IRis

Northeastern University

School of Law Faculty Publications

School of Law

October 01, 2010

Liberation reconsidered: understanding why judges and juries disagree about guilt

Amy S. Farrell
Northeastern University

Daniel Givelber
Northeastern University School of Law

Recommended Citation

Farrell, Amy S. and Givelber, Daniel, "Liberation reconsidered: understanding why judges and juries disagree about guilt" (2010). *School of Law Faculty Publications.* Paper 27. http://hdl.handle.net/2047/d20001194

This work is available open access, hosted by Northeastern University.

0091-4169/10/10004-1549
THE JOURNAL OF CRIMINAL LAW & CRIMINOLOGY Vol. 100, No. 4
Copyright © 2010 by Northwestern University School of Law Printed in U.S.A.

CRIMINOLOGY

LIBERATION RECONSIDERED: UNDERSTANDING WHY JUDGES AND JURIES DISAGREE ABOUT GUILT

AMY FARRELL[*] & DANIEL GIVELBER[**].

A criminal defendant's constitutional right to a jury trial is premised in part on the view that a jury's verdict of guilt or innocence may differ from that of a judge deciding the same case.

Empirical research has confirmed that judges and juries do sometimes disagree about verdicts and that the direction of these disagreements is overwhelmingly in the direction of jury leniency. In their seminal study, The American Jury, *Harry Kalven and Hans Zeisel suggested that when cases are close on the evidence, juries are "liberated" from the dictates of the law, and can—and do—give expression to extralegal values in arriving at verdicts. This explanation feeds the commonly held view that judges decide according to legal rules, but juries make decisions that reflect the values and sentiment of the community, even when those decisions are in opposition to the law. This perspective has been supported by research primarily based on the perceptions of judges about how juries reach their verdicts. Missing from our understanding of why judges and juries disagree is information from jurors about the factors that motivate their*

[*] Amy Farrell, Ph.D., Law, Policy, and Society, Northeastern University, 2001; M.A., Sociology, University of Delaware, 1997; B.A., Beloit College, 1995. Assistant Professor, School of Criminology and Criminal Justice Northeastern University. Professor Farrell's research focuses on the administration and legitimacy of justice institutions with primary emphasis on measuring the effect of race and gender in police, prosecution. and adjudication practices.

[**] Daniel Givelber, J.D., Harvard Law School, 1964, B.A. Harvard College, 1961. Professor Givelber served as dean of Northeastern University School of Law from 1984 until 1993, and was interim dean during the 1998–1999 academic year. He is an expert in the areas of criminal law, criminal procedure and capital punishment, and has been engaged in pro bono death penalty litigation for many years.

verdicts in opposition to judges. Data collected in four jurisdictions by the National Center for State Courts allows us to examine the question of judge and jury disagreement about guilt through a consideration of the views of jurors as well as judges. Using this data, we test in a modern context the hypothesis that the jury's embrace of values—as opposed to its different assessment of the evidence—explains why juries acquit when judges would convict. We find that legal and extralegal factors affect both judge and jury decisions about guilt, that both sets of factors predict disagreement in different contexts, and the pattern of agreement versus disagreement is more complex than suggested by the liberation hypothesis.

I. INTRODUCTION

The constitutional right to a jury trial rests upon values in addition to the interest of accurate fact-finding. When, at the height of the due process revolution, the Supreme Court confronted the question of whether the United States Constitution required states to afford criminal defendants the right to a jury trial in a serious case, the Court did not base its ruling on the view that juries were more accurate fact-finders than judges.[1] Rather, as it noted in *Duncan v. Louisiana*, the constitutional right to a jury trial in a serious criminal case reflected

> a profound judgment about the way in which law should be enforced and justice administered. A right to jury trial is granted to criminal defendants in order to prevent oppression by the Government . . . . If the defendant preferred the common-sense judgment of a jury to the more tutored but perhaps less sympathetic reaction of the single judge, he was to have it.[2]

The Court in *Duncan* found support for its view in empirical research on judge and jury decisionmaking. Responding to the objection that "juries are incapable of adequately understanding evidence or determining issues of fact, and that they are unpredictable, quixotic, and little better than a roll of dice"[3] the Court referenced Kalven and Zeisel's seminal study from *The American Jury*, stating:

> Yet, the most recent and exhaustive study of the jury in criminal cases concluded that juries do understand the evidence and come to sound conclusions in most of the cases presented to them and that when juries differ with the result at which the judge would have arrived, it is usually because they are serving some of the very purposes for which they were created and for which they are now employed.[4]

---

[1] Duncan v. Louisiana, 391 U.S. 145 (1968).

[2] *Id.* at 155–56.

[3] *Id.* at 157.

[4] *Id.* at 157 (citing HARRY KALVEN, JR. & HANS ZEISEL, THE AMERICAN JURY (1966)).

Analyzing survey responses from over 500 judges presiding over 3,000 criminal trials, Kalven and Zeisel concluded that when cases are close on the evidence, juries are "liberated" from the dictates of the law, and can—and do—give expression to extralegal values in arriving at verdicts. No other work of social science relating to jury behavior has been as widely cited or as approvingly referenced by courts.[5] Kalven and Zeisel's research has been cited in twenty-five different Supreme Court decisions (as well as more than 190 decisions of other courts)[6] as support for a proposition concerning the behavior of juries. These citations are a tribute to the eminence of the authors, and to the breadth and sweep of their empirical and analytic work, as well as to the mostly reassuring message that judges are bound by legal rules, but juries can and do make decisions that reflect the values and sentiment of the community, even when those decisions are in opposition to the law.

Kalven and Zeisel drew their conclusions about why juries disagreed with judges entirely from survey responses from judges about their perceptions as to why juries arrived at a different conclusion than they would have in the same case.[7] No attempt was made to verify that the conclusions of the judge about why the jury arrived at its verdict were in fact correct. Additionally, research for *The American Jury* was conducted over fifty years ago at the dawn of the civil rights movement before the composition of police forces, judges, and juries began to reflect more accurately the race and gender of the general population. It was conducted before DNA analysis exposed the vulnerability of previously uncontestable convictions in serious cases. Their data also predated the constitutional revolution in how courts conduct criminal adjudication.[8] The demographics of defendants in felony courts have also changed substantially since the time of Kalven and Zeisel's study.[9] While more recent research on jury

---

[5] Valerie Hans and Neil Vidmar, The American Jury *at Twenty-Five Years*, 16 LAW & SOC. INQUIRY 323 (1991).

[6] This is the result of a search of Westlaw on January 10, 2010 employing the inquiry "Kalven w/3 of Zeisel."

[7] KALVEN & ZEISEL, *supra* note 4, 45–54.

[8] Cases such as *Gideon v. Wainwright*, 372 U.S. 335 (1963) (right to counsel at trial), *Pointer v. Texas*, 380 U.S. 400 (1965) (right to confront witnesses), *Batson v. Kentucky*, 476 U.S. 479 (1986) (illegality of race-based jury challenges), *Taylor v. Louisiana*, 419 U.S. 522 (1975) (holding that a jury must reflect a cross-section of community; cannot exclude women as a class), In re *Winship*, 397 U.S. 358 (1970) (requirement of proof beyond a reasonable doubt), and *Brady v. Maryland*, 373 U.S. 83 (1963) (prosecution must disclose exculpatory evidence), as well as *Duncan v. Louisiana*, 391 U.S. 145 (1968), were all decided after the Kalven and Zeisel survey.

[9] In the cases examined by Kalven and Zeisel, 73% of all defendants were white and the remaining 27% were black. KALVEN & ZEISEL, *supra* note 4, at 195 tbl.60. As of 2002, the felony defendants in the seventy-five largest counties in U.S. were 31% white, 43% black,

decisionmaking has generally supported the notion that the liberation hypothesis is "alive and well" in modern courts,[10] no studies have directly measured whether sentiments as reported by jurors explain judge–jury disagreements, and whether the effect of juror sentiment are most pronounced in cases where the evidence is close.

Contemporary data from a four-city survey of criminal trials collected by the National Center for State Courts (NCSC) affords an opportunity to advance our understanding of judge–jury disagreement beyond Kalven and Zeisel's original findings.[11] Despite the dramatic changes in the American criminal justice system cited above, research using the NCSC data has found essentially the same rates of judge–jury disagreement as those identified by Kalven and Zeisel,[12] but important questions remain unanswered about the factors contributing to judge–jury disagreement about guilt. The NCSC data includes information from both jurors who decided a case and the judge who presided over the trial, which allows us to investigate whether, and to what extent, the jury's embrace of non-legal factors explains why judges and juries disagree about guilt in a modern context.

II. EXPLAINING JUDGE AND JURY DISAGREEMENT

In *The American Jury*, Kalven and Zeisel reported that when judges and juries disagreed, juries were far more likely to be lenient than judges. The authors identified three types of disagreements between judge and jury: (a) disagreement as to whether the defendant was guilty of any of the crimes for which he was on trial (66% of all disagreements), (b) disagreements between the judge and jury as to whether the defendant was guilty of some of the crimes with which he was charged (17% of all disagreements), and (c) disagreements between judge and jury in which the jury hangs as to one or more of the charges against the defendant (17% of all disagreements).[13] Each disagreement was in the same direction: the jury

---

24% Hispanic, and 2% "other." BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS 2002, at 454 tbl.5.52 (2002), *available at* http://www.albany.edu/sourcebook/pdf/t5522002.pdf. The percentage of defendants who are either female or under twenty-one has at least doubled: from 7% women in the 1950s to 18% in 2002, and from 9% of defendants who were twenty or younger in the 1950s to 18% in 2002. *Id.*

[10] For a review see Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 PSYCHOL. PUB. POL'Y & L. 622, 700–01 (2001).

[11] Theodore Eisenberg et al., *Judge–Jury Agreement in Criminal Cases: A Partial Replication of Kalven and Zeisel's* The American Jury, 2 J. EMPIRICAL LEGAL STUD. 171 (2005).

[12] *Id.* at 180–83.

[13] KALVEN & ZEISEL, *supra* note 4, at 109.

was far more likely to be more lenient than the judge. Thus, in 87% of the disagreements about guilt or innocence, the jury was more lenient than the judge was, and in 79% of the disagreements arising from a hung jury, the jury was more lenient than the judge.[14]

Relying solely upon the judge's written explanation for why the jury arrived at a verdict with which he disagreed, Kalven and Zeisel identified five different explanations for disagreement. They cited "evidence factors," "facts only the judge knew," "disparity of counsel," "jury sentiment about the individual defendant," and "jury sentiments about the law."[15] Weighting these factors, Kalven and Zeisel concluded that differing evaluations of the evidence accounted for 54% of all disagreements, sentiments about the law and the defendant for another 40%, and facts that only the judge knew and disparity of counsel the remaining 6%.[16] From these findings, they suggest that the combination of values and evidence explained a significant number of judge–jury disagreements. To understand the influence of these factors on jury decisions, they posited the "liberation hypothesis"[17]—that when the case was close on the evidence, the jury was "liberated" from the dictates of the law and could, and did, give expression to "sentiment"[18] in arriving at its verdict.[19]

Kalven and Zeisel never stated explicitly that when judge and jury disagreed about guilt, the judge was factually correct and the jury in error. Rather, they employed metaphors that suggested this was the case. Thus,

---

[14] *See id.* at 109 tbl.23.

[15] *Id.* at 106.

[16] *Id.* at 115 tbl.29.

[17] By "close" in this context, they apparently meant that the judge identified both value-based and evidence-based reasons to explain why the jury acquitted when the judge would have convicted. KALVEN & ZEISEL, *supra* note 4, at 166. Although they did ask expressly about whether the case was "close" in their second survey (Question 12) involving 1,191 responses, they did not ask this question in the initial survey involving 2,385 trials. *Id.* at 49. Perhaps for this reason, while they employ the "closeness" variable from the second survey to provide a map of the evidence and to demonstrate that jury verdicts follow the evidence, *id.* at 134, 158–59, they do not make substantial use of that variable in their development of disagreement cases and the relative roles of values and evidence. *Id.* at 163–64. In a footnote, *id.* at 164 n.2, they present an apparently mislabeled table (they title it "Normal Disagreements" when the table only makes sense if it includes both "normal" and "cross-over" disagreements) indicating the percentage of disagreements between judge and jury in clear and close cases. They do so, the footnote indicates, to show that disagreement occurs in clear cases as well as close cases. This is not a point they develop at any length in the text.

[18] KALVEN & ZEISEL, *supra* note 4, at 165.

[19] As summarized in *The American Jury*, "The sentiment gives direction to the resolution of the evidentiary doubt; the evidentiary doubt provides a favorable condition for a response to the sentiment. The closeness of the evidence makes it possible for the jury to respond to sentiment *by liberating* it from the discipline of the evidence." *Id.* at 165.

the judge's view of the case was a "baseline representing the law,"[20] while the close cases in which the jury came to a different conclusion constituted a "war with the law," albeit a "modest and subtle" one.[21] They asserted, "[W]hen the jury reaches a different decision from the judge on the same evidence, it does so not because it is a sloppy or inaccurate finder of facts, but because it gives expression to values which fall outside the official rules."[22] This conclusion reflected the popularly understood genius of the jury system that tempers the rigors of the law with the common sense of the community.

Until recently, attempts to replicate Kalven and Zeisel's findings in criminal trial decisions have been quite limited.[23] Most studies investigating the factors that inform judge and jury verdicts have relied upon comparisons of judge and jury verdicts in small samples of cases, interviews with decisionmakers about previous cases and experimental designs employing mock juries. These methodologies have a number of strengths and limitations. Judge–jury verdict comparisons from actual criminal trials necessarily involve the judge rendering a hypothetical judgment while the jury renders a real one. It is possible that judges will be less meticulous in evaluating the evidence when it is not their responsibility to decide a case, and that what the jury has already done may influence the verdict that judges indicate they would render.[24] However, reasoning from

---

[20] *Id*. at 499.

[21] *Id*. at 495.

[22] *Id.* Kalven and Zeisel also suggested "most but not all of the time" when juries *agreed* with judges, the jury was "not importing values of its own" into its decision about guilt or innocence. *Id.* at 494. Further, when juries and judges *disagreed* about guilt, they suggested "two-thirds of the disagreements with judges are marked by some jury response to values." *Id.* at 495.

[23] Many of the studies comparing judge and jury decisions involve civil rather than criminal cases. *See* Jennifer K. Robbenolt, *Evaluating Juries by Comparison to Judges: A Benchmark for Judging?*, 32 FLA. ST. U. L. REV. 469, 477–78 (2005). While these studies are informative, they do not truly represent the experience of judges and juries in the criminal justice system. The burden of proof in a civil case is "more probable than not." In the civil setting, justice is achieved when the parties are afforded a fair process for making their case before a disinterested adjudicator. Criminal justice seeks a higher end: that the court would ideally convict only those who actually committed the crime in question. Additionally, the law in many civil cases invites the decisionmaker to consult values in determining whether the defendant behaved appropriately (for example, was the defendant's behavior unreasonable, was the product unreasonably dangerous?). The very studies that demonstrate that juries are more lenient than judges in criminal cases also indicate that there is no particular direction to leniency in civil cases—when they disagree, it is as likely that the judge is more lenient than the jury as the other way around. *See* Larry Heuer & Steven Penrod, *Trial Complexity: A Field Investigation of Its Meaning and Its Effects*, 18 LAW & HUM. BEHAV. 29 (1994).

[24] *See* Robbenolt, *supra* note 23, at 473–77.

other empirical approaches presents even greater challenges. Archival studies such as post-judgment interviews with judges and juries about their decisionmaking processes often do not collect data from judges and juries in the same cases.[25] Simulations and mock jury experiments also pose challenges since, no matter how realistic the presentation, the participants are aware that their decisions have no genuine consequences.[26] Despite these challenges, research since *The American Jury* has advanced our understanding of the effect of evidentiary and extra-evidentiary factors on jury verdicts.

A. THE ROLE OF EXTRA-EVIDENTIARY FACTORS ON JURY VERDICTS

Extra-evidentiary factors affect jury verdicts, but the contexts in which such factors exert influence are limited. Research conducted in the 1980s using data from thirty-eight sexual assault cases found that juror decisions are dominated by evidentiary factors as opposed to victim or defendant characteristics and that juror attitudes have little explanatory power with respect to case outcomes.[27] Further analysis of the same set of sexual assault cases suggests that when liberation based on juror sentiment did occur, it was only in those cases that were closest on the evidence.[28] Other research has found that "case related" extra-evidentiary influences such as charge severity, pretrial publicity and trial complexity affect jury verdicts, but only when the evidence presented by the prosecution is ambiguous or weak.[29] Measures of defendant characteristics such as race and attractiveness (traditionally categorized as indicators of jury sentiment) did not measurably affect jury verdicts under any evidentiary conditions.[30]

---

[25] Neil Vidmar, *Making Inferences About Jury Behavior from Jury Verdict Statistics: Cautions About the Lorelei's Lied*, 18 LAW & HUM. BEHAV. 599, 607–08 (1994).

[26] Brian H. Bornstein & Sean G. McCabe, *Jurors of the Absurd? The Role of Consequentiality in Jury Simulation Research*, 32 FLA. ST. U. L. REV. 443, 445 (2005); Shari Seidman Diamond, *Illumination and Shadows from Jury Simulations*, 21 LAW & HUM. BEHAV. 561, 564 (1997).

[27] Christy A. Visher, *Juror Decision Making: The Importance of Evidence*, 11 LAW & HUM. BEHAV. 1 (1987).

[28] Barbara F. Reskin & Christy A. Visher, *The Impacts of Evidence and Extralegal Factors in Jurors' Decisions*, 20 LAW & SOC'Y REV. 423, 425–26 (1986). Reskin and Visher measured juror sentiment by (1) assessment of the defendant's attractiveness, (2) any reference to the defendant being employed or unemployed, (3) any negative comment about the victim's moral character, and (4) juror perception of the victim's responsibility for the assault. They classified hard evidence as eyewitness testimony, physical evidence, recovered weapon, and physical injury to the victim. *Id*.

[29] Dennis J. Devine et al., *Strength of Evidence, Extraevidentiary Influence, and the Liberation Hypothesis: Data from the Field*, 33 LAW & HUM. BEHAV. 136, 142–43 (2009).

[30] *Id.* at 136, 144–45.

Recent reanalysis of the data originally collected by Kalven and Zeisel using more sophisticated multivariate regression analysis techniques called into question a central premise of Kalven and Zeisel's hypothesis—the notion that evidentiary strength and juror sentiment were independent phenomena. Reanalyzing Kalven and Zeisel's data, Gastwirth and Sinclair found a strong relationship between a defendant's lack of a criminal record and the perception of judges that jurors were sympathetic to the defendant. They further found that judges' perceptions that jurors were sympathetic to defendants (actually brought about in part by the lack of criminal record) diminished after controlling for the severity of the cases.[31] These findings suggest that the liberation hypothesis may not actually work as Kalven and Zeisel posited. Instead, the presence or absence of particular evidentiary factors may actually bring about juror sympathy for the defendant, necessitating methodologies that control for these factors independently.

Because jurors are more likely than judges are to have demographic characteristics that may lead to identification with—and leniency toward—defendants, extralegal factors such as defendant or juror race might be expected to influence jury decisions.[32] Numerous studies suggest a relationship between juror characteristics and jury leniency toward same-race defendants,[33] but this relationship appears strongest when the evidence supporting guilt is mixed or weak.[34] Other jury research, however, suggests a more complex relationship between juror race, interpretation of evidence

---

[31] Joseph L. Gastwirth & Michael D. Sinclair, *A Re-examination of the 1966 Kalven-Zeisel Study of Judge–Jury Agreements and Disagreements and Their Causes*, 3 LAW, PROBABILITY & RISK 169 (2004).

[32] Data from the Capital Jury Project, a national study of capital jurors' decisionmaking using interviews with more than 1,000 actual jurors from trials in fourteen states, indicates there are significant differences between black and white jurors in terms of their degree of doubt about guilt, perceptions of defendant remorse, and beliefs about the future dangerousness of the defendant. *See* William Bowers et al., *Death Sentencing in Black and White: An Empirical Analysis of the Role of Jurors' Race and Jury Racial Composition*, 3 J. CONST. L. 171, 190–91 (2001); *see also* DONALD BLACK, SOCIOLOGICAL JUSTICE (1993); Norbert L. Kerr et al., *Defendant Juror Similarity and Mock Juror Judgments*, 19 LAW & HUM. BEHAV. 545 (1989).

[33] Sheri L. Johnson, *Black Innocence and the White Jury*, 83 MICH. L. REV. 1611, 1620 (1985); Samuel R. Sommers, *On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations*, 90 J. PERSONALITY & SOC. PSYCHOL. 597 (2006); Denis Chimaeze E. Ugwuegbu, *Racial and Evidential Factors in Juror Attribution of Legal Responsibility*, 15 J. EXPERIMENTAL SOC. PSYCHOL. 133, 136–43 (1979).

[34] VALERIE P HANS & NEIL VIDMAR, JUDGING THE JURY (1991); Robert J. MacCoun & Norbert L. Kerr, *Asymmetric Influence in Mock Jury Deliberation: Jurors' Bias for Leniency*, 54 J. PERSONALITY & SOC. PSYCHOL. 21 (1988).

and liberation,[35] sometimes finding no measurable relationship between defendant and juror race and case outcomes.[36]

Those attempting to apply the liberation hypothesis to other criminal justice decisions, particularly sentencing, have achieved mixed results.[37] Consideration of extralegal factors appears to be constrained when judges sentence offenders convicted of more serious crimes such as murder, rape and robbery, but these same judges appear "liberated" to consider extralegal factors such as race in sentencing decisions in less serious cases.[38] There is only mixed support for the liberation hypothesis in the disposition of murder cases at various stages of the criminal justice process. While victim characteristics affect the processing or murder cases, the effects are not clearly limited to a particular level of case severity.[39]

Principles from the liberation hypothesis have also been used in support of the argument that defendant and victim race significantly affects capital sentencing under conditions of less-than-clear defendant culpability and ambiguous evidence.[40] In an important study cited in *McCleskey v. Kemp*,[41] Baldus, Woodworth, and Pulaski found the race of the victim predicted capital sentences only in those cases where the evidence supporting conviction was neither particularly strong nor particularly weak.[42] Other studies, however, find that the effects of defendant and

---

[35] Nancy Pennington & Reid Hastie, *Evidence Evaluation in Complex Decision Making*, 51 J. PERSONALITY & SOC. PSYCHOL. 242 (1986).

[36] *See* Devine et al., *supra* note 30, at 145 (finding that defendant and foreperson race does not predict acquittal under any type of evidentiary condition).

[37] Thomas J. Keil & Gennaro F. Vito, *Race, Homicide Severity, and Application of the Death Penalty: A Consideration of the Barnett Scale*, 27 CRIMINOLOGY 511 (1989) (finding no support for the liberation hypothesis in their study of prosecutors' decisions to request the death penalty in Kentucky and suggesting that considerations of defendant and victim race were not confined to legally ambiguous cases); *cf.* Eric Baumer et al., *The Role of Victim Characteristics in the Disposition of Murder Cases*, 17 JUST. Q. 281 (2000) (finding only mixed support for the liberation hypothesis in the disposition of murder cases at various stages of the criminal justice process and concluding that while victim characteristics affect the processing or murder cases, the effects are not clearly limited to a particular level of case severity).

[38] Cassia Spohn & Jerry Cederblom, *Race and Disparities in Sentencing: A Test of the Liberation Hypothesis*, 8 JUST. Q. 305, 322 (1991).

[39] Baumer et al., *supra* note 37, at 281.

[40] Arnold Barnett, *Some Distribution Patterns for the Georgia Death Sentence*, 18 U.C. DAVIS L. REV. 1327 (1985).

[41] 481 U.S. 279 (1987).

[42] David Baldus et al., *Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience*, 74 J. CRIM. L.& CRIMINOLOGY 661 (1983).

victim race on decisions to seek the death penalty are not confined to legally ambiguous cases.[43]

B. PREDICTING JURY AND JUDGE DISAGREEMENT

Despite attempts to apply principles of the liberation hypothesis to understand legal outcomes in the criminal justice system, there has been relatively little direct examination of the decisionmaking of judges compared to juries. There are many reasons to anticipate juries would decide cases differently than judges. Juries are "one-shot" versus repeat players.[44] They are unlikely to be aware of the norms and expectations of court actors who have regular interactions and do not have the constraints that accompany repeated interactions. They decide collectively rather than individually, are not legally trained, and are more likely than the judiciary to be ethnically diverse. On the other hand, being novices and under the direction of the judge, they may also respond in their verdict to both verbal and nonverbal cues provided by the judge.[45]

Despite these differences, judges and juries are often in agreement about the outcome of legal cases.[46] When juries do disagree with judges, the research has demonstrated consistently, it is in the direction of greater leniency towards the defendant. The NCSC's four-jurisdiction study of criminal trials found that the rates at which judges and juries agreed today resemble those reported by Kalven and Zeisel fifty years ago.[47] Judges and juries in the NCSC data agreed on conviction in 64% of the cases (compared to 62% agreement on conviction found by Kalven and Zeisel) and agreed on acquittal in 14% of the cases (compared to 13% found by Kalven and Zeisel). The jury acquitted when the judge would have convicted in 19% of the cases (the same proportion identified by Kalven and Zeisel) and the jury convicted when the judge would have acquitted in

---

[43] Keil & Vito, *supra* note 37.

[44] Marc Galanter, *Why the "Haves" Come Out Ahead: Speculations on the Limits of Legal Change*, 9 LAW AND SOC'Y REV. 95, 97 (1974).

[45] Peter D. Blanck, *What Empirical Research Tells Us: Studying Judges' and Juries' Behavior*, 40 AM. U. L. REV. 775, 775 (1991); Ann Burnett & Diane Badzinski, *Judge Nonverbal Communication on Trial: Do Mock Trial Jurors Notice?*, 55 J. COMM. 209 (2005).

[46] Robbennolt observed,

> The most notable conclusion to be drawn from this emerging literature is that the decisionmaking of judges and jurors is strikingly similar. While there is evidence of some differences, there is a high degree of agreement between the groups, they appear to decide real cases quite similarly, and they show a great deal of similarity in responding to simulated cases designed to examine a variety of legal decisionmaking processes.

See Robbenolt, *supra* note 23, at 502.

[47] Eisenberg et al., *supra* note 11, at 180–83.

only 3% of the cases (compared to 6% identified by Kalven and Zeisel).[48] They also found that holding the strength of the case constant, juries were more likely to acquit than judges.[49] This was true whether one employs the judge or the jury's assessment of the strength of the evidence. This finding confirmed Kalven and Zeisel's best-known observation—that judges appear to have a lower conviction threshold than juries—but leaves open the question of why this is so. Are jurors prone to acquittal in the sense that they either are moved by sentiment or are unduly credulous while the judge responds objectively to fact? Alternatively, are judges prone to conviction in the sense that they cannot accurately identify the innocent in any but the most obvious case?[50] Research using the NCSC data indicates that jurors are more attuned than judges to features of the defense case. Whether the defendant produced a witness to support his version of events, whether or not the defendant had a criminal record, and whether he refused to plead guilty on the ground that he is innocent affect juror willingness to acquit, but have little impact on the judge's evaluation of the case.[51]

The NCSC investigation expanded significantly upon the kinds of information collected and analyzed in *The American Jury*. The NCSC survey secured information from individual jurors about the factors that informed their decision—particularly their assessments of the evidence and their sentiments about the law and the defendant. These questions, coupled with those dealing with the nature of the evidence and the strength of the case, provide a basis for attempting to determine whether Kalven and Zeisel's explanation for judge–jury disagreement finds support among today's jurors. More significantly, the NCSC data permits an examination of the factors that lead judges and juries to agree about guilt and innocence as well as those that lead to disagreement. The data cannot tell us whether any of these decisions are factually correct but it can provide a basis for determining whether the factors that are associated with these outcomes are logically related to decisions about guilt or innocence.

---

[48] *Id*. at 181.

[49] *Id*.; Paula L. Hannaford-Agor & Valerie P. Hans, *Nullification at Work? A Glimpse from the National Center for State Courts Study of Hung Juries*, 78 CHI.-KENT L. REV. 1249 (2003).

[50] The NCSC data has also been used in other scholarship to measure the probability that that jury decisionmaking may be prone to either type I (jury incorrectly convicts the innocent) and type II error (jury acquits the guilty). *See, e.g.*, Bruce D. Spencer, *Estimating the Accuracy of Jury Verdicts*, 4 J. EMPIRICAL LEGAL STUD. 305, 305 (2007).

[51] Daniel Givelber & Amy Farrell, *Judges and Juries: The Defense Case and Differences in Acquittal Rates*, 33 LAW & SOC. INQUIRY 31, 39–44 (2008).

C. THE CURRENT STUDY: UNDERSTANDING JUDGE AND JURY DISAGREEMENT

The current study examines the influence of a range of factors, including measures of evidence and jury sentiment, on judge and jury agreement or disagreement about a defendant's guilt. We measure the effect of these factors under conditions in which decisionmakers (examined separately for judges and juries) indicate the case is close on the evidence compared to cases that are not deemed close. In so doing, we are able to test directly in actual criminal trials the "liberation hypothesis" that, in close cases, sentiment moves juries but not judges to find defendants not guilty. Our analysis does not test what Kalven and Zeisel actually found—that judges explain jury disagreements by suggesting that juries are moved by values rather than facts. (We do not test this contention directly because the judge was not asked to explain why she disagreed with the jury verdict; rather, the judge and jurors responded independently to different questionnaires concerning the importance of various evidentiary and non-evidentiary factors in the case that the jurors had just decided.) The liberation hypothesis, however, stands for a proposition grander than how judges explain disagreements.[52] It holds that in close cases juries are more lenient than judges because jurors render decisions based on their feelings about the law and the defendant. In this analysis, we test two hypotheses central to the premises of the liberation hypothesis: first, that jury sentiment distinguishes cases where judges and juries disagree about guilt; and second, that sentiment has the strongest effect on judge and jury disagreement in cases that are close on the evidence. This inquiry is important because the claim that juries are moved by personal sentiment to acquit those whom the judge views as guilty fuels the perception that a "not guilty" verdict is less of an assessment of whether the defendant is actually innocent than it is a response to legally irrelevant factors. This presumption contributes to the view that "they are all guilty of something," even if the state cannot prove it in court.

---

[52] Interestingly, the NCSC data analyzed in the present study suggests judges may not be adroit at predicting how the jury will decide a case, casting some doubt on the ability of judges to understand why juries may disagree with the verdicts they would have rendered. Asking judges to predict how the jury would decide the case and comparing those responses to the verdict the jury actually rendered (excluding hung juries), the NCSC data indicates that judges accurately predicted the jury's decision 81% of the time (214 out of 264 cases). Judges were more proficient at predicting convictions than acquittals. The jury behaved as the judge predicted 86% of the time (161 out of 187 cases) when the judge thought the jury would convict but only 69% of the time (53 out of 77 cases) when the judge predicted the jury would acquit.

III. DATA AND METHODOLOGY

The data used to answer these questions was originally collected and analyzed by the National Center for State Courts as part of a study on hung juries.[53] For the hung jury study, the NCSC collected information from trial courts in the Central Criminal Division of the Los Angeles County Superior Court, California; the Maricopa County Superior Court (Phoenix), Arizona; the Bronx County Supreme Court, New York; and the Superior Court of the District of Columbia between 2000 and 2001.[54] Data was collected on non-capital felony trials in each jurisdiction. The data from this study includes case information about the nature of the charges, demographic information on the offender and victim(s), voir dire, trial evidence and procedures, and jury deliberations for 289 separate criminal cases.[55] In addition to case information, questionnaires were submitted to the presiding judge and to the prosecution and defense attorneys to elicit information about their perceptions of the case proceedings and outcomes. Individual jurors in each case also completed questionnaires about their perceptions of the evidence and testimony presented at trial, the deliberation process, and the outcome of the case.[56] These additional sources of data provide a unique opportunity to examine the sources of judge–jury disagreement, and may even provide some indication of whether the judge or the jury provides the more reliable judgment as to whether the accused is guilty. The data utilized here also overcomes many limitations of previous research, particularly expanding

---

[53] PAULA L. HANNAFORD-AGOR ET AL., INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOC. RESEARCH, EVALUATION OF HUNG JURIES IN BRONX COUNTY, NEW YORK, LOS ANGELES COUNTY CALIFORNIA, MARICOPA COUNTY, ARIZONA, AND WASHINGTON D.C., 2000–2001 (2002), *available at* http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/3689.

[54] According to the authors of the original study, sites were chosen based on a convenience sample. Some attention was given to sites with particular concerns about hung juries. For a detailed description of the sampling and study design, see *id.*

[55] The dataset originally included 394 cases. Thirty-two cases with no case disposition information, twenty-nine cases with no indication of how the judge would have decided the case, and twenty cases with missing data relevant to measures in our present inquiry were excluded from the analysis. An additional twenty-eight cases where the jury hung on all charges were also removed for the present analysis. Consistent with the approach of Kalven and Zeisel, our analysis focuses on the cases in which the jury either acquits the defendant of all charges or convicts him of at least one charge. KALVEN & ZEISEL, *supra* note 4, at 60.

[56] Individual questionnaires were completed by jurors serving in each case. A total of 3,497 juror surveys were completed for the 394 cases (80% response rate). Since the present analysis is conducted at the case level, individual juror responses were averaged for all jurors in each case to create aggregate jury responses.

the analysis to multiple case types across four courts throughout the country.[57]

To test the effect of juror sentiment in those cases where judges and juries disagree on the outcome of the case, we construct a series of multinomial regression models to compare the effects of legal and extralegal variables on judge and jury agreement and disagreement. We measure agreement and disagreement as (1) agreement on guilt; (2) agreement on acquittal; (3) disagreement on acquittal, where a judge would convict but the jury acquits; and (4) disagreement on guilt, where a judge would acquit but the jury convicts. We utilize multinomial logistic regression, an extension for binary logistic regression because our outcome is categorical and has more than two levels. Multinomial regression provides a set of coefficients for each of the comparison groups. The coefficients for the reference group are zeros, similar to the reference group for a dummy coded variable. Through this analysis, we seek to understand if differences between judge and jury outcomes are explained by the notion that judges are guided by "facts" and juries by "sentiments" or whether judges and juries respond to evidence in distinctive ways.[58] To specifically test the liberation hypothesis that when the evidence is close, jury sentiment predicts juror acquittal when judges would convict, we partition our data and present multinomial models predicting judge and jury agreement in cases (a) that judges indicate are close on the evidence compared to cases that judges believe are not close; and (b) that juries indicate are close on the evidence compared to those cases that juries believe are not close.

A. MEASURES

1. Outcome Measures

In the present analysis, we measure the effect of jury sentiment on four possible case outcomes: (a) judge and jury agree on conviction, (b) judge and jury agree on acquittal, (c) judge would convict, but the jury acquits, and (d) judge would acquit, but jury convicts. We created the outcome variable by combining information on the actual verdict rendered by the juries in each of the 289 cases analyzed here with information provided by

[57] Despite this advantage, the four courts analyzed here are all in urban areas limiting our ability to draw conclusions concerning judge and jury disagreements in suburban or rural areas.

[58] The original NCSC data collected case-level information from court records, judges, attorneys, and individual level juror information from each juror participating in a case. Since multiple juror responses are nested in the case-level data in the original NCSC data, we aggregated all juror responses to represent the average juror response for each case to facilitate the use of multinomial regression modeling at the case level.

judges about the verdict they would render if the case had been decided in a bench trial. We then recoded case outcomes into separate dummy variables where one equals the specified judge–jury agreement outcome and zero equals another outcome. Since multinomial regression models predict a set of coefficients for each of the outcome groups compared to a reference category, we use judge and jury agreement on conviction as the reference category against which to test the predictors of when a judge and jury agree on acquittal, when a jury acquits when the judge would convict, and when a jury convicts when judge would acquit.

Consistent with the approach of Kalven and Zeisel,[59] our analysis focuses on the cases in which the jury either acquits the defendant of all charges or convicts him of at least one charge. If jurors are driven by sentiment when they disagree with the judge about guilt in close cases, one would expect to find confirmation of this explanation in those cases in which the disagreement is starkest—when judge and jury come to opposite conclusions as to whether the defendant engaged in any criminal behavior. It is possible that those who are convicted of some but not all charges may be benefitting from jury sentiment, but the data provides no way to characterize which mixed verdicts should be treated as counterfactual "wins" for the defense and which should not. We eschew attempting to characterize some mixed verdicts as "defendant wins" and others as "prosecution wins" and instead treat all cases in which the defendant is convicted of any count as a conviction. For the same reason, we have omitted cases in which the jury hangs on all counts.

2. Independent Measures

Kalven and Zeisel attributed nearly 60% of the disagreements in which the jury was more lenient than the judge to evidentiary factors and 40% to jury sentiment about the law and the defendant.[60] To account for these

---

[59] Kalven and Zeisel treated disagreements between judge and jury with respect to guilt on various charges or penalty as "agreements to convict." KALVEN & ZEISEL, *supra* note 4, at 60. Thus, their famous Table 11 setting forth the basic pattern of judge–jury agreement treats every case in which the jury convicted on any count as representing a case in which the jury agreed with the judge that the defendant should be convicted. *Id.* at 56 tbl.11.

[60] *Id.* at 115 tbl.29. "Sentiments about the defendant" includes all reasons for disagreement "attributable to the personal characteristics of the defendant." "Sentiments about the law" refers to "particular instances of 'jury equity,' reasons for disagreement that imply criticism of either the law or legal result." *Id.* at 107. Noting that it is "surprisingly difficult to give a thumbnail sketch of evidence as a category of judge–jury disagreement," the authors indicate that "closeness of the case," "differing requirements for proof," and "different interpretations of evidence" are all examples of the kinds of reasons assigned to that category. *Id.* at 106. Disparity of counsel and facts known only to the judge accounted for 4% and 2% of disagreements respectively. *Id.* at 115 tbl.29.

important and distinct factors, we created a series of variables that measure aspects of both phenomena.[61]

3. Evidentiary Factors

"Quantity of the evidence" variables provide objective measures of the evidence admitted in trial by either the prosecution or defense. The original NCSC data includes separate counts of the number of witnesses and number of exhibits presented by the prosecution and defense. To measure the magnitude of the evidence presented by either the prosecution or defense, we combined measures of witnesses and exhibits to create additive measures of the total number of prosecution exhibits and witnesses and the total number of defense prosecution exhibits and witnesses.

We also include factors that relate directly to the defendant's case. Recent research suggests that the *quality* of the defense case may be more important than the quantity of witnesses or exhibits presented. When the defense puts forward no witnesses or the defendant testifies alone, the jury is much more likely to convict than when a defense witness or witnesses testify particularly if the defendant also testifies.[62] To account for this variation we include a dummy variable for the quality of the defense case. A weak case is coded 0 and measured as no witnesses testifying (33%) or only the defendant testifying alone (18%). A strong defense case is coded 1 and measured as a defense witness testifying either alone (24%) or in combination with the defendant (28%).

The defendant's criminal history is another important component of the quality of the defense case. Defendants without criminal records are less likely to be convicted than those with criminal records, even when the jury does not learn of the criminal history during the course of the trial.[63] Information on the defendant's criminal history is captured in the original NCSC data in a variable indicating whether the jury learned about a defendant's criminal history (measured as yes, the jury learned about the record, no, the jury did not learn about the record, or no criminal history

---

[61] In Kalven and Zeisel's study, sentiments about the defendant included reasons for disagreement that were attributable to the personal characteristics of the defendant while sentiments about the law referred to "particular instances of 'jury equity,' reasons for disagreement that imply criticism of either the law or the legal result." *Id.* at 107.

[62] The defense offered testimony (either that of the defendant or a witness) in 85% of the cases in which the defendant had no criminal record; the defendant himself testified in 68% of all such cases. Since a defendant is entitled to put his character in issue, the jury would have learned of the lack of a criminal record in all cases in which the defendant testified and would have likely learned of it in the remaining cases in which a witness other than the defendant testified. Givelber & Farrell, *supra* note 51, at 38 tbl.1.

[63] Ronald A. Farrell & Victoria Lynn Swigert, *Prior Offense Record as a Self-Fulfilling Prophecy*, 12 LAW & SOC'Y REV. 437, (1978).

record). We recoded this variable into a dummy variable measuring whether a defendant had a criminal history coded as 0 (regardless of whether the jury learned about the record or not) and 1 when the defendant did not have a criminal record.

The defendant's claim of innocence also predicts disagreements between the judge and jury.[64] While the defendant's insistence upon innocence as a reason for not accepting a plea has no direct analogue in the Kalven and Zeisel typology, it has been included in the present analysis as indicative of the defense case.[65] The NCSC survey asked lawyers to explain why the case had not ended in a plea bargain. This open-ended question was coded into common response categories, including: (1) because the defendant claimed he was innocent, (2) because the defendant refused to plead, (3) because no offer was made, and (4) because the parties could not agree on an appropriate plea. This is information that may or may not come to the attention of the judge if she inquires about the defendant's willingness to plead but would not be presented to the jury as relevant to the question of whether the defendant committed the crime. The reason for the failure of the plea was recoded in the present analysis as a dummy variable coded as 0 when the failure of a plea was any other reason besides the defendant's claim of innocence and 1 when the defendant claimed he was innocent.

4. Jury Sentiment

Kalven and Zeisel explained 40% of the disagreement when juries were more lenient than judges in terms of the jury's beliefs about the defendant and about the law. The NCSC study attempted to explore the latter issue by asking jurors, "How fair do you believe the law was in this case?" and "How fair would you say the legally correct outcome was?" These questions were all measured on seven-point scales, recoded to indicate scores of 1 represented the lowest assessment (i.e. least unfair) and 7 represented the highest assessment (i.e. most unfair).[66] The survey asked

---

[64] *Id.* at 438.

[65] Information on the reason a plea was rejected was missing for 56 of the 289 cases analyzed here. To provide the most conservative measure of the effect of defendants asserting their innocence in refusing the plea, we have recoded the missing values as 0 indicating the defendant did not assert innocence in the decision to reject a plea. It is possible in some of these cases no plea was offered.

[66] Questions about the fairness of the law and the legal outcome were originally coded in the opposite direction where 1 indicated the most unfair and 7 indicated the least unfair. To provide consistency across the sentiment measures, we reverse-coded the fairness measures so the highest scores indicate the feeling that the law was the most unfair in order to parallel the coding of feelings that the defendant was treated too harshly or that the jury felt sympathy for the defendant.

jurors another set of questions designed to measure sentiment toward the defendant. These included: "To what extent were you worried about the consequences to the defendant of a conviction by this jury?" and "How much sympathy did you feel for the defendant?" Questions about juror sentiment toward the defendant were also measured on a seven-point scale where scores of 1 indicated the lowest assessment (i.e. least sympathy, least worried about consequences) and scores of 7 indicated the highest assessment (i.e. most sympathy, most worried about consequences). To assess the effect of individual juror sentiment on case outcomes, we aggregated the data from individual juror questionnaires in each case to represent an average jury response for each case. A score measuring sentiment toward the law was created by taking the mean of the available score for questions about the fairness of the law and fairness of the legal outcome (*alpha*=.75). A separate score measuring sentiment toward the defendant was created by taking the mean of scores for questions about juror sympathy toward the defendant and concern about the consequences of conviction (*alpha*=.72).

5. Race of Defendant and Jurors

Though objective indicators of defendant characteristics do not, by themselves, necessarily result in particular jury sentiments, Kalven and Zeisel included judicial references to the characteristics of the defendants within the category of jury sentiment in their typology of reasons for judge–jury disagreement. Primary among these characteristics were the race, gender, and age of the defendant.[67] Information on the race and gender of the defendant was found in the NCSC case record data. Defendants were classified originally as white non-Hispanic (10%), white Hispanic (25%), black non-Hispanic (55%), black Hispanic (3%), Asian (1%), or another race (5%). We created dummy variables measuring whether a defendant was black (coded 0 for white non-Hispanic, black Hispanic, white Hispanic, Asian, and other racial groups and 1 for black non-Hispanic), which is included in the analysis to control for the effect of the defendants' race. Information was collected on defendant gender, but only 8% of all defendants were female. As a result, given the relatively small number of cases that included female offenders (28 out of 289), there was too little

[67] Defendant race was categorized by Kalven and Zeisel as "White" or "Negro." In cases where judges thought disagreement with the jury was based on jury sentiment toward the defendants, the judge indicated sympathy was felt toward white defendants, female defendants and the youngest and oldest defendants. KALVEN & ZEISEL, *supra* note 4, at 211 tbl.65.

variation in defendant gender to include in subsequent analyses. The NCSC survey did not collect information concerning the defendant's age.

In addition to measuring the effect of a defendant's race on judge–jury agreement, aggregate measures of the race of the jury were included in the analysis. The racial composition of the jury is measured here as the proportion of each jury that identify as being black.

6. Closeness of Case

The NCSC questionnaires to both judge and jury asked the question, "All things considered, how close was the case?" on a seven-point scale ranging from "evidence strongly favors prosecution" to "evidence strongly favors defense." Responses at the beginning of the range (1–2) indicate the evidence favored the prosecution and responses at the end of the range (6–7) indicate the evidence favored the defense. Responses in the middle of the seven-point range (3–5) indicate the case was close—it neither favored the prosecution or the defense. To provide a simplified measure of the closeness of the case we created a dummy variable in which responses 1–2 or 6–7 were coded as 0 (indicating the evidence in the case favored either the prosecution or the defense) and responses 3–5 were coded as 1 indicating the case was close.[68]

7. Severity of Charge

The NCSC data provides information on the type of offense charged for each count in each case. Sixteen offense types were included in this data. To control for the effect of different types of charges, we created an eight-point classification scale corresponding to the most serious offense charged where 1= murder, manslaughter, and attempted murder, 2= rape and robbery, 3= aggravated assault, weapons offense, and child abuse, 4= burglary, 5= drug distribution or sales, 6= drug possession, 7= larceny and theft, and 8= other less severe crimes. Drug offenses were the most common offense, representing the most serious charge in 27% of the cases.

8. Location

Data was collected on criminal cases in four separate courts, the Central Criminal Division of the Los Angeles County Superior Court, California; the Maricopa County Superior Court (Phoenix), Arizona; the

---

[68] Objective measures of the quantity of the evidence and judge or jury perceptions of case closeness do not appear to measure the same phenomena. The correlation between the objective measures of the quantity of the evidence (the number of witnesses and number of exhibits presented) and judge or jury perceptions of closeness of the case is weak and non-significant.

Bronx County Supreme Court, New York; and the Superior Court of the District of Columbia. To control for the differences that may exist between the courts, we created dummy variables indicating whether the case originated or not in each of the four courts.

IV. FINDINGS

Table 1 provides descriptive statistics for outcome and independent measures for the 289 cases analyzed here. Sixty-four percent (185 out of 289) of the cases resulted in agreement between the judge and the jury on conviction, 13% (37 out of 289) resulted in the judge and jury agreeing on acquittal, 17% (49 out of 289) resulted in the jury acquitting when the judge would have convicted and 6% (18 out of 289) resulted in the jury convicting when the judge would have acquitted. As Eisenberg et al. reported, the direction of judge–jury disagreement has remained unchanged since Kalven and Zeisel[69]—juries acquit when judges would have convicted more often than they convict when judges would have acquitted. In more than forty years, there has been practically no difference between the rate of judge–jury agreement as to who is guilty (the jury agrees with the judge when she would have convicted 79% of the time in both the NCSC data and Kalven and Zeisel's study). Today, however, there is a more substantial difference in the rate at which they agree about acquittals. Juries in the NCSC study agree with the judge when she would have acquitted only 67% of the time compared to 86% of the time in the Kalven and Zeisel study.[70]

---

[69] Eisenberg et al., *supra* note 11.

[70] Kalven and Zeisel presented their findings concerning judge–jury agreement in two important tables, Tables 11 and 12, which differed only in terms of their treatment of hung juries. KALVEN & ZEISEL, *supra* note 4, at 56 tbl.11, 58 tbl.12. Kalven and Zeisel's Table 11 showed the extent of judge–jury agreement in cases in which the jury convicted, acquitted, or hung. *Id*. at 56 tbl.11. Their Table 12 collapsed the jury verdict categories from three to two by treating cases in which the jury hung as though half of them ended in convictions and half of them were acquittals. *Id*. at 58 tbl.12. While a table presenting judge–jury agreement with respect to convictions and acquittals only is easier to follow than one including hung juries, Kalven and Zeisel never presented a serious rationale for dividing the hung juries between guilty verdicts and acquittals as opposed to simply omitting them. Kalven and Zeisel made this decision on the advice of an experienced prosecutor who estimated that about half of the time trials that ended in a hung jury eventually resulted in a conviction while the other half were either never retried or resulted in acquittal. *See id*. at 57–58. n.4.

We were unable to find any data concerning what happened to those defendants whose trial ends in a hung jury for the NCSC study sites. Research suggests retrials result mostly in convictions. *See, e.g.*, PAULA L. HANNAFORD-AGOR, ET AL., NAT'L CTR. FOR STATE COURTS, ARE HUNG JURIES A PROBLEM? 26–27 (2002), *available at* http://www.ncsconline.org/WC/Publications/Res_Juries_HungJuriesProblemPub.pdf (finding that in nine counties nearly 70% of the small number of hung cases that were re-tried resulted in conviction). We chose to omit hung juries for the presentation of our findings since the focus of our inquiry is to understand what happens in cases in which judges and juries disagree about guilt and there is

In contrast to the judge–jury agreement reported by Kalven and Zeisel, contemporary jurors are more likely to agree with a judge's decision to convict than they are to agree with a judge's decision to acquit.

The 289 cases examined here are evenly distributed across the four court sites with 21% in Los Angeles, 26% in Maricopa County, 26% in the Bronx, and 27% in Washington D.C. Fifty-five percent of the defendants are black and on average 26% of the jury is black. In half of the cases, the defense presents a witness testifying either alone or in combination with the defendant. Overall, however, the prosecution presents more witnesses and exhibits than the defense. On average, the prosecution presents 22.36 witnesses and exhibits compared to 6.27 on average for the defense. In

no reason to assume that half of the trials that end in hung juries are "actually" acquittals. The table below presents the findings from Kalven and Zeisel's Table 12 as it would look if recalculated to exclude hung juries as well as the comparable data from the NCSC survey. The italicized percentages illustrate the degree that the jury agrees with the judge's assessment of guilt or innocence.

Comparison of Judge–Jury Outcomes in NCSC Data to Kalven and Zeisel's Data (Excluding Hung Juries)

Kalven and Zeisel Data	**Jury Acquits % (N)**	**Jury Convicts % (N)**	**Total % (N)**
Judge Acquits	14.2% (479) *85.8%*	2.3% (79) *14.2%*	16.5% (558) *100%*
Judge Convicts	17.9% (602) *21.4%*	65.6% (2,217) *78.6%*	83.5% (282) *100%*
Total	32.0% (1,081)	68% (2,296)	100% (3,379)

NCSC Data	**Jury Acquits % (N)**	**Jury Convicts % (N)**	**Total % (N)**
Judge Acquits	12.8% (37) 67.2%	6.2% (18) 32.7%	19.0% (55) 100%
Judge Convicts	16.9% (49) 20.9%	64.0% (185) 79.1%	81.0% (234) 100%
Total	29.8% (86)	70.2% (203)	100% (289)

As it turns out, our finding that the direction of judge–jury agreement has changed holds true regardless of whether one excludes hung juries entirely (as we do here) or keeps hung juries in the calculation or divides hung juries between convictions and acquittals as did Kalven and Zeisel.

Table 1

Descriptive Measures of Outcome and Independent Variables (n = 289)

Case outcome	
Agree conviction	64.0% (185)
Judge convicts, jury acquits	17.0% (49)
Agree acquittal	12.8% (37)
Judge acquits, jury convicts	6.2% (18)
Site	
Los Angeles	21.4% (62)
Maricopa	25.9% (75)
Bronx	26.2% (76)
D.C.	26.5% (76)
Defendant race	
Non-black	45.0% (130)
Black	55.0% (159)
Quality of defense	
No witnesses or defendant testifies alone	50.3% (145)
Witness testifies alone or with defendant	49.7% (144)
Criminal record	
Defendant has a criminal record	79.2% (229)
Defendant does not have criminal record	20.8% (60)
Defendant claims innocence in plea refusal	
No	61.6% (178)
Yes	38.4% (111)
Jury thinks case close	
No	50.2% (145)
Yes	49.8% (144)
Judge thinks case close	
No	45.9% (133)
Yes	53.1% (153)

Table 1
(continued)

Variable	Mean	SD	Min	Max
Percent jury black	25.69	25.53	0	100
Severity of charge	3.80	2.21	1	8
Number prosecution witnesses & exhibits	20.36	24.42	1	235
Number defense witnesses & exhibits	6.27	11.99	0	102
Jury sentiment about law	2.39	0.77	1	7
Jury sentiment to defendant	3.21	0.84	1	7

21% of the cases, the defendant does not have a criminal record and in 38% of the cases, the defendant asserts his innocence as the reason that he will not accept a plea bargain.

On a scale of 1 to 7 measuring jury sentiment toward the law (with 1 indicating the law is the most fair and 7 indicating the law is the least fair) and jury sentiment toward the defendant (with 1 indicating the least sympathy or concern for the defendant and 7 indicating the most sympathy or concern for the defendant), juries on average rated the law and legally correct outcome as a 2.39 and their sentiment toward the defendant as a 3.21. Judges are slightly more likely to indicate that a case is close on the evidence than juries. In the 289 cases analyzed here, the jury indicates that the case is close on the evidence 50% of the time while the judge believes this to be true in 53% of the cases.

The distribution of measures varies across cases with different levels of judge and jury agreement. We conducted chi-square tests for independence and analyses of variance to examine differences in the independent variables across the four case outcomes (statistically significant differences are noted in Table 2). Case outcomes varied significantly across courts. Judges and juries agreed on the case outcome (conviction or acquittal) 89% of the time in Los Angeles and 87% of the time in Maricopa, but only 65% of the time in the Bronx and D.C. respectively. Juries in the Bronx were more than twice as likely to have acquitted when the judge would have convicted compared to Los Angeles and four times more likely to do so compared to juries in Maricopa County. Juries in Washington, D.C. were the most likely to acquit when the judge would have convicted,

Table 2

Bivariate Relationships Between Independent Variables and Case Outcomes

	Judge would convict		Judge would acquit	
	Jury convicts (n=185)	Jury acquits (n=49)	Jury acquits (n=37)	Jury convicts (n=18)
Site***				
Los Angeles	78.5%	9.2%	10.8%	1.5%
Maricopa	69.6%	5.1%	17.7%	7.6%
Bronx	51.5%	22.7%	13.6%	12.1%
D.C.	55.8%	31.2%	9.1%	3.9%
Race**				
Non-black	60.9%	13.5%	17.3%	8.3%
Black	66.7%	19.9%	9.0%	4.5%
Quality of defense***				
No witnesses or defendant testifies alone	75.5%	10.8%	7.2%	6.5%
Witness testifies alone or with defendant	53.3%	22.7%	18.0%	6.0%
Prior criminal history***				
Defendant has criminal record	69.8%	13.5%	12.1%	4.7%
Defendant does not have criminal record	43.3%	30.0%	18.3%	8.3%
Defendant claims innocent***				
No	73.0%	14.0%	6.7%	6.2%
Yes	49.5%	21.6%	22.5%	6.3%
Judge thinks case close***				
No	75.9%	9.5%	10.2%	4.4%
Yes	53.3%	23.3%	15.3%	8.0%

* $p < .10$; ** $p < .05$; *** $p < .01$

Table 2
(continued)

	Judge would convict		Judge would acquit	
	Jury convicts (n=185)	Jury acquits (n=49)	Jury acquits (n=37)	Jury convicts (n=18)
Jury thinks case close***				
No	87.8%	3.6%	4.3%	4.3%
Yes	36.9%	30.6%	22.4%	7.5%
	Mean	Mean	Mean	Mean
Percent jury black**	23.72	36.35	22.29	21.92
Severity of charge	3.73	3.69	4.00	4.33
Number prosecution witnesses & exhibits*	23.09	18.12	13.28	14.63
Number defense witnesses & exhibits	6.52	5.36	6.97	4.94
Jury sentiment about law***	2.22	2.80	2.56	2.74
Jury sentiment about defendant***	3.32	2.98	2.90	3.51

with disagreements three times the rate of those in Los Angeles and six times that of Maricopa County. Disagreements in the direction of juror liberation were also more common when the defendant was black (20% of cases resulted in jury acquittal when judge would have convicted compared to only 14% when the defendant was another race) and when a higher proportion of the jury members are black. There were no significant differences in case outcomes with various levels of charge severity.

The quantity of the evidence, measured by the numbers of witnesses presented by the prosecution has a significant, but rather moderate, effect on case outcomes. The average number of witnesses and exhibits presented by the prosecution is higher in cases where the judge and jury agree on guilt (an average of 23 prosecution witnesses and exhibits) than those cases where the judge and jury agree on acquittal (an average of 13 prosecution witnesses and exhibits) or the jury disagrees with the judge (an average of 18 prosecution witnesses and exhibits in liberation cases). The number of witnesses and exhibits presented by the defense is not significantly related

to the case outcome. The *quality* of the defense presented, however, does affect whether or not the judge and jury agree. When the defense presents a weak case (either no witnesses testify or only the defendant testifies as a witness) the judge and jury agreement about guilt is quite robust (76%). When the defense presents a stronger case (defense witnesses testify alone or in combination with the defendant), agreement about conviction drops to 53%. Juries are twice as likely to acquit when the judge would have convicted in cases where the defense presents a witness other than or in addition to the defendant. Not surprisingly, the judge and jury are most the likely to agree about acquittal in those cases where the defendant and a supporting witness testify together.

The findings from the NCSC data reveal that the defendant's lack of a criminal history significantly predicts judge and jury disagreement about conviction. Judges and juries were significantly more likely to agree about the case outcome in those cases in which the defendant had a criminal history (agreeing to convict 70% of the time). When the defendant did not have a criminal history, agreement on conviction dropped to 43%, and 30% of these cases resulted in jury acquittals when the judge would have convicted. In fact, judges disagreed with the jury's decision to acquit defendants without a criminal history two-thirds of the time, suggesting that the lack of a criminal history may have moved juries towards acquittal in many cases in which judges believed the defendant to be guilty.

The NCSC data provides a unique opportunity to explore the relationship between the defendant's reason for refusing to plead guilty and jury outcomes. Although juries were unlikely to know the reason the defendant refused to plead, a defendant's claim of innocence to his or her attorney has a strong positive effect on jury acquittal. Juries acquitted in 44% of all cases in which the defendant did not plead because he insisted he was innocent; judges agreed with the jury that the defendant was not guilty in half of these cases. Perhaps defendants insist that they are innocent in order to encourage their lawyer to behave vigorously on their behalf. Perhaps they insist on their innocence although they know themselves to be guilty because they believe that the state has a weak case. However, a defendant sophisticated enough to claim innocence in order to manipulate his lawyer may well be sophisticated enough to know the risks of insisting upon trial should he be convicted. His lawyer will tell him that if he is convicted after trial, the judge may increase his sentence beyond what he would have received if he pled and may increase it even more if he takes the stand and insists unsuccessfully upon his own innocence.[71] Defendants

[71] *Cf.* Daniel Givelber, *Punishing Protestations of Innocence: Denying Responsibility and Its Consequences*, 37 AM. CRIM. L. REV. 1363, 1368 (2000).

may also decline a plea bargain on the grounds of innocence because they are, in fact, innocent. At a minimum, juries—who are ignorant of plea offers and their refusal—respond more positively to defendants who refuse to plead on this ground.

Jury sentiments about the law and the defendant also differed significantly by case outcomes. Not surprisingly, juries felt the law and legally correct outcome were most fair in those cases where they agreed with the judge on conviction (average rating of 2.22 on a scale of 1 to 7 where 1 is most fair and 7 is least fair). The jury felt the law as well as the legally correct outcome was least fair in liberation cases, when they acquitted and the judge would have convicted (average rating of 2.80). Contrary to Kalven and Zeisel's liberation hypothesis, we find the juries' concern and sympathy for the defendant appears to be reserved for those cases where the defendant is convicted. Juries had the least amount of sympathy for the defendant or concern about the consequences of a conviction (measured on a scale of 1 to 7 where a score of 1 represents the least sympathy or concern and 7 represents the most sympathy or concern) in those cases where they acquitted (either with or without the judge's agreement). It is possible that the jury does not worry about the consequences to the defendant in cases in which they acquit because their action benefits rather than harms the defendant.

The liberation hypothesis rests upon the notion that cases that are close on the evidence free jurors to consider sentiment. Unlike Kalven and Zeisel's data, which only allowed for examination of the judge's perspective of the closeness of a case, the NCSC data provides information from both the judge and the jury about the degree to which they believe the case is close on the evidence. Not surprisingly, jury assessment of case closeness is more predictive of acquittal than judicial assessment. When the jury indicates that the case is close on the evidence, they acquit 53% of the time compared to only 8% of the time when they do not think the case is close on the evidence. When judges think the case is close, the jury acquits in only 39% of the cases compared to 20% when the judge does not think the case is close.

The NCSC data confirms that close cases breed judge and jury disagreement. When judges think cases are close on the evidence, the jury agrees with the judge on the case outcome in only 69% of the cases (compared to 86% agreement when the judge thinks the evidence is clear). Turning to jury assessments of closeness, we find the gap in agreement about case outcomes to be even wider. When juries think the case is close on the evidence, they agree with the judge 59% of the time: in 37% of the cases they agree that the defendant is guilty while in 22% of the cases they agree that the defendant is not guilty. Conversely, when juries do not think

Table 3

Multinomial Regression Predicting Judge and Jury Agreement (n=289)

	Model 1: Evidence and Sentiment			Model 2: Evidence, Sentiment, and Judge Closeness			Model 3: Evidence, Sentiment, and Jury Closeness		
	Judge convicts, jury acquits B/(SE)	Judge acquits, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)	Judge convicts, jury acquits B/(SE)	Judge acquits, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)	Judge convicts, jury acquits B/(SE)	Judge acquits, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)
LA	-1.652** (768)	-.811 (.815)	-2.330* (1.382)	-1.698*** (.794)	-.652 (.816)	-2.247 (1.382)	-2.672*** (.927)	-1.624* (.932)	-2.544* (1.426)
Maricopa	-2.433*** (.927)	-.021 (.845)	-.853 (1.192)	-2.353*** (.926)	.023 (.849)	-.864 (1.186)	-3.385*** (1.081)	-.546 (.932)	-.922 (1.209)
Bronx	-.911 (.615)	-.423 (.734)	.462 (.919)	-.868 (.615)	-.350 (.735)	.468 (.917)	-.913 (.696)	-.490 (.795)	.404 (.950)
Defendant black	-.506 (.532)	-.573 (.519)	-.541 (.764)	-.458 (.540)	-.531 (.522)	-.568 (.760)	-.425 (.597)	-.657 (.576)	-.540 (.777)
Jury black	.005 (.011)	-.012 (.013)	-.020 (.017)	.003 (.011)	-.015 (.013)	-.022 (.017)	-.006 (.012)	-.022 (.014)	-.022 (.017)
Charge severity	.068 (.128)	-.012 (.122)	.211 (.168)	.078 (.131)	.007 (.122)	.209 (.167)	.079 (.146)	-.031 (.133)	.206 (.174)
Prosecution witnesses & exhibits	-.032* (.017)	-.044** (.017)	-.027 (.022)	-.029* (.017)	-.041** (.017)	-.026 (.022)	-.028 (.018)	-.044** (.018)	-.024 (.021)
Defense witnesses & exhibits	-.028 (.040)	-.009 (.029)	-.015 (.038)	-.030 (.051)	-.014 (.029)	-.017 (.038)	-.009 (.048)	.001 (.035)	-.007 (.035)

	Model 1: Evidence and Sentiment			Model 2: Evidence, Sentiment, and Judge Closeness			Model 3: Evidence, Sentiment, and Jury Closeness		
	Judge convicts, jury acquits B/(SE)	Judge acquits, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)	Judge convicts, jury acquits B/(SE)	Judge acquits, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)	Judge convicts, jury acquits B/(SE)	Judge acquits, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)
Quality of defense case	.539* (.532)	1.450*** (.564)	1.325* (.735)	.383 (.549)	1.399*** (.571)	1.265* (.737)	-.372 (.610)	.842 (.612)	1.001 (783)
No criminal history	1.996*** (.549)	.627 (.591)	.766 (.754)	1.789*** (.642)	.499 (.603)	.685 (.761)	2.509*** (.663)	.646 (.683)	.743 (.775)
Defendant claims innocent	.167 (.467)	1.344*** (.475)	-.449 (.700)	.084 (.473)	1.303*** (.476)	-.444 (.705)	.026 (.516)	1.188** (.515)	-.503 (.709)
Jury sentiment about law	.928*** (.311)	.827*** (.314)	1.095** (.393)	.920*** (.314)	.824*** (.314)	1.167*** (.428)	.332 (.349)	.364 (.347)	.852* (.401)
Jury sentiment about defendant	-.967*** (.307)	-1.254*** (.342)	-.059 (.370)	-.949*** (.310)	-1.215*** (.343)	-.058 (.366)	-1.028*** (.342)	-1.285*** (.382)	-.072 (.372)
Judge case close	-	-	-	.711 (.506)	.708 (.491)	.355 (.646)	-	-	-
Jury case close	-	-	-	.-	-	-	3.364*** (.736)	2.703*** (.663)	1.176 (.715)
Intercept	-3.523* (1.921)	2.160 (1.855)	-6.246** (2.981)	-2.905* (1.253)	2.412 (1.879)	-6.003** (3.004)	-1.292 (2.121)	3.302* (.1.971)	-5.475* (3.061)
R-Square (Nagelkerke)	.476			.484			.593		

* p <.10; ** p <.05; *** p <.01; reference category is judge and jury agreement about conviction

cases are close on the evidence they agree with the judge that the defendant is guilty 87% of the time and that the defendant is not guilty 4% of the time (for a total agreement of 91%).

We now turn to the heart of our inquiry—examining if and how evidentiary and non-evidentiary factors explain disagreements between judges and juries. Our particular interest is examining whether "jury sentiment" toward the law and the defendant can be identified as an explanation for judge–jury disagreement in the modern NCSC data. Since case outcome is a multi-category variable representing judge and jury agreement or disagreement about conviction or acquittal, here, we use multinomial regression analysis to help isolate the degree to which variables illustrative of jury sentiment affect outcomes in those cases in which judges and juries disagree on the verdict. In order to arrive at some understanding of the role of sentiment, we also examine the effect of a range of other possible contributors to disagreement. We calculate three multinomial models to predict the independent effect of evidence and jury sentiment on case outcomes. The first model predicts the effect of evidence and jury sentiment on outcomes excluding measures of judge and jury evaluation of whether or not the case was close. The second model examines the effects of evidence and juror sentiment controlling for judge assessment of the case closeness while the third model controls for jury perceptions of case closeness. In all three models, the comparison category for the dependent variable is judge and jury agreement on conviction, meaning that the coefficients for agreement on acquittal and disagreement on conviction or acquittal are measured against agreement on conviction. The results from the multinomial models are set forth in Table 3. To simplify the presentation of results, the outcomes for cases indicative of the liberation hypothesis (where the jury acquits when the judge would have convicted) are highlighted in grey and a majority of the discussion of findings centers around predictors of these "liberation" cases.

Model 1 indicates some difference in judge–jury agreement between study sites. Cases in Maricopa and Los Angeles cases were less likely than cases in Washington D.C. (the reference category) to result in jury acquittals when the judge would have convicted. As demonstrated in previous research,[72] when the defendant had no criminal record juries were more likely to acquit when judges would convict as compared to agreement on conviction (b=1.996, p=.001). While the total number of prosecution and defense witnesses did not strongly predict judge and jury disagreement, judges and juries were significantly more likely to agree on acquittal when the defense presented a strong case in terms of having a supporting witness

[72] Givelber & Farrell, *supra* note 51, at 40–43 & tbl.2.

testify (alone or in combination with the defendant, *b*=1.450, *p*=.010) and when the defendant claimed he was innocent (*b*=1.344, *p*=.005). One possible explanation for both phenomena is that defendants most likely insist upon innocence in those cases where there is strong evidentiary basis for doing so. Moreover, a defense lawyer with a client insisting on innocence and witnesses to help prove his case may be more energized to put on an effective defense. There are no significant effects of defendant race, jury race, or charge severity on differences in case outcomes.

Since we hypothesized that jury sentiment would distinguish cases where judges and juries disagree on the outcome, we should anticipate a significant relationship between jury sentiment variables (belief that the law is unfair and sympathy for the defendant) and the coefficients for disagreement in the multinomial models. Specifically, we should find the strongest positive relationship between both sentiment variables and jury acquittals occurring in liberation cases—where the judge would convict when the jury acquits. Model 1 illustrates that jury sentiment toward the law has a significant positive effect on all examined outcomes. That is, when juries perceive the law or the correct legal outcome to be unfair, it is significantly more likely that (a) the jury will acquit even if the judge would convict (*b*=.928, *p*=.003), (b) the judge and jury will agree on acquittal (*b*=.827, *p*=.008), or (c) the judge will indicate that she would have acquitted even if the jury convicts (*b*=1.095, *p*=.005) than it is that (d) the jury will convict and the judge agrees. Therefore, while sentiment toward the law does predict disagreement, it also significantly predicts judge and jury agreement about acquittal. Contrary to the prediction of the liberation hypothesis, in the cases where the jury expresses sympathy for the defendant, jurors are less likely to acquit when the judge would convict (*b*=-.967, *p*=.002) than they are to agree with the judge that the defendant is guilty. As we suggested earlier, this may be because juries feel less sympathy for a defendant whom they are going to acquit.

Adding judicial perception of case-closeness into the second model, we find that cases are not more likely to result in disagreement when the judge indicates that the case is close on the evidence than otherwise. More importantly, adding judge perceptions about the closeness of the case does not dramatically alter the relationships between jury sentiment and case outcomes observed in Model 1. However, adding jury perceptions of case-closeness in Model 3 both improves the strength of the model and changes the effect of some important predictors of disagreement. Initially, the amount of variance explained in the model increases when jury perceptions of closeness are included (pseudo r-square measures improve from .48 in Model 1 and Model 2 to .59 in Model 3). Jury perception of closeness also significantly predicts judge and jury disagreement (acquitting when the

judge would convict, *b*=3.364, *p*=.000) and jury agreement with the judge on acquittal (*b*=2.703, *p*=.000). Evidentiary factors such as the defendant's lack of a criminal history still strongly predict judge and jury disagreement (*b*=2.509, *p*=.000), but jury sentiment toward the law is no longer a significant predictor of judge–jury disagreement. While jury sentiment toward the defendant remains a significant predictor of judge–jury disagreement, the direction of the effect is opposite to that predicted by the liberation hypothesis.

The liberation hypothesis posits that jurors who disagreed with the judge by acquitting someone the judge believed guilty were responding to sentiment in cases that were close on the evidence. If this proposition holds true today, we should expect that the effects of jury sentiment toward the law and the defendant should be strongest in those cases that are close on the evidence. We test this contention directly by partitioning the NCSC data by both judge perceptions of closeness (close or clear) and jury perceptions of closeness (close or clear) and reexamining the effects of jury sentiment on case outcomes in close and clear cases separately. As we noted in the discussion of Table 2, cases that the jury believes are clear generally generate agreement between the judge and jury and cases that are close generate disagreement. This does not necessarily mean that we can assume "sentiment" is more influential in close cases than it is in clear cases. Indeed, sentiment may affect jury decisionmaking in different ways in cases that either the judge or the jury believes are close compared to those that are believed to be clear. Although nearly one-quarter of the instances of judge–jury disagreements identified in *The American Jury* occurred in cases that the judge believed to be clear on the evidence, Kalven and Zeisel never analyzed these cases separately to determine whether or not jury sentiment also predicted judge–jury disagreement in clear cases.[73] The present analysis attempts to remedy this omission, providing a more complete evaluation of the role of jury sentiments play in predicting judge and jury disagreement in both close and clear cases.

[73] Because the second survey was the one which asked the judge whether the case was "close" or "clear," it seems likely that the figures about disagreement in "clear" cases are based on that survey. They do not discuss disagreements in "clear" cases as a distinct category although they note that as a matter of theory disagreements in "clear" cases should be disagreements based exclusively on values, a phenomenon which they identify as occurring in 24% of all cases in both surveys. They note that the second survey indicated that disagreements occurred in 25% of all cases that the judge designated as clear. *Id*. at 164 n.2. The figure for all disagreements (both normal and cross-over) in Kalven and Zeisel was 22%. *See* KALVEN & ZEISEL, *supra* note 4, at 58 tbl.12.

Multinomial models predicting the effects of evidentiary and jury sentiment measures on case outcomes[74] for cases that judges indicate are "not close" and cases that judges indicate are "close" on the evidence[75] are displayed in the first two columns of Table 4. As suggested by the liberation hypothesis, when the judge considered the case to be close on the evidence, jury sentiment about the fairness of the law predicts that the jury will acquit when the judge would convict (b=.702, p=.027). Sentiment, however, is not the only significant predictor of judge jury disagreement in cases that judges indicate are close on the evidence. The defendant's lack of a criminal record also significantly predicts judge–jury disagreement in the direction of liberation (b=1.101, p=.043), suggesting both sentiment and evidentiary factors influence juries to acquit when judges would convict in cases that judges believe are close on the evidence. Interestingly, jury sentiment about the defendant and the lack of a defendant's criminal history also predict judge–jury disagreement in cases that the judge indicates are not close on the evidence, indicating that evidence and sentiment may also be at work in clear cases as well.[76]

The last two columns of Table 4 present the effect of measures of evidence and jury sentiment on judge–jury disagreement in cases the *jury* thinks are not close compared to cases the jury thinks are close on the evidence. While jury sentiment about the law predicted judge–jury disagreement when the judge thought the case was close, there is no relationship between jury sentiment about the law or the defendant and case

---

[74] To simplify the interpretation the partitioned multinomial model results, we collapsed case outcomes into three categories: judge and jury agree about conviction or acquittal (reference category), judge would convict when jury acquits, and judge would acquit when jury convicts. The strength and direction of key measures are then same when agreement cases are separated into agreement on acquittal and agreement on conviction (as presented in Table 3) and results from these more complex models are available upon request from authors. Models show here are also limited to key variables measuring evidence and jury sentiment.

[75] Judges and juries were each asked to rate the degree to which the evidence favored the prosecution or defense on a 1 to 7 scale (1 indicating evidence favored the prosecution and 7 indicating evidence favored defense). The responses were recoded into not close (coded 0 when the judge or jury indicated an answer at the ends of the scale of 1–2 or 6–7), and close (coded 1 when the judge or jury indicated an answer in the middle of the scale of 3, 4, or 5).

[76] We used a coefficient comparison test to examine differences in the coefficients across partitioned close and not close models for judges and juries (for more information about the z-value test. *See* Raymond Paternoster et al., *Using the Correct Statistical Test for the Equality of Regression Coefficients*, 36 CRIMINOLOGY 859, 859. (1998) (computation of $Z = b_1 - b_2/\sqrt{SEb_1^2 + SEb_2^2}$). While the strength of some evidentiary and jury sentiment measures differed between cases that judges thought were close compared to those cases that judges did not think were close, none of these differences were large enough to be statistically significant. Z-score calculations are available from the authors upon request.

Table 4

Partitioned Multinomial Models Predicting Judge and Jury Disagreement by Jury and Judge Assessments of Case-Closeness

	Judge Perception				Jury Perception			
	Not Close		*Close*		*Not Close*		*Close*	
	Judge convicts, jury acquits B/(SE)	Judge acquits, jury convicts, B/(SE)	Judge convicts, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)	Judge convicts, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)	Judge convicts, jury acquits B/(SE)	Judge acquits, jury convicts B/(SE)
Prosecution witnesses & exhibits	.014 (.014)	-.344* (.171)	-.020 (.018)	.011 (.021)	.029 (.018)	-.064 (.058)	-.012 (.015)	.003 (.023)
Defense witnesses & exhibits	-.019 (.057)	-.048 (.086)	-.042 (.041)	-.022 (.050)	-.029 (.075)	-.058 (.084)	-.029 (.042)	-.015 (.087)
Quality defense case	-.154 (.834)	1.004 (1.269)	.590 (.610)	1.105 (1.062)	-1.777 (1.622)	2.266 (1.684)	-.031 (.551)	-.101 (.924)
Prior criminal history	1.863** (.916)	-.579 (1.531)	1.101** (.477)	-.350 (.946)	2.927** (1.493)	1.249 (1.258)	1.342*** (.526)	.399 (.928)
Defendant claims innocent	-.465 (.811)	-2.680 (2.209)	.062 (.499)	.227 (.862)	-.551 (1.435)	.270 (1.264)	-.402 (.454)	-.1.607* (.902)
Jury sentiment about law	.645 (.561)	1.208 (.762)	.702** (.318)	1.017* (.546)	-.037 (.707)	.900 (.628)	.353 (.312)	.841 (.621)
Jury sentiment about defendant	-.759* (.442)	.413 (.721)	-.401 (.328)	.580 (.489)	-.361 (.542)	-.185 (.692)	-.412 (.304)	.487 (.467)
Intercept	-.912 (2.658)	-7.016* (3.520)	.935 (1.145)	-6.880** (2.865)	-2.304 (3.575)	-3.992 (3.030)	1.518 (1.497)	-6.842** (3.110)
Pseudo R-Square (Nagelkerke)	.342		.214		.342		.212	

* p <.10; ** p <.05; *** p <.01; reference category is judge and jury agreement (agreement about conviction and agreement about acquittal combined).

outcomes in those cases that the jury thinks are close on the evidence. The only variable that predicts judge–jury disagreement in the direction of liberation in cases that the jury thinks are close on the evidence is a defendant's lack of a criminal record (b=1.342, p=.009). Jurors, then, were *least likely* to be influenced by sentiment in the very cases that the liberation hypothesis suggests that sentiment should matter most—when they believed the case to be close on the evidence. While jury sentiments toward the law and to a lesser degree (and in the opposite direction of that predicted by Kalven and Zeisel) toward the defendant help predict cases where the jury and judge disagree about guilt generally, these factors only predict disagreement in a limited number of cases. Specifically, sentiment toward the law only predicts liberation (jury acquittal when the judge would convict) in cases where the judge thinks the evidence is close and sentiment toward the defendant negatively predicts liberation only in cases that judges believe are not close on the evidence. Jury sentiment toward the law or the defendant has no effect on case outcomes when the jury considers the case close. While this result seems logical—sentiments should help explain why judge and jury disagree when one or the other of them considers the case to be clear—it provides little support for a theory that suggests evidentiary uncertainty triggers the jury's retreat to sentiment. Ultimately, the strongest and most consistent predictor of jury acquittal in opposition to a judge's vote for conviction is a defendant without a criminal record. At least in the NCSC data, the evidentiary factor of a criminal history has a much stronger influence than jury sentiment in predicting judge–jury disagreements.

V. CONCLUSION

This study has explored the question "What accounts for judge and jury disagreement about guilt?" with a particular focus upon whether jury sentiments about the defendant and the law offer a significant explanation of the phenomenon. Kalven and Zeisel explained disagreements in part through the liberation hypothesis that juries, in close cases, resorted to sentiment in deciding whether the defendant was guilty. Empirical research over the past forty years has been unable to replicate these findings, offering mixed evaluations of the role to which jury sentiment as opposed to evidentiary factors explain cases in which the jury acquits defendants whom the judge would have convicted. To see whether we could find evidence of this phenomenon at work, we looked at the data from a four-jurisdiction survey of criminal trials, undertaken for the explicit purpose of analyzing hung juries but employing survey instruments and techniques permitting inquiry into a broader set of questions. While earlier research analyzing this data confirmed that contemporary juries, like those studied

by Kalven and Zeisel, were more acquittal prone than judges,[77] our analysis indicates that sentiments, as measured by the NCSC study, do not provide a particularly meaningful explanation of the tendency of the jury to acquit when the judge would convict. We find that contrary to the liberation hypothesis, which indicates juries will resort to sentiment when the evidence is close, jury sentiment toward the law has only has only a limited effect in those cases where the judge believes the evidence is close. Indeed, jury sentiment toward the law or the defendant is least likely to play a role in the very situation in which the theory suggests it would have greatest effect—when the jury finds the case close on the evidence.

The results presented here should not be completely surprising. There should be some reason that a jury refuses to decide a case in the way that the judge considers to be clear. It is interesting, but ultimately unsatisfying, to speculate as to why Kalven and Zeisel did not discuss disagreements in "clear" cases as separate categories. While they explicitly eschew the role of advocate for or against the jury system, their work stands as a powerful and reassuring testament to the justice of the jury system, as the many judicial citations to it testify.[78] Perhaps they considered that an extensive treatment of cases in which a jury's decision apparently flies in the face of evidence would have undercut their message. Whatever the reason, the data presented here reveals, as did Kalven and Zeisel's survey forty years earlier, that there are a small but persistent fraction of jury trials that result in verdicts that the judge considers to be counterfactual. The explanation for these disagreements, however, appears to be more complex than the liberation hypothesis suggests. In the face of evidentiary uncertainty, juries do not appear to retreat to sentiment; rather, a more complex process of weighing the value of evidence, including the criminal record of the defendant appears to explain such disagreements.

This conclusion requires a number of qualifications. First, since the NCSC study did not ask judges why juries disagree with them as to guilt, we do not know whether contemporary judges would offer the same explanation of disagreements that their counterparts offered in the 1950s. If the NCSC asked the same questions as Kalven and Zeisel, perhaps it would have received the same answers. Regardless, the NCSC data provides superior measures of the factors that actually affect jury decisions and the results presented here would still provide no support for the view that it is particularly when the jury considers a case close that it will turn to sentiment to decide the case. Additionally, the NCSC data is limited in a number of ways that constrain the generalizability of our findings. The data

---

[77] Eisenberg et al, *supra* note 11, at 181.

[78] *See supra* note 5 and accompanying text.

include a relatively small number of criminal trials occurring in only four jurisdictions, all of which are in urban settings. As a result, the findings from this study, and concerns raised about the role of jury sentiment, may not be generalizable to different types of locations.

It is possible, of course, that sentiment plays a significant role in leading a juror to consider the case close on the evidence but that the jurors who responded to the NCSC study were unaware of this subtle influence on their perceptions. They answered in light of the vote they had just cast, and they understood their vote in terms of resolving a factual dispute rather than permitting sentiment to outweigh the conclusion that the facts demanded. One could imagine that jurors might be more aware of the role of sentiment when they chose to acquit despite clear evidence of guilt. Accepting both of these premises, the findings presented here would not undermine the liberation hypothesis but would not provide support for the view that factual uncertainty opens a juror's mind to consider sentiment rather than the other way around. We cannot discount the possibility that the juror's failure to acknowledge the role of sentiment in the liberation outcome reflected their inability to understand their own emotions rather than the absence of those emotions. However, we can suggest that, if this is the case, it has yet to be demonstrated convincingly. In our view, the data rather supports the less exciting possibility that, in close cases, jurors acquit because they believe that, on the evidence presented, the state has failed to persuade them of the defendant's guilt. In fact, the evidentiary factor of a defendant without a criminal history is the strongest predictor of judge and jury disagreement in cases that both judges and juries indicate are close on the evidence.

The liberation hypothesis suggests that when juries disagree with judges, they do so out of emotion, invoked in those cases that are close on the evidence. This hypothesis has remained largely untested for over forty years and arguably has supported commonly held impressions about the unique role of the American jury—that in the face of evidentiary uncertainty, jurors employ personal and community sentiment to arrive at acquittals of defendants whom judges believe to be guilty. The findings presented here suggest that that when juries decide to acquit in cases in which judges would have convicted, they do not arrive at this decision because of sentiment, but rather do so because they evaluate the evidence differently than judges. Thus, we should understand differences between judges and juries not as evidence of the jury's "flight from the law," but rather as an indication that judges and juries rationally evaluate the same evidence but arrive at different conclusions based on their unique vantage points.

Ultimately, it is less important to understand why judges and juries disagree than to understand who is correct in those cases where they do

disagree. Even if juries are not moved by sentiment, it is still possible that judges do a better job than juries do at evaluating credibility and weighing the evidence. If so, it may be that the judge's conclusion as to guilt or innocence is the one that reflects the most accurate application of law to fact. There is no way to answer that question definitively. However, we can answer the question of whether the jury arrives at conclusions different from the judge in close cases because they are unusually susceptible to sentiment in such cases. The answer to that question (at least as measured by the data presented here) is "no." This finding has important implications for how we think about the meaning of acquittals. If courts and other criminal justice professionals are going to continue to treat jury acquittals as irrelevant to the question of whether the defendant committed the crime, they are going to need a firmer base than the belief that juries are moved by sentiment and judges by fact.