The University of New Hampshire Law Review

November 2016

The Results of Deliberation

Maggie Wittlin
University of Nebraska College of Law

Follow this and additional works at: https://scholars.unh.edu/unh_lr

Part of the Computer Engineering Commons, Courts Commons, Law and Society Commons, Rule of Law Commons, and the Sexuality and the Law Commons

The Results of Deliberation

MAGGIE WITTLIN[*]

ABSTRACT

When evaluating whether to sue, prosecute, settle, or plead, trial lawyers must predict the future—they need to estimate how likely they are to win a given case in a given jurisdiction. Social scientists have used mock juror studies to produce a vast body of literature showing how different variables influence juror decision-making. But few of these studies account for jury deliberation, so they present an impoverished picture of how these effects play out in trials and are of limited usefulness.

This Article helps lawyers better predict the future by presenting a novel computer model that extrapolates findings about jurors to juries, showing how variables of interest affect the decisions not only of individuals but also of deliberative bodies. The Article demonstrates the usefulness of the model by applying it to data from an empirical study of the factors that influence juror decisions in acquaintance rape cases. This application first elucidates a tension in criminal law: even if a substantial majority of jurors in a community would vote to convict a defendant, a majority of juries might still acquit. It also demonstrates that certain legal reforms will have a meaningful effect in some areas of the country but not others, suggesting that rape law reform should occur at a local, not national, level.

CONTENTS

INTRODUCTION

Trial lawyers frequently need to predict the future. "If I go to trial," the attorney must ask herself, "what are my chances of winning?" Prosecutors aim to maximize their conviction rates: to win frequently and to lose very rarely,[1] so a prosecutor will charge a defendant with a crime only if he believes there is a high probability a jury will convict. When a defense lawyer helps his client decide whether or not to accept a plea bargain, he engages in a complicated cost-benefit analysis, weighing the risk of conviction and the severity of a sentence against the prosecution's offer.[2] Both the prosecutor and the defendant "bargain in the shadow of the jury," deciding which plea offers to make and take based on the chances of a conviction.[3] Civil litigants face similar calculations. In deciding whether to sue in the first place and then whether to settle a case or proceed to trial, parties estimate their odds of victory in front of a jury.[4] Once the lawyer gets to trial, she endeavors to select a jury that gives her the greatest probability of success; and once the jury is empaneled, she tailors her trial strategy to maximize her chances of getting a favorable verdict from that specific jury.[5]

---

[1]   Daniel C. Richman, Old Chief v. United States: *Stipulating Away Prosecutorial Accountability?,* 83 VA. L. REV. 939, 966–69 (1997).

[2]   *See* Margareth Etienne, *Remorse, Responsibility, and Regulating Advocacy: Making Defendants Pay for the Sins of Their Lawyers*, 78 N.Y.U. L. REV. 2103, 2167 (2003).

[3]   *See* Michael A. Dawson, *Popular Sovereignty, Double Jeopardy, and the Dual Sovereignty Doctrine*, 102 YALE L.J. 281, 300 (1992).

[4]   *See* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 4, 7–9 (1984).

[5]   *See* Franklin Strier & Donna Shestowsky, *Profiling the Profilers: A Study of the Trial Consulting Profession, Its Impact on Trial Justice and What, If Anything, To Do About It*, 1999 WIS. L. REV. 441, 453–55 (1999) (discussing the work of trial

Lawyers have a number of tools to evaluate their likelihood of success. They can review the outcomes of similar cases that they have tried, observed, or researched and calculate the fraction won by the party in their position.[6] They can attempt an objective comparison of the evidence of the defendant's wrongdoing to the legal decision standard.[7] And they can predict behaviors of individual jurors, either by looking at studies of how prospective jurors might vote in a similar case or relying on their own (or their consultant's) understandings of which jurors are likely to vote which way.[8]

In this Article, I introduce a new tool and a new way to think about the probability of success: a computer simulation that uses models of *individual juror* voting to predict how a *jury* randomly drawn from a specified community will come out in a given case. Psychologists and legal scholars have produced a large body of "mock-juror" studies that examine how variables of interest influence trial verdicts.[9] But these studies rarely involve actual juror deliberation; more often, the researchers simply present individual subjects with questions about how they would vote in a particular case. A program that extrapolates these results from individuals to deliberative bodies allows lawyers to extract new and valuable information from these studies: it allows them to predict the verdict of a jury drawn from a pool that varies with respect to the variables of interest. If a study models how different types of people react to differences in case strategy, a deliberation simulation could predict chances of victory under different strategies.

Lawyers are not the only ones who can gain useful knowledge from the program. Social scientists who conduct mock-juror studies may strengthen their conclusions and glean new implications of their research by observing how findings about individuals play out in a group context. Scholars frequently qualify the ecological validity of their studies by noting that deliberation may alter their results.[10] The program mitigates that nearly

---

consultants).

   [6]   *See* Stuart S. Nagel, *Lawyer Decisionmaking and Threshold Analysis*, 36 U. MIAMI L. REV. 615, 618 (1982); Robert J. Rhee, *A Price Theory of Legal Bargaining: An Inquiry into the Selection of Settlement and Litigation Under Uncertainty*, 56 EMORY L.J. 619, 673 (2006).

   [7]   *See* Peter Siegelman & John J. Donohue III, *The Selection of Employment Discrimination Disputes for Litigation: Using Business Cycle Effects To Test the PriestKlein Hypothesis,* 24 J. LEGAL STUD. 427, 461–62 (1995).

   [8]   *But see* VALERIE P. HANS & NEIL VIDMAR, JUDGING THE JURY 76 (2001) (quoting Alan Dershowitz as saying, "Lawyers' instincts are often the *least* trustworthy basis on which to pick jurors. All those neat rules of thumb, but no feedback. Ten years of accumulated experiences may be ten years of being wrong.").

   [9]   *See infra* Part II.A.

   [10]   *See, e.g.*, Daniel Krauss & Nicholas Scurich, *The Impact of Case Factors on*

universal concern[11] and allows researchers to argue more forcefully for reforms based on their results.   Legislators and other policymakers accountable to voters will want to pass laws that achieve certain outcomes, including the conviction and sentencing of people who engage in conduct that is offensive to voters.[12]   Legislators, then, will be interested in drafting laws that not only nominally criminalize offenses in accordance with individual voter preference but also translate that preference into post-deliberation jury convictions.

   Although a number of researchers have used more basic computer simulations to mimic deliberation and gain valuable insights into existing research,[13] the program introduced here has several key advantages over earlier models.   First, I employ a recently developed, more nuanced formula for converting initial jury ballots into post-deliberation verdict probabilities.[14] In contrast to earlier schemes, which required the researcher to choose from

---

Jurors' Decisions in a Sexual Violent Predator Hearing, 20 PSYCHOL. PUB. POL'Y & L. 135, 143 (2014); Richard L. Wiener et al., *Anticipated Affect and Sentencing Decisions in Capital Murder*, 20 PSYCHOL. PUB. POL'Y & L. 263, 278 (2014); David L. Schwartz & Christopher B. Seaman, *Standards of Proof in Civil Litigation: An Experiment from Patent Law*, 26 HARV. J.L. & TECH. 429, 471 (2013); Emily C. Hodell et al., *Factors Impacting Juror Perceptions of Battered Women Who Kill Their Abusers: Delay and Sleeping Status*, 18 PSYCHOL. PUB. POL'Y & L. 338, 356 (2012); Harmon M. Hosch et al., *Effects of an Alibi Witness's Relationship to the Defendant on Mock Jurors' Judgments*, 35 LAW & HUM. BEHAV. 127, 139 (2011); Timothy R. Robicheaux & Brian H. Bornstein, *Punished, Dead or Alive: Empirical Perspectives on Awarding Punitive Damages Against Deceased Defendants*, 16 PSYCHOL. PUB. POL'Y & L. 393, 414 (2010); Tisha R. A. Wiley & Bette L. Bottoms, *Effects of Defendant Sexual Orientation on Jurors' Perceptions of Child Sexual Assault*, 33 LAW & HUM. BEHAV. 46, 57 (2009).

[11] The program does not eliminate the concern entirely. Some studies may involve issues where deliberation is likely to have more or less of an effect than in a generalized case.   *See, e.g.*, Schwartz & Seaman, *supra* note 10, at 471 (hypothesizing that deliberation might allay juror confusion).

[12] William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 530–31 (2001).

[13] *See, e.g.*, Joseph W. Filkins et al., *An Evaluation of the Biasing Effects of Death Qualification: A Meta-Analytic/Computer Simulation Approach*, 4 THEORY & RES. SMALL GROUPS 153, 165–71 (2002); Scott Tindale & Dennis H. Nagao, *An Assessment of the Potential Utility of "Scientific Jury Selection": A "Thought Experiment" Approach*, 37 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 409, 412–13 (1986).

[14] The model is derived from Professor Robert MacCoun's new BOP ("balance of pressures" or "burden of (social) proof") deliberation formula.   *See generally* Robert J. MacCoun, *The Burden of Social Proof: Shared Thresholds and Social Influence*, 119 PSYCH. REV. 345 (2012) (developing a deliberation framework based on the hypothesis that there is a shared sense of how much social opposition is necessary before a deliberating person should change views).

several rigid conversion formulas,[15] the formula used herein is flexible, allowing the verdict probability function to vary smoothly based on factors such as the type of case (criminal or civil), decision rule, and jury size. Second, unlike earlier programs, which assumed that all jurors had an equal probability of voting a certain way,[16] the program here more realistically simulates a probability for *each juror individually* based on that juror's characteristics and the researcher's data. In addition, the annotated code in the Appendix allows scholars without a coding background to run deliberation simulations in Stata[17] by tweaking the program to fit their data and question of interest.

    I demonstrate the implications of this program by applying it to a mock juror study of the factors that influence perception of consent in acquaintance rape cases where the woman claims she said "no" but there was no physical resistance. I demonstrate that a "no-means-no" law—a law providing that the word "no" defeats a reasonable perception of consent—would have a more meaningful impact in some locations than in others.[18] This result suggests that a national rape policy might not be ideal: certain states, those with more egalitarian values, could benefit from a "no-means-no" reform, while legislators in states with different cultural values should consider other ways of altering norms before passing this sort of law. I also show how prosecutors with differing levels of knowledge about community makeup could determine whether a jury in their jurisdiction is likely to convict in a case like the one studied. And my results offer a possible explanation for the low rate of prosecution in acquaintance rape cases: even if a majority of individuals in a community would convict, a majority of juries in that community may acquit.

    Part I of this paper reviews past efforts at jury deliberation models, discusses the deliberation framework used here, and develops a flexible computer program that simulates jury verdicts based on prior investigations of the variables that influence individual jurors. Part II discusses the

---

[15] *See generally* James H. Davis, *Group Decision And Social Interaction: A Theory Of Social Decision Schemes*, 80 PSYCHOL. REV. 97 (1973) (introducing the Social Decision Scheme framework, in which the researcher chooses from a small set of social decision matrices that convert initial votes to a probability of each possible verdict).

[16] *See* Kwangbai Park, *Estimating Juror Accuracy, Juror Ability, and the Relationship Between Them*, 35 LAW & HUM. BEHAV. 288, 302 (2011); *see also* Filkins et al., *supra* note 13; Tindale & Nagao, *supra* note 13, at 416 (assuming jurors are chosen from discrete groups with certain voting probabilities).

[17] Stata is a statistical software package that provides a number of tools for data analysis, including multiple linear regression. *See* STATA: DATA ANALYSIS AND STATISTICAL SOFTWARE, http://www.stata.com/ [https://perma.cc/29UG-PGZA] (last visited Sept. 24, 2016).

[18] *See infra* Part III.

implications of this program for the many scholars who have studied mock jurors and the institutional actors who could benefit from their research. Finally, Part III demonstrates the usefulness of the program by applying it to research on acquaintance rape. The Appendix includes sample code that, when manipulated, allows interested persons to specify community demographics and a model of juror voting and, given those specifications, estimate the likelihood of conviction in an acquaintance rape case.

I. MODELING DELIBERATION

In this Part, I discuss the history of deliberation modeling and develop a computer program that calculates how likely a jury from a specified community is to find for one side in a specified case. The model—even before it is applied to mock jury research—yields several interesting results. It reveals that in civil cases, deliberation augments the overall preference of the venire: if just over half of individuals would vote for the plaintiff, decidedly more than half of juries would do so. In criminal cases, deliberation augments deviations from some larger proportion, possibly about two-thirds, voting for conviction: if two-thirds is the threshold, and if 60% of individuals would initially vote for conviction, far fewer than 50% of juries will. What looks like a winning case from an individual perspective, then, may be a losing case to the eye of an experienced prosecutor. The program also shows that for close cases, deliberation exacerbates the differences between communities—if individuals in one community are a little bit more likely to convict than those in another, juries in the first community will be much more likely to convict than juries in the second. (For easy cases, deliberation mitigates the differences between communities in individual preference.) In close cases, then, jury verdicts amplify the unique voice of the community.

A. Efforts at Modeling Deliberation and Social Influence

With the knowledge that initial ballots are a strong but imperfect predictor of verdicts,[19] social scientists have been modeling group deliberation—predicting how initial votes will convert to verdicts—for decades.[20] The most influential system for modeling group deliberation and

---

[19] *See* HARRY KALVEN, JR. & HANS ZEISEL, THE AMERICAN JURY 487–90 (1966); Diane L. Bridgeman & David Marlowe, *Jury Decision Making: An Empirical Study Based on Actual Felony Trials*, 64 J. APPLIED PSYCHOL. 91, 94 (1979); Marla Sandys & Ronald C. Dillehay, *First-Ballot Votes, Predeliberation Dispositions, and Final Verdicts in Jury Trials*, 19 LAW & HUM. BEHAV. 175, 188 (1995).

[20] *See generally* Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 PSYCHOL. PUB. POL'Y & L. 622

decision-making, at least in the jury context, has been James H. Davis's 1973 Social Decision Scheme (SDS) framework.[21] SDS uses a social decision scheme—a group's express and implied rules of decision-making—to translate probabilities of individual preference arrangements into probabilities of final group choice.[22] For example, in a "majority rules" social decision scheme, the probability that choice A wins over choice B is 100% if more than half of the members prefer A, 50% if exactly half of the members prefer A, and 0% if fewer than half of the members prefer A.[23] Therefore, in a group of twelve people operating under a "majority rules" scheme, if there is a 50% chance that eight people prefer A, a 40% chance that six people prefer A, and a 10% chance that four people prefer A, there is a 70% chance overall that choice A wins. Different decision schemes result in different "decision matrices," which translate individual preference arrangement probabilities into group probabilities. For the social decision scheme "truth wins out," if even one group member prefers choice A, and choice A is correct, there is a 100% chance that the group will choose A.[24] SDS research has used several such schemes with predictable decision matrices—including "majority rules" and "truth wins out"—as "benchmarks" against which to measure actual group behavior.[25]

In his work, Davis suggested that researchers could use his framework in conjunction with data on group decision-making behavior to determine actual SDS matrices for decision-making bodies.[26] The SDS matrix would then yield insight into the groups' decision-making processes.[27] Indeed, since Davis published his paper, researchers have used the SDS framework to investigate questions of courtroom procedure, such as the effect of death qualification on the likelihood of conviction in capital cases[28] and the efficacy of "scientific jury selection."[29] Other researchers have expanded on the SDS framework in the jury context. For example, Norbert Kerr, Robert MacCoun, and Geoffrey Kramer derived what MacCoun calls the Kerr

---

(2001) (providing a thorough overview of research on deliberating juries through 1999).

[21]   *See* Davis, *supra* note 15.

[22]   *Id.* at 101.

[23]   *See* MacCoun, *supra* note 14, at 357 tbl.6.

[24]   *See id*.

[25]   *See id.* at 353, 357 tbl.6.

[26]   *See* Davis, *supra* note 15, at 114 (discussing the goal of using "all of the data, group and individual, to obtain an estimate of the social decision scheme matrix itself").

[27]   *See id* at 114, 123.

[28]   *See* Filkins et al., *supra* note 13, at 165.

[29]   *See* Tindale & Nagao, *supra* note 13, at 416.

Influence Model (KIM), an asymmetrical SDS matrix that accounts for judgmental bias.[30]

While SDS and its progeny continue to be the dominant framework for jury deliberation,[31] other models have focused on deliberation processes. For example Helmutt Crott and Joachim Werner's Norm-Information-Distance model[32] predicts the probability an individual will transition from one choice to another by taking into account the "relative subgroup size for the new choice, the informational attractiveness of that choice, and the number of alternatives intermediate between the original and the new opinion."[33] It then translates individual transition probability into group transition probability.[34]

Still other researchers have modeled jury deliberations using computer simulations. In a classic paper, Steven Penrod and Reid Hastie developed DICE, a jury simulation program that allows users to vary several parameters, including jury size, decision rule, and the initial probability that a randomly-selected juror will initially vote to convict.[35] With Nancy Pennington, they designed the JUS program, an advanced version of DICE that accommodated multiple verdict categories, including conviction of a lesser charge.[36]

Other social psychologists have formulated influence models in settings more general than the group deliberation context. Several of these models aim to estimate how likely it is that an individual will conform to pressure from an opposed group. Bibb Latané's Social Impact Theory derives a mathematical model from the established psychological premise that people have diminishing marginal sensitivity to stimuli.[37] Brian Mullen's Other-Total Ratio model stems from the hypothesis that when a person's ingroup is small compared to his outgroup, he will experience self-awareness, and when his ingroup is large compared to an outgroup, he will experience

---

[30]   *See* Norbert L. Kerr et al., *Bias in Judgment: Comparing Individuals and Groups*, 103 PSYCHOL. REV. 687 (1996).

[31]   Devine et al., *supra* note 20, at 625.

[32]   *See* Helmutt W. Crott & Joachim Werner, *The Norm-Information-Distance Model: A Stochastic Approach to Preference Change in Group Interaction*, 30 J. EXPERIMENTAL SOC. PSYCHOL. 68 (1994).

[33]   *See id.* at 68.

[34]   *Id.*

[35]   *See* Steven Penrod & Reid Hastie, *A Computer Simulation of Jury Decision Making*, 87 PSYCHOL. REV. 133, 134 (1980).

[36]   *See* REID HASTIE ET AL., INSIDE THE JURY 175–226 (1983).

[37]   *See* Bibb Latané, *The Psychology of Social Impact*, 36 AM. PSYCHOL. 343, 344 (1981). In Latané's model, each additional source of influence has less impact than the previous source; hence the decreasing marginal effect. *See also* MacCoun, *supra* note 14, at 349 (discussing social impact theory).

deindividuation.[38]   And Sarah Tanford and Penrod have derived the Social Influence Model from their computer simulations.[39]

A different group of researchers have posited logistic "social threshold models" for behaviors that take a binary form, where a person's probability of switching behaviors changes drastically as soon as the level of an opposing group crosses a threshold.[40]   These models were not designed to describe deliberation: Thomas Schelling developed a "tipping point" model of racial segregation, where a person moves if the proportion of her neighbors that are of a different race crosses a certain threshold, say, 50%.[41] For convenience, he assumed in his model that all people had the same internal threshold.[42]   Mark Granovetter developed a similar tipping point model, but allowed thresholds to vary between people; they were either uniformly or normally distributed.[43]   He applied his model in the context of rioting.

In a recent paper, MacCoun develops a family of new social threshold models that can be applied to jury deliberations.[44]   The deliberative models successfully fit both the SDS "benchmarks" and the results of empirical studies of social influence, including mock juror studies.[45]   MacCoun calls this framework the BOP—"burden of (social) proof" or "balance of pressures"—framework.   He shows that a number of previous models— including the Granovetter and Schelling threshold accounts and many of the classic Davis social decision schemes—can be subsumed within BOP as special cases.[46] The BOP framework forms the basis of my computer model.

MacCoun begins his derivation with the well-supported social psychological proposition that people are sensitive to social consensus

---

[38]   *See* Brian Mullen, *Operationalizing the Effect of the Group on the Individual: A Self-Attention Perspective*, 19 J. EXPERIMENTAL SOC. PSYCHOL. 295–98 (1983) (hypothesizing that when a person's ingroup is small relative to the outgroup, that person will be more likely to attempt to conform his behavior to perceived standards); *see also* MacCoun, *supra* note 14, at 349 (discussing Mullen's work).

[39]   *See* Sarah Tanford & Steven Penrod, *Social Influence Model: A Formal Integration of Research on Majority and Minority Influence Processes*, 95 PSYCHOL. BULL. 189, 193 (1984); *see also* MacCoun, *supra* note 14, at 349 (discussing the social influence model).

[40]   *See* MacCoun, *supra* note 14, at 346.

[41]   *See* Thomas C. Schelling, *Dynamic Models of Segregation*, 1 J. MATH. SOC. 143, 181 (1971); *see also* MacCoun, *supra* note 14, at 361 (discussing Schelling's work).

[42]   *See* Schelling, *supra* note 41, at 149.

[43]   Mark Granovetter, *Threshold Models for Collective Behavior*, 83 AM. J. SOC. 1420, 1427 (1978); *see also* MacCoun, *supra* note 14, at 361 (discussing Granovetter's work).

[44]   *See* MacCoun, *supra* note 14, at 345, 361.

[45]   *Id.* at 353 tbl.3, 355 fig.5, 357 tbl.6, 358 fig.6.

[46]   *Id.* at 361.

information and may change positions on an issue if they are faced with opposing social consensus.[47] While many factors, internal and external, may influence an actor's decision to switch positions, MacCoun excludes these factors from the core of the model, stripping it down to a model solely in terms of social influence.[48]

Like earlier threshold models, BOP posits that each person has a threshold of social opposition, and once that threshold is crossed, the person will be much more likely to change positions.[49] MacCoun labels a person's net resistance to social pressure, *b*, for "burden of social proof," and this parameter operates as an internal threshold.[50] If *b* is 0.5, and more than 50% of the group is trying to change a person's mind, that person will be likely to switch.[51] In the group context, *b* can be understood as an average internal threshold.[52]

A second parameter, *c*, operates as an index of "norm clarity." Mathematically, a higher *c* gives the s-shaped probability function a sharper slope (Figure 1), so that crossing the threshold has a more drastic effect on the probability that a person will switch positions. MacCoun discusses two possible psychological interpretations of *c*.[53] First, a low *c* could denote an internal lack of norm clarity: an individual's resistance to social pressure may be uncertain or unstable.[54] Second, a low *c* could signify variation in *b* across members of the deliberating body.[55] An explicit voting standard or standard of proof (such as "beyond reasonable doubt") is likely to increase *c*, because it will signal to members of the deliberating body an approximate value for *b*, the burden of social proof.[56]

---

[47] *Id*. at 346.

[48] *Id*.

[49] *Id.*

[50] *Id*. at 346–47.

[51] *Id.* at 346–48.

[52] *Id.* at 346–47.

[53] *Id.* at 348.

[54] *Id*.

[55] *Id.* In neural network research, a similar sigmoid function may represent "neural firing rate." As electric potential increases, the total firing rate for the neural network increases. The "slope parameter," equivalent to *c* in MacCoun's model, is the inverse of variance in underlying neuronal states: if there is very low variance, the function has a sharp threshold—cross it and all of the neurons start firing. If there is high variance, each increase in potential causes a few more neurons to start firing. This interpretation of the slope parameter is equivalent to interpreting *c* as variance in *b* between group members. *See* André C. Marreiros et al., *Population Dynamics: Variance and the Sigmoid Activation Function*, 42 NEUROIMAGE 147, 149–50 (2008).

[56] *See* MacCoun, *supra* note 14, at 348.

MacCoun combines his psychological premise—that people are responsive to social consensus—and the parameters *b* and *c* to expand from individuals to deliberative bodies**.** He derives a "bidirectional influence" (bBOP) model for the likelihood that one side, here labeled "Sources," will win in a deliberation; the model takes the form of a logistic function:[57]

$$p(SourcesWin) = \frac{1}{1 + e^{-c\left(\frac{S}{N} - b\right)}}$$

In the bBOP model, *S* is the number of sources—the number of group members initially voting for the Source side—and N is the total number of group members.[58] The probability that S will win is therefore dependent on the proportion of group members initially voting for S, the average internal threshold *b*, and the "norm clarity" *c*. The bBOP equation is depicted in Figure 1, with both a higher *c* (c = 20) and a lower *c* (c = 5).

---

[57] *See id*. at 347 tbl.1. For a more mathematically rigorous derivation of the BOP models, see *id.* at 347. MacCoun also demonstrates how BOP can be derived from strict and random utility approaches, both of which lead to a logistic choice model. *Id.* at 370 app. a. Kalven and Zeisel also found that the relationship between first ballot votes and the likelihood of a given verdict took a somewhat sigmoid shape. *See* Hans Zeisel & Shari Seidman Diamond, *The Effects of Peremptory Challenges on Jury and Verdict: An Experiment in a Federal District Court*, 30 STAN. L. REV. 491, 505 (1978).

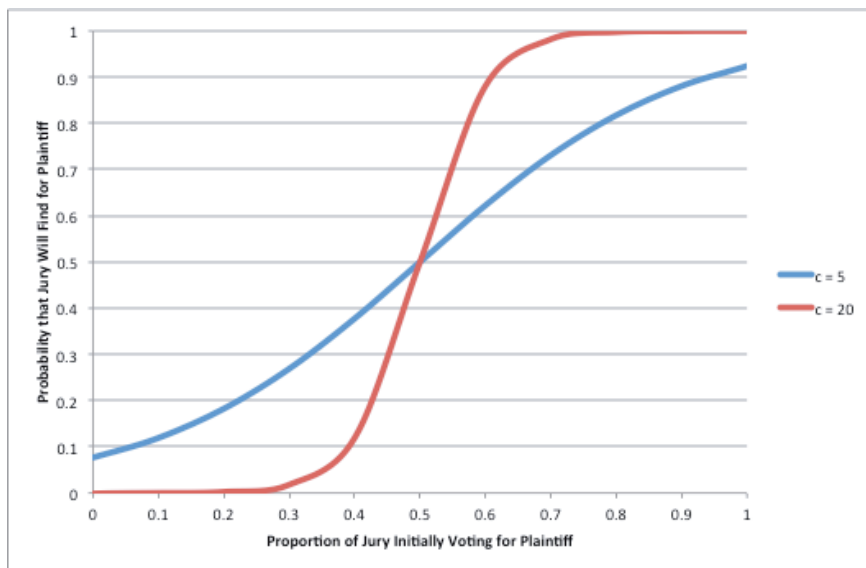[58] MacCoun, *supra* note 14, at 346.

Figure 1. MacCoun's bBOP deliberation model when $b = 0.5$, with two different values of c. When c is high, crossing the threshold of six jurors has a very sharp effect on the probability that a jury will convict. When c is low, the function is almost linear.[59]

With this bBOP model, MacCoun is able to replicate five of the most common SDS benchmarks perfectly or to close approximation.[60]   To replicate "majority rules"—where the initial majority is dispositive, and an evenly-split group is equally likely to come out either way—set b to 0.5 and set c to a very high value, say, 100.[61]   To replicate "truth wins"—where S wins as long as at least one person votes for it initially—keep c high but set b to 0.05.[62]

Additionally, the model can replicate the results of several social influence experiments, including MacCoun's own study of mock criminal juries with Norbert Kerr.[63] Fitting his data to bBOP, MacCoun finds that b is 0.62 and c is approximately 18.[64] It is unsurprising that b in a criminal case would be greater than 0.5, because the "beyond a reasonable doubt" standard

---

[59]   *See id.* at 350 fig.2.

[60]   *See id.* at 357 tbl.6.

[61]   *Id.*

[62]   *Id.*

[63]   Norbert L. Kerr & Robert J. MacCoun, *The Effects of Jury Size and Polling Method on the Process and Product of Jury Deliberation*, 48 J. PERSONALITY & SOC. PSYCHOL. 349, 355 tbl.2 (1985).

[64]   MacCoun, *supra* note 14, at 353 tbl.3.

favors those arguing for acquittal.[65]   A number of studies have found this "asymmetry effect,"[66] although some empirical work has questioned this leniency bias.[67] MacCoun hypothesizes that this evidentiary burden of proof translates into a social burden of proof, because when any substantial minority votes to acquit, others might begin to suspect that there is, indeed, reasonable doubt about the defendant's guilt.[68]

While MacCoun's model is founded on a sound psychological insight, and is backed by empirical research on group decision-making, it does have several limitations. First, it does not account *explicitly* for several factors that could influence group deliberation, even given an initial vote distribution, most notably the size of the jury and the decision rule. However, by adjusting the parameters b and c, a researcher can account for these factors. As noted above, to replicate the decision rule "majority rules," b should be set at 0.5 and c should be set to a high value: if one side has the edge after the first vote, that side inevitably wins. A criminal jury has to reach a unanimous decision that a defendant is guilty beyond a reasonable doubt.[69] To replicate a jury deliberating under this higher burden of proof, b should be set above 0.5—the burden of social proof falls on those advocating conviction—and c should perhaps be set to a somewhat lower value, because even one holdout has a chance of persuading her fellow jurors to come over to her side.[70] Still, c should likely not be *too* low, as studies of actual juries

---

[65]   *See* Robert J. MacCoun, *Modeling the Impact of Extralegal Bias and Defined Standards of Proof on the Decisions of Mock Jurors and Juries*, 46 DISSERTATION ABSTRACTS INTERNATIONAL 700B, 40 (1984).

[66]   *See id.*; Robert J. MacCoun & Norbert L. Kerr, *Asymmetric Influence in Mock Jury Deliberation: Jurors' Bias for Leniency*, 54 J. PERSONALITY & SOC. PSYCHOL. 21, 30 (1988).

[67]   *See generally* Norbert L. Kerr & Robert J. MacCoun, *Is the Leniency Asymmetry Really Dead?: Misinterpreting Asymmetry Effects in Criminal Jury Deliberation*, 15 GROUP PROCESSES & INTERPERSONAL RELATIONS 585 (2012) (discussing work that questions the leniency asymmetry but demonstrating some leniency effect, albeit one that may be less pronounced than originally believed); Paula L. Hannaford-Agor et al., *Are Hung Juries a Problem?* THE NATIONAL CENTER FOR STATE COURTS 67 fig.5.2 (2002), http://www.ncsc-jurystudies.org/What-We Do/~/media/Microsites/Files/CJS/What%20We%20Do/Are%20Hung%20Juries%20 A%20Problem.ashx (showing a slight severity asymmetry).

[68]   *See* MacCoun, *supra* note 14, at 354.

[69]   *In re* Winship, 397 U.S. 358, 361 (1970).

[70]   *See* Norman Rockwell, *The Holdout*, SATURDAY EVENING POST, Feb. 14, 1959, cover illustration. "What could be a most menacing scene is saved by the young woman's look. Her defiant crossing of the arms, cool demeanor, and upright posture leave little doubt that while she may feel harried, it is the beseeching and sermonizing male jurors who are wearing down and eventually must give way to her will." SCOTT E. SUNDBY, A LIFE AND DEATH DECISION: A JURY WEIGHS THE DEATH

indicate that 90% of trials produce the verdict initially favored by a majority,[71] and the "beyond a reasonable doubt" standard provides an explicit decision rule, which might increase norm clarity. Similarly, the model does not account for jury-specific or case-specific factors that may influence deliberation systematically, such as the race and gender composition of the jury[72] or the attractiveness of the defendant.[73] Again, if we both had and wished to employ concrete knowledge about how jury composition affects deliberation, we could code the program to adjust the parameters b and c automatically based on the composition of an individual, simulated jury.

The assumption in the model that group influence is directly proportional to relative group size may be more troublesome. Absolute jury size does, indeed, affect group deliberation, with smaller juries less likely to hang.[74] It seems likely that the parameter c might take a larger value in smaller groups, as it would be easier for jurors to come to a common understanding about an appropriate burden of social proof.

The bBOP model also does not account directly for influence that cannot be predicted from demographic data, such as the presence of more persuasive people on the jury, individual strength of conviction, or the style of deliberation.[75] But this does not pose a problem: like SDS, bBOP is a probabilistic model. These stochastic factors, unique to every jury, explain why one jury with ten people initially favoring conviction might acquit while another might convict. bBOP tells us, overall, how many of these juries will go one way and how many will go the other.

PENALTY 59–60 (2005).

[71]   *See* Devine et al., *supra* note 20, at 690.

[72]   *See* Jonathan M. Golding et al., *The Impact of Mock Jury Gender Composition on Deliberations and Conviction Rates in a Child Sexual Assault Trial*, 12 CHILD MALTREATMENT 182, 185, 188 (2007); 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, 606 (2006).

[73]   Richard R. Izzett & Walter Leginski, *Group Discussion and the Influence of Defendant Characteristics in a Simulated Jury Setting*, 93 J. SOC. PSYCHOL. 271, 275–76 (1974) (finding that before deliberation, mock jurors rendered a harsher punishment against an "unattractive" defendant, but they became more lenient after deliberation).

[74]   *See* Kerr & MacCoun, *supra* note 63, at 359–60 (finding that smaller juries are less likely to hang than larger juries and that very small groups use different deliberation processes than larger groups); *see also* Devine et al., *supra* note 20, at 669–70 (discussing research on jury size and deliberation).

[75]   *See* NEIL VIDMAR & VALERIE P. HANS, AMERICAN JURIES: THE VERDICT 143–44 (2007) (discussing the differences between "verdict-driven" deliberation and "evidence-driven" deliberation); Devine et al., *supra* note 20, at 701 ("Clearly, the evidence-driven style is closer to the normative ideal desired by the courts; in contrast, many juries adopt the verdict-driven style that seems most likely to lead to the rapid delineation of factions and steadily increasing normative pressure.").

Finally, the model contains only a probability of conviction and does not have a separate category for hung juries.[76] Although only a small proportion of juries hang—approximately 6 percent, according to the National Center for State Courts[77]—this somewhat limits the realism of the bBOP model.

Despite these limitations, MacCoun's model provides a workable, theoretically sound, and empirically supported model of jury deliberation. It incorporates the probabilistic nature of SDS research with the psychological insight of social threshold models, while maintaining flexibility to different decision rules and different distributions of internal thresholds within a group. In the next Section, I incorporate the bBOP formula into a computer simulation that calculates the likelihood that a jury from a specified community will find for each side in a case.

B. *A Computer Simulation of Jury Deliberations*

The heart of this paper is a new computer simulation—coded in Stata statistical analysis software—that uses information about what variables influence juror verdicts in a particular case in order to predict how likely a jury drawn from a specified community is to find for one side in that case. The Appendix contains annotated Stata code for the program when it is set up to calculate the likelihood that a jury will convict in an acquaintance rape case, where the woman said "no" but did not physically resist the defendant. This example is discussed in greater detail in Part III.

The program uses Monte Carlo simulations to produce a large sample of juries and have them deliberate according to the bBOP formula. Monte Carlo experiments are a class of computer algorithms that investigate the properties of physical or mathematical systems by drawing random samples from a specified domain, performing calculations on each sample, and analyzing the results of those calculations in aggregate.[78] For example, if you wanted to know the probability of winning a game of solitaire,[79] you might set up a program that draws a thousand random initial distributions of cards, plays through each game, recording whether the game was a win or loss, and reports the proportion of games won. This would be easier than trying to engage in a single analytic calculation of the probability of victory,

---

[76] *Accord* Kerr & MacCoun, *supra* note 67, at 595 (discussing fitting a bBOP model after dropping hung juries).

[77] National Center for State Courts, *A Profile of Hung Juries*, 9 CASELOAD HIGHLIGHTS 1 (2003), http://www.ncsc-jurystudies.org/~/media/Microsites/Files/CJS/What We Do/ caseload highlights hung juries.ashx [https://perma.cc/WN33-75LK].

[78] *See generally* Nicholas Metropolis & Stanislaw Ulam, *The Monte Carlo Method*, 44 J. AM. STAT. ASS'N 335 (1949) (developing the motivation for and description of the Monte Carlo Method).

[79] *See id*. at 336.

an intractable problem.[80]   Statistical simulations are useful not only for calculating point estimates but also for determining the confidence interval of an outcome.   After randomly generating 1000 outcome values that all represent a single unknown quantity, the program could find the 95% confidence interval for that quantity by ordering the outcomes lowest to highest and reporting the value of the 25th and 976th outcome.[81]

The program developed in this paper allows a user to simulate the probability that a jury will convict a defendant in a known case.   In its simplest form, which I call the Simple Simulation, or SimpSim, Program, it calculates this probability in five steps:

1.   The program specifies *V*, a population of jurors, or venire, with desired demographic characteristics (percent male, average income, etc.).
2.   From this venire, the program draws 12 jurors—each with his or her own characteristics.
3.   The program uses a user-specified statistical model to simulate an individual "first-ballot" verdict for each of the 12 jurors. This produces an initial number of jurors, N_C, that favor conviction.
4.   The program uses MacCoun's bBOP formula[82] to calculate the probability that a jury with N_C jurors initially favoring conviction will, indeed, convict the defendant.
5.   The program repeats steps 2 through 5 one thousand times, to simulate one thousand juries. It then averages the probabilities of conviction from all of these juries to find an overall probability that a jury drawn from *V* will convict.

SimpSim requires several user specifications.   First, the user must specify *a model that uses individual characteristics to predict how likely it is that a single juror will convict*[83] in the case. For example, if a researcher has gathered data from research subjects on age, sex, political affiliation, and whether the subject would convict a defendant in an acquaintance rape case, the researcher could ask the program to run a logistic regression on her data—with "convict" as the left hand variable and "age," "sex," and

---

[80]   *Id.*

[81]   *See* Gary King et al., *Making the Most of Statistical Analyses: Improving Interpretation and Presentation*, 44 AM. J. POL. SCI. 347, 349 (2000).   To report confidence intervals here, I use Stata's "centile" function, which estimates specified centiles.

[82]   *See supra* text accompanying note 57.

[83]   This model can be used for either civil cases or criminal cases, and in criminal cases, it can output either likelihood of conviction or likelihood of acquittal. I use "convict" here for convenience.

"political affiliation" as right hand variables—and use the regression results as her predictive model.

Second, the user must specify *the characteristics of the venire*. In the above example, the user would specify the percent of the population that is male, the average age and standard deviation of age, and percent of the population that is Republican, Democrat, Independent, or unaffiliated. The user could also specify correlations between these variables, to more accurately represent that older people are more likely to also be registered Republicans, for example.[84] Specifying the venire can be technically challenging.[85]

Finally, the user should specify *the parameters b and c in the bBOP model*. For a civil case, where the "preponderance of the evidence" standard gives neither side an advantage, *b* should be set to 0.5. In a criminal case, however, where the jurors who want to convict have the burden of proof, *b* should be set at a higher value. The precise threshold in a criminal case is contested. Dennis Devine et al. suggest that in laboratory studies, acquittal is all but inevitable if seven jurors or fewer favor conviction, and conviction is highly likely if at least ten jurors favor conviction; if eight or nine jurors favor conviction, the case is a toss-up.[86] This would mean the threshold *b* is between 0.67 and 0.75. When MacCoun fitted his model to his own mock juror studies, he found that the best-fitting *b* was 0.62.[87] Recently, Kerr and

---

[84] The program, as written, does not account for jury selection and peremptory challenges. If a user knew which classes of jurors would be acceptable or unacceptable to the lawyers in the case, he or she could alter the venire accordingly. *Cf.* Zeisel & Diamond, *supra* note 57 (finding that *voir dire* will sometimes, but not always, significantly affect the jury verdict).

[85] If the user does not wish to specify correlations between variables, specifying the venire is easy. For an 80% white, 20% black jury, each juror would draw a random number between 0 and 1. If the number is less than 0.8, the juror is assigned "white." If it is greater than 0.8, he is assigned "black." This is how I have created venires for specified cities discussed *infra*—Binghamton, Bozeman, Berkeley, and Oxford—and for the cultural model, see *infra* Figure 7.

For the remaining models, I have drawn variables from a normal distribution with specified means and a correlation matrix derived from a nationally-representative survey. I then convert those continuous variables into categorical or binary variables. The Appendix shows an example of this conversion. In each instance, I have verified that the resulting population has, on average, the characteristics I wished to specify, by simulating large populations from the specified distribution and observing summary characteristics.

[86] *See* Devine et al., *supra* note 20, at 692. Devine and his coauthors question the ecological validity of this leniency asymmetry, suggesting that it is insufficiently supported in studies of real juries. *Id*. at 692–93. Kerr and MacCoun's reanalysis of real-world jury data indicates that acquittal bias still likely exists but may not be as strong as it is in mock juries. *See* Kerr & MacCoun, *supra* note 67, at 599.

[87] *See* MacCoun, *supra* note 14, at 353 tbl.3.

MacCoun investigated this "leniency bias," looking at recent field studies, and suggested it might be weaker, or at least more variable, than these studies suggest.[88]   In my own simulations of criminal cases, I use *b* equal to 0.67 unless otherwise noted.[89]   Because this value is highly contested in the literature, the results of the model presented here are primarily illustrative.

The level of norm clarity, *c*, in juries is also non-obvious.   Several factors, such as a unanimous decision rule, would counsel toward having a low c, while others, such as an explicit decision rule like "beyond a reasonable doubt" point to a high degree of clarity.   In MacCoun's analysis of deliberation research, he found *c*s ranging from approximately 2 to 18.[90] The *c* for his study of mock criminal juries was at the high end of that range.[91]   Therefore, unless otherwise noted, I use a *c* of 18 in my own models.   When I discuss my findings, I will also discuss the implications of varying *c*.

SimpSim should accurately produce a point estimate for the likelihood that a jury from a specified venire will convict in a certain case.   However, the algorithm will not accurately calculate the standard error of that estimate. The logistic model that, for each juror, produces a probability of initially voting for conviction will have some error associated with it.   If the model came from a regression on a small sample, or if the model for some other reason has uncertain regression coefficients, there is little reason to be confident in the point estimate it produces.   However, SimpSim uses the logistic model to deterministically calculate the probability of conviction for each juror, without attributing any error to that probability estimate. Therefore, the only error in the entire algorithm that will show up in the final point estimate will be sampling error—the error associated with drawing only a small sample of all possible jurors and juries from the venire.   By increasing the number of juries sampled, a user can shrink that error infinitely.   This should not be: if the regression model is poor, that error should emerge in the program's output.

An altered version of the computer program, which I will call JurySim, allows modeling error to show up in the standard error of the final point estimate.

One way to think about the quality of a regression model is the uncertainty of the parameter estimates, including estimates of the regression

---

[88]   *See* Kerr & MacCoun, *supra* note 67, at 598.

[89]   With a high *c*, the sigmoid function has a fairly steep slope, see *infra* Figure 2, so the probability a jury will convict is somewhat sensitive to the specific choice of *b*.   Additional research on actual criminal juries that pinpoints the precise value of *b* would greatly help the accuracy of this model.

[90]   *See* MacCoun, *supra* note 14, at 353 tbl.3.

[91]   *See id.* (noting *c* for Kerr & MacCoun, *supra* note 63).

coefficients.[92] If uncertainty about the coefficients is low, the model reveals something meaningful about the relationship between the left hand variable and the right hand variables. If the coefficients are very uncertain, this means there is a wide range of possible "true" coefficients, and we have no clear picture of how much the left hand variables and right hand variables relate. A computer program could therefore account for modeling error if it accounted for coefficient uncertainty, incorporating the range of possible coefficients on each variable into the algorithm.

Statistical simulation can do just this. As noted above, simulation allows an analyst to estimate a 95% confidence interval by looking at the 2.5[th] and 97.5[th] centile of results in a simulation of 1000 juries drawn from a specified distribution. Instead of just drawing 1000 juries from a distribution specified by our chosen venire, however, we could draw *1000 models* from a distribution specified by the coefficient estimates and variance-covariance matrix[93] produced by a logistic regression. Clarify, a statistical application designed by Harvard government professor Gary King, simulates 1000 sets of parameters, including regression coefficients, every time it runs a regression.[94] By running SimpSim once on each of 1000 simulated models, obtaining one overall probability of conviction for each simulated model, a researcher could determine a 95% confidence interval for the probability estimate. The 25th smallest probability estimate would form the confidence interval's lower bound, and the 25th highest estimate would delineate an upper bound.

Simulating 1000 juries 1000 times is unnecessarily resource intensive. Stata would take hours to run the program, and comparable statistical power can be achieved with a smaller sample. The number of models should be enough to fill out the full distribution of simulated coefficients, so the model accounts for the full uncertainty of the regression. The number of sampled juries should be enough to fairly represent the population. The standard error from sampling is:

$$SE = \sqrt{\frac{(p)(1-p)}{N}}$$

Assuming an overall probability estimate of 0.5, the error attributable to sampling with 200 jurors per model is 3.5% for that model. For the 40,000

---

[92] *Cf.* King et al., *supra* note 81, at 348–49 (discussing types of uncertainty in the parameters).

[93] *Id.* at 349. The variance-covariance matrix captures the extent to which the parameters, including the regression coefficients, vary, and also the extent to which they vary with each other.

[94] *See* Michael Tomz et al., *Clarify: Software for Interpreting and Presenting Statistical Results*, J. Stat. Software, June 2001, at 5–6 (2001), http://gking.harvard.edu/files/clarify.pdf [https://perma.cc/G395-676M].

juries produced by running 200 juries for each of 200 models,[95] the sampling error is 0.25%. This is very small compared to the uncertainty on the coefficients of nearly any regression model of human behavior and can usually be ignored. The JurySim model, then, is able to produce both accurate point estimates and accurate confidence intervals.

C. Generalized Results

While the most interesting results of the JurySim model come from its application to actual mock juror research, the SimpSim model yields useful general information about how differences in venires interact with the deliberation process. The SimpSim model can begin to tell us under what circumstances deliberation will exacerbate differences between communities and when it might mitigate those differences. Although much of this information is inherent in the logistic shape of the bBOP curve itself, SimpSim serves as a clarifying tool, expressing these results in a clear, accessible way.

1. Civil Cases

Say we have three communities, A, B, and C, and three identical civil cases, say, a false advertising suit[96] against a drug company. In Community A, 60% of the population is initially inclined to find for the plaintiff; in B, 70% is so inclined; and in C, 80% of jurors would find for the plaintiff. In civil cases, the parameter b is set at 0.5. The probability of conviction for a jury drawn from any one of these communities therefore depends only on (1) the proportion of the venire who would initially vote to convict, and (2) the value of the parameter c.[97] SimpSim generates the results (Figure 2; Table 1).

---

[95] Generating 40,000 individual model-jury pairs would achieve similar result. I run multiple juries on a single model to highlight that modeling error and sampling error are distinct, and both remove precision from the estimate.

[96] *See* 15 U.S.C. § 1125 (2012).

[97] This is an analytically tractable problem, so we do not actually *need* a simulation to solve it. Using the binomial distribution and the probability that each juror will vote for the plaintiff, p, we can determine what proportion of juries will have each initial vote count—from 0-12 to 12-0. Then, using the bBOP formula, we can determine, for each possible initial vote count, what proportion of juries with that initial vote count will ultimately find for the plaintiff. The total proportion of the juries that will find for the plaintiff is the sum of the products of these two numbers:

$$Proportion\ Pro\ P = \sum_{j=0}^{12} \binom{12}{j} p^j (1-p)^{12-j} * \frac{1}{1 + e^{-c\left(\frac{j}{12}-b\right)}}$$
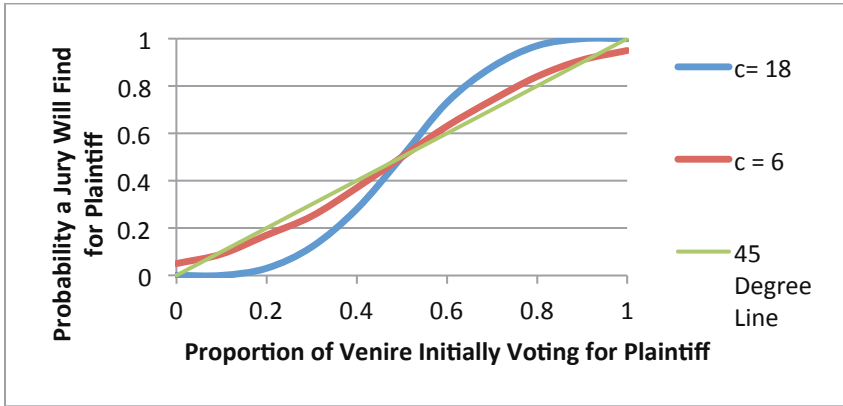
Figure 2. The probability that a jury will find for the plaintiff, given the proportion of the population initially prone to vote for the plaintiff, at two values of c. For both, $b = 0.5$.

Community	Proportion of Community Pro-P	Proportion of Juries Pro-P ($c = 6$)	Proportion of Juries Pro-P ($c = 18$)
A	0.60	0.63	0.73
B	0.70	0.74	0.88
C	0.80	0.84	0.97

Table 1. The proportion of each community that favors the plaintiff and the proportion of juries from each community that find for the plaintiff.

See Eric W. Weisstein, *Binomial Distribution*, *in* MATHWORLD—A WOLFRAM WEB RESOURCE, http://mathworld.wolfram.com/BinomialDistribution.html [https://perma.cc/2MAC-86PH] (last visited Oct. 21, 2016).

Figure 3. The proportion of individuals and juries that favor the plaintiff in each community.
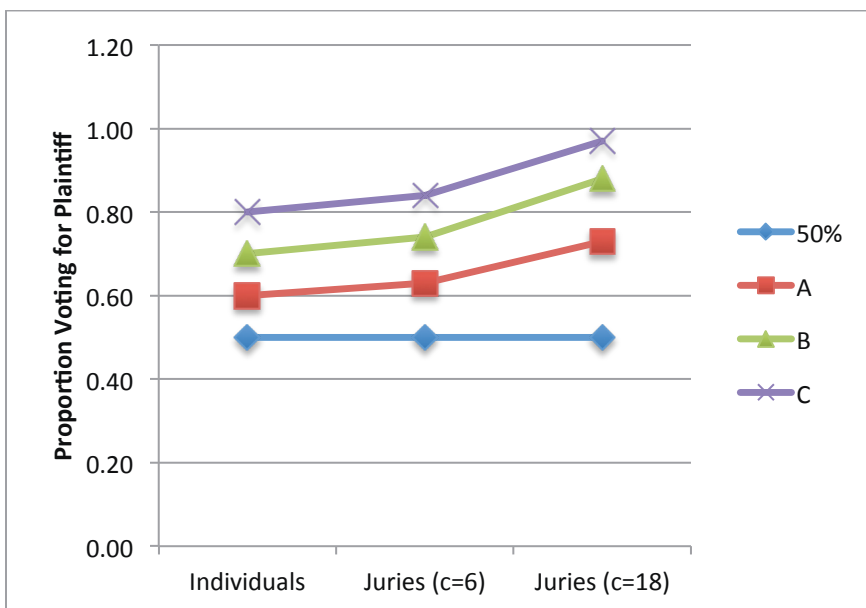
The deliberation process slightly augments the population's preference when c is 6, and it exacerbates the population's preferences to a greater degree when c is 18. When the population is disposed toward the plaintiff, as these three communities are, an even higher percent of juries from that community will find for the plaintiff.

More interestingly, the deliberation process sometimes augments and sometimes attenuates differences *between* communities. We know that if the community is split evenly in a civil case, juries will be evenly split as well.[98] Deliberation therefore augments the difference between a 50% pro plaintiff community and A, a 60% pro plaintiff community, and the augmentation is greater when c is higher. (In the $c = 18$ case, a difference of ten percentage points in individual preference turns into a 23 percentage point difference in jury preference.) But when initial preference is far from the 50% mark, deliberation may mitigate differences between juries. In the $c = 18$ case, a ten percentage point difference between individuals in communities B and C turns into a nine percentage point different between juries in those

---

[98]   Because the threshold parameter b is 0.5, juries split evenly will have a 50% chance of finding for the plaintiff. The function is symmetric around this inflection point, so an evenly split population—where a 4-8 initial vote is just as likely as an 8-4 initial vote—has no greater chance of producing a pro-plaintiff jury than a pro-defendant jury, or vice versa.

communities. And because 97% of juries in C convict, clearly the difference between juries in community C and juries in a community where 90% of individuals are pro-plaintiff cannot be greater than three percentage points.

Deliberation may therefore either accentuate or mitigate differences in population preferences depending on how close to the threshold of 50% both communities initially are. Around 50%, differences are accentuated, and far from 50% differences are mitigated. This means that the most divisive cases—those where the public is evenly split—will appear even more divisive if they are allowed to play out in the courtroom. The extent to which they are augmented or mitigated depends on the value of c, the norm clarity parameter.

2. Criminal Cases

The main structure of the bBOP model is the same for criminal cases as for civil.[99] However, the threshold parameter b is no longer set at 0.5. While the precise value of b in a civil case is disputed, it is almost certainly greater than 0.5 and less than 0.75.[100] I set b at 0.67.
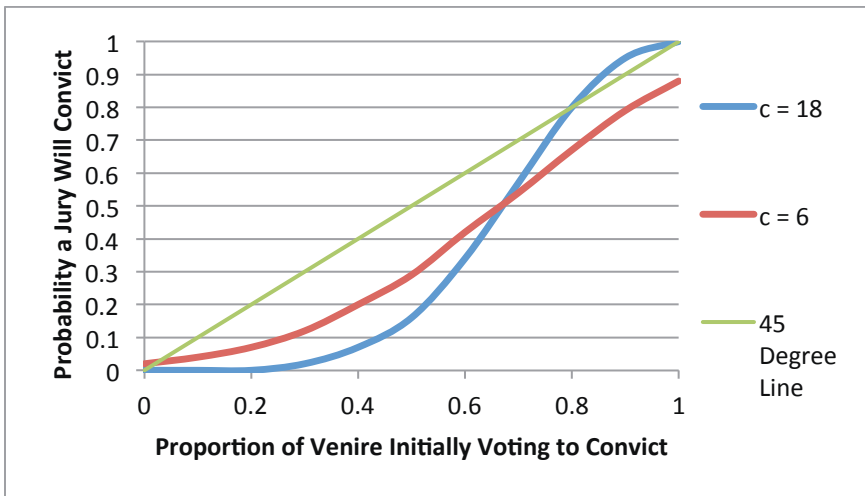

Figure 4. The probability that a jury will convict, given the proportion of the population initially prone to convict, at two values of c. For both lines, $b = 0.67$.

---

[99] *See* MacCoun, *supra* note 14, at 354–55 (discussing social thresholds under both a "beyond a reasonable doubt" and a "preponderance of the evidence" standard).

[100] *See supra* Part I.B.

The main difference between criminal and civil cases is that in criminal cases, differences between communities are exacerbated to the extent the percentage of people in the community who would initially vote to convict is close to this elevated *b*, here 67%. The closest cases in terms of population preference—those that divide the community evenly—are not the closest in deliberation when *b* exceeds 0.5. In the *c* = 18 scenario, if a population is evenly divided on a criminal case, only 16% of juries will convict. And moving from a community where 45% of individuals would initially vote to convict to a community where 40% would initially vote to convict brings the juror conviction rate down from 10.5% to 6.2%, slightly mitigating the difference between individual preferences in those two closely-divided communities.

Two-thirds of the community will need to favor conviction for there to be even a 50% chance that a jury drawn from that community will convict. If a prosecutor requires a much higher chance of victory before he is willing to risk a loss on a case, he will probably want upwards of 80% of the population initially favoring guilt, which will give him a comparable chance of winning the case at trial. What the population at large sees as a "close case" can differ drastically from what a prosecutor sees as a close case.

Prosecutors, having seen many cases, probably have a good sense of whether a case falls to the right of the inflection point, giving them a highly likely win, or to the left of the inflection point, giving them a likely loss. Prosecutors will likely disagree with each other only in a narrow range, right around the inflection point of 67 percent. Slight shifts in the proportion of the venire inclined to convict translate into great differences in the overall probability of conviction.

The next Part steps away from these generalizations to examine how the bBOP model and JurySim could help researchers who have studied the effects of variables of interest on individual jurors. But it will be useful to keep in mind the generalized findings of this Part: when the proportion of a large community that initially favors conviction is close to the bBOP threshold, deliberation will exacerbate differences between sub-communities; and while a criminal case that divides a community evenly may seem like a close case to the people in that community, a prosecutor is likely to recognize the case as a losing bet.

II. GETTING MORE OUT OF STUDYING INDIVIDUALS

In an ideal world, researchers interested in the factors that influence jury decision-making would bring full mock juries into the lab and have them deliberate after viewing or reading about a trial. However, full mock jury studies are expensive and difficult to organize, so researchers rarely study

juries engaged in live deliberation.[101]   Instead, a large body of research has studied mock jurors, individual subjects who are presented with written or visual material and asked to make decisions about how they would vote on a jury. These studies have examined juror responses to everything from the weight jurors afford eyewitness testimony[102] to factors in corporate behavior that influence whether jurors in a civil case will impose punitive damages.[103] While several scholars have expressed concern about the ecological validity of jury simulations that do not allow for deliberation,[104] researchers continue to examine individual mock jurors to draw conclusions about behavior that, in the real world, always has a deliberative component.[105]

The JurySim algorithm gives scholars a new way to understand findings about individual decision-makers when those individuals will actually deliberate before making decisions. Although studies of jurors rely on the intuition that deliberative bodies from a community will mirror individuals from that community, JurySim breaks that intuition by demonstrating how juries sometimes diverge from jurors. It allows us to extrapolate studies of individuals to deliberative bodies in a more realistic, telling way. This Part

---

[101] *See* Samuel R. Sommers & Phoebe C. Ellsworth, *How Much Do We Really Know About Race and Juries? A Review of Social Science Theory and Research*, 78 CHI.-KENT L. REV. 997, 1022 (2003). *But see* Devine et al., *supra* note 20, at 627–65 tbls.1–5 (collecting studies where mock jurors deliberated, along with studies of actual juries).

[102] *See, e.g.*, Paul Skolnick & Jerry I. Shaw, *A Comparison of Eyewitness and Physical Evidence on Mock-Juror Decisionmaking*, 28 CRIM. JUST. & BEHAV. 614, 621-23 (2001).

[103] *See, e.g.*, W. Kip Viscusi, *Corporate Risk Analysis: A Reckless Act?*, 52 STAN. L. REV. 547, 552–59 (2000) (discussing a survey that examined whether mock jurors were more or less lenient on corporations that had conducted cost-benefit analysis on a product prior to an accident).

[104] *See, e.g.*, Brian H. Bornstein, *The Ecological Validity of Jury Simulations: Is the Jury Still Out?*, 23 LAW & HUM. BEHAV. 75, 75–76 (1999); Shari Seidman Diamond, *Illuminations and Shadows from Jury Simulations*, 21 LAW & HUM. BEHAV. 561, 564–65 (1997).

[105] *See, e.g.*, Michael D. Cicchini & Lawrence T. White, *Truth or Doubt? An Empirical Test of Criminal Jury Instructions*, 50 U. RICH. L. REV. 1139, 1162–63 (2016) (discussing lack of deliberation as a limitation but concluding that "having mock jurors deliberate before rendering a verdict is not likely to change the observed pattern of verdicts across conditions"); Casey L. Magyarics et al., *The Impact of Frequency of Behavior and Type of Contact on Judgments Involving a Criminal Stalking Case*, 39 LAW & HUM. BEHAV. 602, 611 (2015) (acknowledging this limitation and noting that Diamond, *supra* note 104, found "individual jurors' beliefs often reflect the entire jury's decision."); Jeremy W. Bock, *Does the Presumption of Validity Matter? An Experimental Assessment*, 49 U. RICH. L. REV. 417, 451 (2015) (recognizing that lack of deliberation might affect the results of mock juror study and suggesting future work ask mock jurors to deliberate).

explores mock juror studies and discusses how computer simulations could extract useful information from existing research.

A.   How Informative Are Mock Juror Studies?

In most mock juror studies, subjects do not deliberate with each other before deciding on a verdict—they simply express an individual preference.[106]   In *The American Jury*, Harry Kalven and Hans Zeisel report interviewing jurors after 225 trials and finding that, "with very few exceptions the first ballot decides the outcome of the verdict."[107]   "And if this is true," they argue, "then the real decision is often made before the deliberation begins."[108]   MacCoun's bBOP formula and my resulting computer simulations do not contradict this generalization.   If *c*, the function's slope parameter, is high, and if *b* is close to 0.5, the vast majority of final verdicts will accord with the first ballot.   Because *b* is greater than 0.5 in criminal cases, juries in which a slight majority initially favor conviction are likely to acquit in the end.[109]

The conclusion that final verdicts usually follow first ballots means that we can gain useful information from learning how individuals will vote.   If most individuals on a jury will favor the plaintiff, given certain evidence, that jury, too, will likely favor the plaintiff.   Even so, mock juror studies still leave us with several open questions that can be answered by computer simulations of deliberation.   First, there is the simple question of extrapolating "percentage of people" to "percentage of juries."   In a community where 40% of the population would initially vote for the plaintiff, only 34% of juries would have at least six jurors initially voting to convict, and less than 16% of juries would have at least seven.   In a "majority rules" decision scheme, then, only about 25% of juries would find for the plaintiff.[110]   Second, how much more likely is a jury to find for the plaintiff

---

[106] *Cf.* David Alan Sklansky, *Evidentiary Instructions and the Jury as Other*, 65 STAN. L. REV. 407, 432 (2013) ("Some of the mock jury experiments on evidentiary instructions assign the subjects to jury panels and have them deliberate before reaching their final decisions. But most do not.").

[107] HARRY KALVEN, JR. & HANS ZEISEL, THE AMERICAN JURY 487, 488 (1966) (emphasis omitted).

[108] *Id*. at 488 (emphasis omitted).

[109] *See* MacCoun & Kerr, *supra* note 66, at 30 ("factions favoring acquittal are more influential than comparably sized factions favoring conviction").

[110] These percentages can be calculated either analytically or via computer simulation.   To calculate the probability that exactly six jurors will vote for the plaintiff, when 40% of the population would initially vote for the plaintiff, use the formula:

$$P(X = 6) = \frac{12!}{6!(12-6)!}(0.4)^6(1-0.4)^{(12-6)}$$

when eight members initially vote that way than when seven members initially vote for the plaintiff? Since juries don't actually vote on a "majority rules" scheme, it is useful to know which cases have a significant chance of flipping during deliberation and which do not. This is particularly relevant in criminal cases where "close" cases usually come out as acquittals.

Third, and most informatively, simulations allow a researcher to define a complex community with specified distributions of characteristics that affect individual juror votes. For a study that examines the influence of a single, binary variable on juror voting, it may be possible to calculate analytically the effect of that variable on an expected jury vote.[111] For a more complicated study that models juror behavior based on traits that are distributed both within and between populations, a simulation more easily allows a researcher to determine how a certain case would come out in a specified community or how much difference a procedural intervention would make in a given community.

While simulations can add information to mock juror studies, those studies inevitably have shortcomings unrelated to lack of deliberation, and these shortcomings will not be solved by computer simulation. For example, one study found that when mock juries hearing a school disciplinary proceeding knew they were participating in a study, they behaved differently from mock juries who believed they actually had the power to expel a student.[112] Other researchers have cited concerns such as the mock juror sample (often subjects are undergraduates), the presentation of trial evidence (often subjects read summaries instead of witnessing trials), and the type of outcome variable (often subjects report a dichotomous judgment where a probability-of-guilt estimate might be more informative).[113] Although research has shown little difference between mock juror studies using different juror samples or trial media,[114] these elements of ecological invalidity counsel against relying on mock juror—or even mock jury—studies as perfect predictions of real-world behavior. Still, by isolating the influence of variables of interest on mock juror verdicts, these studies allow us to forecast how real juror verdicts may vary across different conditions.

See Weisstein, *supra* note 97.

[111] Without Clarify's model simulations, however, such an estimate would not contain an accurate confidence interval.

[112] *See* David L. Breau & Brian Brook, *"Mock" Mock Juries: A Field Experiment on the Ecological Validity of Jury Simulations*, 31 LAW & PSYCHOL. REV. 77, 83, 87 (2007).

[113] *See* Bornstein, *supra* note 104, at 75–76.

[114] *See id*. at 88 ("few differences have been found as a function of either who the mock jurors are or how the mock trial is presented").

B.   *Studies That Could Benefit from Simulation*

A number of studies have looked at how certain variables influence the likelihood that jurors will vote to convict a defendant. One study showed that mock jurors weigh eyewitness testimony more than hearsay evidence.[115] The authors found that after watching a videotape of a trial, 62% of subjects who were presented with eyewitness testimony voted to convict, where only 40% of subjects who heard hearsay evidence would have convicted.[116] With circumstantial evidence alone, 36% of subjects voted to convict.[117]   The authors argued that their results may support hearsay reform.[118] A similar study found that jurors are influenced more by physical evidence than by eyewitness testimony, and found that jurors were more likely to convict after learning of damning physical evidence (84%) than after reading about eyewitness testimony (67%).[119] These researchers could expand their finding from individual jurors to juries, setting *b* to an appropriate value and finding the overall probability of conviction in their study cases under the two conditions.   Authors who support hearsay reform could bolster their recommendation by showing how in a case like the one presented, where a decided minority of subjects would favor conviction based on circumstantial evidence alone, the addition of hearsay evidence would not drastically change the number of juries who would convict; that number would remain low.

Simulations can allow researchers to extrapolate findings of studies that evaluate the effects of personal characteristics of jury members, such as race,[120] from individuals to populations with specified racial distributions. Samuel Sommers found that white jurors' initial votes are influenced not only by their own race and the race of the defendant, but by the races of their fellow jurors.[121]   A simulation could include this effect of the racial

---

[115] *See* Peter Miene et al., *Juror Decision Making and the Evaluation of Hearsay Evidence*, 76 MINN. L. REV. 683, 691–92 (1992).

[116] *Id.*.

[117] *Id.* at 692.

[118] *See id.* at 699–700. (suggesting their findings imply that jurors generally do not overvalue hearsay, and while more research is needed before hearsay reforms are implemented, their results at least have implications for harmless error analysis).

[119] *See* Skolnick & Shaw, *supra* note 102, at 622; *cf.* Brian H. Bornstein et al., *Intuitions about Arousal and Eyewitness Memory: Effects on Mock Jurors' Judgments*, 32 LAW & PSYCHOL. REV. 109, 120 (2008) (finding that mock jurors' beliefs about whether arousal helps subsequent memory retrieval interact with an eyewitness's reported arousal level when they evaluate eyewitness testimony; also finding that a positive main effect exists for eyewitness arousal and perception of credibility).

[120] *See, e.g.*, Sommers & Ellsworth, *supra* note 101, at 1002.

[121] *See id.* at 1028 (discussing Samuel R. Sommers, *Race and Juries: The Effects*

composition of the jury in its prediction of individual juror first ballot votes. It could thereby help the researchers demonstrate how racial bias could affect criminal trials in different communities.

Simulations based on the bBOP deliberation formula are not useful for mock juror studies without dichotomous outcomes. MacCoun's threshold model predicts only the probability that an individual will switch from one of two choices to the other.[122] So while a number of studies have examined factors that affect the level of sanction jurors impose on defendants,[123] JurySim cannot predict how deliberation will act on individual appraisals of appropriate damages. The next section, however, discusses a body of research perfectly styled to benefit from computer simulations based on MacCoun's framework.

C. *JurySim and Cultural Cognition Research*

While the JurySim algorithm is useful for extracting additional information from any mock juror study, the code is particularly useful for studies that investigate how variation in juror characteristics affects individual verdicts. Several studies by the Cultural Cognition Project, a group of researchers that is based at Yale Law School, have done just this,[124] and they provide a fruitful example of how JurySim can enhance research on individual jurors. These studies examine how cultural values influence juror fact perceptions and thereby affect juror verdicts.[125] JurySim can illuminate how these values—and the demographic characteristics they interact with— play out in real-world juror situations, including how jury verdicts will vary between communities. These results have implications for the normative recommendations in cultural cognition studies.

"Cultural cognition" refers to the influence of individuals' group cultural commitments on their factual beliefs.[126] When people sit in judgment in a

---

of Race-Salience and Racial Composition on Individual and Group Decision-Making (2002) (unpublished Ph.D. dissertation, University of Michigan)).

[122] MacCoun, *supra* note 14, at 361.

[123] *See, e.g.*, Viscusi, *supra* note 103, at 556–57 (finding that mock jurors imposed higher damages on a corporation in a tort suit if the corporation had engaged in cost-benefit analysis during product design); Adriaan Lanni, Note, *Jury Sentencing in Noncapital Cases: An Idea Whose Time Has Come (Again)?*, 108 YALE L.J. 1775, 1776 (1999) (noting that people report a desire for harsher penalties in the abstract, but, when acting as mock jurors, suggest penalties below the recommended minimum).

[124] *See infra* notes 130–150, 225–263 and accompanying text.

[125] *See id.*

[126] *See* Dan M. Kahan, *Culture, Cognition, and Consent: Who Perceives What, and Why, in Acquaintance Rape Cases*, 158 U. PA. L. REV. 729, 732 (2010) ("'Cultural cognition' refers to the influence of group values on individuals'

courtroom, they must infer facts—the events that transpired, the mental states of parties at various points—based on the evidence presented. Individuals are psychologically motivated to conform their factual perceptions to their cultural values, confirming that the world works in the way they expect it should.[127] For the same reasons that individualistic, pro-business citizens are less likely to believe that we are at risk from anthropogenic climate change,[128] and people who subscribe to egalitarian values are more likely to believe that widespread gun possession poses a large safety risk,[129] jurors' perceptions of legally consequential facts will reflect their cultural commitments.

A significant body of research investigates how cultural cognition influences juror decision-making. In each study, researchers map subjects' cultural values along two dimensions, derived from the work of anthropologist Mary Douglas.[130] One dimension measures "hierarchy" versus "egalitarianism": does the subject subscribe to a traditional social ordering, where a person's social role is determined by conspicuous characteristics such as sex and class, or does the subject subscribe to a worldview that rejects distinctions in obligations and entitlements based on these fixed traits?[131] The other dimension measures "individualism" versus "communitarianism": does the person value self-sufficiency and resent government intervention, or does the person believe that society should both assist and restrict individual members in pursuit of the collective good?[132]

The studies then evaluate how the subject's positions along these two axes influence his or her perceptions of facts, both on their own and in interaction with his or her demographic characteristics and study manipulations.[133] The research stimulus puts the subject in the role of a

---

perceptions of facts."); Dan M. Kahan & Donald Braman, *Cultural Cognition and Public Policy*, 24 YALE L. & POL'Y REV. 149, 150 (2006) ("Essentially, cultural commitments are *prior* to factual beliefs on highly charged political issues.").

[127] Kahan, *supra* note 126, at 732.

[128] *See* Dan Kahan, *Fixing the Communications Failure*, 463 NATURE 296, 297 (2010).

[129] *See* Dan M. Kahan et al., *Culture and Identity-Protective Cognition: Explaining the White-Male Effect in Risk Perception*, 4 J. EMPIRICAL LEGAL STUD. 465, 481 fig.3, 505 (2007).

[130] *See generally* MARY DOUGLAS, NATURAL SYMBOLS: EXPLORATIONS IN COSMOLOGY (Routledge, 2d ed. 1996).

[131] *See* Kahan & Braman, *supra* note 126, at 153–54.

[132] *See id.* at 153.

[133] *See, e.g.*, Dan M. Kahan et al., *"They Saw a Protest": Cognitive Illiberalism and the Speech-Conduct Distinction*, 64 STAN. L. REV. 851, 883–84, 900 (2012) [hereinafter Kahan et al., *"They Saw a Protest"*] (finding values influenced whether subjects perceived videotaped protesters as engaging in activity that would constitute constitutionally protected "speech" or unprotected "conduct," conditional on whether

juror,[134] asking for a determination of a legally relevant fact or conclusion and finding cultural correlates. Because certain demographic characteristics both correlate with cultural values and influence juror fact perceptions—on their own and in interaction with culture—computer simulations can show how juries are likely to come out on these cases in different communities.

   For example, one study found a correlation between cultural values and whether a subject, after watching a video of a police car chase, perceived that the fugitive—a plaintiff in a civil suit—posed such a threat to public safety that the police were justified in using deadly force to stop him.[135] This investigation arose from a case in the United States Supreme Court, *Scott v. Harris*.[136] Victor Harris, a motorist who fled from the police in a high-speed chase, had sued police officer Timothy Scott, who stopped the chase by ramming his vehicle into Harris's, paralyzing Harris from the neck down. Justice Antonin Scalia, writing for the majority, held that no reasonable jury could fail to find that Harris posed a serious and immediate risk to public safety; Scott was therefore justified in using deadly force against him.[137] Justice John Paul Stevens, in dissent, suggested that reasonable juries *could* differ about whether Harris posed a lethal threat.[138] In support of the majority's holding, the Supreme Court took the unusual step of posting a video[139] to its website. Justice Scalia wrote that the majority was "happy to allow the videotape to speak for itself,"[140] and directed readers to the website.

   A nationally representative group of subjects viewed the video. The researchers then asked the subjects whether Harris's driving put police and

---

they were told the protest was against an abortion provider or a military recruitment facility); Dan M. Kahan et al., *Whose Eyes Are You Going to Believe? Scott v. Harris and the Perils of Cognitive Illiberalism*, 122 HARV. L. REV. 837, 869–79 (2009) [hereinafter Kahan et al., *Whose Eyes Are You Going to Believe?*]; Kahan, *supra* note 126, at 773–93.

    [134] *See, e.g.*, Kahan et al., *"They Saw a Protest"*, *supra* note 133, at 863; Kahan et al., *Whose Eyes Are You Going to Believe?*, *supra* note 133, at 849; Kahan, *supra* note 126, at 765.

    [135] Kahan et al., *Whose Eyes Are You Going to Believe?*, *supra* note 133, at 879–80.

    [136] 550 U.S. 372 (2007).

    [137] *Id.* at 380.

    [138] *Id.* at 396 (Stevens, J., dissenting).

    [139] Videotape: Scott v. Harris (United States Supreme Court 2006), https://www.supremecourt.gov/media/media.aspx   [https://perma.cc/6KEK-TSEU]. While   Scalia's   opinion   points   readers   to   the   address http://www.supremecourtus.gov/opinions/video/scott_v_harris.html [https://perma.cc/Q46A-WH5C], the Court's webmaster apparently failed to provide a redirect page when the Court switched to the www.supremecourt.gov [https://perma.cc/X6KL-SFR9] domain.

    [140] *Scott*, 550 U.S. at 378 n.5.

members of the public at a serious risk of death.[141]   Most subjects agreed with the Court; 73% said they at least moderately agreed with the statement, "During the pursuit, Harris drove in a manner that put members of the public at great risk of death."[142]   The subjects that disagreed, however, were not randomly distributed throughout the population.   Instead, they were disproportionately   African-American,   female,   egalitarian   and communitarian.[143] The authors drew up a profile of "Linda," a black social worker from Philadelphia who is a registered Democrat and self-identified "liberal."[144]   According to their analysis, fewer than one half of the people who share Linda's characteristics would moderately or strongly agree that Harris posed a deadly threat.[145]

The study authors criticize the court for deciding this case, in which different cultural groups may perceive facts in different ways, through summary judgment:

> By insisting that a case like *Scott* be decided summarily, the Court not only denied those citizens an opportunity, in the context of jury deliberations, to inform and possibly *change* the view of citizens endowed with a different perspective. It also needlessly bound the result in the case to a process of decision-making that deprived the decision of any prospect of legitimacy in the eyes of that subcommunity whose members saw the facts differently.[146]

The authors recommended that judges attend to cues that a particular subcommunity might be outraged if judges privilege their own factual perceptions above those of the community.[147] If a judge can conjure a mental image of a dissenter—that person's race, sex, socioeconomic status, or political affiliation—he should evaluate his own perceptions with humility, and consider sending the case to a jury.[148]

My computer simulation expands on the findings of this study and reinforces its normative implications by showing that the deliberative process would likely not mitigate the polarization between white hierarchs and "Lindas."   It is not clear whether any heightened differences after deliberation would actually increase the perception of illegitimacy, as people

---

[141] Kahan et al., *Whose Eyes Are You Going to Believe?*, *supra* note 133, at 857.

[142] *Id.* at 865 fig.2.

[143] *Id.* at 867.

[144] *Id.* at 850.

[145] *Id.* at 875.

[146] *Id.* at 841–42.

[147] *Id.* at 898.

[148] *Id.* at 898–99.

generally experience and voice their opinions outside of the deliberative context. In other words, members of a minority group will not know how a jury from their own community would come out. However, juries—especially those comprised of members from a traditionally underrepresented community—perform several valuable democratic functions. When a minority subcommunity makes a decision that carries the force of law, it contributes to the marketplace of ideas, engages in self-governance, and expresses its values with a rare degree of visibility.[149] In other words, it participates in the democratic process. By denying a jury the opportunity to express itself through a verdict, a judge denies the subcommunity an opportunity to participate in this way.   However, if the deliberative process strongly mitigates the differences between different communities, we might be less concerned about the opportunity denied, which wasn't much of an opportunity at all: a jury from a majority-majority community would be almost as likely to find for the plaintiff as a majority-minority community.

    But JurySim does not assuage these concerns; it may even exacerbate them. Figure 5 compares juries from a majority African-American, majority Democrat, middle-class, majority female, northeastern community with juries from an overwhelmingly white, majority Republican, relatively wealthy, Western, community.[150]

---

[149] *See* Heather K. Gerken, *Dissenting by Deciding*, 57 STAN. L. REV. 1745, 1749 (2005); *see also* Heather K. Gerken, *Second-Order Diversity*, 118 HARV. L. REV. 1099, 1162 (2005).

[150] The "Pro-Defendant" Community is loosely based on Colorado Springs, Colorado, which is 80% white, and 51% female, where 35% of the population has a bachelor's degree, and where the median household income is around $53,000. *See* U.S. Census Bureau, *Fact Sheet: Colorado Springs city, Colorado*, http://www.census.gov/quickfacts/table/PST045215/0816000 [https://perma.cc/7RP3-8CGK]. Registered voters in El Paso County, Colorado, are about two-thirds Republican. *See* El Paso County: Election, http://car.elpasoco.com/election. The "Pro-Plaintiff" Community is loosely based on Baltimore, Maryland, which is 63.4% black and 53.5% female, where 25% of the population has a bachelor's degree, although 76% have graduated from high school, and where the median household income is around $39,000. *See* U.S. Census Bureau, *Fact Sheet: Baltimore city, Maryland*, http://1.usa.gov/iroBe9. African-Americans may be underrepresented in the jury pool. Baltimore, for example, draws its jury pool from voter registration lists and lists of statewide identity card holders, and disqualifies anyone who has convicted of a crime that carries a sentence of six months or more. *See* Robert M. Bell, *Order Adopting Revised Plan for Random Selection of Jurors in Baltimore City* (2010), http://www.courts.state.md.us/juryservice/juryplans/baltimorecity.pdf [https://perma.cc/9PZ5-46KF].   This selection process may disproportionately exclude African-American men, in particular, as they are more likely to have been imprisoned. *See, e.g.*, Bureau of Justice Statistics, *U.S. Dep't of Justice, Sourcebook of Criminal Justice Statistics 2009*, at tbl.6.33.2009,
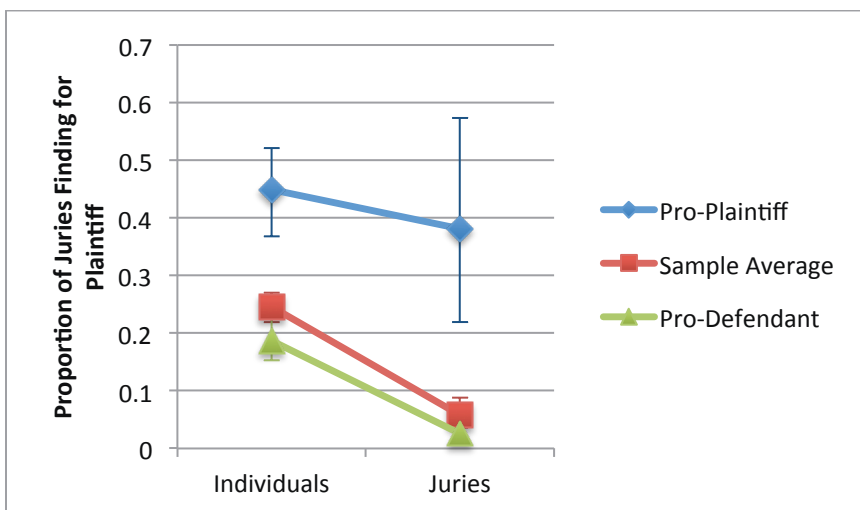
Figure 5.   How deliberation affects the difference between cultural communities in the *Scott v. Harris* case. The "Pro-Plaintiff Community" is 53% female, 60% African-American, 70% Democrat, Northeastern, and has an average education of some college, and an average household income of between $35,000 and $40,000. The "Pro-Defendant Community" is 50% female, 80% white, 70% Republican, Western, and has an average education of some college and an average household income of between $50,000 and $60,000.

Jurors from an egalitarian, African-American community are more likely than jurors from a hierarchical, white community to find for Harris, but more importantly, *juries* from the former community are *similarly, perhaps even more* likely to find for Harris than juries from the latter. Indeed it is *juries*, not *jurors*, who would decide. Because the coefficients on the regression model predicting individual votes have sizeable standard errors—we do not know the independent effects of different traits with a great amount of precision—the program yields large confidence intervals. This is especially true for juries from the pro-plaintiff community: Close to 50% of individuals from that community would find for the plaintiff. As we have seen, the logistic bBOP curve amplifies differences close to *b*—here, 0.5. So small differences in the regression model produce large differences in the proportion of juries we expect will find for plaintiff. But for most of the simulated models, the difference between juries in the pro-plaintiff and pro-

defendant communities is larger than the difference between individuals in those communities. If we believe the law should make room for the voices of these communities, the computer algorithm based on MacCoun's model lends further support to the thesis from the cultural cognition paper: judges should maintain a degree of humility and pause before denying a jury the opportunity to hear a potentially culturally divisive case.

However, the model points to a different conclusion when communities are polarized but both heavily disposed toward the same side (say, 1% favoring the plaintiff in one subcommunity and 20% favoring the plaintiff in another). In both of these cases, very few juries will opt for the plaintiff. While a decision to take the case from the jury could still delegitimize the court in the eyes of the second subcommunity—a valid concern—it cannot realistically be understood to be taking a *decision* out of their hands. This subcommunity would almost certainly find the same way as the court did.

These simulations could be even more useful when applied to other cultural cognition studies. In the next Part, I examine how JurySim, in conjunction with the Cultural Cognition Project's study of acquaintance rape, can generate results useful to scholars, lawyers, and legislators.

III. MODELING ACQUAINTANCE RAPE: A CASE STUDY

In the prototypical case of rape, a woman walking alone at night is approached by a stranger with a weapon, who pulls her out of public view and, overcoming her determined physical resistance, forces her to have sex with him.[151] As recent news stories about sexual assault have highlighted,[152] few actual instances of rape follow this pattern.[153] In 2008, fewer than one-third of female rape victims did not know their assailant before the attack.[154]

---

[151] Louise Ellison & Vanessa E. Munro, *A Stranger In The Bushes, Or An Elephant In The Room? Critical Reflections Upon Received Rape Myth Wisdom In The Context Of A Mock Jury Study*, 13 NEW CRIM. L. REV. 781, 782–83 (2010).

[152] *See* Deborah Tuerkheimer, *Rape on and off Campus*, 65 EMORY L.J. 1, 2, 8 (2015) (noting the new national conversation about non-stranger rape); *see, e.g.*, Noreen Malone & Amanda Demme, *'I'm No Longer Afraid': 35 Women Tell Their Stories About Being Assaulted by Bill Cosby, and the Culture That Wouldn't Listen*, NEW YORK MAGAZINE (July 26, 2015), http://nymag.com/thecut/2015/07/bill-cosbys-accusers-speak-out.html [https://perma.cc/6QRV-54PV]; Matt Hamilton, *Former Stanford Swimmer Convicted of Sexually Assaulting Unconscious Woman on Campus*, L.A. TIMES (March 30, 2016), http://www.latimes.com/local/lanow/la-me-ln-former-stanford-swimmer-convicted-sexually-assault-20160330-story.html [https://perma.cc/FJT5-UUH3].

[153] *See* Michelle J. Anderson, *All-American Rape*, 79 ST. JOHN'S L. REV. 625, 625–27 (2005).

[154] MICHAEL R. RAND, U.S. DEP'T OF JUSTICE, NCJ 22777, NATIONAL CRIME VICTIMIZATION SURVEY: CRIMINAL VICTIMIZATION, 2008 5 tbl.6 (2009),

More commonly, the perpetrator is someone the woman knows.[155]   And sometimes the victim does not put up a forceful physical resistance, and the perpetrator is not carrying a physical weapon.[156]   When a rape follows this latter pattern—where the victim knows her assailant and does not forcefully resist—a jury may be reluctant to convict, because the jurors don't perceive lack of consent. Even though the woman said "no," they reason, she might not have meant "no"; she might have communicated consent through her actions.[157]   If juries are unlikely to convict in these cases, prosecutors may be understandably reluctant to bring charges against perpetrators of acquaintance rape who encountered only verbal resistance.

In this Part, I review the relevant scholarly discussion of the acquaintance rape issue[158] and discuss some of the issues lawyers and

---

http://bjs.ojp.usdoj.gov/content/pub/pdf/cv08.pdf [https://perma.cc/VL6W-4KFP].

[155] *Id.* (noting that 42% of rapes were committed by a friend or acquaintance of the victim, and 18% were committed by an intimate partner of the victim). In this paper, I limit my discussion to rape committed by men against women. Men make up approximately one-fifth of rape and sexual assault victims. *Id.* However, the study I incorporate into my simulations addressed a case of male-on-female rape; therefore, any results that my simulations yield do not necessarily apply to male-on-male or female-on-male sexual assault. It is likely, however, that the mechanisms of cultural cognition would influence perceptions of consent in male-on-male rapes as well as in male-on-female rapes. *See, e.g.*, Damon Mitchell et al., *Attributions of Victim Responsibility, Pleasure, and Trauma in Male Rape*, 36 J. SEX RES. 369, 371–72 (1999) (finding that subjects attributed more responsibility, more pleasure, and less trauma to a homosexual male rape victim than to a heterosexual victim).

[156] Laurie Bechhofer & Andrea Parrot, *What Is Acquaintance Rape?*, *in* ACQUAINTANCE RAPE: THE HIDDEN CRIME 9, 10 (Andrea Parrot & Laurie Bechhofer eds., 1991)

[157] *See* Jacquelyn W. White & John A. Humphrey, *Young People's Attitudes Toward Acquaintance Rape*, *in* ACQUAINTANCE RAPE: THE HIDDEN CRIME 43, 52 (Andrea Parrot & Laurie Bechhofer eds., 1991).

[158] The conversation about acquaintance rape extends far beyond what is addressed here. *See, e.g.*, David P. Bryden & Sonja Lengnick, *Rape in the Criminal Justice System*, 87 J. CRIM. L. & CRIMINOLOGY 1194, 1294–1377 (1997) (focusing on false rape reports, the burden of proof, and victim behavior as "three pervasive issues" relating to leniency in acquaintance rape cases). And academic discussion of rape extends far beyond acquaintance rape. Articles within just the last few years have addressed topics as diverse as preventing prison rape and teaching rape law. *See, e.g.*, Kim Shayo Buchanan, *Engendering Rape*, 59 UCLA L. REV. 1630 (2012) (addressing sexual abuse in prisons perpetrated by women); Helim Kathleen Chun & Lindsey Love, *Rape, Sexual Assault & Evidentiary Matters*, 14 GEO. J. GENDER & L. 585 (2013) (discussing rape shield laws); Jennifer M. Denbow, *The Pedagogy of Rape Law: Objectivity, Identity and Emotion*, 64 J. LEGAL EDUC. 16 (2014) (addressing teaching rape law at law schools); Corey Rayburn Yung, *How To Lie with Rape Statistics: America's Hidden Rape Crisis*, 99 IOWA L. REV. 1197 (2014) (finding that many police departments undercount reported rapes).

legislators face when determining how to prosecute or write laws that enable prosecution of these cases.[159] I then discuss a Cultural Cognition Project study of juror perceptions in acquaintance rape cases and show how the JurySim program can model acquaintance rape juries under a variety of specifications. I derive several substantive conclusions. First, a "no-means-no" reform, where consent is defeated by evidence that the woman said "no," is likely to make a meaningful difference in some jurisdictions and not others, which counsels toward local, not national reform. Second, even in locations where a sizeable majority of potential *jurors* would convict in an acquaintance rape case, fewer than 50% of *juries* may reach the same result. This supports the conventional explanation for why prosecutors are unlikely to bring charges in these cases: they are too likely to lose.

As an Appendix, I include the code for the JurySim algorithm. This program provides a tool for prosecutors and scholars to understand how close acquaintance rape cases will be in different communities. With modification, it can also allow legislators to determine how effective an explicit "no-means-no" law might be in communities of interest within their state.

A. The Conversation About Acquaintance Rape

Much legal scholarship on acquaintance rape has focused on whether laws should be changed to define sex where a woman says "no" but does not physically resist as rape.[160] The traditional, common law definition of rape—which still governs in many states[161]—requires the defendant to have acted not only without the victim's consent, but also with "force or threat of

---

[159] *See* Leigh Bienen, *Rape III — National Developments in Rape Reform Legislation*, 6 WOMEN'S RTS. L. REP. 170, 171 (1980) ("The articulated purposes of the new laws are to increase the number of rape convictions and to ensure that the interests of victims are respected in the criminal justice process."). Increasing prosecutions and convictions is not, however, the only goal of law reform. For example, rape shield laws aim to encourage reporting and protect survivors from embarrassment, see Myka Held & Juliana McLaughlin, *Rape & Sexual Assault*, 15 GEO. J. GENDER & L. 155, 171 (2014), and the Prison Rape Elimination Act addresses a pervasive problem in our prisons, see Karri Iyama, *"We Have Tolled the Bell for Him": An Analysis of the Prison Rape Elimination Act and California's Compliance as It Applies to Transgender Inmates*, 21 TUL. J.L. & SEXUALITY 23, 38 (2012). Even during the reform movement, some suggested a lack of coherence in goals. *See Research into Violent Behavior: Overview and Sexual Assaults, Hearings Before the Subcomm. on Domestic and International Scientific Planning, Analysis and Co-operation of the Comm. on Science and Technology*, 95th Cong., 2d Sess. 427 (1978) (statement of Jan Ben Dor, C.S.W.), *quoted in* Bienen, *supra*, at 177.

[160] *See* Kahan, *supra* note 126, at 745–49.

[161] *See* Tuerkheimer, *supra* note 152, at 15 ("That said, a survey of rape laws shows that many states expressly define rape as requiring force, while others define rape as sex without consent but then include force as a component of non-consent.").

force."[162]   "Force" is defined as behavior that overcomes the physical resistance of the victim;[163] and "threat of force" is behavior that would put a woman in "reasonable fear" of physical injury.[164]   If a woman does not physically resist—if she only says "no," or otherwise verbally resists—the man's actions do not formally fit under the definition of rape. Indeed, courts have enforced these laws, finding that rape does not encompass sexual intercourse where the woman verbally expresses non-consent but does not physically resist.[165]   Also, in most jurisdictions, if the man made a "reasonable mistake" about the woman's consent, he has a defense to the crime of rape.[166]

Scholars have therefore debated whether this standard definition of "rape" should be changed and, if so, what would be the most effective way to reform rape law.   Dan Kahan categorizes these arguments into three positions.[167]   The *standard feminist critique* of traditional rape law says these laws should be changed because they originate from and perpetuate false and harmful sex stereotypes.[168]   These stereotypes hold that in normal sexual relationships, men are aggressors, and women, naturally ambivalent, are aroused by this aggression.[169]   Laws that reinforce these stereotypes reinforce male domination over women and subordinate women's sexual autonomy.[170]   Further, the mistake-of-fact defense privileges a man's perception of consent over a woman's intent to withhold it.[171]   The *conventionalist defense* of the common law definition of rape replies that the law reflects actual behavioral norms: women really do sometimes say "no" when they mean to consent.[172]   Because several studies have shown that some women actually engage in

---

[162] 2 WAYNE R. LAFAVE, SUBSTANTIVE CRIMINAL LAW § 17.1(a) at 605 (2d ed. 2003).

[163] *See id.* § 17.14(a), at 639–40.

[164] *See id*. § 17.3(b), at 624–26.

[165] *See, e.g.*, Commonwealth v. Berkowitz, 641 A.2d 1161, 1164–65 (Pa. 1994).

[166] Beatrice Diehl, Note, *Affirmative Consent in Sexual Assault: Prosecutors' Duty*, 28 GEO. J. LEGAL ETHICS 503, 515–16 (2015); Douglas N. Husak & George C. Thomas III, *Date Rape, Social Convention and Reasonable Mistakes*, 11 LAW & PHIL. 95, 95 (1992).

[167] *See* Kahan, *supra* note 126, at 745–53.

[168] *Id.* at 746.

[169] *See* JOANNA BOURKE, RAPE: SEX, VIOLENCE, HISTORY 67–76 (2007); Kahan, *supra* note 126, at 745.

[170] *See* JOAN MCGREGOR, IS IT RAPE? ON ACQUAINTANCE RAPE AND TAKING WOMEN'S CONSENT SERIOUSLY 62–63 (2005); Kahan, *supra* note 126, at 747.

[171] Catharine A. MacKinnon, *Feminism, Marxism, Method, and the State: Toward Feminist Jurisprudence*, 8 SIGNS 635, 653 (1983); Kahan, *supra* note 126, at 747.

[172] *See* Husak & Thomas, *supra* note 166, at 122; Kahan, *supra* note 126, at 747–48.

"token resistance,"[173] it may be reasonable for a man to perceive consent in the face of verbal resistance. He should not be held criminally responsible if he errs in this perception.[174] Finally, *norm-reconstructionists* assert that while women may actually engage in token resistance, the law should privilege the rights of women who mean to withhold consent, and it should work to change the norms underlying the phenomenon of token resistance.[175] The law should insist that it is per se unreasonable to ignore a woman's words.[176] Legislatures should adopt a strict "no-means-no" rule.[177]

Other scholars have assumed a problem with the current state of acquaintance rape prosecution and have argued for specific legal or procedural reforms. Kahan has previously argued that juries might be more willing to impose civil penalties on acquaintance rapists, and the regular imposition of civil liability should change the norms that currently hinder successful acquaintance rape prosecution.[178] Similarly, Katharine Baker has suggested that by using Title IX to treat rape as a civil wrong, the Department of Education may succeed in changing "the norm of male entitlement,"[179] perhaps eventually enabling criminal enforcement. Ian Ayres and Baker have suggested creating a new crime of "reckless sexual conduct;" a defendant would be guilty of this crime if the jury found that he had sexual intercourse without a condom during a first-time sexual encounter with a woman.[180] Consent to unprotected sex would be an affirmative defense, but the defendant would need to prove consent by a preponderance of the evidence.[181] Stephen Schulhofer has proposed dividing sexual abuse into two offenses: "rape," which would include an element of force, and "sexual abuse" or "sexual misconduct," which would cover interference with

---

[173] *See, e.g.*, Charlene L. Muehlenhard & Carie S. Rodgers, *Token Resistance to Sex: New Perspectives on an Old Stereotype*, 22 PSYCHOL. WOMEN Q. 443, 448 tbl.1 (1998) (Although a majority of female respondents reported engaging in token resistance, only 15% of women produced non-fictitious narratives that met the definition of token resistance).

[174] *See* Husak & Thomas, *supra* note 166, at 123–24; Kahan, *supra* note 126, at 748–49.

[175] *See* Kahan, *supra* note 126, at 750.

[176] *See* SUSAN ESTRICH, REAL RAPE 102–03 (1987).

[177] *Id.*; *see also* Susan Estrich, *Rape*, 95 YALE L.J. 1087, 1182 (1986).

[178] *See* Dan M. Kahan, *Gentle Nudges vs. Hard Shoves: Solving the Sticky Norms Problem*, 67 U. CHI. L. REV. 607 (2000).

[179] Katharine K. Baker, *Why Rape Should Not (Always) Be a Crime*, 100 MINN. L. REV. 221, 222 (2015); *see also* Krista M. Anderson, *Twelve Years Post* Morrison: *State Civil Remedies and a Proposed Government Subsidy to Incentivize Claims by Rape Survivors*, 36 HARV. J. L. & GENDER 223, 257 (2013) (proposing that governments subsidize rape survivors' civil suits).

[180] *See* Ian Ayres & Katharine K. Baker, *A Separate Crime of Reckless Sex*, 72 U. CHI. L. REV. 599, 599 (2005).

[181] *Id.*

a woman's autonomy through nonviolent conduct.[182] Others have argued for an "affirmative consent" standard.[183]

Those who propose procedural reforms try to circumvent juries' resistance to convict acquaintance rapists. Donald Dripps has suggested instituting a juryless sex crimes court to try rape cases where the woman did not consent but did not physically resist.[184] To comply with the Supreme Court's Sixth Amendment jurisprudence, these courts could impose sentences of up to only six months.[185] Baker has suggested that universities could shift social norms by imposing public shaming sanctions on acquaintance rapists: instead of facing jail time, college men who rape would be banned from team activities and forced to wear an article of clothing that would label them as a rapist.[186] Others suggest handling rape cases through a process of restorative justice.[187] All of these suggestions assume that juries, as they stand, are insufficient institutions for acquaintance rape prosecution.

While the most innovative reforms suggested by these scholars have not been put into action, several states have reformed their rape laws, criminalizing intercourse without consent.[188] Wisconsin instituted rape law reform in 1975 with a statute making "sexual intercourse with a person without the consent of that person" a felony.[189] After the *Berkowitz* acquaintance rape case,[190] where the Pennsylvania Supreme Court held that a college student who had sex with a woman despite her verbal protestations could not be convicted of rape,[191] Pennsylvania reformed its rape law. The

---

[182] Stephen J. Schulhofer, *Taking Sexual Autonomy Seriously: Rape Law and Beyond*, 11 LAW & PHIL. 35, 67 (1992).

[183] *See* Diehl, *supra* note 166, at 505.

[184] *See* Donald Dripps, *After Rape Law: Will the Turn to Consent Normalize the Prosecution of Sexual Assault?*, 41 AKRON L. REV. 957, 976 (2008).

[185] *Id*.

[186] *See* Katharine K. Baker, *Sex, Rape, and Shame*, 79 B.U. L. REV. 663, 698 (1999).

[187] *See, e.g.*, Kerry M. Hodak, Note, *Court Sanctioned Mediation in Cases of Acquaintance Rape: A Beneficial Alternative to Traditional Prosecution*, 19 OHIO ST. J. ON DISP. RESOL. 1089, 1099–1101 (2004).

[188] *See* John F. Decker & Peter G. Baroni, *"No" Still Means "Yes": The Failure of the "Non-Consent" Reform Movement in American Rape and Sexual Assault Law*, 101 J. CRIM. L. & CRIMINOLOGY 1081, 1083–96 (2011) (discussing "true non-consent states" and "contradictory non-consent states," in which the prosecution must show force or incapacitation to establish non-consent).

[189] WIS. STAT. ANN. § 940.225(3) (West 2005); *see also* Christina M. Tchen, Comment, *Rape Reform and a Statutory Consent Defense*, 74 J. CRIM. L. & CRIMINOLOGY 1518, 1543 (1983). Michigan's pioneering and better-known 1975 reform statute, MICH. COMP. LAWS ANN. §§ 750.520a–.5201 (West 2014), retains a requirement of "force or coercion."

[190] *See infra* Part III.B.

[191] Commonwealth v. Berkowitz, 641 A.2d 1161, 1163–64 (Pa. 1994).

commonwealth enacted a definition of "forcible compulsion" that includes "intellectual, moral, emotional or psychological force, either express or implied."[192] It also created a new crime, "sexual assault": "sexual intercourse . . . with [another person] without [that person's] consent."[193] While the penalty for sexual assault is less than the penalty for rape, it is greater than the penalty for "indecent assault," the conviction Berkowitz ultimately received.[194] The New Jersey Supreme Court has held that non-consensual penetration satisfies the "physical force" requirement of the state's rape law.[195] California now requires colleges that receive state funding for financial aid to adopt a sexual assault policy that incorporates an "affirmative consent" standard.[196]

Despite these reforms, however, acquittal rates are still unusually high for rape overall and for acquaintance rape in particular.[197] Legislators who write rape reform laws would therefore benefit from additional information on what factors contribute to a law's effectiveness in attaining convictions.[198] Prosecutors, too, will want to know the circumstances under which a trial is likely to result in a conviction. When is a case really a close case, when is it a probable win, and when is it a sure loss? And scholars who analyze the cultural factors that influence acquaintance rape law and recommend law reforms will be interested in the same questions: where and under what circumstances will a jury convict a man who had sex with a woman without her consent? A recent Cultural Cognition Project study begins to answer these questions,[199] but computer simulations of jury deliberation allow lawyers, legislators, and scholars to make more informed estimates about how these cases will play out in court. First, I review the cultural cognition study. I then run through some useful computer simulations of an acquaintance rape case, showing the probability that juries will convict under different conditions and describing who can benefit from this information.

---

[192] 18 Pa. Cons. Stat. Ann § 3101 (West 1972) (Supp. 2013).

[193] 18 Pa. Cons. Stat. Ann. § 3124.1 (West 1995).

[194] *Berkowitz*, 641 A.2d at 1166.

[195] *See* State *ex rel*. M.T.S., 609 A.2d 1266, 1277 (N.J. 1992).

[196] *See* S.B. 967, 2013-14 Sess. (Cal. 2014).

[197] Henry F. Fradella & Kegan Brown, *The Effects of Using Social Scientific Rape Typologies on Juror Decisions to Convict*, 31 Law & Psychol. Rev. 1, 2 (2007).

[198] *See* Bienen, *supra* note 159, at 171, 184.

[199] *See* Kahan, *supra* note 126.

B.   The Berkowitz Case

The Cultural Cognition Project used the facts of *Berkowitz,* the Pennsylvania case, as a stimulus for a study of what causes people to perceive consent in an acquaintance rape case.[200]

1.   Facts of the Case

Pennsylvania's law reform came on the heels of a much discussed acquaintance rape case, *Commonwealth v. Berkowitz*.[201] The defendant was a male college sophomore, Robert Berkowitz, and the complainant was a female sophomore at the same college who had friends in common with Berkowitz.[202] On the afternoon of the non-consensual sex, the victim had a martini and went to a dormitory to meet her boyfriend, with whom she had argued the night before.[203] Seeing that her boyfriend had not arrived, she went upstairs to visit her friend Earl, Berkowitz's roommate.[204] Earl wasn't in the room but Berkowitz was, and they talked for a while, she sitting on the floor and he on the bed.[205] He got off the bed, pushed the victim back, and began kissing her.[206] She said, "Look, I gotta go. I'm going to meet [my boyfriend]," but he persisted.[207] The victim then said "no."[208] She continued to say "no" and "no, I gotta go, let me go," as he touched her breasts and attempted to make her perform oral sex on him.[209] He got up and locked the door.[210] Berkowitz moved the victim to the bed, removed her sweatpants, and had sex with her.[211] After the intercourse, Berkowitz said, "Wow, I guess we just got carried away," to which the victim replied, "No, *we* didn't get carried away, *you* got carried away."[212] The victim left Berkowitz's room and raced to her boyfriend in the dormitory lounge.[213] She began crying, and soon after, she and her boyfriend called the police.[214]

---

[200] *See id.* at 731.

[201] Commonwealth v. Berkowitz, 609 A.2d 1338 (Pa. Super. Ct. 1992) (per curiam), *aff'd in part and rev'd in part*, 641 A.2d 1161 (Pa. 1994).

[202] *Id.* at 1339.

[203] *Id.*

[204] *Id.*

[205] *Id.* at 1339–40.

[206] *Id.* at 1340.

[207] *Id.*

[208] *Id.*

[209] *Id.*

[210] *Id.*

[211] *Id.*

[212] *Id.*

[213] *Id.*

[214] *Id.* A more complete account of the incident can be found in the opinion of

At trial, Berkowitz told the court that while the victim had said no, her actions and her tone had indicated consent, even encouragement, and he had stopped the intercourse as soon as he realized she was unhappy.[215] The complainant maintained that she did not consent.[216] Even though "forcible compulsion" was an element of rape in Pennsylvania, a jury convicted Berkowitz of rape and indecent assault.[217] On appeal, the Superior Court overturned the rape conviction, saying that while the victim's protestations might be sufficient to show that she did not consent, nothing in the record showed forcible compulsion.[218] The Superior Court remanded for a new trial on the indecent assault conviction on evidentiary grounds.[219] The Supreme Court of Pennsylvania affirmed the Superior Court's judgment on the rape conviction but reinstated the jury verdict on the indecent assault charge.[220] The decision set off a war between feminists on one side, who harshly criticized the court's decision,[221] and commentators who supported the court's judgment.[222] The legislature responded by expanding the definition of forcible compulsion to include "intellectual, moral, emotional or psychological force, either express or implied,"[223] and by creating a new intermediate offense, "sexual assault," defined as sexual intercourse without the other person's consent.[224]

2. The Cultural Cognition Study

A recent study by Dan Kahan, as part of the work of the Cultural Cognition Project, examined the factors that led individuals to perceive that Berkowitz reasonably understood the victim to be expressing consent.[225] Subjects in the study read the facts of the *Berkowitz* case taken from the

---

the Superior Court of Pennsylvania.

[215] *Id.* at 1341.

[216] *See id*. (noting appellant's story "differed only as to the consent involved").

[217] *Id*. at 1341–42.

[218] *Id.* at 1347–48.

[219] *Berkowitz*, 609 A.2d at 1352.

[220] *See* Commonwealth v. Berkowitz, 641 A.2d 1161, 1166 (Pa. 1994).

[221] *See, e.g.*, Dale Russakoff, *Where Women Can't Just Say "No,"* WASH. POST, June 3, 1994, at A1 (quoting Cassandra Thomas, President, Nat'l Coal. Against Sexual Assault calling the decision "one of the worst setbacks for the sexual assault movement in the last several years").

[222] Nancy E. Roman, *Scales of Justice Weigh Tiers of Sexual Assault; State May Reform Rape Law*, WASH. TIMES, June 16, 1994, at A8 (quoting Camille Paglia, Professor, Univ. of the Arts, as saying the case was "not even remotely about rape"). *See* Kahan, *supra* note 126, at 741.

[223] 18 PA. CONS. STAT. ANN. § 3101.

[224] 18 PA. CONS. STAT. ANN. § 3124.1.

[225] *See* Kahan, *supra* note 126, at 731–32, 769–71.

Superior Court opinion, with names changed.[226] They were then presented with one of four legal definitions of rape—(1) the common law definition, (2) a strict liability definition where mistake is no defense, (3) a reform definition based on the Wisconsin statute, and (4) a "no-means-no" definition where saying "no" is sufficient to show non-consent—or no definition at all.[227] The subjects were then asked to agree or disagree with statements, such as, "Despite what she said or might have felt after, Lucy really did consent to sexual intercourse with Dave," and "Dave should be found guilty of rape."[228]

The study found that subjects who subscribed to a hierarchical worldview were more likely to agree that the defendant reasonably perceived consent, and less likely to say that the defendant should be convicted of rape.[229] Conversely, subjects with an egalitarian worldview were more likely to say that the defendant could not have reasonably perceived that the victim had consented, and more likely to say that he was guilty of rape.[230] In contrast with popular perception, there was no significant difference between how men and women, overall, perceived the case.[231] Gender did, however, interact with both age and cultural worldview, so older, hierarchical women were more likely to perceive lack of consent than either younger, female hierarchs or older, male hierarchs.[232]

Notably, the legal definition of rape presented to the subject did *not* influence whether or not that person would have found the defendant guilty, with one exception: the "no-means-no" condition did have a significant effect, with more subjects in that condition saying that the defendant should be convicted.[233] However, culture had a substantially greater effect on perceptions than the "no-means-no" law.[234]

Region also had a significant effect on perceptions, with northeastern and far western jurors more likely to convict than jurors from the south or mountain states.[235] Although other variables that are correlated with cultural

---

[226] *See id.* at 765.

[227] *Id.* at 807–12. The study did not test every proposed definitional reform. *Cf.* Michelle J. Anderson, *Negotiating Sex*, 78 S. CALIF. L. REV. 1401, 1407 (2005) (proposing that the law "define 'rape' as engaging in an act of sexual penetration with another person when the actor fails to negotiate the penetration with the partner before it occurs").

[228] Kahan *supra* note 126, at 812–13. "Dave" and "Lucy" were pseudonyms used for the parties in the *Berkowitz* case.

[229] *Id.* at 779 tbl.1.

[230] *Id.*

[231] *Id.* at 782.

[232] *Id.* at 782–83.

[233] *Id.* at 781.

[234] *Id.*

[235] *Id.* at 779 tbl.1.

worldview, such as race, religion, and party affiliation, did not significantly affect guilt votes independently,[236] they are included in some of these simulations to reflect any significant cumulative impact.

C. *Simulating Acquaintance Rape Juries*

JurySim can predict how juries will come out in cases similar to *Berkowitz* under various specified conditions.

1. No-Means-No Law

The study found that a law explicitly stating that the word "no" defeats a reasonable perception of consent did have a significant effect on subjects' decisions.[237] But it is not immediately clear from this that such a law would significantly influence *jury* verdicts. In terms of MacCoun's bBOP model, if the community's initial voting disposition lies far from the value of the threshold parameter, *b*, the effect of such a law might be minimal. If a "no-means-no" law makes jurors ten percentage points more likely to convict, this will hardly matter if only 10% of the community initially wanted to vote for conviction. With a threshold around 67% and a relatively high *c*, few juries would convict in this community, with or without the law. On the other hand, if the boost brings a community from 60% for conviction to 70% for conviction, it could make a large difference in the number of juries willing to convict. Perhaps it could make enough of a difference that prosecutors would be more likely to bring these cases.

Say two acquaintance rapes take place on college campuses two thousand miles apart: one in Binghamton, New York, and the other in Bozeman, Montana.[238] Binghamton is a small city of about 45,000 people in upstate New York. It is approximately 79% white, 9% African-American, and 5% Latino.[239] The average household income is around $47,000.[240]

---

[236] *Id.* at 779 tbl.1, 782.

[237] *Id.* at 781.

[238] I choose these towns because they are culturally different from each other and because they both contain major public universities. I do not suggest that either of these municipalities has a high rate of acquaintance rape.

[239] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – ACS Demographic and Housing Estimates: Binghamton city, New York, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/DP5YR5/1600000US0 606000%7C1600000US2854840%7C1600000US3008950%7C1600000US3606607 ?slice=GEO~1600000US3606607. The Cultural Cognition Project's study was conducted in 2009, and categories such as income matched that data, so I use the demographic measures for these cities from that time.

[240] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – Mean Income in the Past 12 Months (In 2009 Inflation-Adjusted

Bozeman is a small mountain state city of about 37,000 people. It is approximately 90% white, 1% black, and 2.5% Latino.[241] The average household income is about over $58,000.[242] Say the New York and Montana legislatures are each considering passing a "no-means-no" law in response to jury acquittal in these two cases.[243] Will the laws work? Will juries in these college towns reliably convict in cases like *Berkowitz* if the legislature reforms the law?

To answer this question, I simulate juries in both towns. First, the simulation runs a regression on the outcome measure "guilty," "Dave should be found guilty of rape." There are six possible outcomes for "guilty," ranging from "strongly disagree" to "strongly agree." I therefore run an ordered logistic regression instead of a regression that assumes a continuous outcome variable. On the right hand side, I include only variables I can easily determine the values for in both Binghamton and Bozeman: gender ratio, racial distribution,[244] income, urbanicity, and region, along with a variable for whether there is a "no-means-no" law present. I simulate 200 models of this regression, and for each of those models, run 200 juries, each

---

Dollars):   Binghamton   city,   New   York, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/S1902/1600000US060 6000|1600000US2854840|1600000US3008950|1600000US3606607.

[241] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – ACS Demographic and Housing Estimates: Bozeman city, Montana, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/DP5YR5/1600000US0 606000|1600000US2854840|1600000US3008950|1600000US3606607?slice=GEO~ 1600000US3606607.

[242] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – Mean Income in the Past 12 Months (In 2009 Inflation-Adjusted Dollars):   Bozeman   city,   Montana, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/S1902/1600000US060 6000|1600000US2854840|1600000US3008950|1600000US3606607.

[243] New York's statute does not say that the word "no" is sufficient to defeat reasonableness of consent, although it provides that lack of consent may occur when "the victim clearly expressed that he or she did not consent to engage in such act, and a reasonable person in the actor's situation would have understood such person's words and acts as an expression of lack of consent to such act under all the circumstances." *See* N.Y. PENAL § 130.05 (Consol. 2011). Montana defines "without consent" as forcible compulsion or incapacitation. *See* MONT. CODE ANN. § 45-5-501 (1973) (Supp. 2015).

[244] African-Americans, specifically African-American men, may be underrepresented on juries. *See supra* note 150. I do not adjust for this in my calculations, because I do not have information on whether African-Americans are underrepresented in this specific city, and if so, by how much. This could, admittedly, pose a problem in such a sensitive model. Ideally, someone interested in the results of the model would plug in the racial composition of the jury pool, not just the composition of the community.

pulled from a population distribution that mimics either Binghamton's demographics or Bozeman's, to find the overall average probability of conviction under four conditions: Binghamton without a law, Binghamton with a law, Bozeman without a law, and Bozeman with a law. The results follow.
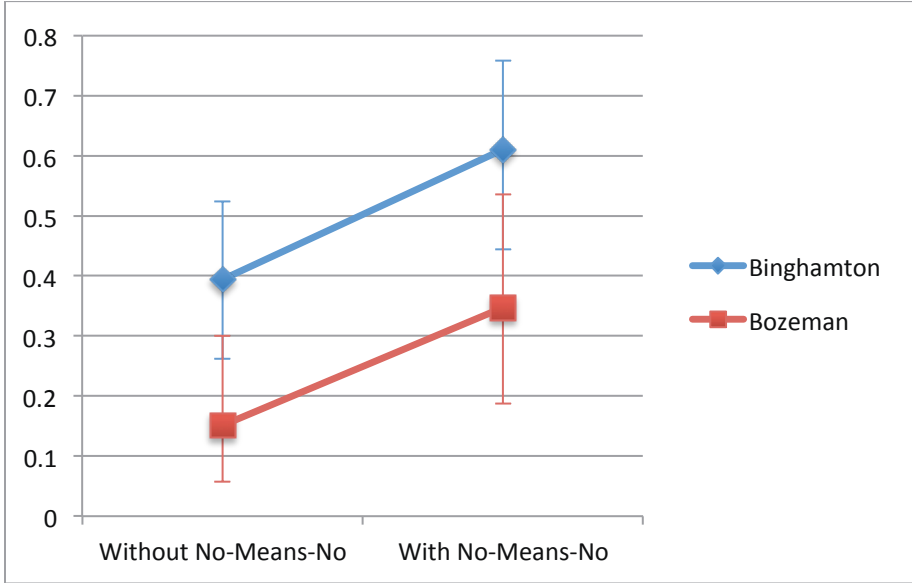

Figure 6. The effect of "no-means-no" law reform on juries in Binghamton, New York and Bozeman, Montana. Before deliberation, an expected 62% of individual Binghamton residents would have convicted, and 71% would have convicted after the reform. In Bozeman, 47% of individuals would have voted to convict before the law, and 59% would likely vote to convict after the law.

If the large confidence intervals in Figure 6 seem to suggest that the model cannot tell us anything useful about how juries will vote after a no-means-no law is passed, a probability density distribution graph better highlights the value of the program. Figure 7 shows the density of simulation results at each probability of conviction.
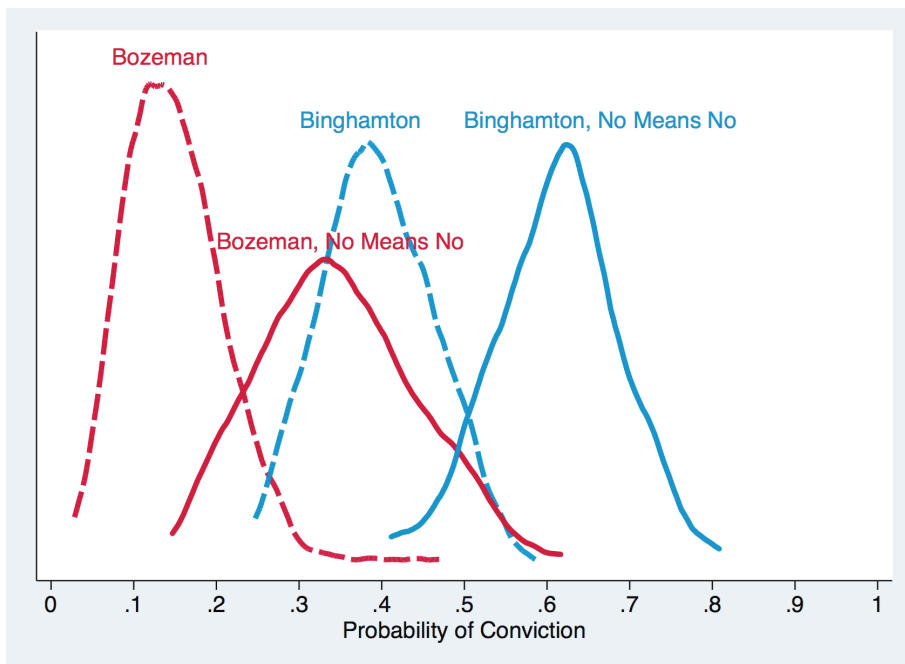
Figure 7. Probability density distributions showing the effect of a "no-means-no" law on juries in Binghamton and Bozeman.

Under the vast majority of Clarify's simulated regression models, after a no-means-no law is passed, a majority of juries in Binghamton and a minority of juries in Bozeman would convict a defendant like Berkowitz.

Without a no-means-no law, even in a northeastern city like Binghamton, it is most likely that a minority of juries would convict. However, after a law is instituted, under most simulated models, a majority would convict. While a prosecutor might still be hesitant to bring charges against those odds, they are much better than the odds without the law reform. In Bozeman, on the other hand, more juries are likely to convict with a "no-means-no" law than they are without one; however, the odds are fairly low under both conditions. Prosecutors might be very hesitant to bring this sort of case to trial in Bozeman.

This distribution hints that perhaps we should not have a national policy on law reform, since its effectiveness depends on the city and state where the charges are brought. Instead, new laws will be most effective in places that are already heavily egalitarian. State legislators should consider their location before passing the no-means-no law. If the law fails, and acquaintance rapists are regularly acquitted, the new law will reflect poorly on the local government, and the effort would have used time and money inefficiently. Therefore, legislators in already hierarchical states should consider one of the other norm reconstruction techniques—perhaps civil

liability or shaming sanctions—to get a state full of constituents who are willing to convict acquaintance rapists. Only then should the legislature introduce a true "no-means-no" model. In places more like Binghamton, however, this sort of law reform could be effective today.

Prosecutors operating under a common law definition of rape, a Wisconsin-style reform statute, or a "no-means-no" statute will be able to anticipate their cases with JurySim. If they must have a minimal expectation of winning before they charge a defendant with acquaintance rape, this simulation could give them an idea of whether they have hit that threshold. Unfortunately, the error bars around JurySim's point estimates are relatively large. The simulation pulls a finite sample of juries from the community, but more importantly, it is based on a limited model, one that does not perfectly predict the likelihood that even an individual will vote for conviction. It may make a very big difference to a prosecutor whether the likelihood of conviction is 60% or 75%, and JurySim will not be able to tell that prosecutor what the precise likelihood of conviction is within a percentage point or two. This is an inevitable limitation of the computer program; it is only as precisely predictive as the underlying individual juror prediction model.

2.   A Cultural Model

The single most powerful influence on perceptions of consent was cultural worldview, with 67% of egalitarians saying that they at least "slightly" agreed that the defendant should be convicted of rape but only 50% of hierarchs responding the same way.[245] The cultural effect also interacted with respondent gender and age: only 45% of female hierarchs over 60 years of age agreed that the defendant should be found guilty of rape.[246] By contrast, 52% of male hierarchs over 60 would have found the defendant guilty, as would have 56% of female hierarchs under 30.[247] On the other end of the spectrum, about 76% of egalitarians under 30 years old agreed the defendant was guilty.[248] A model that predicts a verdict based on culture, gender, age, and an interaction of the three is very powerful.

Most members of the public have not taken a cultural cognition survey,[249] and prosecutors cannot reliably plug their community's average "hierarchy" score into a simulation. However, prosecutors likely understand the kind of community they live in. They have a sense of whether people in the community support or oppose gay marriage and affirmative action,

---

[245] Kahan, *supra* note 126, at 776 fig.3.
[246] *Id.* at 776, 777 fig.4.
[247] *Id.* at 776, 777 fig.4.
[248] *Id.* at 777 fig.4.
[249] Yet.

whether their schools teach comprehensive sex education or abstinence only, whether people gripe about their taxes going to fund welfare programs or gripe about tax cuts going to the highest-earning Americans. Further, jury consultants may perform community attitude surveys to gauge demographic and attitudinal distributions.[250]   They therefore may be able to estimate whether the community is more or less hierarchical than average. If they are able to estimate local levels of hierarchical worldview, they could use a cultural worldview model.

The figure below (Figure 8) shows the effects of deliberation on the probability of convicting an acquaintance rape defendant in two communities. The first is, on average, one standard deviation above the national mean in hierarchy, has an average age of 60 years old, and is 75% female. The second is, on average, one standard deviation below the national mean in hierarchy, has an average age of 30, and is 75% male.
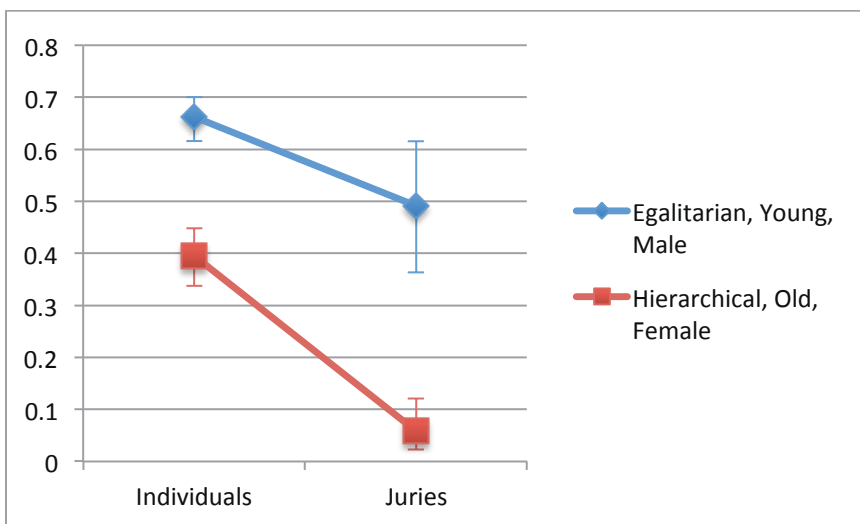

Figure 8. The effects of deliberation on an old, female, hierarchical community and a young, male egalitarian community.

Deliberation exacerbates the difference between these two groups. While about 50% of juries in the egalitarian group will convict, even though about 40% of individuals in the hierarchical group would convict, likely less than 10% of juries from this group agree that the defendant is guilty. What seemed like a close case for individuals is not a close case when extrapolated to juries.

---

[250] *See* Strier & Shestowsky, *supra* note 5, at 452 n.30.   Consultants are infrequently used by prosecutors, however. *See id.* at 451 tbl.4.

This simulation also suggests, more generally, that prosecutors may be reluctant to bring these cases because they have an unacceptably low acquittal rate. Even in an unusually—perhaps unrealistically—young and egalitarian community, only about half of randomly selected juries would convict a defendant similarly situated to Berkowitz. This insight supports the conventional wisdom that prosecutors screen out "unconvictable" or even uncertain cases, so as not to jeopardize their conviction rates.[251]

3. Full Demographic Model

A researcher with access to lots of demographic data or a prosecutor who is less sensitive to culture and more sensitive to these statistics, would benefit from a model that incorporates as many demographic characteristics as possible. Variables not easily accessible from census reports, such as age distribution of the jury-eligible population, education level, and religious and political[252] affiliations all may have effects on the probability that a person will believe the defendant should be convicted. Many of these effects are insignificant when isolated, but when cultural data is unavailable, and these variables are appropriately correlated—a person who is Jewish is more likely to be white,[253] for example—their insignificant effects may add up to a meaningful combined influence.[254]

This model could help the extremely well informed determine the probability of conviction in specific locations, or it could apprise scholars of the chances of conviction in unspecified "pro-conviction" and "pro-acquittal" locations. The figure below (Figure 9), shows the difference between a location high in factors that counsel toward a probable conviction and a location high in factors that counsel toward a probable acquittal. The Pro-

---

[251] *See* Celesta A. Albonetti, *Prosecutorial Discretion: The Effects of Uncertainty*, 21 LAW & SOC'Y REV. 291, 295, 300 (1987) (theorizing that prosecutors seek to avoid uncertainty, and a relationship between victim and defendant may introduce uncertainty into prosecution); Lisa Frohmann, *Discrediting Victims' Allegations of Sexual Assault: Prosecutorial Accounts of Case Rejection*s, 38 SOC. PROBS. 213, 214, 220–21 (1991) (performing an ethnographic study and finding that prosecutors look for "holes" in sexual assault cases and use those to justify case rejection).

[252] Information on voting patterns may be readily available. However, lack of political affiliation is a very strong variable in these particular regressions, so imputing to simulated jurors the political affiliation that corresponds with their voting pattern might be misleading.

[253] *See* PEW RESEARCH CENTER, PEW RESEARCH CENTER SURVEY OF U.S. JEWS, APP'X B 196 (2013), http://www.pewforum.org/files/2013/10/jewish-american-survey-app-b-topline.pdf [https://perma.cc/38HB-YH6J].

[254] Kahan notes that age and education have a significant combined effect. *See* Kahan, *supra* note 126, at 782.

Conviction community is disproportionately African-American, young, educated, low-income, Jewish or non-religious, Democratic or Independent, and entirely from the Northeast.[255]   The Pro-Acquittal community is disproportionately white, old, uneducated, wealthy, Catholic or Protestant, Republican or politically unaffiliated, and entirely from the South.[256]
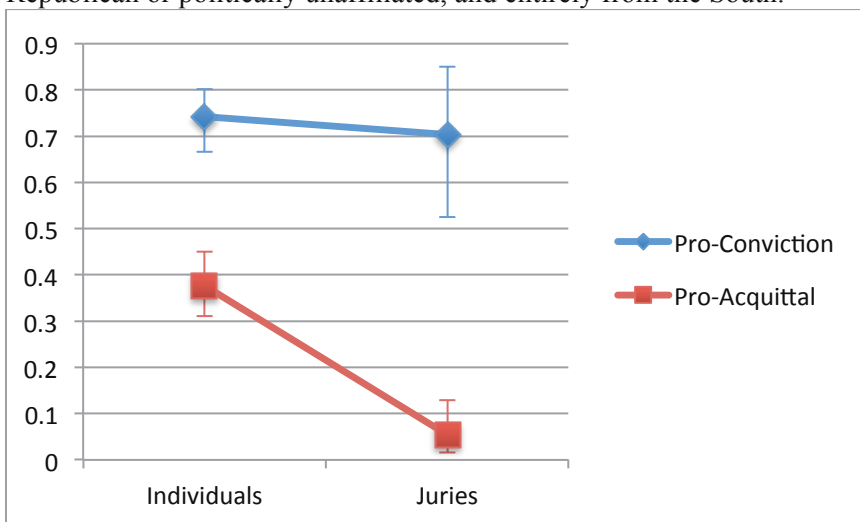

Figure 9. The effect of deliberation on two demographically-specified communities: one approximately one standard deviation above the mean on pro-conviction variables, and the other approximately one-standard deviation above the mean on pro-acquittal variables.

The demographic model does a reasonably good job of replicating the cultural model, with a substantial majority of juries in the pro-conviction group estimated to convict and a small minority of juries in the pro-acquittal community convicting. However, because the percent of individuals voting to convict in the pro-conviction group is so close to *b*, 67%, deliberation exacerbates small differences between the models, yielding large confidence

---

[255] While both the northeast and far west are more pro-conviction than other regions, because I am theoretically modeling a geographic community, I choose one location for each run.

This community is 43% African-American, 20% Jewish and 32% non-religious, 57% Democrat and 31% independent, has an average age of 30.4 years, has an average education of between 2 and 4 years of college, and has an average annual income of approximately $20,000.

[256] This community is 86% white, 35% protestant, 27% Catholic, and 23% other Christian, 58% Republican, has an average age of 59.9 years, has an average education of high school graduate, and has an average income of approximately $80,000.

intervals. Overall, however, this indicates that parties with good demographic information may be able to predict jury behavior nearly as well as those with only good cultural information (and information about gender and age).

4. A Minimal Information Model

If neither information about a community's cultural values nor detailed information about a community's demographics is available, a prosecutor or other interested party can always find basic information through the U.S. Census website.[257] This website shares, inter alia, the average household income, racial distribution,[258] and gender distribution of any municipality in the country. That information, along with the community's geographical region and urbanicity, can be incorporated into a more basic model that predicts the proportion of juries from the community who would convict in an acquaintance rape case.

The figures below (Figures 10 & 11) show the likelihood of conviction for two culturally divergent college towns—Berkeley, California and Oxford, Mississippi—as predicted by the simple demographic model. Berkeley is a relatively well-off city[259] in the far west, where, at the time of this survey, 58% percent of the population was white, 10% was African-American, 17% was Asian, and 10% was Latino "of any race."[260] Oxford is a less wealthy city[261] in the south, where 73% of the population was white, 21% was African-American, 3% was Asian, and 2% was Latino.[262]

---

[257] U.S. Census Bureau, *Fact Sheet*, http://factfinder.census.gov/servlet/ACSSAFFFacts.

[258] The census includes "Hispanic or Latino (of any race)" as a category, which citizens are invited to check off in addition to a "race" category such as "white" or "black or African-American." Our data, however, includes "Hispanic" as a separate race. These models include only non-Latino whites, African-Americans, and Asians in those categories.

[259] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – Mean Income in the Past 12 Months (In 2009 Inflation-Adjusted Dollars): Berkeley city, California, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/S1902/1600000US060 6000|1600000US2854840|1600000US3008950|1600000US3606607.

[260] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – ACS Demographic and Housing Estimates: Berkeley city, California, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/DP5YR5/1600000US0 606000|1600000US2854840|1600000US3008950|1600000US3606607?slice=GEO~ 1600000US3606607.

[261] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – Mean Income in the Past 12 Months (In 2009 Inflation-Adjusted Dollars): Oxford city, Mississippi, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/S1902/1600000US060
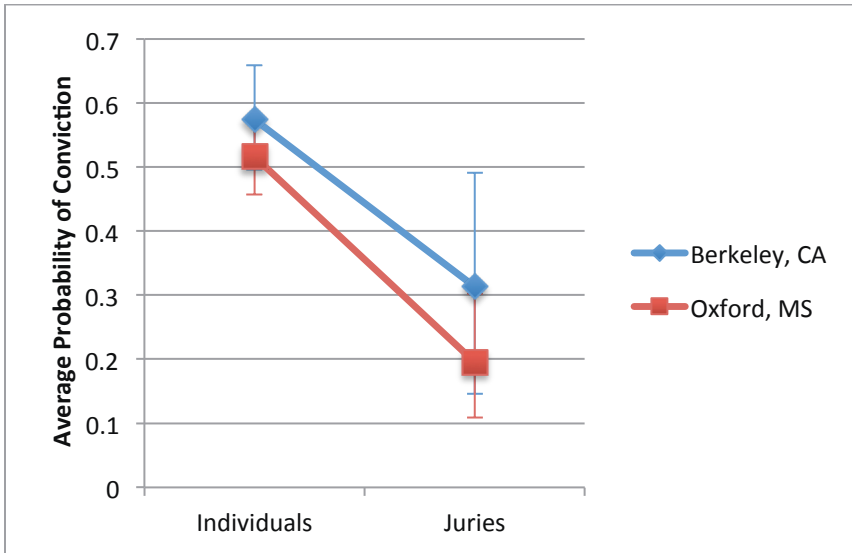
Figure 10. The effect of deliberation on juries from Berkeley, California and Oxford, Mississippi, modeled only with data available on the U.S. Census website.

6000|1600000US2854840|1600000US3008950|1600000US3606607.
    [262] U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates – ACS Demographic and Housing Estimates: Oxford city, Mississippi, http://factfinder.census.gov/bkmk/table/1.0/en/ACS/09_5YR/DP5YR5/1600000US0 606000|1600000US2854840|1600000US3008950|1600000US3606607?slice=GEO~ 1600000US3606607.
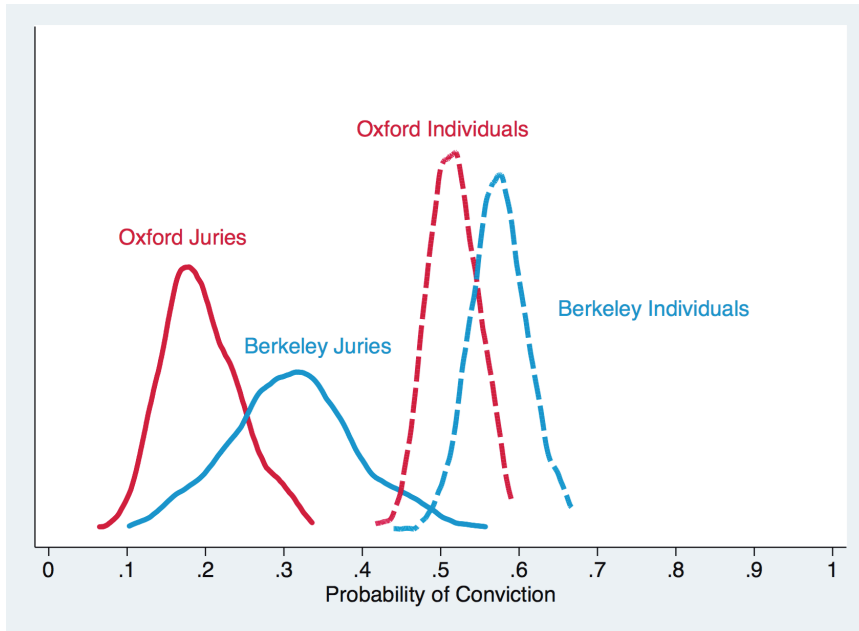
Figure 11. Probability density distributions showing the effect of deliberation on juries from Berkeley and Oxford.

As we would expect, juries in Berkeley are more likely to convict than juries in Oxford. Because the model is based on relatively little information, however, the results could underestimate the differences between these two cities. Berkeley is wealthy—a factor that, on its own, predicts acquittal—and Oxford is fairly racially diverse—a factor that counsels toward conviction. Other than the regional variable, there is little to show exactly how "liberal" Berkeley is. It seems quite possible, then, that the simple model underestimates the likelihood of conviction in Berkeley and therefore underestimates both the disparity in the probability of conviction in these two cities and the polarizing effect over individual differences. However, the simulation does reinforce the conclusion that these cases may be bad bets for prosecutors, even in relatively egalitarian cities. It suggests that prosecutors may not bring acquaintance rape cases because of a reasonable concern that, even where a fair majority of the populace would support conviction, a majority of randomly drawn juries will not convict.

However, this example also demonstrates the current limits on the model for prosecutors: it is sensitive to slight differences in information, and prosecutors, determining whether or not to send a case to trial, would want a fairly precise estimate of the likelihood of conviction. With a better, full-demographic model, a prosecutor can do fairly well. With a less substantial model, a prosecutor may not feel comfortable relying on the results of the computer program. A prosecutor could, however, refine the model over

time, using a machine learning algorithm[263] to tweak the parameters to better replicate past results and predict future verdicts.

CONCLUSION

Lawyers, legislators, and scholars are all interested in how juries are likely to come out under specified conditions. Lawyers may have strong intuitions based on experience with the community. Legislators may have a sense of what the voters want and therefore how the public is disposed. And scholars may have data on factors that influence individual jurors. But all of these people could benefit from a program that extrapolates findings of individuals to jurors.

The program is far from perfect. First, it is only as good as the model on which it is based and the information the user can provide. While a cultural model of the acquaintance rape case produces fairly precise results, a prosecutor may not have precise data on the cultural distribution of her community; and while certain demographic data is easily available, it may give an impoverished picture of the community. Second, its estimates, even given relatively strong models, are not perfectly precise. A "full demographic model" with 200 simulated models and 200 juries per model provides a confidence interval of up to about thirty percentage points. And while the estimates for b and c used in this paper have at least some support, small differences in these numbers can make significant differences in predicted jury verdicts. Additional research, then, into the true values of these parameters would help the model's accuracy.

Even though the point estimates this program produces are imprecise, the ranges it provides may be useful to lawyers and legislators. And with better juror voting models, better model parameter estimates, and better descriptions of the venire of interest, the model's estimates can become both more accurate and more precise. The program also makes headway toward increasing the usefulness of studying individuals. By looking at the more realistic deliberative situation, it begins to tell us how juror-influencing variables play out in the real world.

---

[263] *See, e.g.*, Bernard Widrow & Martin E. Hoff, *Adaptive Switching Circuits*, 1960 IRE WESTERN ELECTRIC SHOW & CONVENTION RECORD, pt. 4, Aug. 23, 1960, at   96–104, *reprinted in* JAMES   ANDERSON   &   EDWARD   ROSENFELD, NEUROCOMPUTING: FOUNDATIONS OF RESEARCH 123 (1988) (developing an adaptive algorithm).

APPENDIX: ANNOTATED SIMULATION CODE

This Appendix includes annotated code for JurySim, the program that calculates the probability a jury will convict in an acquaintance rape case, contingent on the demographic characteristics of a community, as in Figure 9. The code is written for Stata 14 but can be adapted to other versions of the program. It requires two free, user-written programs: Clarify and ice. It simulates acquaintance rape juries from a pro-conviction population using the full demographic model but can be altered to accommodate other datasets and populations. The code proceeds through the following nine steps:

1. The program runs a statistical model—an ordered logistic regression—on the Cultural Cognition Project's acquaintance rape data set, to determine the effects of a set of specified variables on the likelihood that an individual will vote to convict.
2. Clarify simulates 200 versions of this model. It thereby captures the precision of the model within the variance of coefficients in the simulated models. For each loop over steps 3 through 7, the program uses a different simulated model, M_i.
3. The program specifies V, a population of jurors, or venire, with desired demographic characteristics (percent male, average income, etc.).
4. From this venire, the program draws 12 jurors—each with his or her own characteristics.
5. The program uses M_i, the simulated ordered logistic regression model, to simulate an individual "first-ballot" verdict for each of the 12 jurors. This produces an initial number of jurors, N_C, that favor conviction.
6. The program uses MacCoun's bBOP formula to calculate the probability that a jury with N_C jurors initially favoring conviction will, indeed, convict the defendant.
7. The program repeats steps 3 through 6 two hundred times, to simulate two hundred juries. It then averages the probabilities of conviction from all of these juries to find an overall probability that a jury drawn from V will convict using simulated model M_i.
8. The program repeats steps 3 through 7 with another M_i.
9. After looping through all the M_is, the program averages all of the probabilities of conviction, each obtained at a unique M_i. It reports this as the overall probability of conviction.

** STEP 1. Running a regression on the data to derive a model **

** Call up the acquaintance rape data set and set file directory. Drop observations from the "no-means-no" condition. **

```
clear
cd "/file_location/"
use "date_rape_recoded.dta"
set more off
drop if nmnlaw == 1
```

** Derive a correlation matrix, showing how variables of interest correlate within the population. This will allow you to simulate more realistic jurors. For example, low-income jurors will be more likely to be low-education jurors. **

```
corr male female white black other_minority age income educ
urbanicity   jewish   protestant   catholic   other_christian
non_judeochristian no_relig republican democrat independent
no_major_party northeast midwest farwest mountain south
matrix cm = r(C)
```

** Find the means of the variables. This will allow you to pull from a nationally-representative sample or compare a community to the nation as a whole.**

```
mean male female white black other_minority age income educ
urbanicity   jewish   protestant   catholic   other_christian
non_judeochristian no_relig republican democrat independent
no_major_party northeast  midwest farwest mountain south
matrix m = e(b)
```

** Find the standard errors of the variables. **

```
matrix S = diag(vecdiag(e(V)))
matrix sddiag = cholesky(S)
matrix sdp = vecdiag(sddiag)
matrix sd = sqrt(e(N))*sdp
matrix minisd = .55*sd
```

** Specify the mean values of the variables in your desired population. Here, I have two matrices: one for a population likely to convict, and another population likely to acquit. The national mean for each variable is next to the variable name.**

```
matrix   mconv   =   (/*male   0.4787986*/.49,   /*female
0.5212014*/.51, /*%white 0.7385159*/.3, /*%black 0.1157244*/.4,
/*other_min   0.1457597*/.3,   /*age   46.30742*/30,   /*income
7.484982*/3.9,   /*educ   3.35159*/4.76,   /*urban   2.768551*/3,
/*%jew 0.0159011*/.3, /*%protestant 0.2985866*/.10, /*%catholic
0.2164311*/.1,   /*%other_christ   0.1704947*/.1,   /*%non_jc
0.0865724*/.15,   /*%   no_relig   0.2120141*/.25,   /*%repub
0.2941696*/.2, /*%dem 0.4028269*/.55, /*%indep 0.2535336*/.4,
/*%no_party   0.0459364*/.18,   /*%northeast   0.1563604*/1,
```

```
/*%midwest 0.2438163*/0, /*%farwest 0.1634276*/0, /*%mountain
0.0644876*/0, /*%south 0.3621908*/0)

    matrix   macq   =   (/*male   0.4787986*/.51,   /*female
0.5212014*/.49,     /*%white     0.7385159*/.7,     /*%black
0.1157244*/.20, /*other_min 0.1457597*/.20, /*age 46.30742*/60,
/*income   7.484982*/11,   /*educ   3.35159*/2.0,   /*urban
2.768551*/2,     /*%jew     0.0159011*/.1,     /*%protestant
0.2985866*/.32,  /*%catholic 0.2164311*/.29,  /*%other_christ
0.1704947*/.27,   /*%non_jc   0.0865724*/.1,   /*%  no_relig
0.2120141*/.15, /*%repub 0.2941696*/.45, /*%dem 0.4028269*/.2,
/*%indep    0.2535336*/.1,    /*no_party    0.0459364*/.25,
/*%northeast 0.1563604*/0, /*%midwest 0.2438163*/0, /*%farwest
0.1634276*/0, /*%mountain 0.0644876*/0, /*%south 0.3621908*/1)
```

** Use multiple imputation to fill gaps in your data set. This will create several additional data sets with simulated values for absent variable observations. **

```
mi set mlong

     ice guilty male white black age income educ urbanicity
jewish protestant catholic other_christian non_judeochristian
republican democrat independent northeast midwest farwest
mountain, m(5) replace cmd(guilty urbanicity educ:ologit, male
white  black  jewish  protestant  catholic  other_christian
non_judeochristian republican democrat independent northeast
midwest   farwest   mountain:logit,   age   income:regress)
saving(fulldemog_imputed, replace)

    clear

    use fulldemog_imputed
    gen _j= _mj
    gen _i = _mi
    misplit, clear
```

** Run an ordered logit regression on the imputed data sets, predicting the 6-tiered variable "guilty" using specified demographic variables. Clarify will save 1,000 simulated models of this regression. **

```
     estsimp ologit guilty male white black age income educ
urbanicity   jewish   protestant   catholic   other_christian
non_judeochristian republican democrat independent northeast
midwest   farwest   mountain, mi(_mitemp2 _mitemp3 _mitemp4
_mitemp5 _mitemp6)
```

** Store Clarify's simulated models. **

```
    estimates save ologmat, replace
    keep b1 b2 b3 b4 b5 b6 b7 b8 b9 b10 b11 b12 b13 b14 b15 b16
b17 b18 b19 b20 b21 b22 b23 b24
    save clarifybs, replace
```

** STEP 2. Choosing a Clarify model M_i. **

** This step formally occurs later in the code. The next several steps will define a program that you will run through 200 times, each with a different M_i.**

```
program define demogcrimo, rclass
   version 14.1
   drop _all
```

** **STEPS 3 and 4. Simulating jurors from a specified distribution.** **

** Draw jurors from a normal distribution. Specify the demographic variables that will characterize your jurors. Specify the number of jurors, (currently n(12) means there are 12 jurors), the correlation matrix between the variables (currently cm, the matrix you derived above), the means of each variable (currently set at mconv, the pro-conviction jury defined above), and the standard deviations of your variables (currently set at minisd). **

```
   drawnorm male female white black other_minority age income
educ urbanicity jewish protestant catholic other_christian
non_judeochristian no_relig republican democrat independent
no_major_party northeast midwest farwest mountain south, n(12)
corr(cm) means(mconv) sds(minisd)
```

** Replace all of the normal variables with binary or categorical variables. This will take the variable in any mutually-exclusive set with the highest value and set that equal to 1. So if protestant = .56 and catholic = .52, this will change protestant to 1 and catholic to 0. This means that the mean of your normal distribution will not necessarily be the population average of the variable. You may need to play around with the means to replicate the desired population. The algorithm also reassigns the age of anyone under 18 to either 18, 19, 20, or 21.**

```
   gen agez = uniform()
   replace age = 18 if age < 18 & agez < .25
   replace age = 19 if age < 18 & agez >= .25 & agez < .50
   replace age = 20 if age < 18 & agez >= .50 & agez < .75
   replace age = 21 if age < 18 & agez >= .75
   replace male = 1 if male > female
   replace male = 0 if female > male
   replace female = 1 - male
   replace white = 1 if white > black & white > other_minority
   replace black = 0 if white == 1
   replace other_minority = 0 if white == 1
   replace black = 1 if black > white & black > other_minority
   replace white = 0 if black == 1
   replace other_minority = 0 if black == 1
   replace other_minority = 1 if other_minority > black &
other_minority > white
   replace black = 0 if other_minority == 1
   replace white = 0 if other_minority == 1
   replace income = 1 if income < 1.5
```

```
   replace income = 2 if income > 1.5 & income < 2.5
   replace income = 3 if income > 2.5 & income < 3.5
   replace income = 4 if income > 3.5 & income < 4.5
   replace income = 5 if income > 4.5 & income < 5.5
   replace income = 6 if income > 5.5 & income < 6.5
   replace income = 7 if income > 6.5 & income < 7.5
   replace income = 8 if income > 7.5 & income < 8.5
   replace income = 9 if income > 8.5 & income < 9.5
   replace income = 10 if income > 9.5 & income < 10.5
   replace income = 11 if income > 10.5 & income < 11.5
   replace income = 12 if income > 11.5 & income < 12.5
   replace income = 13 if income > 12.5 & income < 13.5
   replace income = 14 if income > 13.5

   replace educ = 1 if educ < 1.5
   replace educ = 2 if educ > 1.5 & educ < 2.5
   replace educ = 3 if educ > 2.5 & educ < 3.5
   replace educ = 4 if educ > 3.5 & educ < 4.5
   replace educ = 5 if educ > 4.5 & educ < 5.5
   replace educ = 6 if educ > 5.5
   replace urbanicity = 3
   **replace urbanicity = 2
   replace jewish = 1 if jewish > protestant & jewish >
catholic & jewish > other_christian & jewish >
non_judeochristian & jewish > no_relig
   replace protestant = 0 if jewish == 1
   replace catholic = 0 if jewish == 1
   replace other_christian = 0 if jewish == 1
   replace non_judeochristian = 0 if jewish == 1
   replace no_relig = 0 if jewish == 1
   replace protestant = 1 if protestant > jewish & protestant
> catholic & protestant > other_christian & protestant >
non_judeochristian & protestant > no_relig
   replace jewish = 0 if protestant == 1
   replace catholic = 0 if protestant == 1
   replace other_christian = 0 if protestant == 1
   replace non_judeochristian = 0 if protestant == 1
   replace no_relig = 0 if protestant == 1
   replace catholic = 1 if catholic > jewish & catholic >
protestant & catholic > other_christian & catholic >
non_judeochristian & catholic > no_relig
   replace jewish = 0 if catholic == 1
   replace protestant = 0 if catholic == 1
   replace other_christian = 0 if catholic == 1
   replace non_judeochristian = 0 if catholic == 1
   replace no_relig = 0 if catholic == 1
   replace other_christian = 1 if other_christian > jewish &
other_christian > protestant & other_christian > catholic &
other_christian > non_judeochristian & other_christian >
no_relig
   replace jewish = 0 if other_christian == 1
   replace protestant = 0 if other_christian == 1
   replace catholic = 0 if other_christian == 1
   replace non_judeochristian = 0 if other_christian == 1
   replace no_relig = 0 if other_christian == 1
   replace non_judeochristian = 1 if non_judeochristian >
jewish & non_judeochristian > protestant & non_judeochristian >
```

```
catholic    &    non_judeochristian    >    other_christian    &
non_judeochristian > no_relig
   replace jewish = 0 if non_judeochristian == 1
   replace protestant = 0 if non_judeochristian == 1
   replace catholic = 0 if non_judeochristian == 1
   replace other_christian = 0 if non_judeochristian == 1
   replace no_relig = 0 if non_judeochristian == 1
   replace no_relig = 1 if no_relig > jewish & no_relig >
protestant & no_relig > catholic & no_relig > other_christian &
no_relig > non_judeochristian
   replace jewish = 0 if no_relig == 1
   replace protestant = 0 if no_relig == 1
   replace catholic = 0 if no_relig == 1
   replace other_christian = 0 if no_relig == 1
   replace non_judeochristian = 0 if no_relig == 1
   replace democrat = 1 if democrat > republican & democrat >
independent & democrat > no_major_party
   replace republican = 0 if democrat == 1
   replace independent = 0 if democrat == 1
   replace no_major_party = 0 if democrat == 1
   replace republican = 1 if republican > democrat &
republican > independent & republican > no_major_party
   replace democrat = 0 if republican == 1
   replace independent = 0 if republican == 1
   replace no_major_party = 0 if republican == 1
   replace independent = 1 if independent > democrat &
independent > republican & independent > no_major_party
   replace democrat = 0 if independent == 1
   replace republican = 0 if independent == 1
   replace no_major_party = 0 if independent == 1
   replace no_major_party = 1 if no_major_party > democrat &
no_major_party > republican & no_major_party > independent
   replace democrat = 0 if no_major_party == 1
   replace republican = 0 if no_major_party == 1
   replace independent = 0 if no_major_party == 1
   replace northeast = 1 if northeast > midwest & northeast >
farwest & northeast > mountain & northeast > south
   replace midwest = 0 if northeast == 1
   replace farwest = 0 if northeast == 1
   replace mountain = 0 if northeast == 1
   replace south = 0 if northeast == 1
   replace midwest = 1 if midwest > northeast & midwest >
farwest & midwest > mountain & midwest > south
   replace northeast = 0 if midwest == 1
   replace farwest = 0 if midwest == 1
   replace mountain = 0 if midwest == 1
   replace south = 0 if midwest == 1
   replace farwest = 1 if farwest > northeast & farwest >
midwest & farwest > mountain & farwest > south
   replace northeast = 0 if farwest == 1
   replace midwest = 0 if farwest == 1
   replace mountain = 0 if farwest == 1
   replace south = 0 if farwest == 1
   replace mountain = 1 if mountain > northeast & mountain >
midwest & mountain > farwest & mountain > south
   replace northeast = 0 if mountain == 1
   replace midwest = 0 if mountain == 1
```

```
   replace farwest = 0 if mountain == 1
   replace south = 0 if mountain == 1
   replace south = 1 if south > northeast & south > midwest &
south > farwest & south > mountain
   replace northeast = 0 if south == 1
   replace midwest = 0 if south == 1
   replace farwest = 0 if south == 1
   replace mountain = 0 if south == 1
```

** **STEP 5. Use one of the simulated ordered logit models to simulate a first-ballot vote for each juror.** **

** Recall the Clarify ologit estimation. **

```
   estimates use ologmat
   append using clarifybs
```

** The next several lines will include a global variable "i." The code loops through 200 values of "i," one for each M_i, the simulated ordered logit models that Clarify generated. **

** Generate a "gscore" for each juror based on the coefficients in the randomly selected line of simulated parameters. This "gscore" is the output of the ordered logit regression in model M_i. **

```
     gen gscore = b1[$i]*male + b2[$i]*white + b3[$i]*black +
b4[$i]*age + b5[$i]*income + b6[$i]*educ + b7[$i]*urbanicity +
b8[$i]*jewish + b9[$i]*protestant + b10[$i]*catholic +
b11[$i]*other_christian + b12[$i]*non_judeochristian +
b13[$i]*republican + b14[$i]*democrat + b15[$i]*independent +
b16[$i]*northeast + b17[$i]*midwest + b18[$i]*farwest +
b19[$i]*mountain
```

** Generate the probability that the juror falls into each of the six guilt "tiers." If the juror falls into the fourth through sixth tier, a vote of "guilty" will be imputed to that juror. These generated probabilities are based on the ordered logit "cuts" in M_i. **

```
   gen prg1 = 1/(1 + exp(gscore – b20[$i]))
   gen prg2 = 1/(1 + exp(gscore – b21[$i])) – 1/(1 +
exp(gscore – b20[$i]))
   gen prg3 = 1/(1 + exp(gscore – b22[$i])) – 1/(1 +
exp(gscore – b21[$i]))
   gen prg4 = 1/(1 + exp(gscore – b23[$i])) – 1/(1 +
exp(gscore – b22[$i]))
   gen prg5 = 1/(1 + exp(gscore – b24[$i])) – 1/(1 +
exp(gscore – b23[$i]))
   gen prg6 = 1 – 1/(1 + exp(gscore – b24[$i]))
```

** Pick a random number. If it is higher than the sum of the probabilities that the juror will fall into "guilt tiers" one through three, then impute an initial vote of "guilty" to that juror. **

```
   gen z = uniform()
```

```
   gen guilty = 0
   replace guilty = 1 if z > (prg1 + prg2 + prg3)

   keep in 1/12
```

** Display and save the percent of jurors in the twelve person jury who initially vote "guilty." **

```
   summarize guilty
   return scalar igmean = r(mean)
```

** End the program that will run 200 times for each M_i. **

```
end
```

** **STEPS 2 & 8. The next section loops through M_is.** It defines a global variable "i," which signifies the model M_i. For each "i" in the loop, the program generates 200 juries, averages their conviction probabilities, and adds that average to a file containing all of the mean conviction probabilities, "avgprobic.dta." **

```
global i 13
while $i <= 212 {
```

** **STEP 6 & 7. Simulate 200 juries and plug them into the bBOP formula.** **

```
   disp `i'
```

** Simulate 200 juries; reps(200) means 200 juries. **

```
   simulate igmean=r(igmean), reps(200): demogcrimo
   gen numconvict = 12*igmean
   gen numacquit = 12 - numconvict
   gen pctconvict = igmean
   gen pctacquit = 1 - igmean
```

** Choose values for the bBOP parameters b and c. **

```
   gen c = 18
   gen b = .67
```

** Use the bBOP formula to generate a probability of conviction. **

```
   gen convprob = 1/(1 + exp(c*(b - pctconvict)))
```

** Summarize "convprob." This will save the average conviction probability. **

```
   summarize convprob
   clear
```

** I call up the file with all of the average conviction probabilities. **

```
    use "avgprobic.dta"
```

** Because it was based on M_i, you have produced the "ith" conviction probability. Store it in the "ith" cell. **

```
    replace avgprobic = r(mean) in $i
    save, replace
```

** Proceed to the next value of "i" in the loop. If you have just used Clarify's eighth simulated model, you will now repeat all of these steps—simulating 200 juries, having them deliberate, finding their average probability of conviction—for the ninth simulated model. **

```
    global i = $i + 1
}
```

** **STEP 9. Report the average probability of conviction.** **

```
summarize avgprobic
centile(avgprobic), centile(2.5,97.5)
```

** Clear your tracks. **

```
program drop demogcrimo
replace avgprobic = .
save, replace
```