, it would, ironically, be firm evidence against

¹⁶ Priest and Klein rely on a coin flip analogy to help explain departures from 50 percent success. *Id.* at 39 n.76.

the view that the outcome of trials behaves like a binomial random variable with $P = 0.5$. The uniform result would be better evidence of the "over-the-shoulder" model, in which someone tells juries systematically to strive for 50 percent success for plaintiffs.

II. REANALYZING THE PRIEST-KLEIN DATA

Consider Priest and Klein's data in light of the suggested emphasis on the distribution of plaintiff success rates. Their original data consist largely of jury verdicts in Cook County, Illinois. Two major tests they supply are by (1) court and year and (2) category of case. This Section analyzes their data using the methods sketched above. The reevaluation suggests that the Priest-Klein data provide, at best, modest support for the 50 percent hypothesis.

A. *Test across Courts*

Initially, they subdivide the data by year and court of disposition. Table 1, which is based on Priest and Klein's table 5, presents their data on plaintiff trial success rates in Cook County over time and across courts.

Based on these data, Priest and Klein conclude, "The . . . figures generally support the 50 percent implication. For all three courts, the composite rate of recovery is close to 50 percent and is extremely close for the Circuit Courts (column [4], 49.54 percent). Indeed, in the Circuit Courts, the proportion of plaintiff victories is close to 50 percent in most individual calendar years."¹⁷ They explore in detail possible explanations of trial results that depart from the 50 percent hypothesis.¹⁸

When one has a limited set of data, as did Baxter, there are few choices about how to analyze it. Checking the data's conformity to a predicted 50 percent success rate may be all he can do. When one has a few more data points, checking each unit of data's success rate against 50 percent is also a logical thing to do. Thus, Priest and Klein understandably measured their data against this standard.

But given the number of data points they had, an alternative method of analysis is available. Priest and Klein's analysis emphasizes plaintiff success rates in individual years or courts as a test of the 50 percent hypothesis. Where success rates over time or across courts are available, testing each year for correspondence with a predicted 50 percent success rate may not be the preferred way to test whether the data in fact support the hypothesis. Instead, as suggested above, under a plausible interpreta-

¹⁷ *Id.* at 33.

¹⁸ *Id.* at 24-29, 37-44.

TABLE 1
PROPORTION OF PLAINTIFF VICTORIES BY COURT OF DISPOSITION,
COOK COUNTY, ILLINOIS, 1959-79

YEAR	MUNICIPAL COURT		CIRCUIT COURT		U.S. DISTRICT COURT	
	% Victories	Cases	% Victories	Cases	% Victories	Cases
1959	66.1	56	48.3	147	38.9	36
1960	42.8	139	58.8	262	37.5	64
1961	44.4	223	44.7	263	57.4	27
1962	42.6	238	47.4	272	48.2	55
1963	52.7	244	56.1	651	57.7	52
1964	57.5	212	51.5	555	43.5	46
1965	66.1	165	48.4	553	55.9	51
1966	53.8	266	50.5	648	48.3	59
1967	45.5	309	45.6	613	60.3	29
1968	43.8	380	48.2	517	73.9	23
1969	39.2	260	50.5	479	68.8	40
1970	44.3	230	46.7	470	75.9	29
1971	46.9	322	47.7	514	40.0	30
1972	32.4	227	47.3	503	66.3	49
1973	39.6	255	51.5	477	53.8	39
1974	45.9	258	49.5	385	67.5	20
1975	47.6	269	56.2	409	48.6	37
1976	41.0	166	50.9	427	28.6	21
1977	34.0	209	38.3	341	63.6	22
1978	39.3	177	46.9	398	51.5	68
1979	54.5	99	53.0	264	50.0	22
Total	45.60	4,704	49.54	9,148	52.99	819

SOURCE.—George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. Legal Stud. 1, table 5 (1984).

tion of the hypothesis, the distribution of plaintiff success rates across measuring units (in this case, time and courts) better tests the hypothesis.

The data over time in Table 1 are based on each year's trials in each of three courts. Treat the outcome of each year's set of trials in each court as determined by N flips of an unbiased coin, where N is the number of trials in the court. The 50 percent hypothesis would not predict a plaintiff's probability of success as 0.5 for each year in each court. Rather, it would forecast a distribution of success rates with mean equal to 0.5 and a calculable standard deviation. To test the hypothesis, one should compare the predicted distribution of plaintiff success rates with the observed distribution.

For example, assume that coin *A* is assigned to the year 1959 in Table 1 for the municipal court. The 50 percent hypothesis could be modeled by flipping the coin 56 times (the number of trials in that court in that year)

and recording its P value (plaintiff success rate).¹⁹ Coin B could be assigned to the year 1960 and the municipal court and flipped 139 times. Coins C , and so on, could be assigned to each successive year and court and flipped a number of times corresponding to the number of trials conducted in each year by each court. There would be a total of 63 coins flipped, each coin corresponding to a given year and court in the Priest-Klein table, and each coin yielding its own success rate after being flipped a number of times equal to the number of trials conducted in the court during the year. A test of the 50 percent hypothesis is the extent to which the distribution of the 63 plaintiff success rates across years and courts corresponds to the distribution of success rates one would predict from the 63 series of hypothesized coin flips.

The distribution of P values over time predicted by the coin flip model is derivable from the number of cases (flips) each year for each court. Using a predicted P of 0.5, one can calculate the predicted standard deviation for each year and court. One can then combine these results into a single predicted mean and standard deviation for the series of cases.²⁰ Using the numbers of cases in the Priest-Klein sample, the coin flip model predicts a normal frequency distribution of P 's over years and across courts with mean 0.5 and standard deviation .079. These numbers can be used to plot the predicted frequency of each range of plaintiff success rates (P values) and to compare the predicted frequency with the observed frequency.

The predicted frequency distribution of success rates is plotted as the line in Figure 1.²¹ Each point on the line represents the number of the 63 series of sets of trials (the vertical axis) that one would expect to fall in the probability range centered on the corresponding point of the x -axis. It predicts, for example, that eight success rates in Table 1 will fall in the probability range .4875–.5125, the range centered on $P = 0.5$; similarly, five will fall in the range centered on $P = 0.55$. One should test the 50 percent hypothesis against the predicted frequency distribution of plaintiff success rates. If the distribution of success rates over the years conforms to the predicted frequency distribution, there is evidence that a random binomial selection process with $P = 0.5$ is at work. An observed P for a particular year or court less or more than 0.5 does not necessarily count as evidence against the hypothesis. Seeking to explain away such

¹⁹ The number of trials used in this article differs from the number used in Priest and Klein's analysis because the observed number of trials differs from the number reported in the Priest and Klein tables. I have adjusted the number of trials downward based on their report of the size and the actually observed sample. See *id.* at 31 n.61, 38 table 7 notes.

²⁰ Jan Kmenta, *Elements of Econometrics* 104–5 (2d ed. 1986).

²¹ The normal distribution of expected success frequencies was constructed using the method suggested in George W. Snedecor & William G. Cochran, *Statistical Methods* 48–49 (8th ed. 1989).

FIGURE 1.—Probability of success at trial: expected vs. observed frequency, Priest-Klein data (George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 *J. Legal Stud.* 1 (1984)).

departures from 50 percent is a hollow endeavor when the departures are consistent with a distribution generated by a 50 percent prediction. The predicted distribution allows for many years and courts with success rates substantially different from 0.5

The vertical bars in Figure 1 plot the observed distribution of success rates from Table 1. They show, for example, that five of the 63 success rates in the table fall in the probability-of-success interval centered on 0.5 (ranging from .4875 to .5125). Visual inspection of the predicted line versus the observed line suggests no clear insight into the validity of the hypothesis. Part of the observed distribution of probabilities seems reasonably close to the predicted distribution of probabilities. Some departures from the ideal predicted distribution are to be expected. Can one make a statistically meaningful statement about the relationship between the observed and predicted distributions? One standard statistical test suggests that it is not possible to reject the null hypothesis that the Priest-Klein data distribution of the probabilities would result from a predicted distribution with the hypothesized mean and standard deviation.²²

²² The Kolmogorov-Smirnov statistic tests the goodness of fit between two distributions. See W. J. Conover, *Practical Nonparametric Statistics* 344-56 (2d ed. 1980). It is a useful test to see if a random sample agrees with some specified distribution. *Id.* at 357. For the Priest & Klein data the Kolmogorov-Smirnov *Z* is .873, corresponding to a two-tailed probability of .431.

TABLE 2
PROPORTION OF PLAINTIFF VICTORIES IN CONTESTED CIVIL CASES TRIED TO JURIES,
BY CASE TYPE, COOK COUNTY, ILLINOIS, 1959-79

CASE TYPE	PLAINTIFF VICTORIES	
	%	Total Cases
1. Traffic	47.4	9,987
2. Common carrier	52.3	879
3. Injury on property	47.8	1,396
4. Street hazard	55.6	399
5. Dramshop	53.8	371
6. Worker injury	66.3	775
7. Product liability	42.8	477
8. Malpractice	39.6	202
9. Assault, dignitary harm, business tort	53.9	1,013
Total traffic	47.79	10,866
Total nontraffic	52.50	4,633
Total all cases		15,499

SOURCE.—George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. Legal Stud. 1, table 7 (1984).

One need not try to explain each individual *P* value that is distant from 0.5. Some values reasonably far from 0.5 are part of the 50 percent theory's prediction. Rather, it might be more profitable to try to explain the peaks or valleys in the observed frequency distribution that differ most substantially from the predicted frequency distribution, that is, the peak slightly to the left of 0.5 and the peaks at the right tail of the distribution.²³

B. Test across Subject Matter Categories

Priest and Klein's second major test of their hypothesis studies cases by subject matter category. They subdivide the cases tried by juries into nine categories and apply the 50 percent hypothesis to each category. Table 2 (their table 7) presents the results of this analysis.

Priest and Klein characterize their results as follows:

²³ Focusing on the distribution of success rates yields two other insights into the Cook County data. First, there is a significant difference in success rates between cases tried in municipal court and cases tried in the federal district court, though the large annual deviations in the district court data counsel caution before making too much of this. Plaintiffs do significantly better in federal court. (I am grateful to Professor C. G. G. Aitken of the University of Edinburgh for pointing out this difference.) Second, there is a near-significant downward trend over time in the municipal court. Plaintiffs were doing much better in the earlier years than in more recent years. Note, however, that movement over time in tried case success rates does not necessarily track movement over time in all disputes. See Theodore Eisenberg & James A. Henderson, Jr., *Is the Quiet Revolution in Products Liability Reflected in Trial Outcomes?* Cornell L. F. (in press).

Some of the results of [the table] suggest equal stakes leading to 50 percent plaintiff victories, others clearly do not. The proportion of plaintiff victories in three of the nine categories (common carriers, property injuries, and dramshop cases) cannot be distinguished from 50 percent. In a fourth category (street hazard), the difference from 50 percent is ambiguous: only significantly different when the total is inflated. In a fifth category (traffic collisions), the rate of recovery is very close to 50 percent (47.8 percent), although it is significantly different because of the large number of cases. Furthermore, the proportion of plaintiff victories in the residual category (assault, dignitary harms, and business torts) is also close to 50 percent (54.1 percent), though little can be concluded because of the peculiar aggregation of cases.

In three other categories, however—worker injury, product liability, and malpractice (rows 6, 7, and 8)—the proportion of plaintiff victories is significantly different from 50 percent.²⁴

They then provide explanations for the case categories that depart from 50 percent.

If a binomial random variable models the outcome of the cases selected for trial, one can again test the distribution of plaintiff success rates for the Priest-Klein case category data against a theoretical predicted distribution. Here, given the small number of data points available to Priest and Klein, not much is lost by examining each case category rather than the distribution of probabilities across categories. Distributional tests of the 50 percent hypothesis by case category are inconclusive. I repeat their case category analysis below with a much larger data set.

III. TESTING THE 50 PERCENT HYPOTHESIS USING FEDERAL DATA

Existing tests of the 50 percent hypothesis are inconclusive. The data used over time and across courts and case categories do not allow for rejection of the null hypothesis that observed trial rates could come from a population of cases in which plaintiffs prevail 50 percent of the time—but they hardly mandate accepting it. Even if existing studies were more conclusive, subject matter limitations would vex their results. Priest criticized Wittman's challenge because of special factors attending rear-end collision cases;²⁵ Baxter formulated his theory in a field, antitrust, with relatively low success rates at trial;²⁶ and Priest and Klein's original data are limited to tort cases tried before juries in one geographical area—yet, there are wide variations in trial outcome across regions in the United States.²⁷

²⁴ Priest & Klein, *supra* note 1, at 39.

²⁵ *Id.*

²⁶ Baxter, *supra* note 1, at 17, 23 tables 1-1, 1-2; Steven C. Salop & Lawrence J. White, Private Antitrust Litigation: An Introduction and Framework, in *Private Antitrust Litigation* 3, 11 (Lawrence J. White ed. 1988); see the Appendix.

²⁷ Eisenberg, *supra* note 1, at 1584-94.

To explore the 50 percent hypothesis across all civil litigation, it is helpful to collect trial success rates for many different subject areas. A success rate near 50 percent in a single subject area or region may not be representative. After discussing the large national data set used to test the 50 percent hypothesis, I reject the hypothesis that plaintiffs prevail in 50 percent of tried disputes.

A. *The Administrative Office Data*

Data collected by the Administrative Office of the United States Courts permit the study of trial success rates in many subject areas and across the entire country. Plaintiffs commencing federal civil actions file a form that the Administrative Office uses to classify their cases.²⁸ The form contains several subcategories of contract, tort, property, social security, civil rights, prisoner, and other cases.²⁹ The Appendix lists the categories and subcategories sizable enough for inclusion in the analysis. When a federal case closes, the district court clerk's office files another form containing the procedural progress and subject matter category of the case and whether a judgment was recorded for plaintiff or defendant. Using the termination data, one can determine the percentage of plaintiff judgments at trial in each subject matter category. The data used here consist of all reported trial outcomes in all nonbankruptcy civil cases concluded in all federal district courts for a period beginning in 1978 and ending in 1985, a total of 64,965 trials.³⁰ Discussion of the strengths and weaknesses of the federal data may be found elsewhere.³¹

²⁸ Plaintiffs must file Form JS-44 with every civil action commenced in federal district court. Admin. Off. of the U.S. Courts, Vol. 11, in *Guide to Judiciary Policies and Procedures*, Transmittal 64 (March 1, 1985), at II-6 (hereafter A. O. Guide). A complete description of Administrative Office data appears at Inter-university Consortium for Political and Social Research (ICPSR), *Federal Court Cases: Integrated Data Base, 1970-1987*, ICPSR 8429 Federal Judicial Center, Principal Investigator (Winter 1989).

²⁹ See A. O. Guide, *supra* note 28 (exhibit J).

³⁰ Of course, the Administrative Office data have limitations (see, generally, ICPSR, *supra* note 28). Prior work suggests that, even if Administrative Office data might not be used to determine the absolute levels of some phenomena, they are useful in making comparisons (see Eisenberg & Schwab, *The Reality of Constitutional Tort Litigation*, *supra* note 3, at 686-87). There is little reason to believe that the Administrative Office data on case terminations are biased in favor of stating defendant or plaintiff success rates. The cases used to establish the plaintiff success rates consist of all cases for which the Administrative Office reports a judgment for plaintiff or defendant. I exclude tried cases that did not lead to clear judgments for plaintiff or defendant. The data here also differ from data in the bulk of litigated cases because most cases are not litigated in federal court. Federal litigation probably is characterized by greater rates of government participation than other litigation and by emphasis on different subject areas.

³¹ For example, James A. Henderson, Jr., & Theodore Eisenberg, *The Quiet Revolution in Products Liability: An Empirical Study of Legal Change*, 37 *UCLA L. Rev.* 479 (1990).

FIGURE 2.—Probability of success at trial: expected vs. observed frequency of districts, all federal civil cases. Source: Administrative Office data, 1978–85.

B. Test across Districts

To test the 50 percent hypothesis, I first divided the litigated disputes by federal districts, a subdivision analogous to Priest and Klein's subdivision of the Cook County data by court. This approach provided 92 data points, each with many trials. I then plotted the frequency of the plaintiff success rates against the expected distribution of success rates if a random binomial process (with $P = 0.5$) were at work. The difference between the predicted and observed frequency distributions ought to lead one to accept or reject the hypothesis. Note that this technique could bolster the Priest and Klein claim by relieving them of explaining the inevitable departures from a 50 percent success rate. Figure 2 presents the results.

Both visual inspection and traditional statistical tests for normality allow for rejection of the hypothesis that tried cases lead to plaintiff victory half the time.³² There is no realistic chance that the observed distribution of success rates could result from the distribution of success rates predicted by the hypothesis.

³² The predicted normal distribution is one with mean = 0.5 and standard deviation = .028. The Kolmogorov-Smirnov Z is 3.095, corresponding to a two-tailed probability of .000.

C. Test across Subject Matter Categories

Priest and Klein disaggregate their Cook County data by case category, a disaggregation that, as noted above, leads to inconclusive tests of the 50 percent hypothesis. One can similarly disaggregate the Administrative Office termination data by case category. The Administrative Office data include dozens of case categories, providing a comprehensive test of the hypothesis. The trial success rates by category are presented in the Appendix. One need neither graphically nor formally statistically assess the null hypothesis that the observed distribution of success rates could result from random fluctuation from the distribution of success rates predicted by the 50 percent hypothesis. The data do not even come close to supporting the hypothesis. The vast majority of case categories have success rates far from 50 percent. Only 16 of 72 subject matter areas show success rates in the range of 45 to 55 percent. The 50 percent hypothesis can again be rejected.

D. Testing Tort Litigation

As a third test of the hypothesis, consider only tort cases, with success rates again divided by federal district. Tort is the subject area from which Priest and Klein draw the most empirical support, and even if the hypothesis proved valid only in tort cases, it would be an impressive result. Tort trials constituted about 30 percent of all cases tried in federal court during the period studied.³³

Figure 3 shows the expected and observed frequency of success rates by federal district for cases classified as tort cases by the Administrative Office. The theoretically predicted normal distribution clusters more tightly around 50 percent than does the observed distribution of success rates.

One can again reject the hypothesis that the observed frequency distribution could result from a normal distribution with 0.5 mean and variance as predicted by the 50 percent hypothesis.³⁴ But there is a reasonable-looking fit between the predicted and observed frequency distributions. Figure 3, when contrasted with the earlier results, suggests that tort may be an area of law in which large numbers of trials are more amenable to being modeled by something like a binomial random variable than are other areas of law.³⁵ Relatively unsuccessful tort categories, such as med-

³³ See the Appendix.

³⁴ The predicted normal distribution is one with mean = 0.5 and standard deviation = .049. The Kolmogorov-Smirnov Z is 1.958, corresponding to a two-tailed probability of .001.

³⁵ I have also analyzed a few subsets of tort litigation to test for conformity with the 50 percent hypothesis. These subsets include diversity cases and cases not involving the federal government. Neither yields results closer to 50 percent than the more general tort category.

FIGURE 3.—Probability of success at trial: expected vs. observed frequency of districts, all federal tort cases. Source: Administrative Office data, 1978–85.

ical malpractice and products liability, might explain the bulge to the left of 50 percent.³⁶ Excluding these categories from the analysis could improve the fit somewhat, though even pure personal injury cases significantly depart from 50 percent success.³⁷

IV. EXTENDING DISTRIBUTIONAL ANALYSIS BEYOND THE 50 PERCENT HYPOTHESIS

I hesitate to press for explanations of why the tort success rate distribution departs from 50 percent. That all of tort law does not show a 50 percent trial success rate is not surprising. Tort law consists of several different subareas with widely varying success rates. Each subarea has its own characteristics, including possibly different stakes to the parties, that shape the expected distribution of trial success rates. For example, products liability and malpractice plaintiffs prevail in under 40 percent of tried cases.³⁸ Plaintiffs in motor vehicle cases prevail in 60 percent of tried

³⁶ These categories may have low success rates because of differential stakes to the parties. See Priest & Klein, *supra* note 1, at 40.

³⁷ The "other personal injury" category shows a plaintiff success rate of .46. See the Appendix.

³⁸ See the Appendix. Products liability plaintiffs, even accounting for their low trial success rate, obtain a premium recovery relative to other tort plaintiffs. See Henderson & Eisenberg, *supra* note 31, at 530 n.197.

cases.³⁹ Both the Priest-Klein data and the federal data used here show wide variations in success across tort categories. That Priest and Klein detected an overall set of tort success rates near 50 percent need not be viewed as solid evidence that 50 percent success describes many subcategories of tort cases. Rather, it could be equally well viewed as evidence that the commingling of many categories of tort cases, some with success rates well above 50 percent and some with success rates well below 50 percent, yields an aggregate success figure near 50 percent. The extremes may just cancel each other out.

To test whether a class of cases satisfies a predicted success rate, prospects are brighter and analysis will be clearer by focusing on homogeneous subcategories, for that approach reduces the possibility of a wide variation in success rates canceling out to yield a misleading aggregate observed success rate. Even in the narrower subcategories, however, available evidence is that 50 percent success rarely describes trial outcomes.⁴⁰ In some of these cases, the deviations from 50 percent could be explained by the asymmetrical stakes in litigation. Defendants are repeat players in product liability cases, where plaintiffs prevail in only 25 percent of tried cases, and in employment discrimination cases, where plaintiffs succeed at trial only 21 percent of the time.⁴¹ But consistent deviation from 50 percent across large numbers of case categories does have a cumulative effect. The 50 percent hypothesis probably does not hold even under ideal conditions. Selection effect theory, however, is not undermined by the data.

Both selection effect theory and observed success rates suggest the need to test non-50 percent hypotheses. The techniques described here can be used to test other hypotheses. One can extend distributional analysis to test the hypothesis that a class of cases has a binomially distributed success rate of any given percent. Assume that several measurements of success rates for the class are available, with each measurement having a different number of trials. Rather than compare each success rate with the predicted percent figure, one should ask whether the observed distribution of success rates could come about from a binomial random process with $P = 0.x$, where $0.x$ is the hypothesized mean success rate, and in which there are the observed number of measurements and the observed number of trials leading to each measurement.

The process of fitting a predicted normal frequency distribution to a

³⁹ See the Appendix.

⁴⁰ *Id.*

⁴¹ Eisenberg, *supra* note 1, at 1578. Other results in a study of civil rights and prisoner trials seem not so readily explained by different stakes analysis. See *id.* at 1579-84, 1593-94.

series of success rates differs when one departs from the 50 percent prediction. To illustrate, consider the area of products liability. Measurements of plaintiff trial success rates in products cases consistently fall well below 50 percent.⁴² These repeated low observations could not result from random fluctuation. Products liability therefore is a prime candidate for non-50 percent analysis.

Given the departure from 50 percent success, there no longer is a theoretically predictable success rate ($P = 0.5$) from which to derive a predicted sample mean and standard deviation. Thus, the technique used to generate the expected frequency of success rates in Figures 1–3 above is not available. If theory does not suggest a success rate, the data are the only other realistic source of a sample mean. The Administrative Office data provide a trial success rate for each district for the 7 years studied. One could take the mean of this series of success rates and ask whether the distribution could arise from a normally distributed sample of success rates with the observed mean and variance predicted by the mean and a binomial process with $P = 0.36$.⁴³

Figure 4 shows the distribution of products liability case success rates for the 64 districts reporting more than 10 products liability trials during the period covered. The expected frequency line results from using the mean of the observed success rates, $P = 0.36$, and the variance associated with the mean for each district, $P \times (1 - P)/n$, where n is the number of trials in the district. As before, one can construct the expected distribution frequency, which has a standard deviation of .096. There is a reasonable fit between the observed and expected distributions. One cannot reject the hypothesis that the observed distribution of frequencies could result from the predicted distribution.⁴⁴

A second issue attending the predicted frequency distribution arises when one allows for different stakes. Suppose one believes that different stakes explain the low success rate but also that different sets of different stakes are at work in different districts. Perhaps one believes that defendants' stakes in California are three times plaintiffs' stakes but that defen-

⁴² Eisenberg & Henderson, *supra* note 23; Priest & Klein, *supra* note 1, at 38 (table 7); the Appendix.

⁴³ The .36 success rate used differs from that in the Appendix (showing personal injury—product liability = .25) because I have excluded one district with extraordinary trial results. See Henderson & Eisenberg, *supra* note 31, at 519 n.159, noting hundreds of Bendectin trials simultaneously resolved against plaintiffs in consolidated actions. See also Eisenberg, *supra* note 1, at 1593 n.93, showing distorting effect of this district on overall tort trial success rates. For discussion of the Administrative Office's various products liability categories, see Henderson & Eisenberg, *supra* note 31, at 521.

⁴⁴ The predicted normal distribution is one with mean = .36 and standard deviation = .096. The Kolmogorov-Smirnov Z is .650, corresponding to a two-tailed probability of .791.

FIGURE 4.—Probability of success at trial: expected vs. observed frequency of districts, federal products liability cases. Source: Administrative Office data, 1978–85 (one district omitted).

dants' stakes in Montana are only twice plaintiffs' stakes. There is evidence of wide regional variation in success rates within categories of cases.⁴⁵ Each district's success rate could then be viewed as a binomially distributed random variable with a distinct success rate and variance. If enough cases reach trial, the binomial distribution is well approximated by a normal distribution.⁴⁶ Since the combination of random variables that are normally distributed is itself a random variable that is normally distributed,⁴⁷ belief in interdistrict differences in stakes should not deflect from a predicted distribution of district success rates that is normal but of unknown variance. If one can make sensible statements about the size of the stakes differential, the observed distribution could be tested against a predicted distribution.⁴⁸

⁴⁵ Eisenberg, *supra* note 1, at 1584–94.

⁴⁶ Kmenta, *supra* note 20, at 85–97.

⁴⁷ *Id.* at 90.

⁴⁸ One must be wary of viewing any normal distribution as supporting predictions of normality (as opposed to predictions of specific means and variances) because the central limit theorem forecasts normality for a wide range of samples. Whatever the shape of the frequency distribution of an underlying population, the frequency distribution of the independently sampled means in repeated samples tends to become normal. See Snedecor & Cochran, *supra* note 21, at 44.

This analysis of products liability is subject to the same qualification that attends analysis of the larger tort category. Products liability consists of many different subcategories of cases, including design defects, failure to warn, and manufacturing defects. The stakes and other characteristics of these subcategories may differ. To the extent they do, it is misleading to search or test for a single success rate across the full range of products liability cases. That the observed distribution of success rates fits the binomial random theory fairly well might be taken as evidence that products liability is, in some measure, a reasonably homogeneous category.

V. CONCLUSION

The study of litigation models has become an important branch of interdisciplinary legal scholarship. The many noteworthy litigation model theories have been accompanied by noticeably fewer empirical tests of those theories. Almost none of the literature examines how to test the theories. This article presents a theory of how to test a major set of litigation models. It is no more complete than are the models themselves. Like the other empirical tests of selection effect theory, the theory sketched and applied here accounts for neither the size of awards at trial nor the possibly different stakes to the parties. The theory uses a simple, readily observable measure of success—who won at trial.⁴⁹ A nominally successful trial is sometimes in fact a loss for the nominal winner. Without observing amounts recovered and expended on the litigation, more sophisticated measures of success must await future research. As the theory of litigation models changes, so will the theory of how to test those models.

Despite these limitations, the results presented here should assist in focusing litigation model efforts. First, the enormous variety in trial success rates across case categories, together with the different attributes of legal subject areas, suggests that the search for a single theory to explain trial or other success rates may be futile. Even a subject matter category consisting of a single traditional doctrinal area, such as tort, shows varying success rates across its many subcategories.⁵⁰ A more fruitful line of analysis may be to dig deeply within subject matter areas before trying to generate a broad, testable litigation model.⁵¹ Each subject matter area will have its own set of factors that potentially influence the outcome of trials. Some subject matter areas, such as civil rights, may be expected to dis-

⁴⁹ For discussions of assessing success, see note 5 *supra*.

⁵⁰ See the Appendix.

⁵¹ Eisenberg, *supra* note 1 for civil rights cases; Henderson & Eisenberg, *supra* note 31 for products liability cases; Priest & Klein, *supra* note 1 for tort cases.

play regional effects due to possible jury prejudice.⁵² Other areas, such as products liability, might be substantially influenced by recent statutory developments that differ across states, as well as by different stakes to the parties.⁵³

Second, distributional analysis can better identify weaknesses in existing theory than can other methods of testing litigation models. Without distributional analysis, it is difficult to decide when an observed set of results departs from theoretical predictions. Reshaping the theoretical model in light of observed data requires that one accurately evaluate the data's conformity with theory. Distributional analysis offers a powerful tool to help with such evaluations.

⁵² Eisenberg, *supra* note 1, at 1584–94. Predicted regional effects in civil rights cases are not found; but, possibly contrary to selection effect theory, regional variation does exist.

⁵³ Henderson & Eisenberg, *supra* note 31.

APPENDIX

TABLE A1

PLAINTIFF SUCCESS RATES AT TRIAL, BY CASE CATEGORY,
FEDERAL DISTRICT COURTS, 1978-85

	Administrative Office Code	N	Success Rate
Contract:			
Insurance	110	2,494	.52
Marine	120	1,089	.69
Miller Act	130	285	.79
Negotiable instruments	140	560	.78
Recovery of overpayments and enforcement of judgments	150	128	.77
Recovery of defaulted student loans	152	19	.84
Recovery of overpayments of veteran benefits	153	51	.80
Other contract actions	190	6,643	.64
Contract product liability	195	252	.57
Real property:			
Unlabeled category	21-	19	.74
Land condemnation	210	4,298	.12
Foreclosure	220	211	.84
Rent, lease, ejectment	230	91	.68
Torts to land	240	297	.64
Tort—product liability	245	44	.30
Other real property actions	290	389	.54
Torts—personal injury:			
Airplane	310	448	.55
Airplane product liability	315	149	.40
Assault, libel, and slander	320	490	.42
Federal employers' liability	330	890	.72
Marine	340	2,429	.59
Marine—product liability	345	46	.37
Motor vehicle	350	3,261	.60
Motor vehicle—product liability	355	392	.33
Other personal injury	360	3,808	.46
Personal injury—medical malpractice	362	697	.38
Personal injury—product liability	365	3,255	.25
Torts—personal property damage:			
Other fraud	370	716	.60
Truth in lending	371	10	.60
Other personal property damage	380	980	.59
Property damage—product liability	385	243	.40
Civil rights:			
Other civil rights	440	4,970	.33
Voting	441	118	.53
Jobs	442	7,165	.21
Accommodations	443	223	.43
Welfare	444	59	.47

TABLE A1 (Continued)

	Administrative Office Code	N	Success Rate
Prisoner petitions:			
Vacate sentence § 2255	510	89	.20
Habeas corpus	530	811	.26
Mandamus and other	540	35	.31
Prisoner—civil rights	550	3,489	.13
Forfeiture and penalty:			
Agriculture acts	610	23	.52
Food and drug acts	620	93	.81
Occupational safety/health	660	14	.57
Other forfeiture and penalty suits	690	644	.86
Labor laws:			
Fair Labor Standards Act	710	610	.63
Labor/Management Relations Act	720	873	.48
Labor/Management Reporting and Disclosure Act	730	137	.44
Railway Labor Act	740	50	.34
Other labor litigation	790	622	.38
(ERISA)	791	279	.53
Property rights:			
Copyright	820	243	.71
Patent	830	473	.48
Trademark	840	407	.67
Social security:			
Social Security (before July 1, 1978)	860	41	.34
Medicare Act Part A (42 USC 1395ff(b))	861	15	.20
Disability Insurance (42 USC 405(g))	863	70	.43
Supplemental Security Income Disability Title XVI	864	12	.50
Other statutes:			
Antitrust	410	586	.43
Bankruptcy trustee	420	118	.69
Bankruptcy transfer	421	34	.62
Bankruptcy Appeals Rule 801	422	224	.52
Banks and banking	430	58	.47
Interstate commerce	450	190	.62
Deportation	460	27	.19
Securities, commodities, exchange	850	833	.48
Tax suits	870	1,313	.65
Internal Revenue Service (IRS)— third party suits (26 USC 7609)	871	160	.92
Other statutory actions	890	1,100	.47
Agricultural acts	891	182	.32
Environmental matters	893	182	.51
Energy Allocation Act	894	55	.84
Freedom of Information Act of 1974	895	46	.33