Empirical Legal Studies Before 1940: A Bibliographic Essay

Herbert M. Kritzer

University of Minnesota Law School, kritzer@umn.edu

Follow this and additional works at: https://scholarship.law.umn.edu/faculty_articles

Part of the Law Commons

Recommended Citation


This Article is brought to you for free and open access by the University of Minnesota Law School. It has been accepted for inclusion in the Faculty Scholarship collection by an authorized administrator of the Scholarship Repository. For more information, please contact lenzx009@umn.edu.
EMPIRICAL LEGAL STUDIES BEFORE 1940: A BIBLIOGRAPHIC ESSAY

Herbert M. Kritzer*

INTRODUCTION

“Empirical Legal Studies” is a term that began to come into vogue around 2000. ELS built on and extended the law and society (socio-legal studies) approach and the law and economics approach, both of which have strong empirical elements. Of course empirical research on law and legal processes predate law and society and law and economics, with a number of well-known studies that were conducted in the 1950s and early 1960s. These mid-century projects include the American jury project (Kalven and Zeisel 1971), the commercial arbitration study (Mentschikoff 1952; Mentschikoff 1961), the court delay study (Zeisel and Callahan 1963; Zeisel, Jr., and Buchholz 1959), the pretrial settlement conference study (Rosenberg 1964), studies of Supreme Court decision making (Kort 1957; Kort 1966; Pritchett 1948; Schubert 1959; Schubert 1963; Schubert 1965; Snyder 1958), studies of the legal profession (Carlin 1962; Carlin 1966; Johnstone and Hopson 1967; Smigel 1964; Zander 1968), studies of compensation for injuries suffered in auto accidents (Connard, Morgan, Pratt, Voltz, and Bombaugh 1964; Hunting and Neuwirth 1962; Linden 1965; Morris and Paul 1962), and studies of the lower courts (Dolbeare 1967; Green 1961; Jacob 1969). Many of these studies of the 1950s and 1960s have framed research agendas that continue to this day.

While the mid-century studies are reasonably well-known, much more obscure is the body of research conducted before World War II. In the 1920s and 1930s, and in a few cases even earlier, one can find a wide range of empirically-oriented research on law. The specific topics of this early research include:

- Appellate courts and appellate decision making
- Automobile accident compensation and litigation
- Banking
- Bankruptcy
- Criminal courts
- Debt
- Divorce

* Marvin J. Sonosky Professor of Law and Public Policy, University of Minnesota. The staff of the Warren Burger Library at William Mitchell College of Law was extremely helpful in locating the information used in this essay; I would particularly thank Margaret James and Dawn Sommers for handling my large number of interlibrary loan requests, and Sonya Huesman for help in locating some of the more obscure materials. Steven Hainlen provided valuable assistance in the early stages of the project. Cyril Glasser was extremely helpful in regards to trying to locate research from the UK.

1 This project produced a number of law review articles in addition to the well-known seminal book (see Kalven 1959; Kalven 1964; Strodtebeck and Hook 1961; Strodtebeck, James, and Hawkins 1957).
A significant portion of the early work was linked one way or another to the legal realist movement, and one can find some description of the research in historical treatments of legal realism (Kalman 1986; Schlegel 1980). In significant part because legal realism was essentially an American movement,\(^2\) almost all of the empirical legal research of the period was done by Americans focusing on the United States.\(^3\)

The purpose of the discussion that follows is to highlight the range of empirical research on law from the pre-World War II period. To locate that research, I examined the tables of content of all law reviews published prior to 1940, reviewed a large number of English government reports (command papers), and followed up on any citations I could find within those works to other sources. While I do not discuss every study I found, I have provided as complete a bibliography of that research as possible. My discussion of the research is organized around a series of topics. While most studies easily fall within a single topic, a few stretch across multiple topics, and appear more than once.

**EMPIRICAL RESEARCH ON CRIMINAL JUSTICE**

Some of the earliest empirical legal research focused on the operation of the criminal justice system; the earliest study I found was published in 1911. However, empirical research related to crime came even earlier, in the mid to late 19\(^{th}\) century with the beginnings of positivist criminology. The first publication of annual crime statistics began in France in the 1820s, with some American states following soon thereafter. These statistics produced the insight that crime

---

\(^2\) See Kritzer (2010) for a discussion of the factors that account for the U.S.-focus of early empirical research on law, and for a discussion of the other factors that likely explain what early empirical research appeared during this particular period.

\(^3\) Due to the author’s limitation, this essay focuses on empirical research in English-speaking (and hence common law) countries. It is very likely that at least some empirically-oriented legal research was conducted in other countries in the pre-World War II period, but that research is not included in the following discussion. Three was empirical research on crime patterns in Europe going back into the 19\(^{th}\) century.
rates, both general and particular, were relatively stable from year to year, and that differences among localities in crime patterns also varied little over time. Arguably this led to the hypothesis that crime was in some sense fundamental to a segment of human society, reflecting a combination of stable social and biological characteristics. Adolphe Quetelet sought to show that crime reflected social conditions such as poverty and lack of education (which in turn led to social behaviors such as alcohol consumption) combined with social characteristics such as age and gender (Quetelet 1833); around this same time, Andre Guerry (1833) was drawing on these same crime statistics to examine patterns of crime in France. In contrast, Cesare Lombroso (1876) proposed that something one might label a “criminal type” existed, and was marked by certain physical characteristics which could be measured).

Early empirical (positivist) criminology was largely focused on the causes of crime. It was not until the early 20th century, by which time statistics on crime and criminal cases in court, were routinely published by many European countries and perhaps half of the American states, that empirical research on the handling of crime by the legal system began to appear. An early impetus, at least in the United States, was the creation of the American Institute of Criminal Law and Criminology in 1909, which the next year started publishing a journal that we know today as the Journal of Criminal Law and Criminology. The Institute grew out of the National Conference on Criminal Law and Criminology organized by the then dean at Northwestern Law School, John Henry Wigmore. The Institute’s objective was “to advance the scientific study of crime, criminal law and procedure” (Rundell 1912, 3).

While it was several years before empirical studies of criminal case processing began to appear in the new institute’s journal, Oliver Rundell, an instructor at the University of Wisconsin Law School, quickly developed plans for a study of the duration of criminal cases in Wisconsin, publishing a report of this study in 1912 in a university publication (Rundell 1912). Rundell and his assistants collected data from courts in three counties, covering five years for the higher volume lower courts and 10 years for the Circuit Courts which had a much lower volume. While slim, this report was to presage a burst of research that was to come in the form of the crime surveys. Specifically, Rundell found that relatively few cases actually went to trial before a jury; this was particularly the case for the lower court. Moreover, while there could be considerable

---

4 A 1933 review of research on criminology (Michael and Adler 1933) considered causation, treatment, prevention, and administration of the criminal law (policing, courts, sanctions). The review cited very few empirical studies from before 1920.

5 See Warner (1934, 35 ) for a list of states publishing court statistics by 1930; most of the series began after 1900, although the earliest 1879 by Indiana, quickly followed by Ohio in 1880.

6 Aside from this one journal, specialist criminology journals did not begin to appear until well after the period considered in this essay. For example, the British Journal of Criminology began publication in 1960, Criminology in 1970, and Criminal Justice in 1963. While there undoubtedly were criminology articles appearing in sociology journals during the period of interest, no effort was made to search out those articles.
delay in cases actually getting to trial, the actual trial of cases consumed “comparatively little
time” and “protracted trials are a rare occurrence” (p. 59).

Another of the earliest studies focused on what is a continuing question: the need to provide
public defenders. Walton Wood, in a study of Los Angeles County published in 1916 compared
cases with private attorneys, pro bono attorneys, or paid public defenders, and found that 70
percent of the clients of public defenders pled guilty compared to only 49 percent of those
represented by private attorneys. The majority of the difference reflected the ability of private
attorneys to obtain dismissals, although 29 percent of private attorneys’ clients went to trial
compared to 22 percent of public defenders’ clients (and 26 percent of those with pro bono
representation). Trials involving public defenders were shorter, and public defenders filed fewer
motions and fewer appeals. Wood’s conclusion, somewhat perverse by contemporary
perspectives, was that the use of public defenders “resulted in a savings to the county” because
of more guilty pleas, shorter trials, and fewer motions and appeals (Wood 1916). 7

A third study from prior to 1920 examined the behavior of magistrates hearing minor
criminal matters in New York City. Drawing on reports from the City Magistrates’ Courts,
Everson (1919) describes the variation in behavior among about 40 magistrates. For cases of
intoxication, there was a high degree of consistency, with most magistrates convicting over 90
percent of those charged; however, one magistrate convicted only about two-thirds and one
convicted only about one-fifth. In contrast, there is a great deal of variation in the sentences
imposed on those who were convicted. The author also describes variation in how disorderly
conduct cases were handled; while the variation was less than in intoxication cases, it was also
quite substantial.

During the 1920s and 1930s, the crime “surveys” carried out in a variety of states and
communities represented the first major burst of empirical research on how criminal courts
handled cases. 8 The Cleveland Crime Survey, directed by Roscoe Pound and Felix Frankfurter
(1922) was the both the first and most extensive of these projects, involving a large team of
researchers and reports/chapters on issues including police administration, prosecution, criminal
courts, corrections, “medical science and criminal justice,” legal education in Cleveland, and
newspaper coverage of crime. The other crime survey reports were more limited in their contents

---

7 It is worth noting that at the time Wood was the Los Angeles County Public Defender.

8 A useful compilation and summary of some aspects of the crime surveys can be found in the Wickersham
Commission’s Report on Prosecution (National Committee on Law Observation and Enforcement 1931). There is a
lengthy section (pp. 186-221) that presents tables drawing on the then extant studies. While I describe these as the
first “burst”, the foregoing discussion should make clear that there were empirical studies of the criminal courts
prior to the crime surveys. In addition to the studies mentioned previously a notable example is the report of the City
Council Committee on crime of the City of Chicago, chaired by political scientist Charles Merriam, which issued a
report in 1915 that included a variety of statistics on how the criminal courts in Chicago were operating. These
statistics were later updated through 1921 by Edith Abbott (1922); for some specific issues, Abbott’s analysis
included comparisons to statistics from England and Canada, perhaps the first cross-national study of criminal
courts.
(I found no other reports with sections on legal education or newspaper coverage), but discussions of policing, prosecution, and disposition of cases were common in many of the reports; some reports included discussions of release on bail (and issues related to professional bondsmen) and others considered how criminal cases fared in the appellate process (discussed below under “appellate courts”).

A central focus in the crime surveys was the disposition of cases, and generally the studies provided tables, often labeled “mortality tables,” showing that seldom more than half, and often many fewer, of felony arrests resulted in convictions, and most of the convictions that did occur came as a result of guilty pleas; jury trials constituted a fraction of dispositions. The exact percentages ultimately convicted varied, and where along the road to trial cases veered off differed from one community to another. For example, in Oregon, New York, and Illinois, the percentage of felony arrests leading to conviction was 18 to 23 percent (Morse and Beattie 1931; see also Morse and Beattie 1932). The Cleveland Crime Survey reports that about 32 percent of felony arrests result in conviction (Pound and Frankfurter 1922, 236). The Missouri Crime Survey found that 38 percent of “warrants issued” resulted in convictions, 82 percent by guilty pleas (Moley and Thomas Hennings et al. 1926, 126-27, 169). The Georgia Crime Survey reports that about 41 percent of defendants are convicted and punished (Georgia Department of Public Welfare 1925, 182). In contrast, the Illinois crime survey included Milwaukee, Wisconsin as a comparison, and reported that 64 percent of arrests in Milwaukee resulted in convictions (Wigmore 1929c, 38). This latter figure for Milwaukee is similar to what was reported in a Minnesota study focusing on the counties where Minneapolis and St. Paul or located; 59 percent of cases getting to a preliminary hearing resulted eventually in convictions (Minnesota Crime Commission 1934, 88). One caution in comparing these figures is that the starting point for tracking “case mortality” varies from study to study. For some it is arrest while for others it is entering court at a preliminary hearing (thus not counting cases not pursued after arrest).

A central concern for the various crime survey authors was this frequent failure of arrests to lead to conviction. To these authors, this failure was indicative of problems in the criminal courts such as overworked prosecutors, loafing judges (Pound and Frankfurter 1922, 299-333), and

---

9 A common feature of studies of criminal justice during the 1920s and into the early 1930s was separate considerations of “liquor cases” which arose due to prohibition. Usually the disposition patterns of “liquor cases” differed in a variety of ways from other criminal cases (often faster, more guilty pleas, and minimal sanctions).

10 A good overview of these studies can be found in the Wickersham Commission’s report on prosecution (Bettman 1931) which drew on studies and statistics from eighteen states or localities.

11 While not part of a crime survey, a particularly extreme result in this regard is reported by Clark and Shulman from their study of Connecticut courts (1937, 189). Their analysis of criminal cases focused on the Superior Court, and appears to include mostly felonies. Only 35 defendants out of 3012 cases examined pleaded not guilty

12 Raymond Moley (1929b, 28) provides a nice summary table showing patterns for eight cities; the percent ultimately convicted ranges from 19 percent to 64 percent.

13 Another issue that at least one author alluded to is inconsistencies among judges, with some judges discharging a majority of certain types of cases while judges discharged a small proportion of the cases.
small groups of defense lawyers specializing in cases in particular courts who knew how to work the system (Moley 1929a, 4008-10). Seldom, if ever, did the authors consider the possibility that the problem was not what happened after cases got into court, but rather what was coming into the courts to start with.

In addition to the crime surveys, there were a variety of other studies of criminal courts in various localities published during this period, many of these in the *Journal of the American Institute of Criminal Law & Criminology*. For example, a study of the police court in San Francisco found an overall dismissal rate of 78 percent, 63 percent for those charged with felonies (Heinzen and Rypins 1927, 81). Arthur Beeley (1927) carried out a detailed study of bail bonds and release prior to trial in Chicago, a theme that was considered in a number of the crime surveys; Beeley sought to determine whether some system other than bail bonds would be effective to insure that defendants, or at least certain defendants, would appear in court.

A team from Johns Hopkins University conducted extensive studies of courts in Ohio and Maryland, producing multiple volumes of analyses of criminal courts, some statewide (Bettman, Marshall, Jamison, and Miles 1932; Gehlke 1936; Marshall 1932a; Martin 1933) and some focused on particular communities including Cincinnati/Hamilton County (Douglass 1932; Douglass 1933) and Columbus/Franklin County (Blackburn 1935). One intriguing finding was a difference between patterns in Ohio and patterns in Maryland (or at least Baltimore). In Maryland, 73 percent of dispositions were convictions, and “only” 62 percent of those convictions came as a result of guilty pleas; however, only 17 percent of trials were jury trials—the rest were bench trials (Marshall 1932b, 10-11, author’s computations). Of those cases not resulting in convictions, 58 percent were formal acquittals, 84 percent of which came through bench trials (p. 15). Overall, 49.5 percent of dispositions in Maryland came through trials. In contrast, in Ohio 86.7 percent of convictions come by way of plea, and overall only 10.9 percent of dispositions come by way of trial, mostly jury trials (Gehlke 1936, 34, 88-91). What makes this contrast particularly interesting is the dominance of bench trials in Maryland, a finding highlighted in a study conducted 40 years later comparing criminal court patterns in Baltimore to those in Detroit and Chicago (Eisenstein and Jacob 1977); the authors of the three-city study titled their chapter on Baltimore “A City of Trials.”

In contemporary discussions of criminal justice processes, it is often asserted that high rates of guilty pleas, often labeled “plea bargaining,” reflect the pressures of heavy caseloads in urban courts. In a 1978 essay, Malcolm Feeley (1978) shows that guilty plea rates do not reflect urban-nonurban differences. In fact, Raymond Moley found the exact same thing 50 years earlier: 20 of

---

14 Moley (1928, 105) provides a table showing the percentage of convictions resulting from pleas for 24 cities (not including Baltimore because the Maryland study had not been completed at the time of Moley’s writing). The percentage of convictions by guilty plea ranged from a low of 33 percent in San Francisco (Atlanta was the next lowest at 47 percent) to a high of 95 percent in St. Paul. In five cities, guilty pleas accounted for 90 percent or more of convictions; in six cities 80 to 89 percent of convictions came by plea; in ten cities, 70 to 79 percent came by plea; in only three cities were less than 70 percent (in fact, less than 60 percent) of convictions by guilty pleas.
24 cities for which he found data had guilty plea rates of 70 percent or more; the same was true of four out of five counties without large urban centers, some rural, some partially rural (Moley 1928, 105-06).

While most of the empirical research on criminal justice processes focused on state courts, not surprising given that the vast majority of crime was, and is, handled by those courts, there were several studies of criminal cases in federal courts. The very first volume of the journal Law and Contemporary Problems in 1933 included a study of federal criminal prosecutions (Rubin 1934). This study drew exclusively on statistics appearing in the Annual Reports of the Attorney General (1922-1933). The compiled tables show the numbers of different types of cases, as well as cases controlling for population; a summary table provides information on outcome patterns. The author finds a great deal of variation from one type of crime to another. While overall the number of federal prosecutions had increased sharply, this was due to offenses related to prohibition, and once those cases were removed, there was a fairly stable pattern over the period examined. The author laments the limited nature of the information available. As part of a law student note calling for the creation of a “federal police court” to handle the large number of routine federal crime cases related to prohibition, Arthur Sutherland collected data on federal criminal cases in Massachusetts for 1913 and 1924 (Sutherland 1926). The number of cases shot up between the two time points, but almost all of the change reflected liquor cases that were resolved without trials. While trials and trial days did increase, the increase was very modest in comparison to the overall increase in cases.

In conjunction with the Wickersham Commission, the American Law Institute undertook a major study of federal crime cases, with Charles Clark taking leadership for the endeavor (American Law Institute 1934a). This study was quite massive in its scope, collecting detailed data on 35,761 criminal cases not concerning prohibition and less detailed data on another 37,065 cases connected to prohibition. These data came from 13 of the then existing federal districts, and were compiled by teams of field researchers trained and dispatched to the districts. As one would expect, the aggregate data showed sharp increases in overall caseloads reflecting the impact of prohibition. More generally, the data showed the dominance of guilty pleas for case disposition. Leaving out liquor cases, convictions were obtained in 77.9 percent of cases, and 90.5 percent of those convictions came on pleas of guilty (p. 52). Interestingly, the high guilty plea rate developed after 1915; until then only about half of the federal convictions came through guilty pleas. Importantly, the increase preceded prohibition, with most of the change coming between 1915 and 1918 (the prohibition amendment was ratified in 1919 and came into effect in 1920). There were substantial variations among the districts in how cases were handled. The percentage of convictions coming through trial varied from less than one percent to almost one third; five of the thirteen districts had no bench trials, most of the rest had rates in the 90 to 99 percent range, and in one district, the one in which almost a third of the convictions were through trial, only 49 percent of the trials were before a jury (p. 76).
While the focus in the above discussion has been on disposition patterns, there are many other themes in the early empirical research on criminal justice processes. A common element in most of the crime surveys was the length of time cases required for disposition; the focus in these discussions is largely descriptive statistics. Many of the studies provide summaries of data on sentencing patterns, although little is usually done to explain those patterns other than the natural control for offense. At least two studies do document that choosing to go to trial could lead to a more severe sentence. For example, in the extensive study of judicial administration in Ohio, 65 percent of those convicted by a jury were sentenced to prison compared to something less than 50 percent of those who pled guilty; it is less clear whether the length of sentences for those choosing jury trials was longer than for those pleading guilty (Marshall 1932c, 183, 197). The ALI study of federal cases reports that 82.1 percent of nonliquor defendants were imprisoned after conviction at trial compared to 75.6 percent of those pleading guilty (American Law Institute 1934a, 86). More generally, while the exact phrase “plea bargaining” does not appear in any of the empirical studies, one does find discussions of “justice by compromise” (Moley 1929, 166-192), criminal cases “being settled by the striking of a ‘bargain’ between the defendant and the prosecuting officer” (Grant 1931, 980), “the guilty plea technique” (American Law Institute 1934a, 12), or “bargaining by the prosecutor with the criminal for pleas of guilty” (Morse 1931b, 247-59).

The study of magistrates in New York City identified differences in how judges handled cases, or what in modern terminology has come to be described as problems of disparity. Frederick Gaudet and colleagues collected data on the sentencing behavior of six judges from one county in New Jersey; they obtained information on almost 7,500 cases covering a period of nine years focusing only on types of cases for which there were likely to be sufficient numbers for comparison. The harshest of the six judges sent 57.7 percent of those convicted before him to prison while the most lenient sentenced only 33.6 percent to imprisonment (Gaudet, Harris, and St. John 1933, 816). There was similar variation in the median length of incarceration imposed, ranging from a low of 6 months (for two judges) to a high of 12 months (Gaudet, Harris, and St. John 1934, 679). Interestingly, there was no relationship between a judge’s tendency to imprison and the length of imprisonment the judge imposed ($r = .009$, calculation by the author).

Another contemporary issue that one can find reflected in the early research is the impact of the type of representation on outcomes. The analysis by the then Public Defender in Los Angeles County is described earlier in this section. The issue appears in a number of the studies of the period. For example, in their section on criminal courts in the Cleveland crime survey report, Reginald Heber Smith and Herbert B. Ehrmann report virtually identical rates of guilty pleas; appointed counsel are somewhat less successful in securing acquittals that are retained counsel

---

15 The modern development of systems of sentencing guidelines and mandatory sentencing grew in part from concerns about sentencing disparity.

16 Details on data collection can be found in Gaudet (1938, 9-10).
This analysis contradicts the “impression in Cleveland that the appointed counsel usually induces his client to plead guilty and pockets his modest fee for the persuasion” (p. 310). In a study of waiver of juries (discussed in more detail below), Abraham Goldberg (1929, 169) found minimal difference between the likelihood of waiver between appointed and retained counsel (58.7 percent vs. 62.1 percent). However, retained counsel were much more likely to obtain acquittals at trial (49.8 percent) than were appointed counsel (27.4 percent).17 In the study of Ohio’s criminal courts (Gehlke 1936, 132), 84.8 percent of those with assigned counsel were convicted compared to 71.4 percent of those with retained counsel; however, only 41.8 percent of those with no counsel were convicted, presumably reflecting dismissals early in the process.

While most of the research appearing in legal journals concerned how cases were processed, there were two studies that looked at the issue of predicting recidivism and/or parole violation. One study, by Sam Bass Warner (1923), concerned decisions to release on parole and the parolees success. Warner examined a wide variety of factors including demographics, prior record, type of crime, habits, physical and mental condition, behavior in prison, and assessment by an “alienist” (a psychiatrist). A total of 64 pieces of information was assembled for each inmate; very little was found to be predictive of success. Sheldon Glueck and Elinor Glueck (1928) were more successful in predicting recidivism among former prison inmates. They found a recidivism rate of 60 percent, with prior record measured and prior dependability as a worker being notable predictors; for example, 67 percent of those with serious prior records “failed,” 57 percent of those with “frequent minor” prior offenses failed, 35 percent with occasional minor prior offenses, and 21 percent of those with no prior record.

Several articles consider impact of changes to the legal system. In a series of three papers Abraham Goldberg (1929; 1930; 1931) examines how disposition patterns changed in Michigan after the passage of a statute permitting criminal defendants to waive a jury trial and be tried by the court. He found few waivers outside of the Detroit Recorder’s Court; in that court, the conviction rates for bench and jury trials were virtually the same. A problem with the study is that he examined only the first year after the change, and it is very possible that it took the players in the system a period of time to become accustomed to the change. It is worth noting that some states already allowed waiver of jury trials. Kenneth Martin (1933) examined waivers in Ohio (and to a lesser degree in Maryland). His study included a survey of judges, prosecutors and defense lawyers in Ohio as well as an examination of the results of jury and nonjury trials in the two states. Respondents were asked what would motivate a waiver, the relative efficiency of jury versus nonjury trials, and general attitudes toward waivers.

A second focused on differences arising in case processing disposition when prosecution was commenced based on a grand jury indictment versus a prosecutor’s information. Raymond

---

17 Author’s computation (from Goldberg 1929, 168); acquittal rates for retained counsel were virtually identical for bench and jury trials while for appointed counsel, acquittal came in 30.6 percent of jury tries and 25.5 percent of bench trials.
Moley (1931) compared five states using prosecutor’s informations and four states relying on grand jury indictments; the basis of comparison was a data set of 9,362 cases collected from information states for the study which were compared to data from 30,307 cases collected by previous crime survey projects in indictment states. The analysis showed higher conviction rates on average in information states (67 percent) than in grand jury states (52 percent). Additionally, more convictions came by plea in information states (84 percent) than in indictment states (75 percent). Possibly at least due in part to Moley’s analysis, Wayne Morse, a law professor at University of Oregon (and a future United States senator perhaps best known for his early opposition to the Vietnam War) undertook a massive study of the grand jury (Morse 1931a; Morse 1931b; Morse 1931c). The study had two components: data on actual grand jury decisions (7,414 cases from 21 states during the fall and winter of 1929-30) and a postal survey of judges around the United States (responses were obtained from 545 judges). The study of actual cases showed that grand juries agreed with the recommendation of the prosecutor in the vast majority of cases. Not surprisingly, judges tended to be happy with the system used in their state.

Criminal Justice Research in England

The preceding discussion has considered only research conducted and published in the United States. In England there appears to have been much less research. There were many fewer law journals and the English legal academy at the time had even less interest in empirical research than did academics across the Atlantic. Robert Jackson (1940) produced the first edition of his seminal work, The Machinery of Justice in England in the late 1930s, although due to the start of the war it did not appear until 1940. Jackson was attentive to empirical information, and the first edition of Machinery contained some scattered data showing such things as the number of convictions and the length of sentences in summary courts (p. 101, 181), statistics on juvenile cases (p. 159), and statistics on the Poor Prisoners’ Defence Act 1930 (p. 123). This information was largely, perhaps exclusively, gleaned from tabular statistics contained annual reports published on actions in courts, policing (e.g., arrests), and the prison service. Previously Jackson had published two articles on juries. In the first, he documents the sharp drop in the percentage of cases having jury trials, dropping from well over 50 percent to around 10 percent by 1923 (Jackson 1937). In the second article he has a discussion of eligibility for jury duty; as part of that discussion he reports a study of eligibility among the residents of one ward in Cambridge. About 5,000 residents were eligible to vote in Parliamentary elections and 3,500 in local elections; of those, only 187 met the qualifications to serve on a jury, 5 of whom were known to be Labour Party supporters, 62 Conservative Party supporters, with the rest unknown as to their political preference (Jackson 1938, 371-72).

In 1939, Leon Radzinowicz, who was later to establish the Institute of Criminology at Cambridge University, published an analysis of changes in sentencing patterns in England over
the period 1900 to 1936 in *Law Quarterly Review* (Radzinowicz 1939). He looked specifically at 1900, 1910, 1924, and 1936. His analysis documented the decline in severity of penal sanctions, including the sharply increased use of conditional release, and the virtual disappearance of corporal punishment. Much of the change he attributed to the Probation Act 1907.\(^{18}\)

As noted above, there were a variety of routine government publications containing statistical information on court activities, policing, crime, and prisons. In additional there were occasional departmental committee reports and royal commission reports where data were compiled or collected. For example, the Royal Commission on Police Powers and Procedure (1929) conducted a survey of senior policy officials, all recorders (criminal court judges), stipendiary magistrates,\(^{19}\) chairmen of Quarter Sessions (criminal courts), senior officials in certain government departments, and a number of “representative bodies.” The report of the Commission listed the questions asked; the replies were used in a nonsystematic fashion throughout the report, and it is unclear whether any systematic analysis was carried out. The Report of the Departmental Committee on Persistent Offenders (1932) contained an analysis showing the prior records of those sentenced to imprisonment, including breakdowns showing the number of prior imprisonments. The Report of the Departmental Committee on Corporal Punishment (1938) contained a variety of statistics concerning the incidence of corporal punishment over time, as well as an analysis examining the deterrent effect of corporal punishment for the specific offense of “robbery with violence.” Short of examining each and every report, there is no ready way to identify the use of empirical research in law-related government reports. Many volumes of such reports were examined and the three discussed above were the only criminal justice-related reports found that had an empirical element. Undoubtedly there are others, perhaps many others, although it is likely that there is little or nothing in the way of more sophisticated analyses, even according to the standards of the time.

**EMPIRICAL RESEARCH ON CIVIL JUSTICE**

**Delay**

The earliest academic study of criminal justice processes appears to have been Rundell’s study of “The Time Element in Criminal Prosecutions in Wisconsin”. Interestingly, the earliest study of civil justice processes appears to be Sam Bass Warner’s brief article, “Procedural Delays in California” (Warner 1920). This was the first report from a project that could have produced a wide ranging set of insights. Warner, then a member of the law faculty at the University of Oregon, collected data on a total of about 20,000 cases filed in 1913 in the superior court of Oregon.

\(^{18}\) The following year, Radzinowicz published another piece discussing the importance of criminal statistics for criminal policy and criticizing the then state of English criminal statistics [Radzinowicz, 1940 #14455, 1919 Magistrates staff the lowest level of criminal courts in England. While some magistrates are law-trained, until recently called “stipendiary magistrates” (now called “district judges”), most magistrates are (and were) laypersons.
courts of three California counties, the “Justices’ Court” of San Francisco, and the Federal
District for Northern California; this constituted all cases filed during that year. He analyzed the
time from filing to issue and from filing to judgment. Focusing on personal injury cases, he
compared the time required in court to that required for workplace accident cases before the
Industrial Accident Commission, and chastises the court process as taking “a ridiculously long
time”: 8.5 months to judgment by the courts compared to 2 months by the Industrial Accident
Commission. Bass’s discussion related the time taken to costs, although there was no statistical
analysis directly making the linkage, and ultimately concluded, “In nearly all actions, the cases
take far too long and cost a great deal too much.” This analysis hinted at the potential in his data;
regrettably, there appears to be only one other article published based on the massive data set
Bass assembled, and that focuses on divorce (discussed below).

The question of the duration of litigation as a common theme in the broad studies of civil
litigation that will be discussed below. In addition, there were several focused studies in the
literature of the period. Robert Slayton and Philip Brown (1939) examined the duration of
litigation at the trial, intermediate appellate, and supreme court level in Texas, using an annual
sample of 50 to 150 cases for the period 1905 to 1934. They described their study as “purely
factual and objective [with] no attempt to ascertain or deduce causes.” Herman Oliphant and
Theodore Hope (1932) reported an analysis of “day calendars” as part of a larger study of
litigation in New York City, noting the variation among calendars and commenting on the
difficulty of predicting when a case on the calendar would actually be reached. In some ways the
most interesting study was conducted by a team looking at the Milwaukee trial court
(Berkanovic, Connolly, Harvey, Heller, Mooberry, Porter, and Tonjes 1934). This study, which
was supported with funds from the depression-era Civil Works Administration, involved
observation in the court over an 11 week period. The kinds of findings presented could only be
obtained by the observational technique:

- The average judge spent just under 22 hours a week on the bench
- About a 100 minutes a week was lost due to late arriving attorneys
- Virtually no lawyers presented a trial brief to the judge, something that would have
  expedited trials
- For 148 out of 383 non-expert witnesses called to testify it appeared that the lawyer
  had not previously interviewed the witness.
- Of 140 attorneys who argued motions, 20 percent seemed completely unprepared
  and 61 percent had no briefs.

A general finding of the study dealt with the apparent lack of preparation by many, perhaps most,
attorneys appearing in court.
General Studies of Civil Litigation

The person most associated with early empirical research on civil justice is probably Charles Clark, professor and dean at Yale, “father” of the Federal Rules of Civil Procedure (Ebersole 1980, 53), and long time federal appeals court judge. Clark undertook empirical research, first in Connecticut (Clark and Shulman 1937) and then of the federal court system (American Law Institute 1934b), 20 with a motivation of “assisting in the solution of the most pressing problem in the law by scientific study of all procedure in its functional, comparative, and historical aspects.”21 His study of civil cases in Connecticut was modeled in many ways on the Cleveland Crime Survey:22 obtain information on all civil suits in the state courts for at least one year. As the study turned out, Clark and his collaborators obtained data on 14 years of cases in New Haven and 10 years of cases in Waterbury, a total of 27,596 cases. The analyses of these cases revealed that divorce and negligence constituted about a quarter each, and debt (including foreclosure) about a third (Clark and Shulman 1937, 8). A large proportion of cases involved no actual court action. Negligence, auto and other, constituted almost three quarters of trials, even though they were a bit less than a quarter of the overall docket. Where there were judgments (including defaults), over 70 percent were for the plaintiff; although this varied by type of case (i.e., 75 percent for auto negligence, 54 percent in other negligence, and 82 percent in contract including breach and debt). Awards were modest; in auto cases, only about a third of the awards exceeded $2,000. Few judgments were appealed (3 percent), and about a third of appeals produced reversals, regardless of whether the appellant was the plaintiff or defendant. The overall thrust of the analysis showed that most cases were routine, large numbers involved little or no action by the court (and for many where there was court action, that action was pro forma or magisterial),

More detailed analyses were provided for divorce cases and auto negligence cases. In the former, 80 percent of the cases were uncontested. Grounds cited were predominantly desertion (48 percent) and cruelty (41 percent). Three quarters of the cases involved children. In the auto negligence cases, an interesting finding was the dominant role of a small number of attorneys. A total of 3613 appearances were filed in these cases involving a total of 408 attorneys; 22 attorneys appeared in 53.7 percent of cases constituting 42.3 percent of all of the appearances.

Parallel to the American Law Institute study of federal criminal cases, Clark’s study of federal civil cases was carried out in the same thirteen districts. Resource and time constraints largely limited the data collection coverage to a single fiscal year (July 1929 through June 1930),

---

20 Clark is listed as the chair of the project advisory committee.

21 Schlegel (1995, 83) describes Clark’s Connecticut courts study as “the first of its kind” (p. 91). It is not clear that this is actually true. Warner’s study in California appears to have been somewhat similar, involving the collection of data on some 20,000 or so cases; regrettably, so little was published from that study (only brief articles on divorce and on delay) that we do not know its full scope.

22 Clark and Shulman’s Connecticut study also included an analysis of criminal cases, although that analysis was less detailed than the analysis of civil cases.
although for a small number of districts two years worth of data were collected. Early chapters in the report trace the rise in the number of cases during the period covered by the Attorney General’s report. Analyses reported from the original data collection include the basis of jurisdiction, frequency of objections to jurisdiction, types of legal issues involved, how cases are terminated, prevailing party, size of money recoveries, nature of nonmonetary relief obtained, duration of cases, residence of parties in diversity cases, removal of cases from state court, appeals, and the use of three-judge courts. One central theme is the variation among districts on many of the issues. For example, two districts disposed of 81 percent of their cases within three months while another district disposed of only 9 percent during that length of time; in one district, only 2 percent of cases lasted more than a year while for another 54 percent required more than a year. One enduring aspect of this study, which in turn flowed from the Connecticut study, was the development of a form for recording data about cases; the form developed served as the basis of the forms that became the standard collection vehicle for information on federal civil cases, which in turn is used for preparing annual reports (and is today made available to researchers for their own analyses).

There were two studies of civil litigation in New York City. The first (Wherry 1931) was a small scale study based on data compiled by a judge from what transpired in his courtroom supplemented by the annual statistical reports published for the court as a whole. The analysis covers use of trials, the impact of rules changes (requiring a positive election of jury trial and payment of a fee rather than a jury trial being the default procedure), outcomes of trials, and size of awards. The author points out differences among types of cases on variables such as duration, likelihood of winning, and amounts of awards. One interesting point made by the author is that of the “cases analyzed in this study, 8 percent were palpable fakes, and there were two cases in addition in which the defense was a fake. All of these cases were in the personal injury class, except one, which was a malpractice suit” (p. 423).

The Johns Hopkins Institute of Law undertook what appears to have been planned as a massive study of civil litigation in New York County. The goal was to examine in detail some 100,000 cases in 1927 and 1930 (Oliphant 1931). This study appears to have been motivated at least in part about a concern about the economic costs of litigation featuring prominently a quote from Herbert Hoover when he was Secretary of Commerce:

“The aggregate economic loss to the United States through the necessity of referring to courts the trial of disputes arising in the course of trading would stagger the imagination if it could be accurately compiled, both as to the actual cost and the indirect drain upon natural resources. Next to war, commercial litigation is the largest item of preventable loss in civilization.”

---

23 See Shafroth [, 1948 #14402, 208-09] for samples of the early forms used for collecting information on federal cases for statistical reporting purposes.
The brief preliminary report from which this quote is taken shows the data collection forms that were developed and provides some very sketchy data, although how much data had been actually collected is not clear. The only other analysis that appears to have been produced as part of this study is the analysis of day calendars discussed above. Whether the study was abandoned before the data collection was complete, or the researchers obtained the data but ran out of money or energy is not clear.

Auto Accident Litigation and Compensation²⁴

The early 20th century saw the emergence of litigation over injuries resulting from auto accidents. The general studies of litigation described above showed the significant portions of courts’ civil caseloads made up by such cases. The Committee to Study Compensation for Automobile Accidents (1932) undertook the first empirical study of the compensation process. Data were assembled on 8,849 accidents, through a combination of personal interviews and an examination of court records. One central finding was that compensation as a percentage of loss declined as loss increased (p. 67, 91), a theme that continues to be widely discussed today with regard to the American tort system (see, for example, Abraham 2008, 91-92). In a period prior to mandatory insurance laws and before the development of “uninsured motorist” coverage (whereby tort victims can be compensated by their own insurance company if the tortfeasor is uninsured), a key finding was the relationship between insurance and receipt of compensation: compensation was paid in 69 percent of temporary disability cases and 96 percent of permanent disability cases when the tortfeasor was insured but only 21 percent and 17 percent when uninsured (p. 78, 81). Very few cases went to trial, and those that did produced plaintiffs’ verdicts in about two-thirds of the cases; the median award in most locations was less than $1,000 (p. 93).

One aspect of the study was an intensive examination of compensation from all sources in Connecticut accidents, including employer-provided insurance, life insurance, or the victim’s own insurance. A number of modern studies have similarly sought to examine compensation broadly, although such studies have not typically been limited to auto accidents (see also, Abraham 2008; Harris, Maclean, Genn, Lloyd-Bostock, Fenn, Corfield, and Brittan 1984; Hensler, Marquis, Abrahamse, Berry, Ebener, Lewis, Lind, MacCoun, Manning, Rogowski, and Vaiana 1991; Royal Commission on Civil Liability and Compensation for Personal Injury 1978). The Connecticut analysis shows that in two thirds of cases, part or all of the cost of the accident was born directly by the victim or the victim’s family. This finding, combined with (1) the central role of insurance and (2) the inverse relationship between size of loss and the degree of compensation led the authors of the study to propose that auto accident injuries be compensated

²⁴ While some government reports by police agencies in England provide statistics on road accidents and injuries, I could find no research during this period on how compensation issues were handled in the wake of such accidents.
by a no-fault plan similar to workers’ compensation. The theme of the need for some alternative system for compensating injuries arising from automobile accidents recurs in the other studies of those accidents conducted during this period (Feinsinger 1934; May 1932; Patterson 1933; see also Simon 1997-1998).

Two studies conducted about the same time in Wisconsin are worth noting. One examined the closed claim files of four insurances, assembling data on 1081 claims (Feinsinger 1934). The closed claim study has become one of the standard research approaches for studying injury compensation, particularly compensation for injuries in automobile accidents (see, for example, Insurance Research Council 2003). Strikingly, 72 percent of claims were closed within two months, with 40 percent settled for $100 or less and 77 percent for $500 or less. The median settlement was around $165. That $165 settlement figure translates to about $2,250 in 2001 dollars. Compare these to a national study of auto accident claim files closed in 2002 (Insurance Research Council 2003, 160) which reported a median tort payment of $6,555 for claimants with attorney representation $1,564 for claimants without attorney representation (roughly equal numbers of represented and unrepresented). Given that medical costs drive tort payments, the translation of the $165 figure to 2001 dollars is probably fails to provide a close approximation to the 2002 figure, but the comparison is nonetheless interesting.25

About the same time the study of closed insurance files, another study focused on court cases in 18 Wisconsin counties (Brown 1935). The study reports a variety of interesting points. Cases were less likely to go to trial in Milwaukee County than in the other counties (10 percent versus 25 percent), but plaintiffs were more likely to be successful outside Milwaukee (86 percent plaintiffs’ verdicts outside Milwaukee versus 70 percent in Milwaukee). The median verdict in Milwaukee County was about $3,000 compared to $1,000 outside Milwaukee. This pattern suggests that insurers dealt differently with claims in the urban area compared to the other areas. More interesting is that the $3,000 median verdict in Milwaukee is the equivalent of $40,869 in 2001 dollars; a study of jury verdicts from large U.S. counties in 2001 reported a median for auto accident cases of $16,000, and a median of $27,000 for all cases with plaintiffs’ verdicts; the median of all tort awards (97 cases) in Milwaukee County was $19,000 (Cohen 2004); for the 36 auto injury cases, the median in 2001 was only $10,500.26 Thus, adjusting for inflation (again noting that ideally one would use a medical cost inflator which is not available for this long of a time period), it appears that tort awards in Milwaukee County are today only about one quarter of what they were around 1930. The study also collected data on settlement in these cases, and found that the median settlement, both in Milwaukee and elsewhere, was about $1,000. The author made several other observations that are as applicable today as they were 80 years ago:

---

25 Note that these figures exclude fatalities and permanent disability cases; however, it is unlikely that including such cases would alter the median to a significant degree.

26 This figure is from the author’s analysis of the 2001 data.
The plaintiff starts high and the defendant low. … If liability is clear the defendant’s counsel will seldom allow the case to go to trial. If there is a ‘fifty-fifty’ chance of a favorable verdict for the plaintiff there is a tendency to settle for one half of the actual estimated damages.... It was said that one insurance company habitually raises the amount it sets aside as its reserve for each claim filed whenever it appears that a certain able Wisconsin attorney appears for the plaintiff in the case.27

As was true of other article on auto accident compensation, this one concludes with a discussion of a compensation system modeled on workers’ compensation.

During this same time, the Commonwealth Club of California undertook an inquiry into the system of compensation for injuries suffered in automobile accidents, asking whether the existing system should be replaced by a compensation commission resembling that in place for industrial accidents. As part of this inquiry, the Commonwealth Club’s Research Service carried out a statistical study of 200 consecutive personal injury actions in the Superior Court of San Francisco plus 200 contested industrial accident cases arising in San Francisco that were litigated before the Industrial Accident Commission (IAC); for both forums the cases examined were filed in the first quarter of 1928 (May 1932).28 The reported analysis shows amounts recovered, total medical expenses, total attorney fees, total net compensation (amount recovered less attorney fees and medical expenses). For court cases, the attorneys’ fees averaged about one third of the recovery. In IAC cases attorneys fees were very modest, averaging something around $30 which by my estimate was somewhere between 5 and 10 percent of the recovery, and often considerably less than that (e.g., in a case which settled for $2,001, the attorney’s fee was only $25). One notable fact about injury cases, consistent with the COSSA study, was that of the 116 cases in which compensation was obtained, 88.8 percent of the defendants were insured while of the 54 cases in which no compensation was paid, only 40.7 percent of defendants were insured.29

---

27 See Kritzer (2004, 219-20) for a similar observation about a contemporary plaintiffs’ lawyer in Wisconsin.
28 Included in this report (p. 451) is a table showing the distribution of types of civil cases in the Superior Court during the three month period examined. Personal injury constituted 6.9 percent, damages 2.5, divorce and annulment 29.7, debt-related (money, notes, liens, goods sold, foreclosure, unlawful detainer) 39.5, contract 1.6.
29 Recomputing the percentages, one can say that if the defendant was insured, compensation was obtained in 29.9 percent of cases compared to only 28.9 percent of cases when the defendant was not insured.
Divorce

A second specialized area that received significant attention from early empirically-oriented legal scholars was divorce. The only article from Warner’s study of litigation in California, other than the one discussed above on delay, dealt with divorce, and it signaled the themes that would be central in the other studies (Warner 1921). Major studies, each leading to a book-length monograph, were carried out in Maryland (Marshall, May, Marquard, and Reticker 1932) and Ohio (Marshall, May, Marquard, and Reticker 1933), both by the Institute of Law at Johns Hopkins University, and both part of larger studies of court processes in those two states. The California, Maryland, and Ohio studies all involved collection of information from court records; a fourth study, completed in Wisconsin (Feinsinger 1932), relied on a combination of court records and interviews with lawyers and judges. The two book-length monographs on Ohio and Maryland provided a wealth of detail about the cases, including length of marriages, presence and number of children, grounds for divorce, initiator of the proceedings, attorneys involved (particularly the concentration, or lack of concentration, of the work among a small group), financial relief sought and obtained (and by whom), duration of the process, etc. A finding central to all of the studies was that most divorces were effectively consensual even if the formal legal framework did not envision it as such. The studies typically discuss the need to confront the reality of the divorce process and adopt procedures consistent with that reality. While the nature of the divorce process from the perspective of the husband and wife were consistent across localities, the handling of the divorce does vary reflecting legal differences and differences in practice. The most prominent of these was whether there was a heavy concentration of the work among a small number of practitioners. In Baltimore, for example, 2,000 attorneys showed up in divorce cases, but a group of only 44 attorneys were involved in 60 percent of the cases, with one lawyer handling 188 cases and six lawyers or firms handling 526. In contrast, there was no such concentration in Cleveland, although there was concentration in Toledo and one other Ohio county (Marshall, May, Marquard, and Reticker 1933, 135).

Small Claims

One of the innovations of the early part of the 20th century in the United States was the creation of small claims courts with simplified procedures to allow unrepresented litigants to obtain redress in small damage cases. One such court was created in Hartford, Connecticut. Soon after its creation, it became the subject of a study by Charles Clark. With a collaborator, he

30 While the judicial statistics reports for the English judiciary includes information on divorce cases, there appears to be no research employing those statistics, or looking in more detail at divorce cases, perhaps because of the confidentiality surrounding court records in England where they are not public records in the same sense as in the United States. The Departmental Committee on the Social Services in Courts of Summary Jurisdiction (1936) did conduct a survey of on the matrimonial work of magistrates’ courts but this was in connection with the social services provided in connection with the matrimonial work of those courts, and contains nothing of particular interest.
assembled data on the court’s first 18 months of operation. This particular court’s jurisdiction was limited to matters involving a maximum of $100. During its first 19 months, it disposed of 5,236 cases, most of which appear to have involved debts (“work and labor”, “goods furnished”, “notes and loans”, and “rent” together comprise 92.4 percent of the cases); only 4.6 percent involve negligence claims, with the remaining 3 percent categorized as “miscellaneous.” There were judgments in 2,507 of the likely debt categories, all but 73 of which were for the plaintiff (and 1,918 of which were default judgments). Overall in negligence cases, there were 107 judgments; 74 for the plaintiff (18 by default) and 33 for the defendant. About two thirds of the awards were $35 or less (Clark and O’Connell 1929).

Prior to the creation of small claims courts, some of the lowest level courts such as “mayors’ courts” and justice of the peace courts handled small civil cases. One part of the Johns Hopkins Institute of Law study of judicial administration in Ohio was a study of these lower level courts with a particular focus on the courts in Hamilton County (Douglass 1932; Douglass 1933). The civil jurisdiction of these courts included original jurisdiction for cases involving up to $100 and jurisdiction concurrent with the common pleas courts in cases involving between $101 and $300. The study found that while the mayors legally had civil jurisdiction, few of the mayors in Hamilton County actually exercised it although statewide somewhere between a a third and a half did hear civil cases (Douglass 1933, 45). In contrast, justices of the peace in Hamilton County did handle civil cases. Data on almost 6,000 such cases handled between 1925 and 1929 showed that in the median case the plaintiff sought about $50, and only 22 percent sought $100 or more. Unfortunately, the information on types of cases is not very detailed: 58 percent were “ordinary civil suits”, 20 percent attachments, 7 percent forcible entry and detainer, 6 percent actions on notes, 5 percent actions of replevin, and 4 percent actions for damages; it is likely that the vast majority of these cases were debt-related. No information is provided on actual outcomes (Douglass 1932, 78-80). Twenty-five percent of the cases were disposed of without a judgment (14 percent settled and 11 percent “dismissed without prejudice”). Of the remaining 75 percent, almost 95 percent were in favor of the plaintiff, and over 60 percent of those by default (and another 15 percent “by confession”); that is, overall, it would appear that only about 21 percent of civil cases in the justice of the peace courts were resolved after a contested proceeding (p. 83). The pattern of disposition was similar for most types of cases; the notable exception was damage suits; omitting settlements and dismissals, for damage suits, the odds of odds of a plaintiff prevailing versus a defendant prevailing was about 2.5 to 1 compared to 15 to 1 or more for the other types of cases (p. 84). There was also a significant imbalance in representation with 66 percent of plaintiffs represented by counsel compared to only 20 percent of defendants; this probably reflected the debt collection use of the courts, with businesses hiring attorneys to handle to number of cases at a given court (p. 85). As is true in most debt collection proceedings, the court order was simply a step on the way; obtaining satisfaction (i.e., enforcement) was a significant problem for the users of the justice of the peace courts (p. 87).
Jurisdiction and Procedure

The debate over limiting or abolishing the diversity jurisdiction of U.S. courts has come and gone over the years. Two studies during this period sought to provide empirical information to inform this debate. As part of his study of federal cases for the Wickersham Commission, Charles Clark looked specifically at diversity cases, and produced a short article detailing what he found in terms of type of issue, frequency of trial, size of monetary awards, nature of the parties, and frequency of appeal (Clark 1933). A second study, by Hessel Yntema and George Jaffin (Yntema and Jaffin 1931; see also Yntema 1933), examined caseload trends using material from the Attorney General’s annual reports, and an analysis based on cases in ten volumes of the Federal Reporter; this latter component appears to be the first use of reported cases for lower courts (trial and/or intermediate appellate) as a vehicle for empirical research. The analysis showed that private cases based solely on diversity constituted 11 percent of all cases in the reporter, 14 percent of all civil cases, and 31 percent of all private civil cases.

Summary judgment has been an issue in several contemporary empirical studies. In 1933, Leonard Saxe published a study of summary judgment in New York. At that time the summary judgment procedure was restricted to contract cases, and was primarily a tool used by the plaintiff when there was no real defense to a contract claim over a debt. Based on his analysis showing the effectiveness of summary judgment to discourage litigation and dispose of cases involving “sham defenses” thus saving an estimated 13,300 hours of court time, the author lauded the procedure “for its marked contribution to the cause of speedy justice and the alleviation of the economic waste of unnecessary and protracted litigation” (Saxe 1934, 254).

A study related to the rules of evidence was perhaps the first quasi-experimental study in the empirical legal studies literature. Steuart Britt selected 154 rules of evidence from Wigmore, put each on a separate slip, and mailed sets of slips to members of four groups: professors of evidence (n=34), other law professors (n=34), practicing lawyers (n=34), and psychologists (n=34). Each was asked to rate each rule on a scale: +2 good, +1 probably good, 0 doubtful whether good or bad, -1 probably bad, -2 bad. Britts analysis was one of the first to utilize what today are standard statistical tools including correlations and what amounted to a t-test for comparing means. He found fairly high within groups consistency, with the highest for law professors who taught evidence. He found a number of significant differences among groups, particularly between psychologists and evidence experts. Generally, however, he found that all groups judged the rules as fairly good (Britt 1940).31

31 Probably the first field experiment where conditions were manipulated and responses observed was a study of compliance with parking and traffic regulations done by Underhill Moore and several collaborators (described in Schlegel, 128-141). These experiments, which involved the manipulation of signage and regulations, did not involve random assignment or formal controls, but were reasonably good experimental designs by the standards of the time. A variety of factors, including limited funding, changing collaborators, difficulty in the analysis, the complexities of arriving at a suitable theoretical framing, other scholarly projects and opportunities, and the start of the Second World War, delayed the publication of the results of the research until 1943 (Moore and Callahan 1943).
Civil Juries

One theme in recent empirical work has been the “vanishing trial,” particularly in civil cases. In the 1920s and 1930s, one finds discussion of the vanishing jury, both in the United States and in England. Clark and Shulman reported that between 1919 and 1932 only four percent of civil cases in the Superior Court of New Haven was a jury trial even begun (Clark and Shulman 1934, 869); roughly an equal proportion was tried to the court (p. 872). Silas Harris showed some pattern of decline in jury requests in Connecticut in the wake of the imposition of a fee requesting a jury (Harris 1930). Robert Jackson tracked the drop in civil jury trials in England in the wake of changes in 1918 and 1925 that allowed most such jury trials solely at the discretion of the judge (Jackson 1937).

A contemporary issue about the civil jury in the United States, such juries have all but disappeared in England and are increasingly uncommon in other common law jurisdictions, is how the jury decides cases in contrast to how judges do (see Clermont and Eisenberg 1992; Eisenberg, Hannaford-Agor, Heise, LaFountain, Munsterman, Ostrom, and Wells 2006; Eisenberg, Laffountain, Ostrom, Rottman, and Wells 2002). This question was addressed by research 75 years ago. Clark and Shulman (1934, 883) found little difference in the likelihood that a jury would find for the plaintiff in a negligence case compared to the likelihood that a judge would find for the plaintiff (61.6 percent for juries, 61.6 percent for judges). Interestingly, that were some difference in breach of contract and debt cases, with the court more likely to find for plaintiff (75.7 percent) than a jury (67.5 percent). Carpenter looked at this in several ways. He found that in two of four jurisdictions with sufficient bench trials for comparison, jury trials yielded higher average awards; however he also found that bench trials were more likely than jury trials to yield plaintiffs’ verdicts in two of seven jurisdictions. He also compared appeal and reversal depending on whether there had been a jury or a bench trial and found no difference after controlling for type of case (Carpenter 1929). In their seminal study of the American jury, Kalven and Zeisel briefly report two “studies” by judges comparing their views with jury decisions in civil cases. Judge Philip J. McCook of the New York Supreme Court, tracked this for 114 cases over 10 trial terms for the cases in his courtroom. According to Kalven and

32 A majority (55.1 percent) of those trials were in automobile accident cases, and another quarter were in other types of negligence cases; thus about four-fifths (80.7 percent) were in negligence cases of all types.

33 Sixty-three percent of negligence cases were tried to juries compared to only 16.3 percent of debt and contract cases and 27.8 percent of other types of cases.

34 Outside the United States, civil juries are used with some frequency only the Australian state of Victoria and the Canadian provinces of Ontario and British Columbia (see Vidmar 2010).

35 The authors do not perform any statistical tests; I did a test of significance and determined that the difference was not statistically significant.
Zeisel,\textsuperscript{36} he reports that he disagreed with the jury’s verdict in 28 percent of cases, 23 percent in which he would have found for the defendant while the jury found for the plaintiff and 5 percent in which he would have found for the plaintiff while the jury found for the defendant. He also observed that in 32 of the cases where he agreed with the jury in finding for the plaintiff, he disagreed with the amount of damages awarded, and in two thirds of those cases he would have awarded more. Kalven and Zeisel also quote from the memoirs of a judge who sat during this time period, “… by actual count [civil] jury verdicts are practically identical with the verdicts I should reach in over seventy-five percent of the cases I hear” (Kalven and Zeisel 1971, 521-23).

**DEBT AND BANKRUPTCY**

It should not be surprising that there was a flurry of research related to bankruptcy and debt in the late 1920s and early 1930s. Some well known names of American legal history authored some of this research: William Douglas as a law professor at Yale and Abe Fortas as a Yale law student.

Douglas headed up a study based at Yale Law School. The project, initiated soon after Douglas arrived at Yale in 1928, was initially planned to focus on “business failures” (Clark, Douglas, and Thomas 1930; Douglas and Thomas 1931). For various reasons, the project evolved into a study of bankruptcy in general (see Schlegel 1995, 98-105). The bankruptcy study produced a flurry of articles, mostly based on data from New Jersey and Boston.\textsuperscript{37} One of the themes noted in this work was the excess use of credit by both individuals and by businesses, and a frequent problem of poor record keeping by businesses that got into financial trouble (Douglas 1932; Douglas and Marshall 1932). Several aspects of the project focused specifically on individual (“wage earner”) bankruptcies and debt issues. As reported by Douglas, the individuals who found themselves in bankruptcy proceedings often had “abused” consumer credit opportunities. Other important factors included medical bills, unemployment, and tort judgments against the debtor. Douglas also notes that some bankrupts had gotten into trouble through real estate speculation (Douglas 1933).

Douglas’s observations built very heavily on studies by Fortas (1933) in Chicago and by Peter Nehemkis (1933) in Boston. Fortas looked at wage assignments (i.e., garnishment);\textsuperscript{38} he collected data (658 cases) from companies that cashed pay checks and that withheld money to

\textsuperscript{36} My effort to track down the originals of the early studies to which Kalven and Zeisel refer was unsuccessful; the reports of those studies appeared in a legal periodical, but the citation provided by Kalven and Zeisel was not sufficient to access the reports.

\textsuperscript{37} Another article from the project examined aggregate on the number of bankruptcies by state for 1921 and 1931, years chosen to allow for comparisons to census data. Among other things, the article shows which states had the highest and lowest numbers of wage earner bankruptcies (see Sturges and Cooper 1933).

\textsuperscript{38} Wage assignment loans, which first evolved in the 1880s, were a common practice in the 1920s (see Robinson and Nugent 1935, 43-45).
pay the check-cashers creditors, and provided tabular summaries showing debt obligations, by
incomes, family size, number of creditors, and types of creditors. Drawing on his analysis, he
expressed concern that there had been an over expansion of consumer credit, and the response of
employers whose employees had been subject to garnishment (e.g., dismissals and refusals to
honor wage assignments). Fortas also looked at the 432 case files from the Legal Aid Bureau of
Chicago of clients who sought the Bureau’s assistance in connection with wage assignments. He
found that in most of the cases the creditors were installment sellers, and that the Bureau’s
investigations had concluded that 42 percent of the installment contracts were legally
unenforceable. Finally, Fortas carried out three case studies of employers to see how they dealt
with wage assignments, including the nature of the assignments (most involved installment
sellers—i.e., consumer credit) and the employers’ policies, particularly with regard to dismissal,
 i.e., were employees subject to dismissal if the number of wage assignments exceeded some limit
(Fortas 1933).

Peter Nehemkis’s study (1933) focused on the Boston debtor court. He collected data on
4,866 cases closed between January 1, 1931 and September 15, 1932. He found that the average
debt was $261, with about a quarter under $50 and the median right around $100; only 4 percent
were $1,000 or more. The majority (56 percent) of creditors were retailers, with about 10 percent
each being landlords and individuals who had made personal loans; only six percent were loan
companies and credit unions. In almost two thirds (64 percent) of the cases where summonses
were issued, the debtor failed to respond. Debtors were routinely ordered by the court to pay the
debt, and were subject to being found in contempt of court if they failed to pay. In fact, bench
warrants (a “capias”) was issued in 1,831 cases, although only 17 debtors were actually jailed for
contempt during the period studied; typically the warrants was used as leverage to extract at least
some payment from the debtor. Overall collection was successful in only 22 percent of cases,
largely because many debtors simply could not pay what they owed, and there was no procedure
for ascertaining ability to pay.

As part of the study of administration of justice in Ohio undertaken by the Johns Hopkins
Institute of Law, Thomas Billig (Billig 1932) produced a monograph on equity receiverships
(i.e., the administration of insolvent estates) in Franklin County. Equity receiverships involved a
simpler process than bankruptcy, and had less detailed requirements (i.e., it simply required that
the debtor could not pay his or her debts as they fell due). Billig assembled data on 1,054
receivership cases in 1927 and 1928. His analysis looked at the types of businesses, nature of the
debts and assets, the time taken for the process, and the costs involved. He found that few
debtors filed responses to the initiation of receiverships, and that the “outstanding feature of the
Columbus receivership is the ease with which it may be instituted and operated, rather than any
meticulous devotion to theory” (p. 132). Given the failure of debtors to respond, it is not

---

39 Fortas references a study of wage assignments in New York conducted by the Russell Sage Foundation, although
he does not provide a specific citation; I believe that this study became the basis of a book by Louis Robinson and
Rolf Nugent (1935) on small loan practices.
surprising that Billig found that the receivership process moved quite quickly (p. 134). Perhaps most troubling was the high cost of administration in equity receiverships; administration consumed about 40 percent or more of assets for estates with assets of less than $7,000, and something less than 30 percent for estates between $8,700 and $18,000. In contrast a sample of 648 bankruptcies closed in the twelve months ending June 30, 1927 consumed only 18 percent of the total assets distributed (pp. 138-140).40

The second government report focused on a study of bankruptcies in the southern district of New York directed by Judge Thomas D. Thacher, who was an advocate of bankruptcy reform (Thacher 1931). Thacher’s study employed a variety of data collection strategies and sources:

- Statistical forms completed by 200 bankruptcy referees covering all cases closed during a 10 week period in 1930 (n=3,771)
- 1,000 replies (20 percent response rate) to questionnaires sent to business houses and banks
- Data on 1,000 bankrupts (“case histories”) out of a sample of 3,700 with information actually collected by “local associations”
- 47 replies (of 75 sent) from questionnaires sent to the presidents of credit men’s associations
- Observation of bankruptcy hearings and interviews with judges and others completed during a two month field trip to 21 cities plus conversations at conventions of key player
- 50 responses (300 sent out) to attorneys; 4 identified in each of the 78 largest cities.
- Questionnaires sent to adjustment bureaus (sent to about 175)

Thus, while Thacher’s focus was on New York, his data were not limited to that location. The reported analyses covered a range of topics including distributions of assets and debts, costs of administration, occupations of bankrupts, duration of proceedings, services provided by attorneys in bankruptcy cases, data on trustees and referees (including fees received), a separate analysis of wage earner bankruptcies and assets in small bankruptcies (Thacher 1932).

Two government reports contain studies of bankruptcy processes. One, which appears in a Congressional report, seems to be what some refer to as the “Donovan Report” (Committee Report 1931). Much of this report is a distillation of data from previous Attorney General annual reports showing a variety of characteristics of bankruptcy proceedings, with a focus on volume of proceedings, duration of proceedings (including the time required for various stages), size of estates, and the expenses of the proceedings (ranging from 25 percent of assets to 52 percent of assets).41

---

40 Billig was not the first to raise concerns about the costs associated with equity receiverships. Apparently the New York World conducted a survey of equity receiverships in the federal court for southern New York sometime around 1925 and found that the administrative costs were very high. As described in an article that mentioned the study (I was not able to obtain a copy of the study itself), “If the machinergy of the bankruptcy court is set in motion the ‘legal expenses’ of liquidation are notorious” (see Sturges 1925, 490).
assets). The report provided somewhat more detailed looks at subsets of districts (e.g., those with the lowest expenses, those with large cities, the largest districts). In addition to the data drawn from the Attorney General’s reports, the author of the study collected more detailed data on 350 bankruptcies (liquidations) in the southern district of New York; for these cases, there is a table laying out expenses by categories.

**BANKING**

As part of the legal realist empirical research enterprise at the Yale Law School, Underhill Moore undertook an extensive study of banking practices as related to how transactions of various types were handled and how closely those practices tracked the governing legal statutes and legal decisions. In the words of one contemporary commentator, “Moore’s working hypothesis is that current decisions are more closely related to existing patterns of relevant behavior in the jurisdiction where the case originates than to any comparability with past decisions” (Thomas 1931, 219). a question that presages contemporary research on local legal culture (Moore and Sussman 1931a; Moore and Sussman 1931b; Moore and Sussman 1931c; Moore and Sussman 1931d; Moore and Sussman 1931e; Moore, Sussman, and Brand 1933; Moore, Sussman, and Corstvet 1933a; Moore, Sussman, and Corstvet 1933b). Moore (and his collaborators) studied at least two banking practices (handling orders to stop payment on checks and debiting of direct discounts) in a variety of jurisdictions (see Schlegel, 1995, 115-128, for a good description of the studies). The resulting articles are difficult to digest. One contemporary described the research as “very obscure in both writing and method” (Schlegel 1995, 118, quoting and paraphrasing Dorothy Swaine Thomas). Schlegel (1995, 119) describes them as marked by “convoluted, opaque prose dictated by the convoluted, opaque method” employed in the research.

**APPELLATE COURTS**

Perhaps the single best known empirical study of the pre-World War II period is Frankfurter and Landis’s *The Business of the Supreme Court* (1928). The focus of this book is on how the jurisdiction of the Supreme Court evolved and how the caseload, both volume and content, shifted along with the jurisdictional changes and as both the law and the country as a whole evolved. For example, one finds a discussion of the beginnings of the discretionary docket, and how both the number of petitions and the number of cases heard on this docket doubled between 1916 and 1925 (p. 295). The authors analyze the content of the docket by taking snapshots in 1825, 1875, and 1925, showing how between 1875 and 1925, cases dealing with the Commerce

---

41 Much, perhaps most, of the material in the book previously appeared as a series of eight articles in *Harvard Law Review* in 1925 and 1926.
Clause, due process and regulation of economic enterprises, and taxation increased while categories showing decreases included common law topics dropped (sharply from 81 cases to 11 cases), cases regarding land legislation, and suits by states (p.302). While the focus is largely on the Supreme Court, the authors also consider at places other appellate courts. They show how the volume of cases heard by the federal Court of Appeals increased over time from their creation in 1891 (p. 308). They also compare the caseloads of selected state supreme courts to that of the British House of Lords and Privy Council (pp. 303-305).

Frankfurter’s empirical work on the Supreme Court did not end with the publication of The Business of the Supreme Court. Rather, in 1929, one finds the first of what became annual statistical reports of the work of the Supreme Court that have continued to this day to appear in annually in the Harvard Law Review. Frankfurter coauthored these reports (with varying coauthors) until his own appointment to the Supreme Court. The versions published through 1940 focused on sources of cases, disposition, subject matter of both petitions for certiorari and decisions on the merits; there was no consideration of decisional patterns of individual judges or interrelationships of such decisions which we find in today’s versions of these reports. Perhaps the most interesting pattern shown is the evolution of the role of mandatory versus discretionary jurisdiction in the wake of the 1925 Judges’ Bill. From the series of reports, one sees that mandatory review cases started out as a bit over 80 percent of decisions on the merits before 1925, declined rapidly to 50 percent, in 1929, rebounded to 62 percent and then declined back to 50 percent by the end of the period (Frankfurter and Hart 1934, 243; Frankfurter and Landis 1929, 39; Hart 1940a, 588). Meanwhile, the number of petitions for certiorari stabilized at 800 to 850 by 1934 (Frankfurter & Landis 1929, 45; Frankfurter & Hart, 1934, 248; Hart 1940, 590).42

Empirical research on appellate courts during this period was by no means limited to the federal courts. The annual statistical reports concerning the work of the Supreme Court in the Harvard Law Review inspired similar reports for at least one state. Starting in 1933 (Campbell 1933), and continuing until the late 1940s, the Wisconsin Law Review published a brief statistical report on the work of the Wisconsin Supreme Court covering dispositions, topics covered, who on the Court wrote opinions on which topics, and the frequency of dissent by individual justices.

Probably the most extensive empirical study of state appellate courts was carried out as part of the Johns Hopkins Institute of Law study of judicial administration in Ohio (Harris 1933). This study considered appeals at all levels: the Ohio Supreme Court, the Courts of Appeals, and the appellate business of the courts of common pleas. The study covered a limited period, six years for the supreme court, and only six months for the courts of appeals and the courts of common pleas. The study looked at the basis of jurisdiction, the duration of the appellate process, and disposition (dismissed, overruled, or allowed), with perhaps the most attention on

42 While the first study looking explicitly at the behavior of individual Supreme Court justices did not appear until 1941 (Pritchett 1941), political scientists were thinking about such research as early as the 1920s (see Haines 1922, 96n1; Hall 1926).
the duration. The study includes an interesting table comparing the workloads of 22 state supreme courts and the U.S. Supreme Court (p. 56).  

There were at three studies of the disposition of criminal appeals by state supreme courts, in California (Vernier and Selig 1928), Illinois (Hamo 1929), and Missouri (Moley and Thomas Hennings et al. 1926) (the latter two appeared as parts of the crime surveys for those two states). The California study assembled data for the period 1850 to 1926, and showed that the rate of reversal had declined over that period (from about 50 percent to around 15 percent) at the same time that the number of such appeals was growing sharply (from about 100 in the first decade to over a thousand in the last seven years considered). The decline in reversal rates was particularly sharp in the last 16 years considered, a pattern the authors attribute to a constitutional change in 1911 limiting reversals to cases where “error in light of all the evidence resulted in a miscarriage of justice” (p. 42). The authors also present an analysis of the basis of the reversal. The analyses in the crime surveys focus shorter periods (10 years for both Illinois and Missouri) and document the types of crimes involved and reversal rates and patterns.  

Perhaps the broadest study of a state appellate court was an analysis of five years of the Maryland Court of Appeals (Brune and Strahorn 1940). This study showed that only five percent of the issues on that court’s docket was criminal compared to almost 15 percent negligence and 9 percent estates. Importantly, the study distinguished between issues and cases, recognizing that a case could raise multiple types of issues. Focusing on cases rather than issues, 42 percent of the cases arose through equity, 39 percent civil, 8 percent administrative, and 6 percent criminal. About 4 percent of the cases resulted in reversals, and plaintiffs and defendants were equally successful as appellants. Other analyses showed the number of judges sitting for cases, the dissent and concurrence behavior of individual justices, and the number and length of opinions written by each justice.  

Political scientist Rodney Mott (1936) anticipated a variety of modern studies of state supreme courts in a study of “judicial influence.” His study employed a survey of law professors to rate the reputation of the supreme courts of the then 48 states, an analysis of case books to determine the number of opinions of each state’s supreme court appearing in such books, and a count of citations to opinions of each state’s court by other state supreme courts and by the United States Supreme Court. Mott combined his various measures to produce an overall prestige ranking. That ranking correlates .52 with a set of prestige rankings created about 50 years later by political scientist Gregory Caldeira (1983).

---

43 This table is attributed to a study done by Rodney Mott, but no citation is given.
44 The Illinois study (Hamo 1929, 115, 181) also includes a table comparing reversal rates in criminal and civil cases for a one year period (46 percent criminal and 39 percent civil) and a table showing the ultimate disposition of cases reversed and remanded (only about 16 percent were ultimately convicted or pleaded guilty.)
LEGAL NEEDS

Yet another contemporary topic that one finds addressed in the early empirical legal work is that of legal needs, particularly the legal needs of those with low incomes. Perhaps the preeminent work on this topic was Reginald Heber Smith’s *Justice and the Poor* published in 1919 by the Carnegie Endowment for the Advancement of teaching. Smith drew in part on Wood’s study of the Los Angeles public defender (Wood 1916) and summarized Wood’s statistics on the criminal cases in Los Angeles. He also looked at the work of legal aid organizations including the number of such organizations, their budgets, their source of clients, and their total caseloads. For aid organizations in six cities, he reported the disposition of over 8,000 cases. Interestingly, Wood himself reported on civil legal aid; it is not clear whether these cases were handled by the Public Defender as part of that organizations mission or some other organization. In 1914 there were 7,872 applicants for aid in Los Angeles County, 55 percent of whom sought only advice (the majority of these were women in domestic trouble such as desertion by their husbands). Thirty percent were claims that the office accepted for “adjustment,” two thirds of which were adjusted out of court. Of the claims accepted for adjustment, 70 percent were wage claims (i.e., nonpayment of wages). Few of the one third of claims not adjusted out of court resulted in law suits, often because the claimant could not pay filing and court fees. Only 115 suits were filed (76 of which were wage claims); six were cases in which office defended. In addition, 1,093 cases were referred to members of the bar (Wood 1916, 232-35).

The first study that resembles the modern “legal needs” or “justiciable problems” survey was conducted by Charles Clark and social scientist Emma Corstvet (1938). Using support from the Federal Emergency Relief Administration, they hired young unemployed lawyers and professionals and students in need of financial assistance to conduct interviews. Interviews were completed with 412 residents and 61 businesses. Among the residents, 385 legal needs, to use the modern term, were identified, split fairly evenly between “preventative” and contentious matters. The authors looked at the likelihood of consulting an “advisor” (not necessarily a lawyer, although only 14 percent of those seeking advice consulted someone who was not a lawyer) by problem type. Looking only at those needs occurring at least 10 times, the percentage consulting such an advisor varied from under 10 percent to 70 percent or more; such advice was more often sought for contentious matters (45 percent) than for transactional matters (19 percent). The authors sought to compare the use of advisors by those with incomes less than or equal to $2,000 to those with higher incomes; they found a slightly higher use by the higher income group for contentious matters, but for transactional matters, it was the lower income group that high a higher probability of seeking advice; however, the samples were small, and the differences not large which led the authors to caution drawing conclusions from their analysis. The authors also report on the interviews with businesses which identified 242 legal needs, three quarters of which were transactional (“preventative”). The businesses were slightly more likely to consult advisers for contentious matters than were individuals (52 percent versus 45 percent), but they
were considerably less likely to seek such consultations in transactional matters (6 percent versus 19 percent). For both individuals and businesses, respondents were asked their level of satisfaction with the resolution of contentious matters; both groups were more satisfied when an advisor was consulted, and businesses were more satisfied than individuals both when an advisor was consulted and when an advisor was not consulted.

LEGAL PROFESSION

The first surveys of lawyers appear in the 1930s. The Clark and Corstvet study discussed in the previous section included a small scale survey of lawyers (n=50). From these respondents the interviewers obtain information on income, structure of practice, fees charged, unpaid work, sources of cases, and types of matters handled. The latter was obtained by asking for detail about the last five items completed by the lawyers, with the lawyers encouraged to consult their diaries to insure accuracy. The lawyers typically worked in an office sharing arrangement or entirely alone. The four most common matters completed by the lawyers were torts, property, estates, and other contracts. Most (about 80 percent) had some unpaid work, although not all described this as having been done as free legal work; half of those who did some unpaid work did five or fewer hours per week. The lawyers were asked how the matters completed had originally come to them; 39 percent involved existing or former clients, 15 percent were existing friends for whom they had not previously worked, 40 percent came through referrals (“recommendations”) from clients, other attorneys, friends, with the balance coming through social contact or local reputation.

Today it is common for state bars in the various U.S. states to do periodic surveys of their members, often focused on the economic issues facing legal practitioners.45 The first such survey I found was done by the New York County Lawyers’ Association in 1934 (Committee on Professional Economics of the New York County Lawyers' Association 1936). Questionnaires were sent to 19,000 persons, with 5,000 responses received from what was estimated post hoc to be a population of about 15,000. Using these data a report was prepared profiling the local bar along a variety of dimensions, gender (2 percent women), years in practice (58 percent 10 years or less) income (median $2,990), practice organization (38 percent solo, 41 percent firm or equivalent, 21 percent employed by a firm), type of legal education (55 percent full time law school), prelaw education (53 percent college graduates), overhead (average of about 35 percent), type of clientele, race (one half of one percent “Negro”), and nativity (81 percent native born). In addition a variety of tables examined variation in income by training, length of time in

---

45 Such studies are also regularly done by lawyers’ organizations in other countries. I am aware of such a survey done by the English Law Society in the late 1930s; however, the report of that survey, which I am told is entitled “National Service: Report of the Special Committee,” was never made public. From the title, I infer that at least part of the purpose of the survey, done soon before the outbreak of World War II, was to assess the potential contribution of solicitors to the war effort that was thought to be imminent.
practice, and the like. Also included was a discussion of the change in lawyers’ incomes over
time (the median dropped about 25 percent after the stock market crash). The article concluded
with a call to reduce the number of lawyers in practice, preferably by getting rid of “ambulance
chasers” and a discussion of how income of lawyers might be increased.

Lloyd Garrison conducted a study of the Wisconsin bar with support the Civil Works
Administration (CWA), a depression relief organization (Garrison 1935). While the article is
described as a “survey of the Wisconsin bar,” no survey of lawyers was carried out. Instead
Garrison dispatched his assistants to every courthouse in Wisconsin to come up with figures
showing the numbers of lawyers practicing in each county in each year since the state’s founding
in 1848. A central question he sought to answer was whether the bar was “overcrowded.” In
addition, the research staff sought to obtain a range of indicators in each county of the amount of
level of legal activity or legal work in each county. Somehow, Garrison was also able to access
income tax information for individuals, and to match this up with performance in law school (for
University of Wisconsin Law School graduates). Using this wealth of data, Garrison reaches a
variety of conclusions: lawyer’s incomes had dropped but physicians incomes had dropped more
(and generally lawyers had higher incomes than physicians), Milwaukee had the highest
percentages of lawyers with very high incomes and very low incomes, incomes were highest for
lawyers during middle age and late middle age, and there was a tendency for those with stronger
law school records to have higher incomes although that did not mean that someone with a good
record will have a high income or that someone with a weak record could not still achieve a high
income.

Legal Education

Related to research on the legal profession was a body of work on legal education.46 Much
of this work focused on methods of selecting students for admission; that is, what are the best
predictors of success in law school (Eagleton 1932a; Eagleton 1932b; Murray 1938), and
whether additional predictors, such as general or specialized tests, could be devised? Central to
much of this literature was whether legal aptitude tests, the forerunners to today’s Law School
Aptitude Test (LSAT) would be better predictors than other information available. At least three,
and probably more, such tests were developed, the Yale Legal Aptitude Test, the California
Legal Aptitude Test, and Ferson and Stoddard Law Aptitude Examination (Crawford 1932;
Crawford and Gorham 1940; Gaudet and Marryott 1930; Stoddard 1927; Wigmore 1929a;

46 In discussing research on legal education during this period, one must take note of Alfred Reed’s 1921 report for
the Carnegie Foundation for the Advancement of Teaching, Training for the Public Profession of Law. This report
contained a wealth of information about the history, development, and current state of legal education in the United
States. It included an appendix with a number of statistical tables compiled from various sources on number of
lawyers, number of law schools, numbers of members of other professions, numbers of law graduates compared to
numbers of graduates of other types of professional schools, characteristics of entrance requirements, enrollments at
law schools, all for multiple time points, some dating back to the early 19th century.
Articles in various law reviews attempted to show that the tests either did, or did not, improve upon other predictors. Other studies focused on the predictive quality of traditional indicators such as number of years of undergraduate education and undergraduate grade point averages (Dickinson 1940; Dickinson 1939; Husband 1939). Some of these studies employed what are now standard statistical tools such as correlation and regression.

A second issue concerning legal education that appeared in the empirical literature focused on grading. One study, by John Grant (Grant 1929), focuses on grade distributions in law school courses, and showed that grade distributions differed substantially by course. While the study did not provide information on variation by instructors, the author argued that the differences he found reflected differing standards applied by individual instructors. He speculated that students chose courses with an eye to grading standards. The article concluded with a position that contemporary law faculty will find very familiar: law faculties should adopt a standard grade distribution to insure comparability across courses and that distribution should approximate a normal distribution (7.4% As, 29.8% Bs, 39.7% Cs, 15.4% Ds, 7.7% Fs). A study of grading at Columbia Law School (Wood 1924; Wood 1925; Wood 1927) showed that the correlation of the marking of exams by two professors was usually in the range of .50, and that the average correlation of grades across courses was only .50. The study also estimated the reliability of grades using a standard measure (Spearman-Brown) and found it to be .70, which were consistent with the reliability found for essay exam grading in other fields. Interestingly, the author was able to conduct a small experiment where exams included a combination of true-false questions and essay questions; the correlation between the two parts ranged from .39 to .57; the average reliability for the true-false was .66 compared to .54 for the essay. Not surprisingly, the author of the study recommends that large classes (60 or more students) use exams that combine true-false with traditional essay questions.47

JUDICIAL STAFFING AND SELECTION

A contemporary controversy in the United States concerns methods of selecting judges. This is by no means a new issue; it has been a point of debate periodically throughout American history (Armitage 2002, 628-33). Interestingly, the first empirical study that relates to the judicial personnel was set not the United States but in England. Harold Laski (1926) examined the backgrounds of English judges appointed between 1932 and 1906, a total of 139 judges.48 He found that 80 were members of Parliament at the time of appointment and 11 others had been

---

47 In the 1925 article, Wood reports that the corrected split-half correlation for the true-false questions is .81, and provides a variety of statistics in support of his contention that true-false tests are better than essay tests. In the 1927 article, Wood provides additional statistics such as reliability figures for exams in various specific courses; he also provides some student opinion vis-à-vis “objective” exams, although these opinions are not from law students.

48 This article was reprinted as Chapter VII of Laski’s 1932 book, Studies in Law and Politics.
candidates for Parliament; 66 of those appointed from Parliament were members of the party in power at the time of their appointment. Thirty-six of those who had served in Parliament had held position as Attorney General or Solicitor General, and every Chief Justice save one in the preceding 60 years had previously served as Attorney General. Eight of the judges appointed from Parliament became Lord Chancellor.

The first U.S study, by Rodney Mott, Spencer Albright, and Helen R. Semmerling (1933), focused on the characteristics and reputations of state supreme court judges and federal trial and appellate judges. The authors assembled a variety of characteristics for over 1,000 such judges, including demographics (nature of education, marital status, age at attainment of office, number of children, religion, wealth, military service), years of judicial service, organizational memberships, political experience, and a published rating of legal ability. Sixteen pieces of information were combined to create a personnel index for each state, and the state’s were ranked with Massachusetts highest and Kentucky lowest. They also conducted a survey of law professors asking them to rate the prestige of each state’s supreme court, and found that the average ratings correlated .72 with their personnel index. The also correlated the personnel index with measures of frequency of citation by other state supreme court and by the U.S. Supreme Court, and found correlations of .49 and .55. Regarding the question of impact of selection system, they found what they described as a slightly higher average personnel index for states with justices appointed by the governor compared to states using elections, either popular or legislative (92 vs. 91); however, the difference was so slight that it is likely that if the authors had conducted what today are standard tests of statistical significance, they would have concluded that no difference was statistically discernible.

The second study focused on judicial selection in Chicago. Edward Martin (1936a; 1936b) looked at judicial elections in Chicago, with a particular focus on what impact the bar had on outcomes of those elections. Part of Martin’s analysis was an assessment of the candidate factors (e.g., age, education, experience, professional affiliations, political activity, length of service on the bench, evaluation of the Bar Association’s Committee on Candidates, etc.) that correlated with both the evaluation of judicial candidates through a poll of members of the bar and the votes the candidates ultimately received in primaries and general elections. Generally the correlations were lower for the electorate than for the bar poll; they were most similar for length of service on the bench and most disparate for the rating by the bar’s Committee on Candidates (Martin 1936a, 133). Bar-endorsed candidates were more successful for the Circuit and Superior Courts (81 percent elected) than for the Municipal Court (60 percent elected). The candidates endorsed by the bar association were more likely to be Republicans than Democrats, although over time it seemed to become more balanced in its endorsements (id., p. 150-51). No “Negro” candidates obtained the Bar’s endorsement (p. 155). On the general question of the bar’s influence, Martin concluded that the bar influenced a minority of voters, at most 30 percent and often much less. The amount of influence the bar had varied among voting constituencies depending significantly on the influence of the party organizations, and possibly the level of education.
CONCLUSION

The purpose of this review is to bring to the attention of contemporary empirical legal studies scholars the work of the true pioneers in this field. The early work is both rich in detail and somewhat limited in breadth. Modern ELS scholars pursue research about a much wider range of topics, albeit some of that increased breadth reflects the rise of the modern administrative state and the growth of law and regulation needed to implement government activities in a diverse set of areas.

Perhaps the most interesting aspect of the review presented above is that continuity of research findings. Many of the contemporary themes “discovered” by contemporary empirically-oriented scholars were documented eighty or more years ago. There is a potentially interesting question for sociologists of knowledge to be found here: why do scholars, who are supposed to be steeped in the research in their fields, not know of the work of preceding generations? Or, put more crudely, why do contemporary scholars keep reinventing the wheel? That of course raises the question of whether knowledge gleaned in recent years will be remembered 50 or 100 years from now, or will some future generation of scholars “discover” anew many of the interesting findings of today (and yesterday)?

REFERENCES

Early Empirical Research


Billig, Thomas Clifford (1932) Equity Receiverships in the Common Pleas Court of Franklin County, Ohio in the Years 1927 and 1928. Baltimore: Johns Hopkins Press.


Brune, Herbert M., Jr. and John S. Strahorn, Jr. (1940) “The Court of Appeals of Maryland, a Five Year Case Study.” 4 Maryland Law Review 343-389.

Burgess, Kenneth Farwell (1914) “A Survey of the Ninth Judicial District of Wisconsin.” No. 634 Bulletin of the University of Wisconsin


Dickinson, Edmund C. (1940) “Predictability of Success in Law School.” 46 West Virginia Law Quarterly 105-121.


Harris, Silas A. (1933) *Appellate Courts and Appellate Procedure in Ohio*. Baltimore: Johns Hopkins Press.


Rundell, Oliver (1912) “The Time Element in Criminal Prosecutions in Wisconsin.” No. 512 Bulletin of the University of Wisconsin


Other References


Kort, Fred (1957) “Predicting Supreme Court Cases Mathematically: A Quantitative Analysis of the 'Right to Counsel' Cases.” 51 American Political Science Review


Snyder, Eloise (1958) “The Supreme Court as a Small Group.” 36 *Social Forces* 232-238.


