Bayesian Detection of Bias in Peremptory Challenges Using Historical Strike Data

Sachin S. Pandya Xiaomeng Li Timothy E. Moore

2021-09-21

Abstract

US law bars using peremptory strikes during jury selection because of prospective juror race, ethnicity, gender, or membership in certain other cognizable classes. Here, we propose a Bayesian approach for detecting such illegal bias. We show how to use historical data on an attorney's use of peremptory strikes in past cases to estimate, in a current trial, how likely an attorney has struck prospective jurors because of cognizable-class membership. We also use the *power prior* to adjust the weight of such historical data in the analysis. We show the utility of our model on simulated data. Finally, we extend this with a prototype software application to detect bias in peremptory-challenge in real time during jury-selection. We demonstrate the app using actual historical strike data from a convenience sample of cases from the federal district court of Connecticut.

1 Introduction

In the US, individuals selected for jury-service appear in court as scheduled and are questioned by the parties' attorneys and the trial judge. During this process, a prospective juror, if not excused by the trial judge for cause, may still be dismissed if a party's attorney uses one of their limited number of peremptory challenges against them. By asserting a peremptory challenge, a party can declare a prospective juror ineligible ("strikes" that juror) for a seat on the jury without the burden of explaining why.

Since *Batson v. Kentucky* (1986), a party violates the Equal Protection Clause of the US Constitution by using peremptory challenges if motivated by the prospective jurors' race, ethnicity, gender, or membership in another cognizable class. The party bringing a *Batson* challenge bears the burden of proving such illegal bias (LaFave et al. 2020, § 22.3). For similar challenges under State law, a few States require only that an "objective observer" or an "objectively reasonable person" would find that race, ethnicity or another cognizable-class membership was a "factor" in that party's use of strikes [Wash. Rule 37(e); Calif. Code of Civil Procedure § 231.7(d)(1)]. Peremptory-challenge procedure, and thus the task of proving illegal bias, varies not only by State, but also by court, including the number of peremptory challenges assigned to each side and the order in which each party uses those strikes (National Center for State Courts: Center for Jury Studies 2018; Williams 2017).

In any case, evidence of such illegal bias may include data on a party's use of peremptory challenges in past cases. (*Flowers v. Mississippi* (2019), p. 2243; Wash. Rule 37(g)(v) (n.d.); Calif. Code of Civil Procedure § 231.7(d)(3)(G)). Prior studies have collected historical strike data in past cases and reported the observed difference in strike rates by race or gender. Typically, they test for the probability of observing a non-zero difference in strike rates by race or gender, given repeated sampling from a hypothetical population of peremptory strikes with zero such difference (e.g., Eisenberg 2017; Grosso and O'Brien 2012; for discussion, see Gastwirth and Xu 2014; Gastwirth and Yu 2012). Prior studies have also modeled how much a prospective juror's race affected the odds of being struck. For different modeling approaches using the same historical strike data from Mississippi, see Craft (2018); DeCamp (2021); and Dunn and Zhuo (2021).

In this paper, we extend a Bayesian approach to *Batson* and similar challenges (Kadane 2018) to incorporate historical strike data. In this approach, we specify a model of the peremptory-strike process in the court of interest that includes a bias parameter b, to which we assign an initial prior distribution. Then, we use an attorney's strike data from past trials, as generated by that same peremptory-strike process, to estimate a posterior distribution for b. In so doing, we use the *power prior* (Ibrahim et al. 2015) to control how much the strike history from past cases affects the posterior distribution of the bias parameter b. Finally, we use that posterior distribution as the prior for b in the present case and update the posterior based on data on the target attorney's strikes in that present case.

Here, we demonstrate this approach, including the results of simulation studies to test model performance. Then, we present a software prototype that encodes the same approach with actual historical strike data to help attorneys and others detect bias in peremptory-challenge use in real time during jury selection.

2 Methods

2.1 Statistical Procedure

Following Kadane (2018) and Barrett (2007), we model a "struck-jury" peremptory-challenge process in which each party strikes prospective jurors in an alternating sequence. Under the "struck jury" procedure, the trial judge rules on all challenges for cause before the parties exercise any peremptories, so that the number of potential jurors who remain is the sum of number of seats on the jury (plus alternates, if any) and the number of peremptory challenges allotted to both sides. Then, the parties exercise their strikes on those potential jurors, either submitting strikes simultaneously or in an alternating sequence. Once all strikes are used or waived, the remaining prospective jurors are assigned to seats on the jury.

Accordingly, for any given case i in which jury selection occurs, let j denote a peremptory strike used, and let δ_{ij} denote whether or not a party used that strike on a person who belongs to a "cognizable class." If "race" is the bias type of interest, the cognizable class is racial minority jurors ($\delta_{ij} = 1$, 0 for White jurors). If "gender" is the bias of interest, the cognizable class is female jurors ($\delta_{ij} = 1$, 0 for male jurors). In turn, let c_{ij} denote the number of cognizable class members subject to strike; and m_{ij} denote the number of cognizable class non-members subject to strike, such that $c_{ij} + m_{ij}$ is the total number of jurors potentially subject to strike. If there is no bias, the probability is $\frac{c_{ij}}{c_{ij}+m_{ij}}$ for striking a cognizable class member, and $1 - \frac{c_{ij}}{c_{ij}+m_{ij}}$ for striking someone who does not belong to that class.

By adding one parameter w, we can measure *bias* by different values of w by defining the probability of a cognizable class member being struck to be $\frac{wc_{ij}}{wc_{ij}+m_{ij}}$. To avoid making the weight of the non-cognizable class be the reciprocal of the weight of cognizable class, let $b = log(w)$.

Accordingly, for any given value of the bias parameter b, the probability of strike of a member from either class, or $Pr(\delta_{ij})$, is such that:

$$Pr(\delta_{ij}|b) = \begin{cases} \frac{(e^b)c_{ij}}{(e^b)c_{ij}+m_{ij}} & \text{for } \delta_{ij} = 1 \\ \frac{m_{ij}}{(e^b)c_{ij}+m_{ij}} & \text{for } \delta_{ij} = 0 \end{cases} \tag{1}$$

This Equation (1) is equivalent to

$$Pr(\delta_{ij}|b) = (\frac{(e^b)c_{ij}}{(e^b)c_{ij} + m_{ij}})^{\delta_{ij}}(\frac{m_{ij}}{(e^b)c_{ij} + m_{ij}})^{1-\delta_{ij}} \tag{2}$$

Given the strike data we have, i.e., δ_{ij}, c_{ij}, and m_{ij}, by estimating the value of b, we can measure bias when a party is striking potential jurors. If $b = 0$, there is no bias, and the probability of strike is simply a function of the share of cognizable members (non-members) in the pool of prospective jurors that could be struck. If $b > 0$, we infer that the the party has bias favoring a strike against a juror falling within the cognizable class (e.g., the juror is a racial minority). Where $b < 0$, the party has a preference for a juror within the cognizable class.

The likelihood function of b is

$$L(b|\delta) = \prod_{i=1}^{n_i}\prod_{j=1}^{n_j}(\frac{(e^b)c_{ij}}{(e^b)c_{ij} + m_{ij}})^{\delta_{ij}}(\frac{m_{ij}}{(e^b)c_{ij} + m_{ij}})^{1-\delta_{ij}} \tag{3}$$

where n_i is the total number of jury selections (trials); and n_j is the total number of peremptory strikes.

2.2 Incorporating historical strike data

We incorporate data on strikes in past cases and allow for adjustment of the weight of that historical strike data on the posterior distribution of the bias parameter. To do this, we introduce the power prior:

$$\pi(b|D_0, a_0) \propto L(b|D_0)^{a_0}\pi_0(b) \tag{4}$$

where $0 \leq a_0 \leq 1$ is the parameter controlling the weight of the historical information; D_0 is the observed historical data; $L(b|D_0)$ is the likelihood function of b given the historical data; and $\pi_0(b)$ is the initial prior before the historical data is observed.

In this paper, we assume a normal prior for $\pi_0(b)$ with mean 0 and standard deviation 2. Unlike Kadane (2018), who assigned b a prior of Uniform$(-6, 6)$, we assume that the law for *Batson* and similar challenges entails a weakly-informative prior. For our initial prior of $N(0, 2)$, we let $\mu = 0$, because the law assigns the burden of proof in a *Batson* challenge to the party bringing the challenge. Thus, if the challenging party produces no relevant evidence of illegal bias, the law requires a trial judge to reject the challenge as unproven. This is tantamount to treating zero as the most-likely value of the bias parameter, absent any data. Moreover, we take the law to imply that, absent any data, one must assume that higher degrees of illegal bias are less likely than lower degrees of bias. For this reason, we use a normal (Gaussian) distribution with $\sigma = 2$.

After including the historical information through the power prior, the posterior distribution of b is proportional to the product of likelihood function of b and the power prior of b is

$$L(b|\delta) = L(b|\delta)(L(b|\delta_0))^{a_0} \exp(-\frac{b^2}{8}) \tag{5}$$

where

$$L(b|\delta_0) = \prod_{i=1}^{n_{0i}} \prod_{j=1}^{n_{0j}} (\frac{(e^b)c_{0ij}}{(e^b)c_{0ij} + m_{0ij}})^{\delta_{0ij}} (\frac{m_{0ij}}{(e^b)c_{0ij} + m_{0ij}})^{1-\delta_{0ij}} \tag{6}$$

is the likelihood function of b given historical data; n_{0i} is the total number of jury selections (trials) in the historical data; n_{0j} is the total number of peremptory strikes in the historical data; and δ_{0ij} denotes whether or not a party used that strike on a person who belongs to a cognizable class in the historical trials.

The posterior distribution does not have a closed form. Accordingly, we used the Metropolis-Hasting algorithm to sample b from the posterior distribution. To measure convergence, we used trace plots, the Gelman-Rubin convergence diagnostic and Geweke's diagnostics (Gelman and Rubin 1992; Brooks and Gelman 1998; Geweke 1992).

2.3 Model Performance on Simulated Data

We conducted a simulation study to measure model performance under different values of the bias parameter b and different amounts and kinds of historical information.

To do this, we generated strike data for a current trial using seven different values of the bias parameter ("current" b = $\{-3, -2, -1, 0, 1, 2, 3\}$. As introduced above, $b > 0$ ($b < 0$) represents bias against (for) a prospective juror within (outside) the cognizable class, while $b = 0$ denotes no bias.

We also generated historical strike data of three sizes (same, double, and triple the size of current data, i.e., data on one previous trial, two previous trials) using seven different values of the bias parameter for generating that historical data ("historical b" $= \{-3, -2, -1, 0, 1, 2, 3\}$).

We let historical b differ from current b (our parameter of interest) to simulate noise associated with historical strike data that is incomplete not at random. Alternatively, this difference simulates a real difference in the bias parameter that depends on a feature that the current trial has but that all the past trials in the historical data do not. To illustrate, suppose prosecutor with a bias value conditional on defendant race: $b > 0$ if the criminal defendant is Black, $b = 0$ if not. If our historical strike data contains only in cases with non-Black defendants, accounting for that historical data may lead us to underestimate prosecutor's bias (b) in the current case with a Black defendant.

Because we use the power prior, we can control how much we account for the historical strike data by modifying the weight parameter a. If $a = 1$, the historical strike data is equally weighted with the data on strikes in the current trial. If $a < 1$, the historical data matters proportionally less. We can assign $a < 1$ to evaluate how sensitive the posterior of the bias parameter is to the historical data on strikes. Accordingly, in the simulation study, we tried different values of the weight parameter ($a = \{0.1, 0.3, 0.5, 0.7, 1\}$) to show how that parameter affected modeling accuracy.

We considered 735 scenarios in total, i.e., the combination of 7 different bias parameters of current data and of historical data, 3 different amounts of historical data and 5 different values of the power prior weight parameter a. For each scenario, we generated 1000 datasets and fit the model on those datasets. We recorded the mean value of the posterior mean and credible intervals of one thousand model fits for each scenario. In turn, for three credible interval levels (80%, 90% and 95%), we recorded the coverage rates, i.e., the proportion of times those intervals actually contained the true value of the (current) bias parameter b.

For instance, a 90% coverage rate means that in 900 of 1000 model fits, a given credible interval contained what we fixed as the value of current b. Put another way, if we set $b > 0$, we calculate the proportion of the 90% credible intervals that lie to the right of zero among the 1000 model fits.

The following figures depict the coverage rates of b under an 80%, 90%, 95% credible interval, respectively. In these figures, *historical b* denotes the bias parameter for the generated historical strike data, and current b means the bias parameter for the current data. The values on the right side represent the size of the historical data. The values on the top of the plots represent the power prior weight parameter a. For these figures, a darker color denotes a higher coverage rate, and vice versa.

Figure 1 depicts the coverage rate with an 80% credible interval. As the first column ($a = 0.1$)

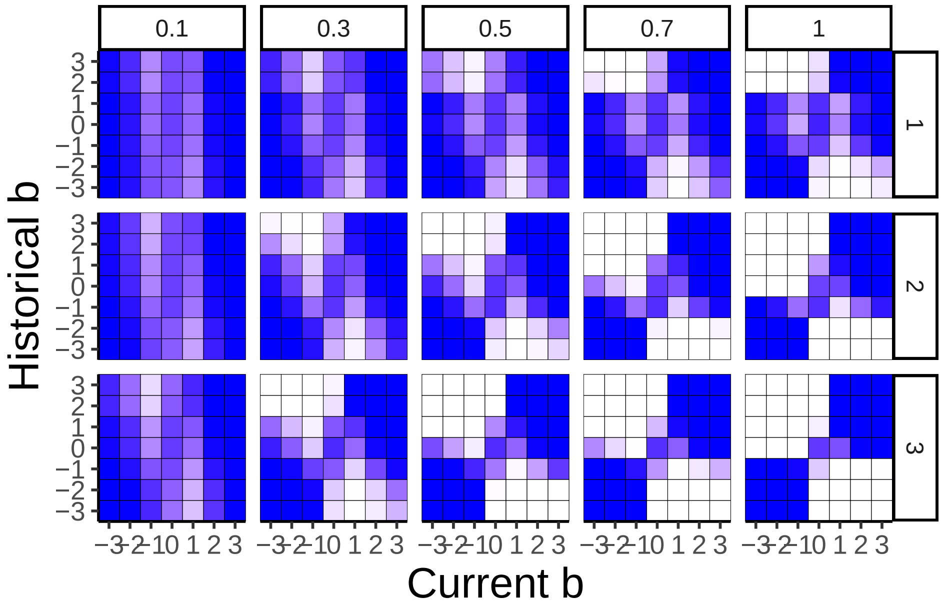

Figure 1: 80% credible interval

depicts, when we place low weight on the historical strike data, we mostly rely on current information. When the bias parameter is 3, 2 or -2, -3, the coverage rate is high, very close to 1.

However, when the bias parameter (“current b”) is 1, the coverage rate drops below 0.8. If the bias is not strong, our model can detect it with a lower accuracy. However, as we place more weight on the historical strike data (as a increases), when the historical strike data is consistent with current information, the accuracy of detecting the bias is also close to 1 when the bias parameter is 1. When the bias parameter is 0 (no bias), however, the coverage rate is also around 80%. Thus, if we use the 80% credible interval, we have around 20% chance of erroneously detecting non-zero bias.

Using a wider credible interval lowers this chance of error. For example, with a 90% credible interval (Figure 2), when $b = 2$, the color is lighter (lower coverage rate) than with a 80% credible interval; when $b = 0$, the color is darker (higher coverage rate). The tradeoff is lower accuracy. With a 90% credible interval, when the bias parameter is 2, the model is less accurate in detecting bias. When the bias parameter is 3, however, the coverage rate is still high.

Figure 3 depicts coverage rates when using a 95% credible interval. Here, given $b = 0$, the coverage rate is very close to 1. This implies that we are very unlikely to conclude a bias when there is none. When the bias parameter is 3, the coverage rate is still high. This implies that we can still accurately detect the bias when it is strong. However, when the bias

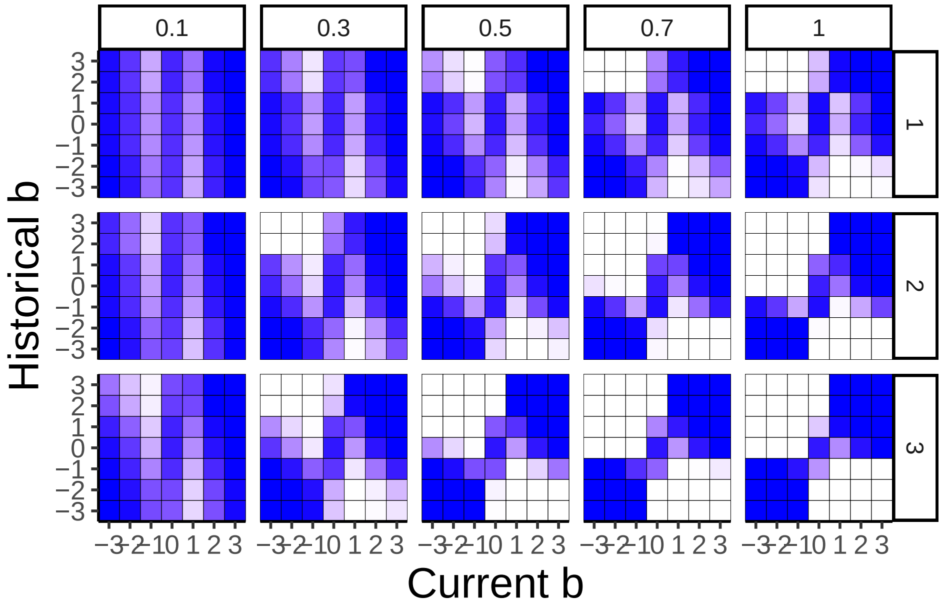

Figure 2: 90% credible interval

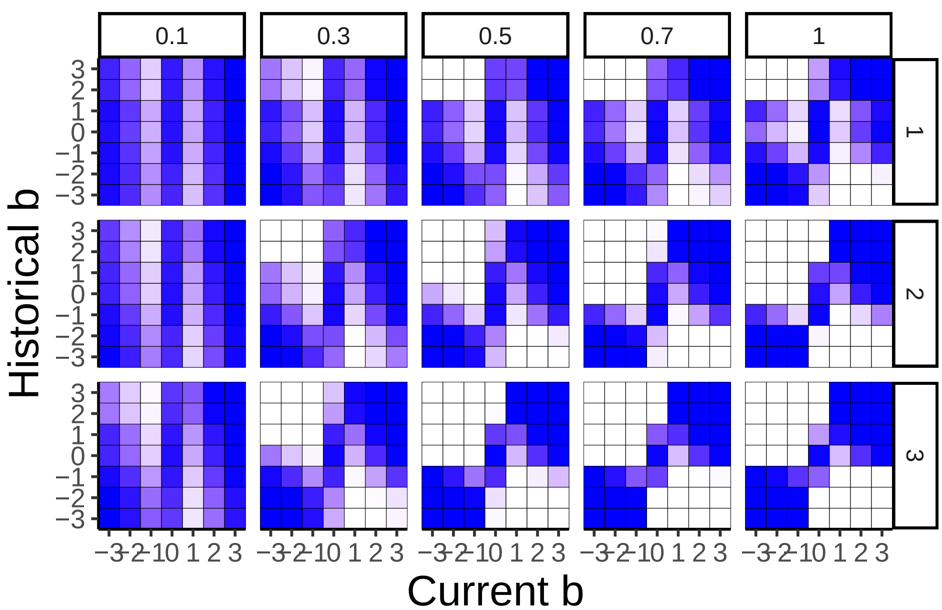

Figure 3: 95% credible interval

is moderate ($b = \{2, -2\}$, the coverage rate is only high when we have consistent historical information.

Overall, these simulation results indicate that our model can always accurately detect strong bias. As for moderate bias, detection accuracy increases with more historical strike information. The choice of credible interval ultimately depends on how conservative we want to be when inferring bias.

2.4 Extension: Simultaneous Strikes

In some courts, both parties simultaneously exercise their peremptory challenges on the prospective jurors in a venire. This simultaneous-strikes process can be modeled as a special case of our model of an alternating-strikes process above (see Equation (3)), i.e, as equivalent to one party engaging in an uninterrupted sequence of strikes against a subset of prospective jurors eligible to be struck. The premise: Regardless of the order in which a party announced those strikes, the posterior for the bias parameter would be the same.

```
## Warning: NAs introduced by coercion
```

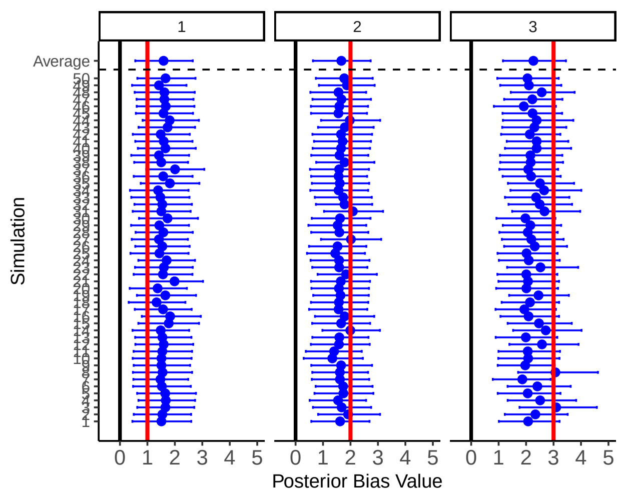

Figure 4: Shuffled Sequence Simulation

To test this premise, we conducted the following simulation study. We first generated a single trial in which one attorney used fifteen strikes in an uninterrupted sequence against 32 prospective jurors, 10 of which were member of a cognizable-class (e.g., Black jurors). Then, we shuffled the order of those strikes to generate 50 trials with the same proportion of struck cognizable-class members but different orders. We then used the model to fit the 50 trials to see if the estimated bias parameters are close to each other. We conducted the simulation study under three scenarios from moderate bias to strong bias ($b = \{1, 2, 3\}$). Figure 4 depicts the results.

The blue points represent the estimated posterior mean for the bias parameter and the blue lines represent the 95% credible interval. The red vertical line indicates the true value of the bias parameter.

We find the estimated bias parameter of the 50 trials with different strike orders are close to each other for all of the three scenarios. None of their credible intervals include zero. We therefore infer that the order of strikes does not influence the estimate of the bias parameter. This confirms that the simultaneous-strikes process can be modeled as a special case of our model of an alternating-strikes process.

2.5 The Software Prototype

We describe here a prototype software application ("app") that implements the approach described above and that attorneys and others can use in real time to detect bias in the use of peremptory challenges. We built this app using R version 4.1.1 (2021-08-10) and shiny (R Core Team 2020; Chang et al. 2021).

To show how this app might be used by lawyers in real cases, we loaded this app with real strike data from a convenience sample of attorneys who appeared during jury selection in criminal cases in the federal district court for Connecticut during fiscal years 2013 through 2017. To collect this historical strike data, one of us filed a request with the U.S. District Court of Connecticut based on 28 U.S.C. § 1868. Thereafter, we received copies (pdf files) of certain jury-selection records associated with [INSERT] trials in that court during this period. These included strike sheets that indicated the ID number of prospective jurors who were struck by peremptory challenge, the order in which they were struck, and which side (prosecutor or defense) struck which juror. Such records also included a tally of answers to juror questionnaires that asked each prospective juror to report their race and gender.

These records, however, often did not indicate the identity of the attorneys exercising the strikes. While the standard forms included a signature line for the attorney, many were left blank or filled with illegible signatures. Accordingly, we turned to the publicly-available docket sheets for each case for the names of the lawyers who appeared in the case on behalf of the prosecution (the US Attorney's office) or the criminal defendant(s) on the date(s) of jury selection. Where only a single attorney represented a party during jury selection, it was easy to attribute the pattern of strikes to that attorney. Where multiple lawyers appeared for one side, we attributed to each of them that side's pattern of peremptory challenges in that case. In such cases, neither the jury-selection documents nor the docket sheets indicated any hierarchy among multiple lawyers or any other basis to attribute strikes to only one attorney among them. After generating a dataset based on these documents, we kept only strikes where a criminal defendant was represented by an attorney. Finally, to de-identify this dataset, we excluded defendant names and substituted fictitious names for the attorneys using the charlatan package (Chamberlain and Voytovich 2020).

With the app, the user uses the pull-down menus to select the cognizable class (top left) and enters by hand the strike information in the case before them in the *strike tally* table (bottom left). For demonstration purposes, the prototype app comes pre-loaded with a completed strike tally with values that can be changed and rows that can be added or deleted. In the strike tally, `round` denotes the order of strikes, `num_cog` denotes the number of prospective jurors that could be struck that belong to the cognitive class; `total` denotes the total number of prospective jurors that could be struck; `cog` indicates whether the prospective juror actually struck in that round was a member of the cognizable class (1 = yes, 0 = no); and `party` indicates which side used the strike (PP = prosecutor, PD = defense attorney).

For the prototype app, we set two cognizable-class options: race and gender. This was because the standard juror questionnaire did not ask about other possible cognizable classes. The data also did not expressly indicate the race or gender of the criminal defendant(s).

In the default setting, the pull-down menus for prosecutor and defense are set to "None." As a result, the app ignores any historical strike data and estimates the posterior distributions of the bias parameter for prosecutor and defense based only on the strike tally data and the initial prior ($b = N(0, 2)$).

Table 1: Hypothetical Strike Tally

round	num_cog	total	cog	party
1	5	21	0	PP
2	5	20	1	PD
3	4	19	1	PP
4	3	18	0	PD
5	2	17	0	PP
6	2	16	1	PD
7	2	15	1	PP
8	2	14	0	PD
9	2	13	0	PP
10	2	12	0	PD

To illustrate, suppose the strike tally in Table 1 depicts the pattern of strikes in the present case with a prosecuting attorney Aaron Waelchi and defense attorney Lawrance Klocko V (both aliases for actual attorneys in the historical strike data). After the user enters this strike tally, the app initially displays two graphs – one for the prosecution and the defense. Each graph depicts the prior density plot (colored grey) and posterior density plot (blue and red for defense and prosecution, respectively) for the bias parameter (Figure 5).

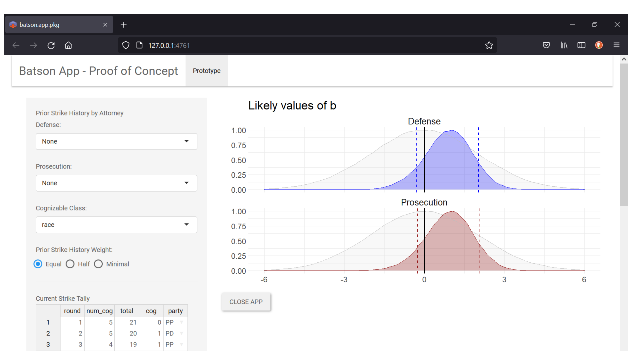

Figure 5: Density Plots without Historical Strike Data

The two vertical dotted lines depict the 80% credible interval. Here, that credible interval includes zero, indicating no credible to infer bias, given the current strike tally.

To use historical strike data, the user selects a name of the prosecutor or defense attorney from the pull-down menus. If an attorney's name cannot be found, the app has no historical strike data for that attorney. Once selected, the prior and posterior density plots automatically update to account for the pre-loaded historical strike data for that attorney. For the weight to assign that attorney's historical strike data, the default is set to equal weight of historical information and current information ($a = 1$). The user has two other options: half weight of historical information ($a = 0.5$) and minimal weight of historical information ($a = 0.2$).

In our illustration, we select the names of the prosecutor and defense attorney from their respective pull-down menus; and leave the weight setting to "Equal." The density plots update accordingly (Figure 6).

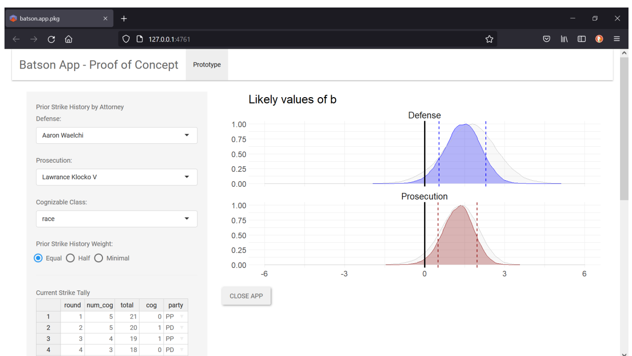

Figure 6: Density Plots with Historical Strike Data (Race)

Now, the credible intervals clearly exclude zero. Thus, we have a credible basis to infer bias against racial-minority prospective jurors in how these attorney use their peremptory challenges in the present case.

Next, suppose the same scenario, except now the strike tally indicates the pattern of strikes by these attorneys against female prospective jurors. If so, we select "gender" as the cognizable class, and the density plots update accordingly (Figure 7).

Here, most of the prosecutor's density curve of both prior and posterior are to the left of zero, indicating the possible bias against male jurors. However, because the 80% credible interval includes 0, inferring such bias from the strike data alone is unjustified.

Finally, if we assign minimal weight to the historical strike data, the density plot updates accordingly (Figure 8).

We can find both the prior and posterior density become flatter, and the credible intervals plainly includes zero.

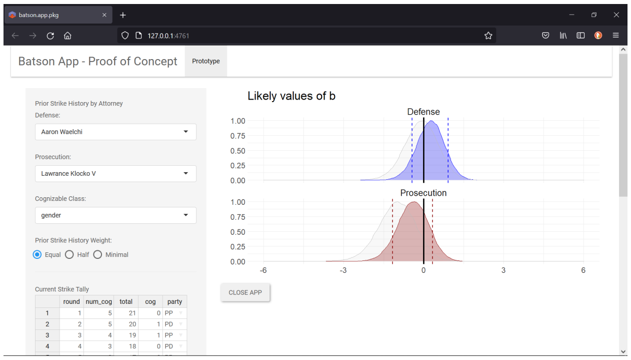

Figure 7: Density Plots with Historical Strike Data (Gender)

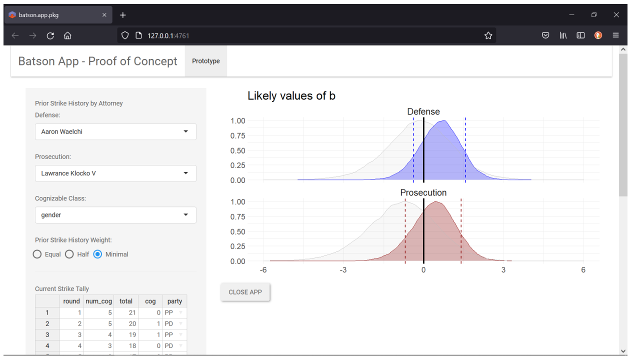

Figure 8: Density Plots with Historical Strike Data (Gender, Minimal Weight)

3 Discussion

Statistical methods for *Batson* and similar changes are intended to measure how much prospective jurors' race, gender, or other cognizable-class membership matters to how likely an attorney will use peremptory strikes against them. The Bayesian approach here accounts for the extent of the available historical strike data when estimating the posterior for the bias parameter. By incorporating the power prior, one can easily assess how sensitive that posterior estimate is to the available historical strike data.

And as our prototype app shows, if the historical strike data is collected and loaded into the app in advance, this approach can be easily used by attorneys and others in real time during jury selection. Despite *Batson* and similar law, racial bias persists in the use of peremptory challenges (Diamond and Rose 2018), perhaps in part because attorneys fear that *raising* a *Batson* challenge raises with it the embarrassment and discomfort associated with accusing another attorney of being racist (O'Brien and Grosso 2019, 29; Offit 2021). Whether an app like the prototype presented here might reduce that and other barriers to bringing *Batson* challenges is a question worth studying further.

At the same time, the Bayesian approach here also requires more from those who would collect the appropriate historical strike data, given the choice of how to model a court's particular peremptory-challenge process. For example, given the model we used, we required data on the cognitive-class composition of the prospective jurors not just at the time of the first strike, but at each time either party used a strike. Moreover, suppose an attorney of interest exercised strikes in two courts, each with different strike procedures. In some cases, one court's strike procedure can be modeled as a special case of a more general model that applied to other court's procedure (as we showed above). If not, that attorney's historical strike data in one court could not used to estimate the posterior of that attorney's bias parameter, given a current trial in the other court.

Finally, for the approach here, like any bias-detection method relying on historical strike data, the ideal data, therefore, is, for each attorney, not just who they struck in the current trial, but a complete history of who they struck in all their past cases. The problem: Jury-selection records in past cases are hard to get because of court practices and otherwise require painstaking effort to compile (Grosso and O'Brien 2017; Wright, Chavis, and Parks 2018). Possible causes include that some jurisdictions do not require recording relevant or detailed enough information about jurors who were or could have been struck (e.g., their race or gender). This forces researchers to turn to other data sources (e.g., voter registration rolls), if any, to find that information. Finally, judges and court staff vary in how often they fully complete or maintain the information they are required to record. This presents an issue for inferring bias if historical strike data is missing not at random. These data challenges, however, might decrease over time, as judges, attorneys, and jury clerks realize that more complete historical strike data makes it easier, with the approach described here, to credibly infer or disprove bias under *Batson* and similar law.

4 Acknowledgements

4.1 CRediT Author Statement

Pandya: Conceptualization, Investigation, Software, Writing; Li: Conceptualization, Formal Analysis, Software, Writing - original draft; Moore: Conceptualization, Supervision, Visualization.

Based on categories in https://casrai.org/credit/.

References

Barrett, Bruce E. 2007. "Detecting Bias in Jury Selection." *The American Statistician* 61 (4): 296–301. https://doi.org/10.1198/000313007X243629.

Brooks, Stephen P., and Andrew Gelman. 1998. "General Methods for Monitoring Convergence of Iterative Simulations." *Journal of Computational and Graphical Statistics* 7 (4): 434–55. https://doi.org/10.1080/10618600.1998.10474787.

Chamberlain, Scott, and Kyle Voytovich. 2020. *Charlatan: Make Fake Data*. https://CRAN.R-project.org/package=charlatan.

Chang, Winston, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. 2021. *shiny: Web Application Framework for R*. https://CRAN.R-project.org/package=shiny.

Craft, Will. 2018. "Peremptory Strikes in Mississippi's Fifth Circuit Court District." https://features.apmreports.org/files/peremptory_strike_methodology.pdf.

DeCamp, Whitney. 2021. "Gender and the Peremptory Challenge: Separating the Effects of Race and Gender in Jury Selection." *Women & Criminal Justice* 31 (2): 159–73. https://doi.org/10.1080/08974454.2020.1785372.

Diamond, Shari Seidman, and Mary R. Rose. 2018. "The Contemporary American Jury." *Annual Review of Law and Social Science* 14 (1): 239–58. https://doi.org/10.1146/annurev-lawsocsci-110316-113618.

Dunn, Jack, and Ying Daisy Zhuo. 2021. "Detecting Racial Bias in Jury Selection." http://arxiv.org/abs/2103.11852.

Eisenberg, Ann M. 2017. "Removal of Women and African Americans in Jury Selection in South Carolina Capital Cases, 1997-2012." *Northeastern University Law Review* 9: 299–345.

Gastwirth, Joseph L., and Wenjing Xu. 2014. "Statistical Tools for Evaluating the Adequacy of the Size of a Sample on Which Statistical Evidence Is Based." *Law, Probability and Risk* 13 (3-4): 277–306. https://doi.org/10.1093/lpr/mgu010.

Gastwirth, Joseph L., and Binbing Yu. 2012. "Case Comment: Appropriate Statistical Methodology Yields Stronger Evidence of Discriminatory Peremptory Challenges in

North Carolina: Application to the Randolph County data in North Carolina v. Rouse and Related Cases." *Law, Probability and Risk* 12 (2): 155–64. https://doi.org/10.1093/lpr/mgs030.

Gelman, Andrew, and Donald B. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science* 7 (4): 457–72. https://doi.org/10.1214/ss/1177011136.

Geweke, John. 1992. "Evaluating the Accuracy of Sampling-based Approaches to the Calculation of Posterior Moments." In *Bayesian Statistics*, edited by J. M. Bernado, J. O. Berger, A. P. Dawid, and A. F. M. Smith, 4:169–93. Oxford University Press.

Grosso, Catherine M., and Barbara O'Brien. 2012. "A Stubborn Legacy: The Overwhelming Importance of Race in Jury Selection in 173 Post-Batson North Carolina Capital Trials." *Iowa Law Review* 97: 1531–59.

———. 2017. "A Call to Criminal Courts: Record Rules for Batson." *Kentucky Law Journal* 105: 651.

Ibrahim, Joseph G., Ming-Hui Chen, Yeongjin Gwon, and Fang Chen. 2015. "The Power Prior: Theory and Applications." *Statistics in Medicine* 34 (28): 3724–49. https://doi.org/10.1002/sim.6728.

Kadane, Joseph B. 2018. "Statistics for Batson Challenges." *Law, Probability and Risk* 17 (1): 1–13. https://doi.org/10.1093/lpr/mgx016.

LaFave, Wayne R., Jerold H. Israel, Nancy J. King, and Orin S. Kerr. 2020. *Criminal Procedure*. 4th ed. Vol. 6.

National Center for State Courts: Center for Jury Studies. 2018. "Comparative Data: Peremptory Challenges." 2018. https://public.tableau.com/shared/7QWPYSBZD.

O'Brien, Barbara, and Catherine M. Grosso. 2019. "Jury Selection in the Post-Batson Era." In *Criminal Juries in the 21st Century: Contemporary Issues, Psychological Science and the Law*, edited by Cynthia J. Najdowski and Margaret C. Stevenson. Oxford University Press.

Offit, Anna. 2021. "Race-Conscious Jury Selection." *Ohio State Law Journal* 82 (2): 201–47.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Supreme Court of the United States. 1986. "Batson v. Kentucky." In *United States Reports*, 476:79. https://www.loc.gov/item/usrep476079/.

———. 2019. "Flowers v. Mississippi." In *Supreme Court Reporter*, 139:2228.

Supreme Court of Washington. n.d. "General Rule 37." https://www.courts.wa.gov/court_rules/pdf/GR/GA_GR_37_00_00.pdf.

Williams, C. J. 2017. "Proposing a Peremptory Methodology for Exercising Peremptory Strikes." *American Criminal Law Review* 54: 277–302.

Wright, Ronald F., Kami Chavis, and Gregory S. Parks. 2018. "The Jury Sunshine Project: Jury Selection Data as a Political Issue." *Univeristy of Illinois Law Review*, 1407.