Exclusion of Extreme Jurors and Minority Representation: The Effect of Jury Selection Procedures

Moro, Andrea and Van der Linden, Martin

Vanderbilt University, Emory University

12 February 2021

Exclusion of Extreme Jurors and Minority Representation: The Effect of Jury Selection Procedures*

Andrea Moro and Martin Van der Linden

February 12, 2021

Abstract

We compare two established jury selection procedures meant to safeguard against the inclusion of biased jurors that are also perceived as causing minorities to be under-represented in juries. The Strike and Replace procedure presents potential jurors one-by-one to the parties, while the Struck procedure presents all potential jurors before the parties exercise vetoes. In equilibrium, Struck more effectively excludes extreme jurors than Strike and Replace but leads to a worse representation of minorities. Simulations suggest that the advantage of Struck in terms of excluding extremes is sizable in a wide range of cases. In contrast, Strike and Replace only provides a significantly better representation of minorities if the minority and majority are heavily polarized. When parameters are estimated to match the parties' selection of jurors by race with jury-selection data from Mississippi in trials against black defendants, the procedures' outcomes are substantially different, and the size of the trade-off between objectives can be quantitatively evaluated.

JEL Classification: K40, K14, J14, J16

Keywords: Jury selection, Peremptory challenge, Minority representation, Gender representation

*Moro: Vanderbilt University, andrea@andreamoro.net. Van Der Linden: Emory University martin.van.der.linden@emory.edu

1 Introduction

In the U.S. legal system, it is customary to let the parties involved in a jury trial dismiss some of the potential jurors without justification. These dismissals, known as *peremptory challenges*, are meant to enable "each side to exclude those jurors it believes will be most partial toward the other side" thereby "eliminat[ing] extremes of partiality on both sides".[1] In the last decades, however, peremptory challenges have been criticized, mainly because they are perceived as causing some groups — in particular minorities — to be under-represented in juries.[2]

The procedure used to let the parties exercise their challenges varies greatly across jurisdictions and is sometimes left to the discretion of the judge.[3] Two classes of procedures are most frequently used in the U.S. In the Struck procedure (henceforth: *STR*), the parties can observe and extensively question *all* the jurors who could potentially serve on their trial *before* exercising their challenges (this questioning process is known as *voir dire*). In contrast, in the Strike and Replace procedure (henceforth: *S&R*), smaller groups of jurors are sequentially presented to the parties. The parties observe and question the group they are presented with (sometimes a single juror) but must exercise their challenges on that group *without* knowing the identity of the next potential jurors.

The goal of this paper is to shed light on a debate that emerged in the legal doctrine over the relative effectiveness of *STR* and *S&R* at satisfying the two objectives of excluding extreme jurors and ensuring adequate group representation. Bermant and Shapard (1981, pp. 93-94), for example, argues that, by avoiding uncertainty, *STR* "always gives advocates more information on which to base their challenges, and, therefore, [...] is always to be preferred". Bermant further notes that "a primary purpose of peremptory challenges is to

[1] *Holland v. Illinois*, 493 U.S. 474, 484 (1990).

[2] For examples of this line of argument against peremptory challenges, see Sacks (1989), Broderick (1992), Hochman (1993), Marder (1994), and Smith (2014). Despite these attacks, the U.S. has so far resisted abandoning peremptory challenges altogether (unlike other countries, like the U.K. where they were abolished in 1988). Peremptory challenges remain pervasive in all U.S. jurisdictions and have been affirmed by the U.S. Supreme Court as "one of the most important rights secured to the accused;" (*Swain v. Alabama* 380 U.S. 202 (1965), see LaFave et al., 2009).

[3] For example, in criminal cases in Illinois, "[State Supreme Court] Rule 434(a) expressly grants a trial court the discretion to alter the traditional procedure for impaneling juries so long as the parties have adequate notice of the system to be used and the method does not unduly restrict the use of peremptory challenges" (*People v. McCormick*, 328 Ill.App.3d 378, 766 N.E.2d 671, (2d Dist., 2002)).

eliminate extremes of partiality on both sides" and that "the superiority of the struck jury method in accomplishing this purpose is manifest."

Others have argued that, by revealing the identity of all potential jurors before challenges are exercised, *STR* facilitates the exclusion of some groups from juries. Although in *Batson v. Kentucky* and *J. E. B. v. Alabama* the Supreme Court found it unconstitutional to challenge potential jurors based on their race or gender,[4] proving that a challenge is based on race or gender is often difficult and the Supreme Court's mandate is notoriously hard to implement.[5] Interestingly, in response, judges themselves have turned to the design of the challenge procedure and the use of *S&R* as an instrument to foster adequate group representation. For example, in a memorandum on judges' practices regarding jury selection, Shapard and Johnson (1994) reports about judges believing that by "prevent[ing] counsel from knowing who might replace a challenged juror" *S&R* procedures "make it more difficult to pursue a strategy prohibited by *Batson*".

To inform this debate, we extend in Section 2 the model of jury selection proposed in Brams and Davis (1978) by allowing potential jurors to belong to two different groups. In the model, each potential juror is characterized by a probability to vote in favor of the defendant's conviction. This probability is drawn from a distribution that depends on the juror's group-membership. The group distributions are common knowledge but the parties to the trial, a plaintiff and a defendant, only observe their realization for a particular juror upon questioning that juror. The parties have opposing goals: the plaintiff wants to maximize the probability of conviction, whereas the defendant seeks to maximize the probability of acquittal.

A jury must be formed to decide the outcome of the trial and the parties can influence its composition by challenging (i.e., vetoing) a certain number of potential jurors. Challenges are exercised according to *S&R* or *STR* procedures which, as explained above, differ mainly in the timing of jurors' questioning (and, as a consequence, in the parties' ability to observe

[4] 476 U.S. 79 (1986) and, 511 U.S. 127 (1994). In terms of legal procedures, the response to these decision has consisted in allowing the parties to appeal peremptories from their opponent, allowing them to nullify a peremptory *if* they can *show* that it was indeed based on race. These appeals are known as *Batson appeals*.

[5] See Raphael and Ungvarsky (1993): "In virtually any situation, an intelligent plaintiff can produce a plausible neutral explanation for striking Pat despite the plaintiff's having acted on racial bias. Consequently, given the current case law, a plaintiff who wishes to offer a pretext for a race-based strike is unlikely to encounter difficulty in crafting a neutral explanation." See also Marder (2012) or Daly (2016) for why judges rarely rule in favor of Batson appeals.

the conviction probability of potential jurors).

We ask how these two procedures perform in achieving the objectives of excluding extreme jurors and ensuring adequate group representation. In Section 3, we provide some intuition for our main result by introducing an illustrative example where a single juror must be selected and the defendant and plaintiff have a single challenge available. In this example, we show that *STR* is more effective than *S&R* at excluding jurors from the tails of the conviction probability distribution, but is less likely to select minority jurors.

The rest of the paper is devoted to characterizing conditions under which these results extend beyond the illustrative example of Section 3. In Section 4 we call a juror *extreme* if its conviction probability falls below (above) a given threshold. We prove that there always exists a low enough threshold such that *STR* is more likely than *S&R* to exclude extreme jurors. Moreover, we show that *STR* always selects fewer extreme jurors than a random selection would, but that there are some (admittedly somewhat unusual) circumstances in which *S&R* would not. Simulations assuming a wide range of conviction probability distributions reveal that, in terms of excluding extreme jurors, the advantage of *STR* over *S&R* can be substantial, even for relatively high thresholds.

Section 5 compares procedures according to their ability to select minorities and identifies conditions under which *S&R* selects more minority jurors than *STR*. Our proof uses a limiting argument showing that the result holds when the minority is vanishingly small and the distributions of conviction probabilities for each group minimally overlap (i.e., groups are polarized). However, simulations again suggest that the result remains true when the size of the minority is relatively high and the overlap between distributions is significant. In Section 6, we explore how changing the number of challenges affect the results of Sections 4 and 5.

Depending on the extent to which jurors of different races have polarized preferences for conviction, the model has different empirical implications for the selection of jurors by race. In Section 7 we exploit peremptory challenge data on a version of *STR* adopted in Fifth Circuit Court District of Mississippi to estimate the groups' distributions of conviction probabilities, and to simulate the outcomes of counterfactual procedures. Results show that groups appear to be substantially polarized in their preferences for convictions, and that the choice of procedure affects both exclusion of extreme jurors and minority representation substantially.

In Section 8 we show how our main theoretical results results extend to a different definition of extreme juries (i.e., a jury in which the *highest (lowest)* conviction-probability juror is below (above) a given threshold). We also explore how the procedures compare in selecting members of groups that are about equal size (such as male and females, as opposed to minorities which induce groups of unequal sizes).

Related Literature

This paper belongs to a relatively small literature formalizing jury selection procedures. Brams and Davis (1978) model *S&R* as a game and derive its subgame-perfect equilibrium strategies which we use in our theoretical results and simulations. Perhaps closest to our paper is Flanagan (2015) who shows that, compared to randomly selecting jurors, *STR* increases the probability that all jurors come from one particular side of the median of the conviction probability distribution (because *STR* induces correlation between the conviction probability of the selected jurors). To our knowledge, this literature is silent on the implications of jury selection for group representation and on the trade-off between excluding extreme jurors and ensuring adequate group representation induced by using different procedures. These are the focus and main contributions of this paper.

While the group-composition of a jury has been shown to influence the outcome of a trial (Anwar et al., 2012; Flanagan, 2018), legal scholars often argue in favor of representative juries regardless of their effect on verdicts. Diamond et al. (2009) for example argue that "unrepresentative juries [...] threaten the public's faith in the legitimacy of the legal system". In an experiment on jury-eligible individuals, they show that participants rate the outcome of trials as significantly fairer when the jury is racially heterogeneous than when it is not. This motivates us to consider group-representativity itself as a desirable feature of jury selection procedures.[6]

[6]One might also be interested in the impact of group-representation on the conviction of defendants who themselves belong to different groups. Without taking groups into account or attempting to compare procedures, Flanagan (2015) studies the impact of jury selection procedures on conviction rates. His results in terms of conviction rates require to assume that the parties have correct beliefs about the probability that jurors eventually vote for conviction (as well as about these probabilities are independent of one another). In contrast, our results about group-representation and exclusion of extremes do not require that the parties' belief at the moment of jury selection be accurate (at least if we are concerned with extremes *as perceived by the parties*, as the U.S. Supreme Court seems to be when saying that the main purpose of peremptory

The empirical literature on jury selection has also identified systematic patterns of group-specific challenges from the parties, with the plaintiffs being almost always more likely to remove minority jurors than defendants (Turner et al., 1986; Rose, 1999; Diamond et al., 2009; Anwar et al., 2012; Craft, 2018; Flanagan, 2018). This justifies our assumption that, at least from the perspective of parties' beliefs, jurors from different groups tend to have different probabilities of voting for conviction.

Diamond et al. (2009) show that for a fixed number of challenges endowed to the parties, larger juries tend to be more representative of the pool's demographic.[7] In Section 6, we show that limiting the number of challenges (while keeping the number of selected jurors fixed) can have a similar effect, though at the expense of a less effective exclusion of extreme jurors.

2 Model

There are two parties to a trial, the **defendant**, D, and the **plaintiff**, P. The outcome of the trial is decided by a jury of j jurors who must be selected from the population. The parties share a common belief about the probability that a juror i will vote to convict the defendant. We denote this probability $c_i \in [0,1]$. Jurors draw this probability independently from the same random variable C, with probability distribution $f(c)$. We denote its cumulative with $F(c)$ and its expected value with μ. Throughout, we assume that C is continuous. To simplify the notation, we also assume that the boundaries of the support of C are 0 and 1.[8]

To address the issue of group representation, we assume that jurors belong to one of two groups a or b. The parties have common beliefs about the probability that jurors from each group vote to convict the defendant. We index the distributions representing these beliefs and their averages with subscript $g \in \{a,b\}$: $f_g(c)$, $F_g(c)$, and μ_g.[9] The corresponding

challenges is to enable "each side to exclude those jurors *it believes* will be most partial toward the other side", see Footnote 1 and associated quote).

[7] Diamond et al. (2009) take advantage of a feature of civil cases in Florida where juries are made of six jurors unless one of the parties requests a jury of twelve jurors and pays for the costs associated with such a larger jury.

[8] This assumption is without loss of generality and all our results hold if C is re-scaled in such a way that $F(c) = 0$ or $[1 - F(1-c)] = 0$ for some $c \in (0,1)$.

[9] Empirical evidence, including the one we report in Section 7 shows that that parties use their challenges unevenly across groups; see also the Related Literature section of the Introduction.

random variables are denoted by C_a and C_b. Although throughout conviction probabilities and their distributions across groups should only be viewed as representing the parties common-*beliefs*, we henceforth lighten the terminology and speak directly of conviction probabilities (rather than parties' *beliefs* about conviction probabilities).

We let r denote the proportion of group-a jurors in the population, and when discussing group representation, we assume that C is obtained by drawing from C_a with probability r and from C_b with probability $(1-r)$ (in particular, $f(c) = rf_a(c) + (1-r)f_b(c)$).

Following the majority of the literature (Brams and Davis, 1978; Flanagan, 2015), we assume that, at the level of jury selection, the parties do not account for the process of jury deliberations and — perhaps as a way to cope with the complexity of jury selection — view the probabilities that jurors votes for conviction as independent from one another. Since conviction in most U.S. trials requires a unanimous jury, the parties then consider that a jury composed of jurors with conviction probabilities $\{c_i\}_{i=1}^{j}$ convict the defendant with probability $\Pi_{i=1}^{j} c_i$. The defendant, therefore, aims at minimizing $\Pi_{i=1}^{j} c_i$ while the plaintiff wants to maximize the same product.

To influence the composition of jury, the defendant and the plaintiff are allowed to challenge (veto) up to d and p of the jurors in a **panel** of $n = j + d + p$ potential jurors randomly and independently drawn the **population** (sometimes also called the **pool**).[10] To avoid trivial cases, we assume throughout that $d, p \geq 1$. The parties use these challenges in the course of a **veto procedure** M (formally, an extensive game-form). The jury resulting from the procedure is called the **effective jury**.

The two veto procedures we study are the **STRuck** procedure (STR) and the **Strike And Replace** procedure ($S\&R$). For comparison, we also consider the **Random** procedure (RAN) which simply draws j jurors independently at random from the population. In all procedures, we assume that once a potential juror i is presented to the parties, the parties observe realized value of c_i for that juror.[11] The two procedures however differ in the timing

[10] In the legal literature, what we call "panel" is sometimes called "*venire*" (though terminology varies and the latter term is sometimes used to speak of what we call the population).

[11] This is motivated by the practice of letting parties extensively question every juror they are presented with, a process known in the legal terminology as *voire dire*. In turn, the fact that the parties have the same assessment of the probability a juror will vote for conviction is motivated by the fact that *voir dire* occurs in the presence of both parties, and that the parties therefore and have access to the same information about the jurors' demographics, background, and opinions.

with which jurors are presented to the parties.

Under *STR*, the entire panel of $j + d + p$ potential jurors is presented to the parties *before* they have the opportunity to use any of their challenges. Each party, therefore, observes the value of c_i for every juror in the panel. The defendant and the plaintiff then choose to challenge up to d and p of the jurors in the panel, respectively. In practice, there are several types of *STR* procedures that differ in the way the parties exercise their challenges after having questioned the jurors in the panel. For concreteness and tractability, we focus in this paper on the *STR* procedure in which the parties have a single opportunity to exercise their challenges on the whole panel. In equilibrium, this leads the plaintiff to challenge the p jurors in the panel with lowest conviction probabilities, and the defendant to challenge the d jurors with highest conviction probabilities.[12] Whether these challenges happen simultaneously or sequentially has no impact on the equilibrium and our results for *STR* apply in either case.[13]

In contrast, under *S&R*, groups of potential jurors are randomly drawn from the population and sequentially presented to the parties. In contrast with *STR* procedures, the parties must exercise their challenges on jurors from a given group *without* knowing the identity of jurors from subsequent groups. There is variation among *S&R* used in practice in the size of the groups that are presented in each round.[14] Again, for concreteness and tractability, we focus in this paper on the *S&R* procedure in which jurors are presented to the parties *one at a time*. The defendant and the plaintiff start the procedure with d and p challenges left, respectively. After each draw, the plaintiff and the defendant observe the potential juror's conviction probability and, if they have at least one challenge left, choose whether or not to challenge the juror. If a juror is not challenged by either party, it becomes a member of

[12] Alternative methods used in the field include procedures in which the parties to challenge sequentially out of subgroups of jurors from the panel only. As long as the procedure remains of the struck type (i.e., the entire panel — and not only the first subgroup — is questioned before the parties start exercising their challenges), the equilibrium effective jury is often the same as under the *STR* procedure we consider here. Other outcome-irrelevant aspects of the equilibrium might, however, be different such as the number of challenges used by the parties (e.g., if the first group is made of the j "middle" jurors in the panel, they may in some cases be selected as effective jurors without the parties exercising any of their challenges).

[13] Since C is continuous, the probability that two jurors in a panel have the same conviction probability and one of the parties does not use all of its challenges in equilibrium is zero and this eventuality can therefore be neglected.

[14] As well as in the ability of the parties to challenge, in a later round, potential jurors who were left unchallenged in previous rounds, a practice known as "*backstricking*".

the effective jury. Any challenged juror is dismissed and the number of challenges available to the challenging party is decreased by one. The process continues until an effective jury of j members is formed.

The (subgame perfect) equilibrium of *S&R* was characterized by Brams and Davis (1978) and takes the form of threshold strategies. In every subgame, D challenges the presented juror i if c_i is above a certain threshold t_D, P challenges i if c_i is below some threshold t_P, and neither of the parties challenges i if $c_i \in [t_P, t_D]$.[15] We will sometimes refer to these values as *challenge thresholds*. As Brams and Davis (1978) show, in any subgame, $t_P < t_D$ and even if the challenges happen simultaneously and both parties are charged for their challenges when they both decide to challenge the presented juror, the latter (i.e., a challenge by both parties) never occurs in equilibrium. The equilibrium is therefore unaffected by the timing of challenges in each round and our results for *S&R* apply regardless of this timing.[16]

In our description of *S&R*, Nature moves in each round to draw a new potential juror from the population to present to the parties. To facilitate conditional comparisons between *STR* and *S&R* based on a particular fixed panel, it will sometimes be useful to consider an equivalent description of *S&R* in which Nature first draws a panel of n jurors $\{c_1, \ldots, c_n\}$ (which the parties are not aware of) and in each round k presents juror c_k to the parties. For similar purposes, it will sometimes be useful to view *RAN* as first drawing a panel of n jurors and then (uniformly at random) selecting j jurors among these n to form the effective jury.

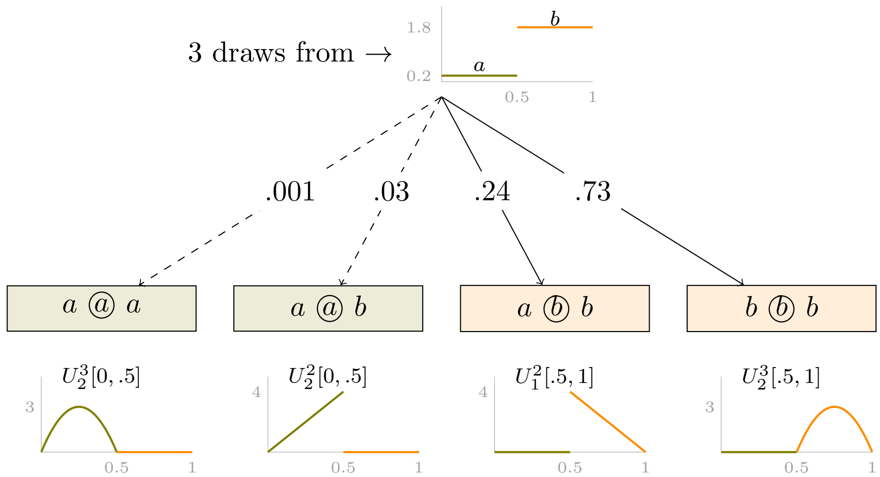
3 Excluding extremes and representation of minorities: An illustrative example

To illustrate the differences between the two procedures, consider the simple case $d = p = j = 1$ together with distributions $C_a \sim U[0, 0.5]$ and $C_b \sim U[0.5, 1]$. Also, suppose that $r = 0.1$, i.e., there is a minority of 10% of group-a jurors in the population.

[15] Each subgame can be characterized by the number of jurors κ that remain to be selected, the number of challenges left to the defendant δ, and the number of challenges left to the plaintiff π. The parties threshold in subgame (κ, δ, π) are a function of the value of subgames $(\kappa - 1, \delta, \pi)$, $(\kappa, \delta - 1, \pi)$, and $(\kappa, \delta, \pi - 1)$ (which can result from the parties action in (κ, δ, π)) and the distribution of C, see Brams and Davis (1978).

[16] By "timing", here, we mean the order (potentially simultaneous) in which the parties decide whether or not to challenge the presented juror.

Figure 1: Illustrative example, equilibrium outcomes under *STR*

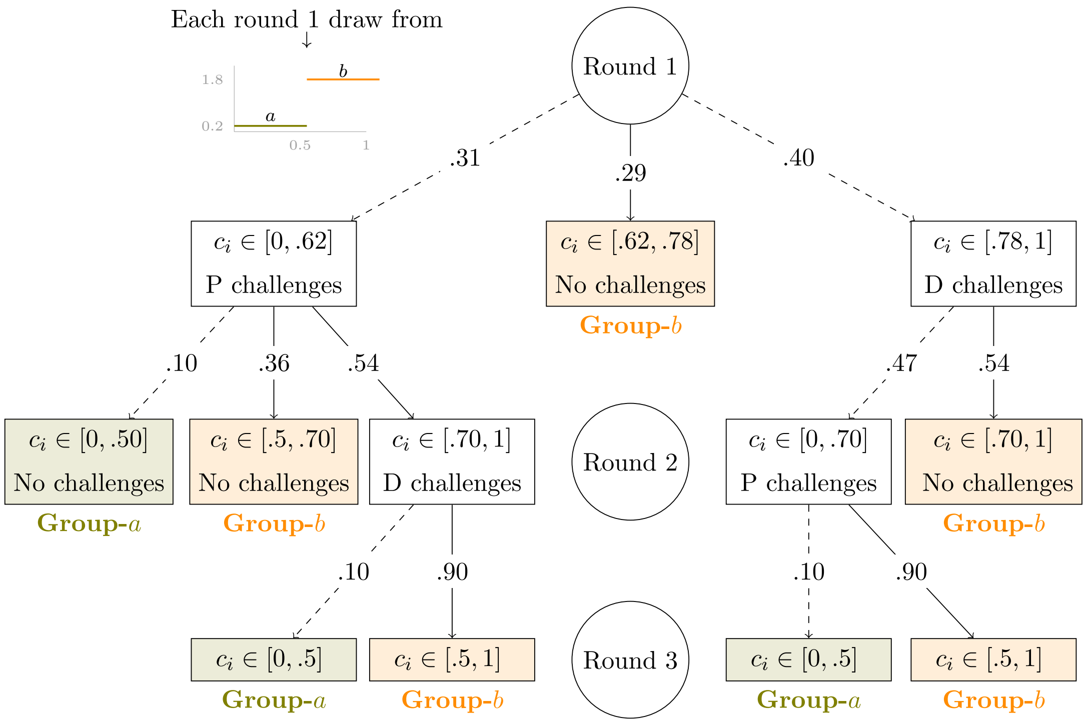


Note: The figure describes the equilibirum of STR assuming $j = p = d = 1$, $C_a \sim U[0, 0.5]$, $C_b \sim U[0.5, 1]$, and $r = 0.10$. The initial node illustrates distribution $C = 0.10 \cdot C_a + 0.9 \cdot C_b$. The numbers on each arrow indicate the probability of drawing a panel with the group-composition represented in the pointed boxes (conditional on each panel composition, the circled letter in the box corresponds to the group-membership of the selected juror). Dashed arrows correspond to outcomes that lead to the selection of a group-a juror and the graph underneath each box shows the distribution of conviction probabilities for the selected juror.

Let $U_x^n[a, b]$ denote the x-th order statistic for a $U[a, b]$ random sample of size n. With this notation, Figure 1 shows the group-membership and distribution of conviction probability for the juror selected under STR, conditional on the composition of the panel. Observe that in this example, if there are group-a jurors in the panel, one of them is systematically challenged by the plaintiff. Therefore, for a group-a juror (i.e., a minority juror) to be selected under STR, there need to be at least two group-a jurors in the panel of $n = 3$ presented to the parties. This occurs with probability 0.03.

In contrast, a group-a juror can be selected under *S&R* even if the panel contains a single group-a juror. To understand why, consider the equilibrium of *S&R* which is illustrated in Figure 2. If a group-b prospective juror with a sufficiently low conviction probability ($c_i \in [0.5, 0.62]$) is presented first, then it will be challenged by the plaintiff. This leads to a subgame in which only the defendant has challenges left and a group-a juror is more likely to be selected than if a juror was randomly drawn from the population. In particular, any group-a juror presented at the beginning of this later subgame is left unchallenged by the defendant and selected to be the effective juror (even if this juror is the only group-a juror

Figure 2: Illustrative example, equilibrium strategies and outcomes under *S&R*



Note: The figure describes the equilibrium strategies conditional on the conviction probability of the juror drawn in each round for the case $j = d = p = 1$, $C_a \sim U[0, 0.5]$, $C_b \sim U[0.5, 1]$ and $r = 0.10$. Dashed arrows correspond to paths that may lead to the selection of a group-a juror. The numbers on each arrow indicate the probability of the path conditional on reaching the previous node. The second row of text inside boxes indicates an equilibrium action, whereas bold text below boxes indicates the group of the selected juror in the game outcome. In round 3, challenges from both parties are exhausted and the parties do not take any action.

in the panel because the third juror — who, in this case, is never presented to the parties — happens to be a group-b juror). This course of action follows from P's choice to challenge a group-b juror with low conviction probability in the first round, which leaves P without challenges left in the second round. This choice of P is optimal from the perspective of the first round of *S&R* (*before* the plaintiff learns that the second juror in the panel is a group-a juror), but suboptimal under *STR* where, having observed the conviction probability of all jurors in the panel, the plaintiff would have challenged the group-a juror instead.

Considering only the branch of the *S&R* game-tree that starts with a challenge from P, the probability of selecting a group-a juror is almost $0.05 = 0.31 \cdot (0.54 \cdot 0.1 + 0.10)$. Adding the possibility that a minority juror is selected after D challenges in the first round followed by a challenge from P in the second round (which happens with probability $0.4 \cdot 0.47 \cdot 0.1 \approx 0.02$), the probability of selecting a minority juror under *S&R* is 0.067.[17] This is larger than the probability under *STR*, 0.03, yet smaller than under *RAN*, 0.10.

In this example, the better representation of minority jurors produced by *S&R* comes at the expense of selecting more extreme jurors. Suppose for the sake of illustration that jurors are considered extreme if they come from the top or bottom 5th percentile of C. In our example, the bottom and top 5th percentile corresponds to conviction probabilities below 0.25 and above 0.94, respectively. The selected juror is within the bottom range with probability 0.015 under *STR* versus 0.033 under *S&R*, and in the top range with probability 0.076 under *STR* versus 0.083 under *S&R*.

To understand the source of these differences, let us consider the bottom 5th percentile $[0, 0.25]$ (a symmetric explanation applies to the *top* 5th percentile). As indicated in Figure 1, when *STR* selects a group-a juror — the type of juror whose conviction probability could possibly be in the bottom 5th percentile — the distribution of that juror's conviction probability follows the middle or upper order-statistics of a random sample from C_a. These order-statistics are unlikely to result in the selection of a juror with conviction probability in the bottom 5th percentile. In contrast, as Figure 2 illustrates, all paths leading *S&R* to select a group-a juror result in the juror's conviction probability being drawn from $U[0, 0.5]$ itself,which makes *S&R* more likely to select a juror in the bottom 5th percentile than *STR*.

In the next two sections, we investigate the extent to which the advantages of *S&R* in terms of minority-representation and of *STR* in terms of exclusion of extreme generalizes beyond this illustrative example.

4 Exclusion of extremes

In the United States, one of the objectives of the jury selection process is to guarantee an impartial jury as dictated by the Sixth Amendment of the Constitution. In this respect, the peremptory challenge procedures implemented in U.S. jurisdictions are often viewed as a way to foster impartiality by preventing extreme potential jurors from serving on the

[17] These are the only cases in which a minority juror can be selected under *S&R*. In particular, jurors accepted in the first round are always group-b jurors ($c_i \in [0.62, 0.78]$). So are jurors accepted in the second round following a challenge from D is the first round ($c_i \in [0.70, 1]$).

effective jury.[18] In the context of our model, we interpret this goal as that of limiting the presence in the jury of jurors from the tails of the distributions of conviction probabilities.

In Sections 4 to 6, we refer to a juror i as *extreme* if its conviction probability c_i lies below or above given thresholds (we refer the reader to Section 8 for results under an alternative definition). For brevity, we will focus on jurors who qualify as extreme because their conviction probability lies *below* some threshold $\underline{c} > 0$. This is without loss of generality and all our results about extreme jurors apply symmetrically to jurors whose conviction probability lies *above* a given threshold $\overline{c} < 1$.

In our illustrative example, jurors in the bottom 5th percentile of C are selected less often under *STR* than *S&R*. This is not true in general. Fixing a particular threshold $\underline{c} > 0$ — or percentile of C — to characterize jurors as extreme, there always exists distributions of C and values of d, p, and j such that *S&R* selects fewer extreme jurors than *STR*. However, our first result shows that regardless of the distribution and value of the parameters, there always exists a threshold sufficiently small such that, if jurors are called "extreme" below that threshold, the probability of selecting extreme jurors is greater under *S&R* than under *STR*.

Let $\mathbb{T}_M(x;c)$ denote the probability that there are at least x jurors with conviction probability *smaller or equal* to c in the jury selected by procedure M.

Proposition 1. *For any $x \in \{1,\dots,j\}$, there exists $\underline{c} > 0$ such that $\mathbb{T}_{STR}(x;c) < \mathbb{T}_{S\&R}(x;c)$ for all $c \in (0,\underline{c})$.*

All proofs are in the appendix. A symmetric statement, which we omit, applies for extreme jurors at the right-end of the distribution. Note that Proposition 1 can be rephrased in terms of stochastic dominance. Let N^c_M denotes the expected number of jurors of type $c_i \leq c$ in the jury selected by procedure M. Then, Proposition 1 says that there exists $\underline{c} > 0$ and such that $N^c_{S\&R}$ has first-order stochastic dominance over N^c_{STR} for all $c \in (0,\underline{c})$. A direct corollary of Proposition 1 is therefore that the *expected number* of extreme jurors is larger under *S&R* than under *STR*.

For some intuition about Proposition 1, consider the case $x = 1$. As illustrated in Section 3, the panel must be composed of more than one extreme juror for *STR* to select

[18] See Footnote 1 and its associated quote. For legal arguments in favor of peremptory challenges based on the Sixth Amendment, see, among others, Beck (1998), Biedenbender (1991), Bonebrake (1988), Horwitz (1992), and Keene (2009).

at least one such juror (since, if there is a single extreme juror in the panel, that juror is systematically challenged by the plaintiff). In contrast, even in panels with a single extreme juror, the extreme juror can be part of the effective jury resulting from *S&R*. This happens, for example, if the extreme juror is presented to the parties after they both exhausted all their challenges. The single extreme juror can also be accepted by both parties if its conviction probability is sufficiently close to $\underline{c}$ and it is presented after the plaintiff used most of its challenges on non-extreme potential jurors.[19] The proof then follows from the fact that, as $\underline{c}$ tends to zero, the probability that the panel contains more than one extreme juror goes to zero faster than the probability the panel contains a single extreme juror.[20]

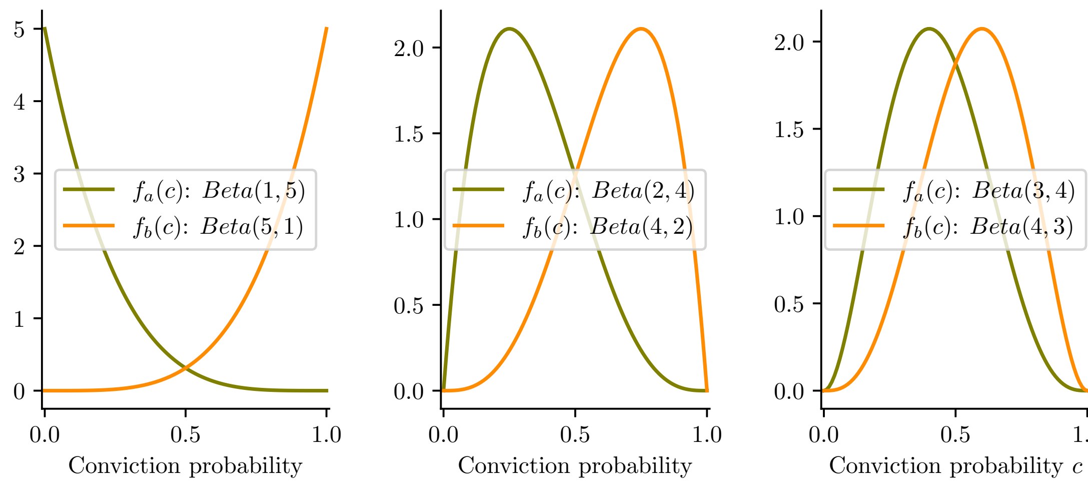
Proposition 1 is silent about the value of the threshold $\underline{c}$ below which *STR* selects fewer jurors than *S&R*, as well as the size of $\underline{\mathbb{T}}_{S\&R}(x;c)-\underline{\mathbb{T}}_{STR}(x;c)$ for $c<\underline{c}$. These values depend on the models' parameters. To illustrate, we simulate $\underline{\mathbb{T}}_{STR}(1;c)$ and $\underline{\mathbb{T}}_{S\&R}(1;c)$ using $j=12$, $d=6$, and $p=6$, a typical combination of jury size and number of peremptory challenges in U.S. jurisdictions. For the distribution of conviction probabilities in the population, we use symmetric mixtures of beta distributions that represents a population made of two groups with polarized views, which allows easier comparison with the results from Section 5 in which we study group-representation. We provide simulation results for three mixtures of the distributions illustrated in Figure 3, which are meant to represent extreme (Panel (a)), moderate (Panel (b)), and mild levels of polarization (Panel (c)). Additional simulations results using $U[0,1]$ instead are reported in Appendix B.

Using these parameters, *STR* is found to exclude more extreme jurors than *S&R* even when the threshold for defining jurors as extreme is relatively high. As illustrated in Figure 4, the difference between the propensity of *STR* and *S&R* to select extreme jurors is sizable. For example, in all three sets of simulations, only about 1% of juries selected by

[19]Subgames in which the defendant has more challenges left than the plaintiff can lead the plaintiff to be conservative and accept jurors who are "barely extreme" ($c_i \approx \underline{c}$) in order to save its few challenges left for "very extreme" jurors ($c_i \approx 0$).

[20]Proposition 1 crucially depends on averaging across all possible panels and does *not* state that *STR* rejects more extreme jurors than *S&R* for *any* particular realization of the panel. The latter would obviously imply Proposition 1 but turns out to be false in general. For a counterexample, let $j=d=p=1$. Consider a panel of three jurors with $c_2<c_3<\underline{c}$ and $c_1>\underline{c}$ and the index of the jurors indicating the order in which they are presented under *S&R*. For this panel, *STR* always leads to the selection of extreme juror 3. In contrast, provided c_2 falls between the challenge thresholds of the defendant and the plaintiff in the first round (which happens with positive probability), *S&R* selects non-extreme juror 2.

Figure 3: Distributions of conviction probabilities by group under extreme, moderate, and mild group-polarization



(a) Extreme **(b) Moderate** **(c) Mild**

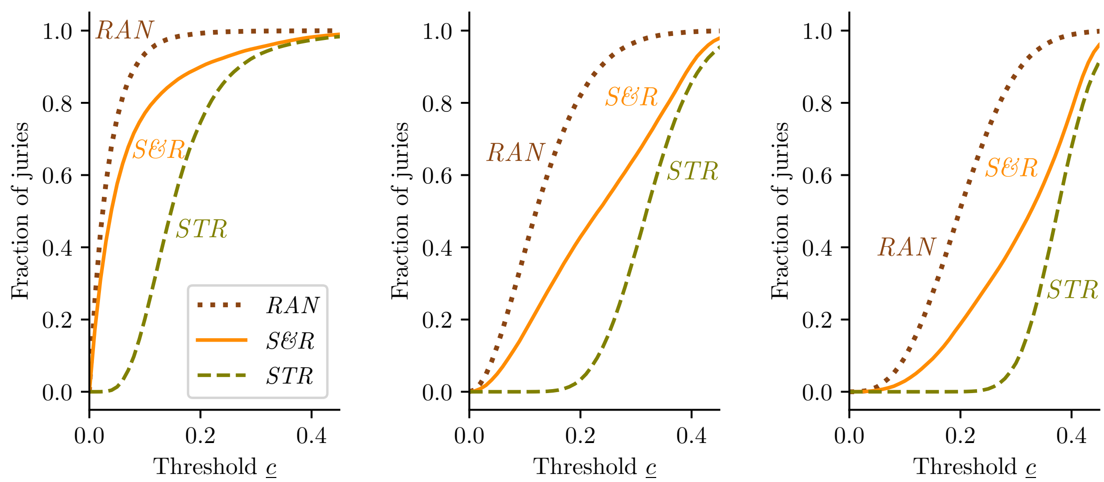
STR feature at least one juror with conviction probability below the 10th percentile of the distribution (the 10th percentile corresponds to 0.01 under the extreme polarization distribution, 0.17 under moderate polarization, and 0.25 under mild polarization). Under *S&R*, the proportion of juries with at least one juror below the 10th percentile rises to 56% with extreme polarization, 35% with moderate polarization, and remains quite high at 30% even under mild polarization. For comparison, a random selection would have resulted in about 73% of the juries featuring at least one such juror.

In these simulations, both procedures select fewer extreme jurors than a random draw from the population. Somewhat surprisingly, this is not true in general. There exist distributions and values of the parameters d, p and j for which *S&R* selects *more* extreme jurors than *RAN*, no matter how small the threshold below which a juror is considered as extreme. In contrast, as we show in the next proposition, *STR* always selects fewer extreme jurors than *RAN*.

Proposition 2. *For any $x \in \{0, \ldots, j-1\}$, there exists $\underline{c} > 0$ such that $\mathbb{T}_{STR}(x; c) < \mathbb{T}_{RAN}(x; c)$ for all $c \in (0, \underline{c})$.*[21]

[21] Proposition 2 generalizes Theorem 2 in Flanagan (2015) which shows that there always exists $\underline{c} > 0$ such that $\mathbb{T}_{STR}(n; c) < \mathbb{T}_{RAN}(n; c)$ for all $c \in (0, \underline{c})$.

Figure 4: Fraction of juries with at least one extreme juror

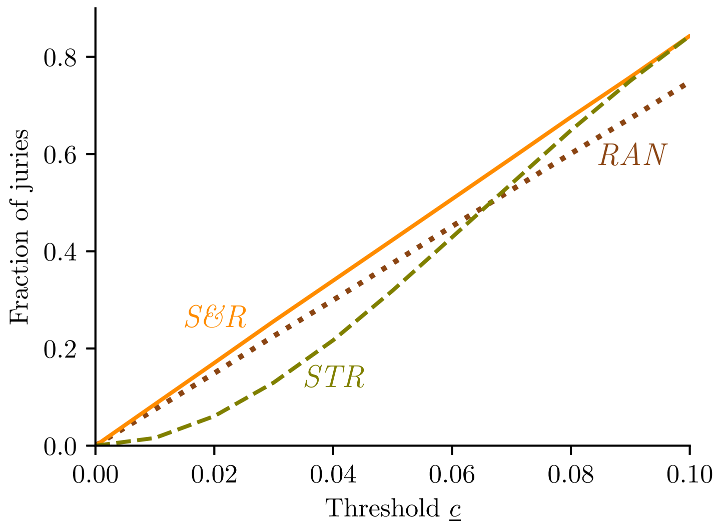


(a) Extreme **(b) Moderate** **(c) Mild**

Note: For each set of parameters, results on the vertical axis are averages across 50,000 simulated jury selections, fixing $j = 12$, $d = p = 6$, and $C \sim 0.5 \cdot C_a + 0.5 \cdot C_b$ throughout (with the distributions for C_a and C_b illustrated in Figure 3). Each line illustrates the fraction of juries with at least one extreme juror, where a juror is considered extreme if her conviction probability falls below the threshold $\underline{c}$ corresponding to the value on the horizontal axis.

Figure 5 illustrates Proposition 2 and the fact that a similar statement does not hold for *S&R*. For the simulations in the figure, we let $j = d = p = 1$ and adopt an extremely polarized distribution of conviction probabilities with $C \sim 0.75 \cdot U[0, 0.1] + 0.25 \cdot U[0.9, 1]$. In this case (as in others), *STR* excludes extreme jurors more often than *RAN* because, for any realization of the panel, the juror with the lowest conviction probability is never selected under *STR* (whereas the same juror is selected with positive probability under *RAN*). Under *S&R*, however, if the distribution is sufficiently right-skewed, the plaintiff is more likely than the defendant to challenge in the first round. A challenge by the plaintiff in the first round leads to a subgame in which only the defendant has challenges left and the selection of an extreme juror is more likely than under a random draw. When they are sufficiently large (i) the added probability of selecting an extreme juror when the defendant has more challenges left than the plaintiff, coupled with (ii) the probability of a challenge by the plaintiff in the first round can, as in the simulation depicted in Figure 5, lead to *S&R* selecting more extreme jurors than *RAN*.

Figure 5: Fraction of juries with at least one extreme juror (case in which *S&R* is more likely to pick extreme jurors than *RAN*)



Note: For each set of parameters, results on the vertical axis are averages across 50,000 simulated jury selections, fixing $j = d = p = 1$, and $C \sim 0.75 \cdot U[0, 0.1] + 0.25 \cdot U[0.9, 1]$ throughout. Each line illustrates the fraction of juries with at least one extreme juror, where a juror is considered extreme if her conviction probability falls below the threshold $\underline{c}$ corresponding to the value on the horizontal axis.

We could not fully characterize the situations in which *S&R* selects more extreme jurors than RAN, and we never observed such a situation in simulations where C is a symmetric mixture of beta or uniform distributions. The example in Figure 5 (as well as other examples we found) requires extreme skewness in the distribution, which may be viewed as unlikely. In this sense, situations in which *S&R* selects *more* extreme jurors than RAN might represent worst-case scenarios for *S&R*'s ineffectiveness at excluding extreme juror (rather than ordinary situations).

5 Representation of minorities

In this section, we study the extent to which STR's tendency to exclude more extreme jurors than *S&R* impacts the representation of minorities under the two procedures. Without loss of generality, we let group-a be the minority group. Since the parties do not care intrinsically about group-membership, any asymmetry in the use of their challenges arises from heterogeneity in preferences for conviction between groups. In our simulations, we assume that group-a is biased in favor of acquittal in the sense that C_b first-order stochastically dominates C_a.[22]

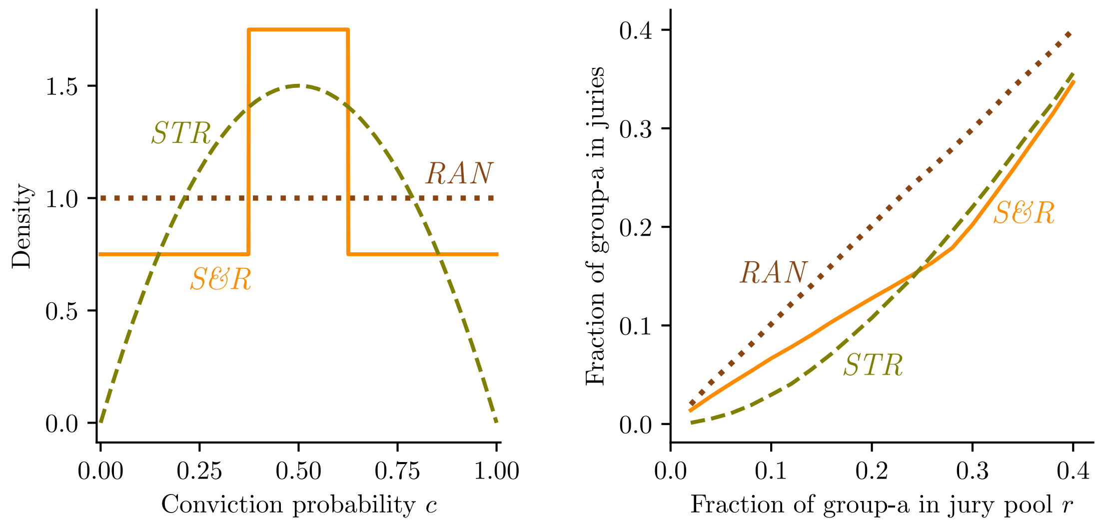
As suggested by Proposition 1, which procedure better represents minorities strongly depends on the polarization between the two groups, and the concentration of minority jurors at the tails of the distribution of conviction probabilities. To illustrate, suppose that $d = p = j = 1$ and $C \sim U[0, 1]$. For this case, the distributions of conviction probabilities for the juror selected under RAN, STR, and $S\&R$ are displayed in Figure 6(a). Consistent with Proposition 1, below some threshold $\underline{c} \approx 0.25$, the probability of selecting a juror i with $c_i < \underline{c}$ is lower under STR than under $S\&R$. If the two groups are polarized and the distribution of C_a is sufficiently concentrated below $\underline{c}$, it follows that STR selects a minority juror less often than $S\&R$. But the same is not true if the distributions lack polarization or the minority is too large. For example, decompose C as follows: $C \sim U[0, 1] = rU[0, r] + (1 - r)U[r, 1]$. Since the parties only care about a juror's conviction probability and not about its group-membership *per se*, the value of r in these decompositions does not affect the distributions of conviction probabilities for the juror selected under RAN, STR, or $S\&R$. Then, letting $C_a \sim U[0, r]$ and $C_b \sim U[r, 1]$, Figure 6(b) illustrates how low values of r — which concentrate minorities at the bottom of the distribution — make $S\&R$ select more minorities than STR, whereas higher values of r — which spread the minority over a larger range of conviction-types — make STR select more minorities than $S\&R$.

From this example, we see that non-overlapping group-distributions are not sufficient to guarantee that $S\&R$ selects more minority jurors than STR. Neither is making the minority arbitrarily small. For example, regardless of the size of the minority r, concentrating the support of the minority distribution inside the interval $[0.2, 0.3]$ would result in STR selecting more minority jurors, as can be seen from Panel 6(a). However, combining a small minority with group-distributions that minimally overlap concentrates the distribution of group-a at the tails which, as suggested by Proposition 1, makes $S\&R$ select more minorities than STR.

Formally, consider a sequence of triples $\{(C_a^i, C_b^i, r^i)\}_{i=1}^{\infty}$. If,

[22]We also simulated the scenario in which the minority is biased towards conviction, the results, which we report in the Appendix, are symmetrically very close).

Figure 6: Jury selection and minority representation in size-1 juries



(a) Distribution of c for selected juror **(b) Minority representation in juries**

Note: For each set of parameter, results on the vertical axes are averages across 20,000 simulated jury selections, fixing $j = 1$, $d = p = 1$, and $C \sim r \cdot U[0, r] + (1 - r) \cdot U[r, 1]$ throughout. The distribution in panel (a) is independent of r whether the lines in panel (b) interpolate results from 20 values of r.

(i) $r^i \in (0, 1]$ for all $i \in \mathbb{N}$ with $\lim_{i\to\infty} r^i = 0$, and

(ii) C_a^i and C_b^i converge in distribution to C_a^* and C_b^*, with either $\mathbb{P}(C_a^* < C_b^*) = 0$ or $\mathbb{P}(C_a^* > C_b^*) = 0$,

then we say that **there is a vanishing minority and group-distributions that do not overlap in the limit**. For any such sequence, let $\mathbb{A}_M^i(x)$ denote the probability that there are at least x minority jurors in the jury selected by procedure M when group-distributions are C_a^i and C_b^i and the proportion of minority jurors in the population is r^i.

Proposition 3. *Suppose that, under $\{(C_a^i, C_b^i, r^i)\}_{i=1}^{\infty}$, there is a vanishing minority and group distributions that do not overlap in the limit. Then for all $x \in \{1, \ldots, j\}$, there exists j sufficiently large such that $\mathbb{A}_{S\&R}^i(x) > \mathbb{A}_{STR}^i(x)$ for all $i > j$.*[23]

[23]Note that, despite the argument presented in the motivating example illustrated in Figure 6, Proposition

Table 1: Representation of Group-a when Group-a is a minority of the pool

Polarization	Extreme		Moderate		Mild		(All)
Procedure	*S&R*	*STR*	*S&R*	*STR*	*S&R*	*STR*	*RAN*
Average fraction of minorities	0.10	0.08	0.18	0.16	0.23	0.23	0.25
Standard deviation	0.11	0.11	0.12	0.12	0.12	0.12	0.12
Fraction of juries with at least 1	0.57	0.45	0.88	0.84	0.96	0.95	0.97

(a) Group-a represents 25% of the jury pool

Polarization	Extreme		Moderate		Mild		(All)
Procedure	*S&R*	*STR*	*S&R*	*STR*	*S&R*	*STR*	*RAN*
Average fraction of minorities	0.02	0.00	0.05	0.04	0.09	0.08	0.10
Standard deviation	0.04	0.01	0.07	0.06	0.08	0.08	0.09
Fraction of juries with at least 1	0.17	0.02	0.47	0.38	0.67	0.64	0.72

(b) Group-a represents 10% of the jury pool

Note: The rows report the average number and standard deviation of group-a jury members, and the percent of juries with at least one group-a jurors, out of 50,000 simulations of jury selection with parameters $j = 12$ and $d = p = 6$. Conviction probabilities are drawn for from $Beta(5,1)$, $Beta(1,5)$, respectively for Group-a, Group-b jurors (Extreme), from $Beta(4,2)$, $Beta(2,4)$ (Moderate), and from $Beta(4,3)$, $Beta(4,3)$ (Mild); see Figure 3 for the shape of these distributions.

Given the result in Proposition 3, it is natural to wonder how small the minority and the overlap between the group-distributions must be for *S&R* to select more minority jurors than *STR*. When the latter is true, one may also wonder about the size of $\mathbb{A}_{S\&R}(x;r) - \mathbb{A}_{STR}(x;r)$ is. Again, the answer naturally depends on the model's parameters. To inform these questions, we ran a set of simulations with $d = p = 6$ and $j = 12$ using the distributions displayed in Figure 3, where the green lines in each panel represent f_a and the yellow lines f_b.

3 does not follow directly from Proposition 1. The reason is that, unlike in the motivating example, most of the sequences $\{(C_a^i, C_b^i, r^i)\}_{i=1}^{\infty}$ covered by Proposition 3 are such that $C^i = r^i C_a^i + (1 - r^i) C_b^i$ varies across the sequence (i.e., $C^j \neq C^h$ for most $j, h \in \mathbb{N}$).

The results of our simulations, displayed in Table 1, suggest that *S&R* might select more minority jurors than *STR* even when the size of the minority is relatively high (as high as 25%) and the overlap between the group-distributions significant. However, without stark polarization across groups,[24] differences between the procedures' propensities to select minority jurors appear to be small. For example, under the distributions we labeled as "extreme group heterogeneity" and with minorities representing 10% of the population, only 2.3% of juries selected by *S&R* include at least one minority juror whereas this number rises to 17.1% under *S&R* (random selection would generate over 70% of such juries). However, under the distributions we labeled as "mild group heterogeneity", the same numbers become 66.5% under *S&R* and 64.5% under *STR* (random selection would generate over 71.9% of juries with at least one minority juror in this second case).

6 Changing the number of challenges

So far, we have compared *STR* and *S&R* assuming that the number of challenges the parties can use, d and p, was the same under each procedure. This was motivated by the fact that judges often have a lot of freedom in selecting the procedure through which the parties use their challenges (see Footnote 3). In contrast, the number of challenges that the parties can use are typically specified more rigidly by state rules of criminal procedure.

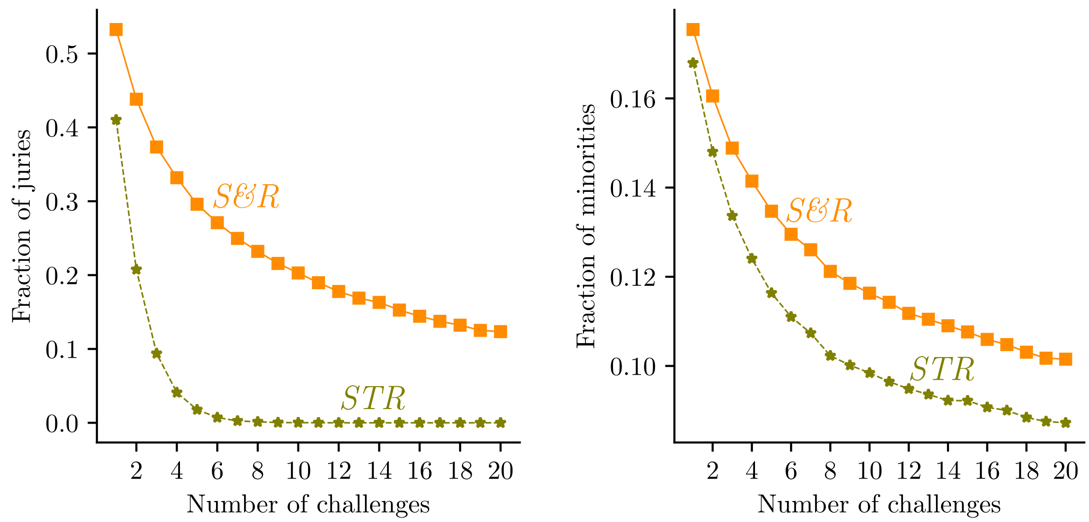
In the last decades, several states have, however, reduced the number of challenges the parties can use.[25] In some instances, these reforms also clarify or alter the jury selection procedures used in the state.[26] In the context of such broader reforms, it is natural to ask how the ability to change *both* the number of challenges the parties are entitled to *and* the procedure through which the parties exert their challenges affect the trade-off between the

[24]Recall that C_a and C_b represent the parties' *beliefs* that randomly drawn group-a or group-b jurors eventually vote to convict the defendant. Polarized C_a and C_b, therefore, corresponds to groups that are *perceived* by the parties to have different probabilities of voting for conviction (whether or not this materializes when jurors actually vote on conviction at the end of the trial).

[25]Examples include California's Senate Bill 843, passed in 2016, which reduces the number of challenges a criminal defendant is entitled to from 10 to 6 (for charges carrying a maximal punishable of one year in prison, or less).

[26]Examples include the 2003 reform of jury selection in Tennessee where some aspects of the jury selection procedure were codified to apply uniformly across the state, while the number of peremptory challenges was also slightly reduced (see Cohen and Cohen, 2003).

Figure 7: The effect of varying the number of challenges



(a) Fraction of extreme jurors **(b) Fraction of minority jurors**

Note: Fraction of juries with at least one juror below the 10th percentile (left panel) and fraction of minority jurors (right panel). For each set of parameters, results on the vertical axes are averages across 50,000 simulated jury selections, fixing $j = 12$ and $C \sim 0.2 \cdot C_a + 0.8 \cdot C_b$ throughout (with the distributions for $C_a \sim Beta(2,4)$ and $C_b \sim Beta(4,2)$, see Figure 3(b)). The values of $d = p$ are on the horizontal axes.

exclusion of extreme jurors and the representation of minorities.

Throughout this section, we fix an arbitrary value of j and consider varying $d = p$. For any procedure M, let M-y denote the version of M when $d = p = y$. The notation for the two previous sections then carries over, with $\mathbb{T}_{M\text{-}y}(x; c)$ denoting the probability that at least x jurors with conviction probability below c are selected under M-y, and $\mathbb{A}_{M\text{-}y}(x)$ the probability that at least x minority jurors are selected under M-y.[27]

For illustration purposes, we first consider the case $C \sim 0.2 \cdot C_a + 0.8 \cdot C_b$, $C_a \sim Beta(2,4)$ and $C_b \sim Beta(4,2)$ (C_a and C_b are illustrated in the Figure 3(b)), and consider a juror as extreme if its conviction probability falls in the bottom 10th percentile of C (which here equals 0.27). Unsurprisingly, the fraction of juries with at least one *extreme* jurors *decreases*

[27] Again, in the case of extreme jurors, we focus on jurors who qualify as extreme because their conviction probability falls *below* a certain threshold $\underline{c}$, though all of our results hold symmetrically for jurors who qualify as extreme because their conviction probability lies *above* a certain threshold $\overline{c}$,

as the number of challenges awarded to the parties increases, regardless of the procedure that is used (Figure 7(a)). Conversely, the fraction of *minority* jurors *decreases* with the number of challenges under both procedures (Figure 7(b)). In other words, for both *STR* and *S&R*, more challenges lead to fewer extreme jurors being selected at the expense of a worse representation of minorities.

As Figure 7(a) illustrates, however, increasing the number of challenges decreases the selection of extreme jurors much faster under *STR* than under *S&R*. As a consequence, for all values of $y \in \{2, \dots, 18\}$, there exists $w < y$ such that *STR-w* performs better than *S&R-y* in terms of *both* objectives.[28]

The latter is not true in general. Even when there exists w such that *STR-w* better represents minorities than *S&R-y*, *STR-w* might still exclude fewer extreme jurors than *S&R-y* if jurors are considered extreme when their conviction probability falls below an arbitrary $c > 0$. However, an extension of Proposition 1 shows that when such a w exists, there also exists $\underline{c} > 0$ such that if jurors are considered extreme when their conviction probability falls below $\underline{c}$, *STR-w* performs better than *S&R-y* in terms of both objectives.

Proposition 4. *Consider any $x \in \{1, \dots, j\}$ and any $y \geq 1$. Suppose that there exists $w \geq 1$ such that $\mathbb{A}_{STR\text{-}w}(x) > \mathbb{A}_{S\&R\text{-}y}(x)$. Then for some $\underline{c} > 0$, we also have $\mathbb{T}_{STR\text{-}w}(x; c) < \mathbb{T}_{S\&R\text{-}y}(x; c)$ for all $c \in (0, \underline{c})$.*

7 Empirical evidence

As emphasized in the analysis so far, group asymmetries in jury representation exist to the extent that groups have polarized preferences for conviction. In this section, we use jury selection data to estimate the distribution of conviction probabilities and provide quantitative evidence of the effect of jury selection procedures and their differences.

Jury selection data is to our knowledge relatively scarce.[29] For the purposes of this Section, we exploit data from Craft (2018) on peremptory strikes in the Fifth Circuit Court

[28] Specifically, in this example, for any $y \in \{2, \dots, 18\}$, there exists $w \in \{1, \dots, y-1\}$ such that $\mathbb{A}_{STR\text{-}w}(1) > \mathbb{A}_{S\&R\text{-}y}(1)$ and $\mathbb{T}_{STR\text{-}w}(1; 0.27) < \mathbb{T}_{S\&R\text{-}y}(1; 0.27)$).

[29] Besides the data used in this section, another important source is the data of jury selection in North Carolina described in Wright et al. (2018) and analyzed in Flanagan (2018). We do not use this source because the jury selection procedures adopted in these jurisdictions do not conform to the rules we study in this paper.

District of Mississippi from 1992 to 2017, where a version of STR was used to select jurors.[30]

For the vast majority of trials, 12 jurors were selected with 6 challenges available to each party.[31] For each trial, the data reports the race and gender of the potential jurors, whether a juror was struck by the defendant or the state, and the race and gender composition of the seated jury and alternate jurors. This allows the computation of jury composition by race, and the computation of challenges by race for each party.

We limit our analysis to the juries' racial composition focusing on Black and White jurors only[32]. Assuming that the distributions of conviction probabilities in each group belong to the class of beta distributions, the model parameters are five: the fraction of whites in the jury pool, $1 - r$, which we directly observe in the data, and the four parameters of f_{Blacks}, f_{Whites}. The data we observe does not allow to identify both of these distributions. Given r, for any given $f_{\text{Blacks}} = Beta(\alpha_a, \beta_a)$, it is always possible to find $f_{\text{Whites}} = Beta(\alpha_b, \beta_b)$ that replicates the same proportion of whites struck by the defendant and by the State of Mississippi, the plaintiff (which in turn determine the fraction of whites in the jury). Intuitively, the reason behind this lack of identification is that it is possible to shift some mass of both distributions to the right without changing, on average, the racial composition of the juries.[33]. While this shift would cause the conviction frequency at trial to change, using this moment for identification would not change the outcomes we focus on in this paper for STR (see Footnote 34).

In Table 2 we report some summary statistics from the data. The sample contains 292 trials, of which 229 include black defendants. We exclude all jurors dismissed by the judge for causes that are not the focus of our analysis. Hence, we define the size of the panel as the sum of the number of jurors, alternate jurors, and jurors dismissed by either the state or the

[30] While the adopted procedure differs in some details from the stylized version we analyzed in this paper, we assume that in equilibrium, its outcome conforms to that of STR. In addition, the number of jurors selected and the number of challenges available sometimes differ by type of trial.

[31] As we explain below, there is some variation to the number of jurors in the data and to the number of challenges used by parties (due to variation in the kind of offenses being prosecuted as well as in judges decisions in the allocation of additional challenges for the selection of alternate jurors). However, the moments we use for identification rely only on race ratios and are relative stable across juries of different size.

[32] The full sample includes almost 15,000 jurors, of which 26% are Black, 42% are White, 32% are of unknown race, and only 3 Latinos and 1 Asian which we pool with the Whites.

[33] With beta distributions, matching these two moments also matches the proportion of juries with x jurors of a given race, for all $x \in \{0, \ldots, j\}$, making it impossible to use higher moments for identification.

Table 2: Summary statistics

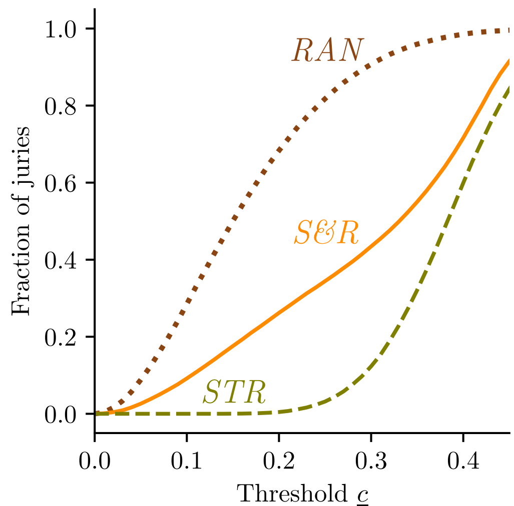
Sample selection	(1)	(2)	(3)	(4)	(5)
Defendant	White	Black	Black	Black	Black
Size of jury pool	Any	Any	≤ 27	Any	≤ 27
Include unknown race jurors	Yes	Yes	Yes	No	No
N. of Trials	63	229	162	131	99
Trial statistics					
Average size of jury pool	26.1	26.9	23.7	26.2	23.5
(std)	(5.0)	(5.8)	(2.5)	(5.7)	(2.6)
Average size of jury	12.0	12.0	12.0	12.0	12.0
(std)	(0.3)	(0.4)	(0.4)	(0.2)	(0.2)
% with unknown race in jury pool	31.2	30.7	26.9	0.0	0.0
Percentage of whites*					
in jury pool	63.1	62.7	63.1	66.5	65.9
in jury	61.0	66.8	67.8	70.5	69.7
among struck by the defendant	86.2	91.4	92.3	93.1	92.9
among struck by the state	40.8	23.6	21.4	23.5	21.6

Standard deviation in parenthesis. *Percentage of white jurors in samples (1), (2), and (3) computed among jurors that have been classified as either whites or blacks

defendant. There is some variation in the size of both the juries and the panel, in part due to the fact that the process of selecting alternate jurors is separate. Unfortunately, the data does not distinguish between jurors who were dismissed in the course of selecting regular jurors, or in the course of alternates. We present data for 5 samples that vary depending on the race of the defendant, the size of the panel, and whether or not we include panels containing jurors of unknown race. These show that the racial composition of juries and challenges is affected by the the race of the defendant but is only weakly affected by the way we select our sample.

The average size of the jury (excluding alternates) is 12 in all samples, though the panels are slightly over 24, mainly because they include potential alternate jurors (and because, in

Figure 8: Counterfactual analysis: Juries with at least one extreme juror

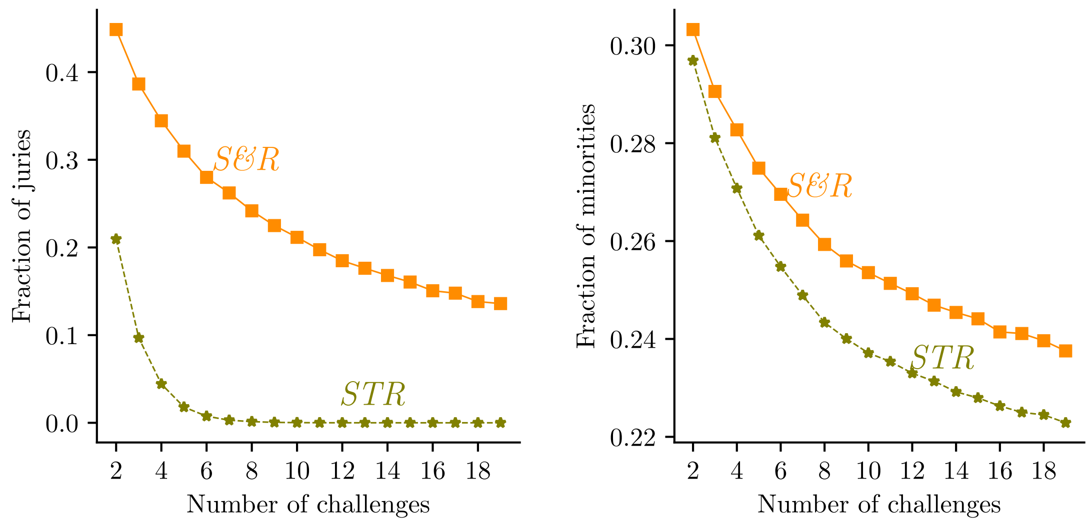


Note: For each set of parameters, results on the vertical axis are averages across 50,000 simulated jury selections, fixing $j = 12$, $d = p = 6$, and $C \sim 0.341 \cdot C_a + 0.659 \cdot C_b$ throughout (with the distributions for $C_a \sim Beta(2, 4)$ and $C_b \sim Beta(5.80, 3.95)$). Each line illustrates the fraction of juries with at least one extreme juror, where a juror is considered extreme if her conviction probability falls below the threshold c corresponding to the value on the horizontal axis.

some cases, judges may grant additional challenges to the parties). Challenge behavior is affected by the race of the defendant: Juries with black defendants have a higher percentage of whites than the panel does, whereas juries with white defendants include fewer whites. When the defendant is black the defense challenges a higher fraction of white jurors, and the state a higher fraction of black jurors. Variation in the size of the jury pool has little impact on the racial composition of the juries or challenged jurors (for either party). Focusing on trials with Black defendants, the fraction of whites in the pool is quite stable across all 5 samples (between 62.7 and 66.5 percent). This is predicted by our theory when jurors have polarized views that favor defendants of their own race. The behavior of the parties differ substantially by race: in sample (5), which we use to estimate our model, 93% of the jurors struck by the defendant are white, whereas only 22% of the jurors struck by the state are White. We use these two moments to estimate the distribution of conviction probabilities.

We proceed by assuming $f_{\text{Blacks}} = Beta(2, 4)$, and compute the parameters of f_{Whites} to match the fraction of white jurors struck by the defendant and the plaintiff (the last two

Figure 9: Counterfactual analysis: Number of challenges



(a) Fraction of extreme jurors **(b) Fraction of minority jurors**

Note: Fraction of juries with at least one juror below the 10th percentile (left panel) and fraction of minority jurors (right panel). For each set of parameters, results on the vertical axes are averages across 50,000 simulated jury selections, fixing $j = 12$, $d = p = 6$, and $C \sim 0.341 \cdot C_a + 0.659 \cdot C_b$ throughout (with the distributions for $C_a \sim Beta(2, 4)$ and $C_b \sim Beta(5.80, 3.95)$). Values of $d = p$ are on the horizontal axes.

moments of Table 2) using sample (5). The estimated parameters of f_{Whites} are (alpha = 5.80, beta = 3.95), with standard errors (1.00, 0.85).[34]

Figure 8 reports the results of simulations computed with the estimated parameters. The figure reveals that the procedure adopted by this jurisdiction — a version of *STR* where each party is allowed 6 challenges — is much more effective at excluding extreme jurors than a counterfactual *S&R*. The adopted procedure excludes nearly every juror below the 10-th percentile, $c = 0.21$, whereas *S&R* with the same number of challenges would produce about 27% juries with at least one juror more extreme than 0.21.

Figure 9 however suggests that a change to *S&R* could improve the representation of

[34] Standard errors computed by bootstrapping 200 replications of the data set. We also tried assuming a left-skewed $f_{\text{Blacks}} = Beta(10, 6)$, which results in $f_{\text{Whites}} = Beta(23.37, 6.71)$. For *STR*, the results obtained with these alternative distributions are almost identical to the ones reported in Figure 9. *S&R* selects about the same number of minorities across all number of challenges, but is capable of excluding fewer jurors below the 10th percentile by about 7 percentage points.

minorities. Keeping the number of challenges at 6, *S&R* would include 6% more minorities than *STR* (about 27% vs 25%) and would produce a jury with 4 black jurors (about the same as the black representation in the jury pool) 12% more often (about 41% vs 37%). To reach a similar representation, the number of challenges in *STR* would have to be reduced to 4, though this would increase the fraction of juries with jurors below the 10th percentile from almost zero to 4.4%.

This analysis suggests that the data is consistent with the parties believing in a distribution that makes the two procedures significantly different in their ability to exclude jurors. The data is also consistent with beliefs in sizeable heterogeneity between juror-groups which, in turn, implies that the procedures also differ in their ability to select of minorities as well.

8 Extensions

8.1 Excluding unbalanced juries

The primary purpose of jury selection is to prevent extreme potential jurors from serving on the effective jury (see Footnote 1 and its associated quote). In our model, it seems natural to interpret this goal as that of limiting the selection of jurors coming from the tail of the distribution. This is the interpretation of extreme that we have studied thus far.

Although it is perhaps less clear that it aligns with the goals of practitioners, another approach could be to consider the extremism of juries *as a whole*. For example, extreme juries could be viewed as juries in which the juror with the highest or lowest conviction probability is extreme. Through variants of the arguments in the proofs of Propositions 1 and 2, one can show that, in that sense too, *STR* is more effective than both *S&R* and *RAN* at excluding extreme juries.[35]

Another measure of juries' extremism, proposed by Flanagan (2015), is whether a jury is excessively "unbalanced" in the sense of featuring a disproportionate proportion of jurors coming from one side of the median of C. Interestingly, Flanagan shows that *STR* introduces correlation between the selected jurors, which leads the procedure to select more unbalanced juries than *RAN*. Even though panels are the result of independent draws from

[35] Specifically, for any $x \in \{0, \dots, j-1\}$, there exists $\underline{c} > 0$ and $\bar{c} < 1$, such that (a) for every $c \in (0, \underline{c})$, the probability that the lowest conviction-probability in the jury is smaller than c is larger under *S&R* and *RAN* than under *STR*, and (b) for every $c \in (\bar{c}, 1)$, the probability that the highest conviction-probability in the jury is larger than c is larger under *S&R* and *RAN* than under *STR*.

the population, jurors selected under STR have conviction probabilities between that of the lowest and highest challenged juror. For example, the selection of two jurors with conviction probabilities 0.25 and 0.75 indicates that challenges were used on jurors with conviction probabilities outside the $[0.25, 0.75]$ range. The latter makes it more likely that STR selected additional jurors in the $[0.25, 0.75]$, thereby introducing a correlation between the selected jurors.

This intuition is formalized in Corollary 2 of Flanagan (2015) which shows that, even when the parties have the same number of challenges ($d = p$), the probability that *all* selected jurors come from one side of the median is *larger* under STR than under RAN. Our next proposition generalizes this result. Using a new proof technique, we show that for *any* x larger than half the jury-size, the probability of selecting at least x jurors from one side of the median is larger under STR than under RAN. Similar to Section 4, we focus for brevity on the probability that the selected jurors are *below* the median. All our results are however symmetrical and apply identically to the probability of selecting jurors above the median. Let $med[C]$ denote the median of C.

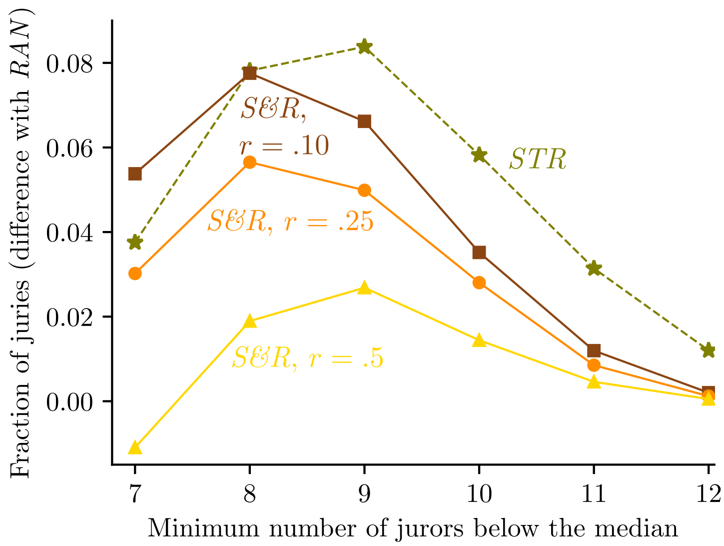
Proposition 5. *If $d = p$, then for any $x \in \{n/2 + 1, \ldots, n\}$ if n is even, and any $x \in \{n/2 + 1.5, \ldots, n\}$ if n is odd, we have $\mathbb{T}_{STR}(x; med[C]) > \mathbb{T}_{RAN}(x; med[C])$.*

Figure 10 illustrates Proposition 5 and that a similar statement does not hold for *S&R*. For $M \in \{STR, RAN\}$, the value of $\mathbb{T}_M\,(x; med[C])$ can be computed analytically and does not depend on the distribution of C.[36] For $M =$ *S&R*, the value of $\mathbb{T}_M\,(x; med[C])$ depends on the distribution in a complex fashion and it is not possible to generally compare *S&R* with the two other procedures in terms of $\mathbb{T}_M\,(x; med[C])$. As the figure illustrates, the probability to select at least x jurors below $med[C]$ can, in some cases (in the figure, $x = 7$ and, barely, $x = 8$ jurors), be larger under *S&R* than under both RAN and STR. In other cases, however, the same probability is lower under *S&R* than under both RAN and STR.

Figure 10 displays the result of simulations when the distribution of C is highly polarized (a mixture of $Beta(1, 5)$ and $Beta(5, 1)$) In Appendix B we present additional simulations for less polarized distributions. These additional simulations suggest that high levels of polarization are required for *S&R* to more often select a majority of jurors below the median

[36] Specifically, $\mathbb{T}_{RAN}(x; med[C]) = \mathbb{P}(Bi[j, 0.5] \geq x)$ whereas $\mathbb{T}_{STR}(x; med[C]) = \mathbb{P}(Bi[j + d + p, 0.5] \geq x + p)$.

Figure 10: Selection of jurors below the median



Note: Fraction of juries with a at least given number of jurors below the median of C under STR (green dashed line) and $S\&R$ (continuous lines) relative to the same fraction under RAN (i.e. $\mathbb{T}_M(x; med[C]) - \mathbb{T}_{RAN}(x; med[C])$). Throughout, we fix $j = 12$, $d = p = 6$ and $C \sim r \cdot Beta(1, 5) + (1 - r) \cdot Beta(5, 1)$ (for $r \in \{0.1, 0.25, 0.5\}$) whereas the number of jurors below the median is on the horizontal axis. For each set of parameters, results for $S\&R$ are averages across 50,000 simulated jury selections, whereas values for RAN and STR are computed analytically and are independent of r (see Footnote 36).

than STR. Also, for lower levels of polarization, $S\&R$ more often selects fewer juries made of a majority of jurors below the median than RAN.[37]

8.2 Representation of balanced groups

Concerns about the effect of jury selection on group-representation often focus on the representation of racial minorities. Thought the U.S. Supreme Court initially banned challenges based on race in *Batson v. Kentucky* (1986), it later also banned challenges based on *gender* in *J.E.B. v. Alabama* (1994). In this context, it is natural to ask whether the advantages of $S\&R$ in terms of minority representation comes at the cost of a worse representation of gender groups.

[37] Because the parties' actions under $S\&R$ are influenced by the mean of the distribution but not in any clear way by the median (and because of the complexity of the game tree), we were unable to formalize the effect of polarization on these comparisons in terms of the model parameters.

Unlike minorities which correspond to groups of unequal sizes represented by small values of r, gender-groups can be thought of as even-sized groups and are better modeled using $r \approx 0.5$. With groups of similar sizes, both procedures almost always select at least a few members from either group. It is therefore more interesting to compare procedures in terms of the proportion of group-a jurors they select (than in terms of the probability of selecting *at least* x members from group-a, as we did before).

In this last section, we let $r = 0.5$ and study the expected proportion of group-a jurors selected under STR and $S\&R$. We denote these proportions r_{STR} and $r_{S\&R}$ and focus on how close r_{STR} and $r_{S\&R}$ are from the 50% of group-a jurors that prevail in the population.[38]

As in the last two sections, it is not possible to generally compare STR and $S\&R$ in terms of the procedures' ability to select an even proportion of group-a and group-b jurors. In some cases, r_{STR} can be further away from 50% than $r_{S\&R}$, and the converse may be true in other cases. For example, with $d = p = 6$ and $j = 12$, if $C_a \sim U[0, 1]$ and $C_b \sim Beta(1, 5)$, simulations reveal that $r_{STR} = 43.7\%$ whereas $r_{S\&R} = 45.8\%$. In contrast, when $C_a \sim Beta[4, 2]$ and $C_b \sim Beta(1, 5)$, $r_{STR} = 50.3\%$ whereas $r_{S\&R} = 52.2\%$.

These two examples however suggest that, as the group distributions become more symmetrical, r_{STR} get closer to 50% . Our next proposition confirms this pattern. If the group-distributions are symmetric or if they do not overlap, and if $d = p$, then $r_{STR} = 50\%$ whereas $S\&R$ does not necessarily select an even proportion of jurors from each group. The latter follows from the fact that, even when $r = 50\%$ and distributions are symmetrical, the multiplicative utility function that the parties use to assess the value of a jury (which is itself a consequence of the fact that juries must reach unanimous decisions) creates asymmetries in the use of challenges under $S\&R$.[39]

We say that random variables C_a and C_b are **symmetric** if $f_a(c) = f_b(1 - c)$ for all $c \in [0, 1]$.

[38] Previous results are stronger in the sense that they establish a first-order stochastic dominance between the number of jurors with certain characteristics (extremism or group-membership) selected under STR and $S\&R$. As we explain after Proposition 1, showing, for example, that $\mathbb{T}_{STR}(x; c) < \mathbb{T}_{S\&R}(x; c)$ for all $x \in \{1, \ldots, j\}$ directly implies that the expected proportion of selected jurors with conviction probability $c_i < c$ is lower under STR than under $S\&R$ (whereas the converse is not true).

[39] Flanagan (2015) shows that, in this symmetrical case, the asymmetry of the payoffs still forces the defendant to be more conservative than the plaintiff when using its challenges, hence leading to an uneven selection of jurors from the two groups.

Proposition 6. *Suppose that $r = 0.5$ and $d = p$. If (a) the two group distribution do not overlap,*[40] *or (b) C_a and C_b are symmetric, then $r_{STR} = r_{RAN}$.*

Table 3(a) illustrates Proposition 6 and the fact that a similar statement does not hold for *S&R*. Unlike *STR*, *S&R* can select unequal numbers of group-a and group-b jurors even when distributions are symmetrical across groups. Therefore, as a consequence of Proposition 6, $r_{S\&R}$ can in these cases be further away than r_{STR} from the 50% of group-a jurors that prevail in the population.

Table 3(a) however suggests that these differences may be quantitatively small, and that sizable differences may require high levels of polarization between groups. Table 3(b) and 3(c) also report the results of simulations in which the symmetries required for Proposition 6 to hold are slightly relaxed. These indicate that the advantage of *STR* in the representation of balanced groups established in Proposition 6 (i.e., the fact that r_{STR} is closer to 50% than $r_{S\&R}$) may not be robust to even mild relaxations of these symmetries. In particular, when $r = 0.45$, r_{STR} is consistently closer than $r_{S\&R}$ to the 55% of group$-a$ that prevail in the population (see Table 1). Also, when $r = 0.5$ but the group-distributions are slightly asymmetric, $r_{S\&R}$ are identical except in the most polarized case.

9 Conclusion

In this paper, we study the relative performance of two stylized jury-selection procedures. Strike and Replace presents potential jurors one-by-one to the parties, whereas the Struck procedure presents all potential jurors before they exercise vetoes. When jurors differ in their probability of voting for the defendant's conviction, and on group membership, we show that when groups have polarized views Strike is more effective at excluding jurors with extreme views, but generally selects fewer members of a minority group than Strike and Replace, leading to a conflict between these two goals.

Sociologists Small and Pager (2020) argue that systemic factors may lead to disparate outcomes even in the absence of taste-based or statistical discrimination, the traditional explanations provided in Economic Theory. This paper formalizes an example in which the pursuit of one legitimate objective — preventing extreme jurors to serve on juries — may

[40] That is either $\mathbb{P}(C_a > C_b) = 0$ or $\mathbb{P}(C_b > C_a) = 0$. The same result would apply if the two distributions did not overlap *in the limit as in Proposition 3.*

Table 3: Representation of Group-a jurors with balanced group sizes

Polarization	Extreme		Moderate		Mild		(All)
Procedure	*S&R*	*STR*	*S&R*	*STR*	*S&R*	*STR*	*RAN*
Average fraction of minorities	0.48	0.50	0.49	0.50	0.50	0.50	0.50
Standard deviation	0.18	0.20	0.16	0.17	0.15	0.15	0.14

(a) Group-a proportion $r = 0.5$, group distributions as in Figure 3.

Polarization	Extreme		Moderate		Mild		(All)
Procedure	*S&R*	*STR*	*S&R*	*STR*	*S&R*	*STR*	*RAN*
Average fraction of minorities	0.39	0.40	0.42	0.42	0.45	0.44	0.45
Standard deviation	0.18	0.20	0.16	0.17	0.15	0.15	0.14

(b) Group-a proportion $r = 0.45$, group distributions as in Figure 3.

Polarization	Extreme*		Moderate*		Mild*		(All)
Procedure	*S&R*	*STR*	*S&R*	*STR*	*S&R*	*STR*	*RAN*
Average fraction of minorities	0.47	0.50	0.49	0.48	0.49	0.48	0.50
Standard deviation	0.18	0.20	0.15	0.16	0.15	0.16	0.14

(c) Group-a proportion $r = 0.5$, group distributions slightly asymmetric*

In panel (c) Extreme corresponds to $C_a \sim Beta(1,5)$ and $C_b \sim Beta(5,2)$, Moderate* to $C_a \sim Beta(2,4)$ and $C_b \sim Beta(4,3)$, and Mild* to $C_a \sim Beta(3,4)$ and $C_b \sim Beta(4,4)$.

Note: The rows report the average number and standard deviation of group-a jury members out of 50,000 simulations of jury selection with parameters $j = 12$ and $d = p = 6$.

lead to group disparities.

A Appendix: Proofs

A.1 Preliminary technical results

A.1.1 Limit of a ratio of binomial probabilities

Lemma 1. *For all $\eta \in \mathbb{N}$ and any $k \in \{1, \ldots, \eta - 1\}$,*

$$\lim_{\pi \to 0} \frac{\mathbb{P}[Bi(\eta, \pi) = k]}{\mathbb{P}[Bi(\eta, \pi) > k]} = \infty.$$

Proof. Using the standard formula for the p.d.f. of a binomial and the representation of the c.d.f. of the binomial with *regularized incomplete beta function*, we can re-write the ratio as

$$\frac{\mathbb{P}[Bi(\eta, \pi) = k]}{1 - \mathbb{P}[Bi(\eta, \pi) \leq k]} = \frac{\binom{\eta}{k} \pi^k (1 - \pi)^{\eta - k}}{1 - (\eta - k) \binom{\eta}{k} \int_0^{1-\pi} x^{\eta - k - 1} (1 - x)^k dx} \tag{1}$$

As $\pi \to 0$, both the numerator and the denominator tend to 0. We use L'Hopital's rule to complete the proof:

$$\begin{aligned}
&\frac{(\partial/\partial\pi) \ \binom{\eta}{k} \pi^k (1-\pi)^{\eta-k}}{(\partial/\partial\pi) \ \left(1 - \left[(\eta - k)\binom{\eta}{k} \int_0^{1-\pi} x^{\eta-k-1}(1-x)^k dx\right]\right)} \\
&= \frac{\binom{\eta}{k} \cdot \left[k\pi^{k-1}(1-\pi)^{\eta-k} + \pi^k(\eta-k)(1-\pi)^{\eta-k-1}\right]}{-(\eta-k)\binom{\eta}{k}\left[(-1)\cdot(1-\pi)^{\eta-k-1}\pi^k\right]} \\
&= \frac{k\pi^{k-1}(1-\pi)^{\eta-k}}{(\eta-k)(1-\pi)^{\eta-k-1}\pi^k} + \frac{\pi^k(\eta-k)(1-\pi)^{\eta-k-1}}{(\eta-k)(1-\pi)^{\eta-k-1}\pi^k} \\
&= \frac{k(1-\pi)}{(\eta-k)\pi} + 1 \quad \xrightarrow[\pi \to 0]{} \quad \infty
\end{aligned}$$

□

A.1.2 Continuity of challenge thresholds in *S&R* as C^i converges in distribution

Lemma 2. *Consider a sequence of random variables $\{C^i\}_{i=1}^{\infty}$ that converges in distribution to some random variable C^*. Let $t_I(\gamma, C^i)$ denote the challenge threshold used by party $I \in \{D, P\}$ in an arbitrary subgame γ of S&R when the distribution of conviction probabilities is C^i. For any such subgame γ, we have $\lim_{i\to\infty} t_I(\gamma, C^i) = t_I(\gamma, C^*)$.*

Proof. In any subgame $\tilde{\gamma}$, $t_I(\tilde{\gamma}, C^i)$ is the ratio of the value of continuation subgames if I challenges the presented juror, or if both parties abstain from challenging (Brams and Davis, 1978). Therefore, $\lim_{i\to\infty} t_I(\gamma, C^i) = t_I(\gamma, C^*)$ follows directly if we show that the value of any subgame, which we denote $V(\gamma, C^i)$, converges to $V(\gamma, C^*)$ as i tends to infinity.[41]

The latter follows directly from the recursive characterization of $V(\gamma, C^i)$ in Brams and Davis (1978). Recall that each subgame γ can be characterized by the number of jurors κ that remain to be selected, the number of challenges left to the defendant δ, and the number of challenges left to the plaintiff π. With this notation, the recursive proof that for all $\kappa, \delta, \pi \geq 0$, $V\big([\kappa, \delta, \pi], C^i\big)$ converges to $V\big([\kappa, \delta, \pi], C^i\big)$ as i tends to infinity can be decomposed in a number of cases. Let $F^i(c)$ denote the the c.d.f. of C^i, $F^*(c)$ the c.d.f. of C^*, and $F(c)$ the c.d.f. of an arbitrary distribution C, with μ^i, μ^*, and μ^j being the corresponding expected values. In each step, the initial formula for $V\big([\kappa, \delta, \pi], C^i\big)$ is taken from Brams and Davis (1978).

Case 1: $\kappa = 0, \delta \geq 0, \pi \geq 0$**.** In this case, $V\big([0, \delta, \pi], C\big) = 1$ for all C and the convergence of $V\big([0, \delta, \pi], C^i\big)$ to $V\big([0, \delta, \pi], C^*\big)$ follows trivially.

Case 2: $\kappa > 0, \delta = 0, \pi = 0$**.** In this case, $V\big([\kappa, 0, 0], C\big) = \mu^\kappa$ for all C and the convergence of $V\big([0, \delta, \pi], C^i\big)$ to $V\big([0, \delta, \pi], C^*\big)$ follows from the fact that C^i converges in distribution to C^*.

Case 3: $\kappa > 0, \delta = 0, \pi > 0$**.** In this case, for all C,

$$V\big([\kappa, 0, \pi], C\big) = V(\kappa - 1, 0, \pi) * \left[1 - \int_{t_I([\kappa,0,\pi],C)}^{1} F(c)\ dc\right],$$

and $t_I([\kappa, 0, \pi], C) = V\big([\kappa, 0, \pi-1], C\big)/V\big([\kappa-1, 0, \pi], C\big)$. The convergence of $V\big([\kappa, 0, \pi], C^i\big)$ to $V\big([\kappa, 0, \pi], C^*\big)$ then follows recursively from the previous cases and from C^i converging in distribution to C^*.

Case 4: $\kappa > 0, \delta > 0, \pi = 0$**.** In this case, for all C,

$$V\big([\kappa, \delta, 0], C\big) = V\big([\kappa, \delta - 1, 0], C\big) - V\big([\kappa - 1, \delta, 0], C\big) * \int_{0}^{t_D([\kappa,\delta,0],C)} F(c)\ dc,$$

where $t_D([\kappa, \delta, 0], C) = V\big([\kappa, \delta-1, 0], C\big)/V\big([\kappa-1, \delta, 0], C\big)$. The convergence of $V\big([\kappa, \delta, \pi], C^i\big)$ to $V\big([\kappa, \delta, \pi], C^*\big)$ then follows recursively from the previous cases and from C^i converging in distribution to C^*.

[41]Because we assume that all distributions of conviction probabilities are continuous, there are no issues related to the possibility for the bottom of one of these ratios to converge to zero.

Case 5: $\kappa > 0, \delta > 0, \pi > 0$**.** In this case, for all C,

$$V\big([\kappa,\delta,\pi],C\big) = V\big([\kappa,\delta-1,\pi],C\big) - V\big([\kappa-1,\delta,\pi],C\big) * \int_{t_I([\kappa,\delta,\pi],C)}^{t_D([\kappa,\delta,\pi],C)} F(c)\ dc,$$

where $t_D([\kappa,\delta,\pi],C) = V\big([\kappa,\delta-1,\pi],C\big)/V\big([\kappa-1,\delta,\pi],C\big)$ and and $t_I([\kappa,\delta,\pi],C) = V\big([\kappa,\delta,\pi-1],C\big)/V\big([\kappa-1,\delta,\pi],C\big)$. The convergence of $V\big([\kappa,\delta,0],C^i\big)$ to $V\big([\kappa,\delta,0],C^*\big)$ then follows recursively from the previous cases and from C^i converging in distribution to C^*.

■

A.1.3 Comparative statics of probabilities from a symmetric binomial

Lemma 3. $\mathbb{P}[Bi(\eta+2,0.5) \geq k+1] > \mathbb{P}[Bi(\eta,0.5) \geq k]$ if and only if $k > \frac{\eta}{2} + \frac{1}{2}$.

Proof. We can decompose $\mathbb{P}[Bi(\eta+2,0.5) \geq k+1]$ in terms of $Bi(\eta,0.5)$ and $Bi(2,0.5)$:

$$\begin{aligned}
& \mathbb{P}[Bi(\eta+2,0.5) \geq k+1] \\
= \; & \mathbb{P}[Bi(\eta,0.5) \geq k+1] \; + \; \mathbb{P}[Bi(\eta,0.5) = k] \; * \; \mathbb{P}[Bi(2,0.5) \geq 1] \; + \\
& \mathbb{P}[Bi(\eta,0.5) = k-1] \; * \; \mathbb{P}[Bi(2,0.5) = 2] \\
= \; & \mathbb{P}[Bi(\eta,0.5) \geq k+1] \; + \; \mathbb{P}[Bi(\eta,0.5) = k] \; * \; 0.75 \; + \; \mathbb{P}[Bi(\eta,0.5) = k-1] \; * \; 0.25
\end{aligned}$$

Also,

$$\mathbb{P}[Bi(\eta,0.5) \geq k] = \mathbb{P}[Bi(\eta,0.5) \geq k+1] + \mathbb{P}[Bi(\eta,0.5) = k].$$

Together, the last two equalities imply that $\mathbb{P}[Bi(\eta+2,0.5) \geq k+1] \; > \; \mathbb{P}[Bi(\eta,0.5) \geq k]$ if and only if

$$\begin{aligned}
\mathbb{P}[Bi(\eta,0.5) = k] * 0.75 + \mathbb{P}[Bi(\eta,0.5) = k-1] * 0.25 &> \mathbb{P}[Bi(\eta,0.5) = k] \\
\mathbb{P}[Bi(\eta,0.5) = k-1] * 0.25 &> \mathbb{P}[Bi(\eta,0.5) = k] * 0.25 \\
\mathbb{P}[Bi(\eta,0.5) = k-1] &> \mathbb{P}[Bi(\eta,0.5) = k] \\
\binom{\eta}{k-1} 0.5^{k-1} 0.5^{\eta-(k-1)} &> \binom{\eta}{k} 0.5^{k} 0.5^{\eta-k} \\
\frac{\eta!}{(\eta-[k-1])!(k-1)!} &> \frac{\eta!}{(\eta-k)!k!} \\
\frac{(\eta-k)!}{(\eta-[k-1])!} &> \frac{(k-1)!}{k!} \\
\frac{1}{\eta-k+1} &> \frac{1}{k} \\
k &> \frac{\eta}{2} + \frac{1}{2}
\end{aligned}$$

$\blacksquare$

A.1.4 Relationship between order statistics of symmetric distributions

For any number of draws w and any $k \leq w$, let $C_g^{k,w}$ denote the k-th order statistic out of w draws from distribution C_g, and $f_g^{k,w}(x)$ the corresponding probability density function.

Lemma 4. *Suppose that C_a and C_b are symmetric. Then, for any $w \in \mathbb{N}$ and any $k \in \{1, \ldots, w\}$, we have $f_a^{k,w}(c) = f_b^{w-k+1,w}(1-c)$ for all $c \in [0,1]$.*

Proof. Recall that, by definition, C_a and C_b being symmetric implies $f_a(c) = f_b(1-c)$ for all $c \in [0,1]$, which, in turn, implies $F_a(c) = F_b(1-c)$ for all $c \in [0,1]$. We therefore have,

$$\begin{aligned}
f_a^k(c) &= k\binom{w}{k} f_a(c)[F_a(c)]^{k-1}[1-F_a(c)]^{w-k} \\
&= k\binom{w}{k} f_b(1-c)[1-F_b(1-c)]^{k-1}[1-(1-F_b(1-c))]^{w-k} \\
&= k\frac{w!}{(w-k)!k!} f_b(1-c)[1-F_b(1-c)]^{k-1}[f_b(1-c)]^{w-k} \\
&= (w-k+1)\frac{w!}{(w-k+1)!(k-1)!} f_b(1-c)[(1-F_b(1-c)]^{k-1}[F_b(1-c)]^{w-k} \\
&= (w-k+1)\frac{w!}{(w-k+1)!(w-(w-k+1)!} f_b(1-c)[1-F_b(1-c)]^{k-1}[F_b(1-c)]^{w-k} \\
&= (w-k+1)\binom{w}{w-k+1} f_b(1-c)[1-F_b(1-c)]^{k-1}[F_b(1-c)]^{w-k} \\
&= f_b^{w-k+1}(1-c)
\end{aligned}$$

$\blacksquare$

A.2 Section 4: Effectiveness at excluding extremes

A.2.1 Proof of Proposition 1

Consider an arbitrary $c \in (0,1)$ and let us refer to jurors with conviction probability no larger than c as *extreme jurors*. Let $\mathbb{T}_M(x; c|k)$ denote the probability that at least x extreme jurors are selected by procedure M *conditional* on there being exactly k of extreme jurors in the panel of n. By the Law of Total Probability,

$$\mathbb{T}_M(x;c) = \sum_{k=x}^{n} \mathbb{P}\Big[Bi\big(n, F(c)\big) = k\Big]\ \mathbb{T}_M(x;c|k). \tag{2}$$

Consider first the STR procedure. Note that for all c, we have $\mathbb{T}_{STR}(x;c|x)=0$ because if there are exactly x extreme jurors in the panel, one of them is necessarily challenged by the plaintiff under STR (recall that $p\geq 1$). Therefore, by (2), we have

$$\mathbb{T}_{STR}(x;c)=\sum_{k=x+1}^{n}\mathbb{P}\Big[Bi\big(n,F(c)\big)=k\Big]\;\mathbb{T}_{STR}(x;c|k)\leq \mathbb{P}\Big[Bi\big(n,F(c)\big)>x\Big], \tag{3}$$

where the last inequality follows from the fact that $\mathbb{T}_{STR}(x;c|k)\in[0,1]$ for all k (as $\mathbb{T}_{STR}(x;c|k)$ is a probability).

Next, consider procedure *S&R*. Our goal is to construct a lower bound for the probability of selecting an extreme juror and show that, as $c\to 0$, this lower bound does not converge to 0 as fast as (3). To do so, we introduce an decreasing function $\sigma(c)>0$ such that, when c is sufficiently small, $\mathbb{T}_{S\&R}(x;c|k)\geq\sigma(c)$ for any $k\geq x$. To construct σ, consider the restricted sample space in which there are k extreme jurors in the panel.

Let $\underline{t}_P$ be the lowest challenge threshold used by the plaintiff in any subgame of the *S&R* procedure. Clearly, $\underline{t}_P>0$.[42] Henceforth, we focus on $c\in(0,\underline{t}_P)$.

We first consider the function $\alpha(c)$ defined as the probability that $c_j\in(c,\underline{t}_P)$ for *all* the $(n-k)$ non-extreme jurors in the panel. Because C is continuous and 0 is the lower-bound of its support, there exists $y>0$ sufficiently small such that $\alpha(c)>0$ for all $c\in[0,y]$.[43] Also, $\alpha(c)$ is weakly decreasing in c.

By construction of $\underline{t}_P$, for such panels (with k extreme jurors and $c_j\in(c,\underline{t}_P)$ for all the $(n-k)$ non-extreme jurors), the plaintiff uses all its challenges on the p first jurors it is presented with, and the defendant never uses any challenges.[44] Therefore, for these panels, the probability that all k extreme jurors are selected is the probability that none of these jurors are among the p first jurors presented to the parties, i.e., $\binom{n-p}{k}/\binom{n}{k}$. Overall, for

[42]Formally, if Γ denotes the set of subgames of *S&R* and $t_P(\gamma)$ the plaintiff's challenge threshold in any subgame $\gamma\in\Gamma$, then $\underline{t}_P=\min_{\gamma\in\Gamma}t_p(\gamma)$ (the minimum is well-defined since Γ is of finite size). In any subgame γ of *S&R*, there is always a conviction probability $c>0$ low enough such that if the juror who is presented to the parties in the first round of γ is of type c, the plaintiff will challenge that juror. Therefore, $\underline{t}_P>0$.

[43]By definition of the support, because 0 is the lower-bound of the support, $\mathbb{P}(C\in[0,\epsilon])>0$ for all $\epsilon>0$. Because C is continuous, there must therefore exists some $\delta>0$ such that $\mathbb{P}(C\in[\delta/2,\delta])>0$. We then have $\alpha(c)>0$ for all $c<\delta$.

[44]The latter follows from the fact that, in any subgame, the threshold used by the defendant is always higher than the threshold used by the plaintiff (in equilibrium, the defendant and the plaintiff never both want to challenge the presented juror).

$c \in (0, \underline{t}_P)$, we have $\mathbb{T}_{S\&R}(x;c|k) \geq \alpha(c) \cdot \binom{n-p}{k}/\binom{n}{k}$, and $\sigma(c) \coloneqq \alpha(c) \cdot \binom{n-p}{k}/\binom{n}{k}$ has the desired property.

Applying $\mathbb{T}_{S\&R}(x;c|k) \geq \sigma(c)$ to (2) with $M = S\&R$, we obtain for all c sufficiently small (specifically $c \in (0, \underline{t}_P)$)

$$\mathbb{T}_{S\&R}(x;c) \geq \sum_{k=x}^{n} \mathbb{P}\Big[Bi\big(n, F(c)\big) = k\Big] \ * \ \sigma(c) \geq \mathbb{P}\Big[Bi\big(n, F(c)\big) = x\Big] \ * \ \sigma(c). \tag{4}$$

Overall, combining (3) and (4) yields

$$\lim_{c\to 0} \frac{\mathbb{T}_{S\&R}(x;c)}{\mathbb{T}_{STR}(x;c)} \geq \lim_{c\to 0} \frac{\mathbb{P}\Big[Bi\big(n, F(c)\big) = x\Big] \ * \ \sigma(c)}{\mathbb{P}\Big[Bi\big(n, F(c)\big) > x\Big]} = \infty, \tag{5}$$

where the last equality follows from Lemma 1 and the fact that $\sigma(c) > 0$ is decreasing in c.[45] In turn, $\lim_{c\to 0} \mathbb{T}_{S\&R}(x;c)/\mathbb{T}_{STR}(x;c) = \infty$ and the fact that $\lim_{c\to 0} \mathbb{T}_{S\&R}(x;c) = \lim_{c\to 0} \mathbb{T}_{STR}(x;c) = 0$ together imply implies that there exists some $\underline{c} > 0$ small enough such that $\mathbb{T}_{STR}(x;c) < \mathbb{T}_{S\&R}(x;c)$ for all $c \in (0, \underline{c})$.

A.2.2 Proof of Proposition 2

Using the same notation as in the proof of Proposition 1, we have

$$\mathbb{T}_{RAN}(x;c) \quad \geq \quad \mathbb{P}\Big[Bi\big(n, F(c)\big) = x\Big] \ * \ \mathbb{T}_{RAN}(x;c|x). \tag{6}$$

Note that $\mathbb{T}_{RAN}(x;c|x)$ is the probability that an Hypergeometric random variable with x success, $n - x$ failures, and j draws, results in the draw of exactly x successes. Therefore, $\mathbb{T}_{RAN}(x;c|x) > 0$. Finally, combining (6) and (3) yields

$$\lim_{c\to 0} \frac{\mathbb{T}_{RAN}(x;c)}{\mathbb{T}_{STR}(x;c)} \geq \lim_{c\to 0} \frac{\mathbb{P}\Big[Bi\big(n, F(c)\big) = x\Big] \ * \ \mathbb{T}_{RAN}(x;c|x)}{\mathbb{P}\Big[Bi\big(n, F(c)\big) > x\Big]} = \infty,$$

where the last equality follows from Lemma 1 and the fact that $\mathbb{T}_{RAN}(x;c|x) > 0$. The result then follows as in the proof of Proposition 1.

[45] To apply Lemma 1, note that because C is continuous and the lower-bound of the support of C is 0, we have $F(c) > 0$ for all $c > 0$ and $\lim_{c\to 0} F(c) = 0$.

A.3 Section 5: Representation of minorities

A.3.1 Proof of Proposition 3

The structure of the proof is similar to that of the previous propositions. We focus on the case we analyzed in the main paper, where the minority uniformly favors the defendant, i.e., $\lim_{i\to\infty}\mathbb{P}(C_a^i > C_b^i) = 0$. The proof for the other case is symmetrical.

For now, consider arbitrary C_a^i, C_b^i, and r^i. Similar to the previous proofs, for any triple (C_a^i, C_b^i, r^i), we first decompose $\mathbb{A}^i_{STR}(x)$ and $\mathbb{A}^i_{S\&R}(x)$ by conditioning on the number of minority jurors in the panel.

First, consider STR and let us decompose $\mathbb{A}^i_{STR}(x)$ conditional, on the one hand, on the panel containing more than x minority jurors — which occurs with probability $\mathbb{P}\big[Bi(n,r^i) > x\big]$, and on the other, on the panel containing exactly x minority jurors — which occurs with probability $\mathbb{P}\big[Bi(n,r^i) = x\big]$. In the first case (i.e., more than x minority jurors in the panel), the probability that the panel contains at least x minority jurors is an upper bound on the probability that STR selects them. In the second case (i.e., exactly x minority jurors in the panel), STR selects at least x minority jurors provided that none of the minority jurors in the panel are challenged. This occurs with a probability no larger than the probability that the lowest conviction-probability among minorities is larger than the p-th conviction probability among majority jurors (since the latter is required for the plaintiff not to challenge any of the minority jurors in the panel). Recall that for any number of draws w and any $k \le w$, we let $C_g^{k,w}$ denote the k-th order statistic out of w draws from group $g \in \{a, b\}$. With this notation, we therefore have,

$$\mathbb{A}^i_{STR}(x) \le \mathbb{P}\big[Bi(n,r^i) > x\big] + \mathbb{P}\big[Bi(n,r^i) = x\big] * \mathbb{P}\big([C_a^i]^{1,x} > [C_b^i]^{p,n-x}\big). \tag{7}$$

Note that because $\lim_{i\to\infty}\mathbb{P}(C_a^i > C_b^i) = 0$, we have $\lim_{i\to\infty}\mathbb{P}\big([C_a^i]^{1,x} > [C_b^i]^{p,n-x}\big) = 0$.

Second, consider *S&R*. Clearly, $\mathbb{A}^i_{S\&R}(x)$ is no smaller than the probability for *S&R* to select at least x minority jurors when there are exactly x minority jurors in the panel. The latter is equal to $\mathbb{P}\big[Bi(n,r^i) = x\big] * \sigma(x; r^i, C_a^i, C_b^i)$, where $\sigma(x; r^i, C_a^i, C_b^i)$ denotes the probability that *S&R* selects x minority jurors *conditional* on having x minority jurors in the panel, as a function of r^i, C_a^i, and C_b^i. In summary, with this notation, we have,

$$\mathbb{A}^i_{S\&R}(x) \ge \mathbb{P}\big[Bi(n,r^i) = x\big] * \sigma(x; r^i, C_a^i, C_b^i). \tag{8}$$

We now show that $\lim_{i\to\infty} \sigma(x; r^i, C_a^i, C_b^i) > 0$. For all $i \in \mathbb{N}$, let $C^i = r^i C_a^i + (1 - r^i) C_b^i$. Observe that because $\lim_{i\to\infty} r_i = 0$ and because C_b^i converges in distribution to C_b^*, C^i converges in distribution to C_b^*. By Lemma 2, this implies that for any subgame γ of *S&R* and both $I \in \{D, P\}$, we have $\lim_{i\to\infty} t_I(\gamma, C^i) = t_I(\gamma, C_b^*)$. Note that $t_I(\gamma, C_b^*)$ lies in the interior of the support of C_b^* for both $I \in \{D, P\}$. Also recall that in the limit, the supports of C_a^i and C_b^i do not overlap as we have $\mathbb{P}(C_a^* > C_b^*) = 0$. Therefore, in the limit, the defendant never challenges a minority juror, which in turn implies that

(a) as i tends to infinity, the probability that the defendant challenges one of the x minority jurors in the panel tends to zero.

Because $t_I(\gamma, C_b^*)$ lies in the interior of the support of C_b^* for both $I \in \{D, P\}$, there is also a range of conviction probabilities $[\underline{c}, \overline{c}]$ low enough inside the support of C_b^* such that $P(C_b^* \in [\underline{c}, \overline{c}]) > 0$ and P challenged the juror presented in subgame γ if her conviction probability lies within $[\underline{c}, \overline{c}]$. Furthermore, the probability that a juror with conviction-probability in $[\underline{c}, \overline{c}]$ is a majority juror is strictly positive (and tends to one as i tends to infinity). Overall, in the limit,

(b) the probability that the plaintiff challenges a majority juror presented in subgame γ is strictly positive.

Combining (a) and (b), in the limit and given a panel containing x minority jurors, there is a positive probability that p majority jurors are presented first, are all challenged by P, and are followed by the x minority jurors which are left unchallenged by the parties (resulting in a jury composed of at least x minority jurors). That is, $\lim_{i\to\infty} \sigma(x; r^i, C_a^i, C_b^i) > 0$.

We are now equipped to complete the proof. Combining (7) and (8) yields

$$
\begin{aligned}
&\lim_{i\to\infty} \frac{\mathbb{A}^i_{STR}(x)}{\mathbb{A}^i_{S\&R}(x)} \\
&\leq \lim_{i\to\infty} \frac{\mathbb{P}\big[Bi(n,r^i) > x\big] + \mathbb{P}\big[Bi(n,r^i) = x\big] * \mathbb{P}\big([C^i_a]^{1,x} > [C^i_b]^{p,n-x}\big)}{\mathbb{P}\big[Bi(n,r^i) = x\big] * \sigma(r^i, C^i_a, C^i_b)} \\
&= \lim_{i\to\infty} \frac{\mathbb{P}\big[Bi(n,r^i) > x\big]}{\mathbb{P}\big[Bi(n,r^i) = x\big] * \sigma(r^i, C^i_a, C^i_b)} + \frac{\mathbb{P}\big[Bi(n,r^i) = x\big] * \mathbb{P}\big([C^i_a]^{1,x} > [C^i_b]^{p,n-x}\big)}{\mathbb{P}\big[Bi(n,r^i) = x\big] * \sigma(r^i, C^i_a, C^i_b)} \\
&= \lim_{i\to\infty} \frac{\mathbb{P}\big[Bi(n,r^i) > x\big]}{\mathbb{P}\big[Bi(n,r^i) = x\big]} * \frac{1}{\sigma(r^i, C^i_a, C^i_b)} + \frac{\mathbb{P}\big([C^i_a]^{1,x} > [C^i_b]^{p,n-x}\big)}{\sigma(r^i, C^i_a, C^i_b)} \\
&= \underbrace{\lim_{i\to\infty} \frac{\mathbb{P}\big[Bi(n,r^i) > x\big]}{\mathbb{P}\big[Bi(n,r^i) = x\big]}}_{=0,\text{ by Lemma 1}} * \underbrace{\lim_{i\to\infty} \frac{1}{\sigma(r^i, C^i_a, C^i_b)}}_{<\infty,\text{ by } \lim_{i\to\infty}\sigma(x;r^i,C^i_a,C^i_b)>0} + \underbrace{\lim_{i\to\infty} \frac{\mathbb{P}\big([C^i_a]^{1,x} > [C^i_b]^{p,n-x}\big)}{\sigma(r^i, C^i_a, C^i_b)}}_{\substack{=0,\\ \text{by } \lim_{i\to\infty}\mathbb{P}([C^i_a]^{1,x}>[C^i_b]^{p,n-x})=0,\\ \text{and } \lim_{i\to\infty}\sigma(x;r^i,C^i_a,C^i_b)>0}} \\
&= 0
\end{aligned}
$$

In turn, $\lim_{i\to\infty} \mathbb{A}^i_{STR}(x)/\mathbb{A}^i_{S\&R}(x) \leq 0$ and the fact that $\lim_{i\to\infty} \mathbb{A}^i_{STR}(x) = \lim_{i\to\infty} \mathbb{A}^i_{S\&R}(x) = 0$ together imply that there exists some j sufficiently large such that $\mathbb{A}^i_{S\&R}(x) > \mathbb{A}^i_{STR}(x)$ for all $i > j$.

A.4 Section 2: Changing the number of challenges

A.4.1 Proof of Proposition 4

The structure of the proof is similar to that of the previous propositions. Observe that (3) and (4) are true regardless of the number of challenges awarded to the parties in *STR* or *S&R*. That is, by the same arguments as in the proof of Proposition 1, the following two inequalities hold regardless of the values of w, y, $\mathbb{A}_{STR\text{-}w}(x)$, or $\mathbb{A}_{S\&R\text{-}y}(x)$,[46]

$$
\begin{aligned}
\mathbb{T}_{STR\text{-}w}(x;c) &= \sum_{k=x+1}^{n} \mathbb{P}\Big[Bi\big(n,F(c)\big) = k\Big]\, \mathbb{T}_{STR\text{-}w}(x;c|k) \leq \mathbb{P}\Big[Bi\big(n,F(c)\big) > x\Big], \\
\mathbb{T}_{S\&R\text{-}y}(x;c) &\geq \sum_{k=x}^{n} \mathbb{P}\Big[Bi\big(n,F(c)\big) = k\Big] \, * \, \sigma(c) \geq \mathbb{P}\Big[Bi\big(n,F(c)\big) = x\Big] \, * \, \sigma(c).
\end{aligned} \tag{9}
$$

The proposition then follows from the same argument as in the proof of Proposition 1 (in particular, see (5)).

[46] Recall that the proposition assumes $w, y \geq 1$.

A.5 Section 8: Extensions: Unbalanced juries and representation of balanced groups

A.5.1 Proof of Proposition 5

The probability that STR selects at least x jurors with conviction-probability above the median is the probability that at least $x+d$ of the jurors in the panel have conviction-probability above the median (since d of these jurors are challenged by the defendant). Because $d=p$, for any $x \in \{1,\dots,n\}$, we therefore have

$$\mathbb{T}_{STR}\big(x; med[C]\big) = P[Bi(j+d+p, 0.5) \geq x+d] = P[Bi(j+2d, 0.5) \geq x+d]$$

In contrast, we have

$$\mathbb{T}_{RAN}\big(x; med[C]\big) = P[Bi(j, 0.5) \geq x].$$

Therefore, by repeated application of Lemma 3, $x > (n/2)+(1/2)$ implies $\mathbb{T}_{STR}\big(x; med[C]\big) > \mathbb{T}_{RAN}\big(x; med[C]\big)$. Since n is integer-valued, the last inequality corresponds to $x \geq n/2+1$ if n is even and $x \geq n/2+1.5$ if n is odd.

A.5.2 Proof of Proposition 6

Part (a). Under STR, since the group-distributions do not overlap, each party first uses all of its challenges on one of the two groups before challenging the lowest conviction probability jurors from the other group. For concreteness and without loss of generality, suppose that group a favors the defendant (i.e., $\mathbb{P}(C_a > C_b) = 0$). Let m denote the number of jurors from group-a in the panel.

Note that because $r = 0.5$, the probability that $m = k$ is the same as the probability that $m = n-k$ for all $k \in \{1,\dots,\lfloor n/2 \rfloor\}$. Also, because $d = p$, the number of group-a jurors who are selected when $m = k$ is equal to the number of group-b jurors who are selected when $m = n-k$.[47] Therefore, the expected number of group-a jurors in the jury selected by STR is exactly $j/2$.

[47] First, suppose that $k \leq p$. Then, if $m = k$, no jurors from group-a (and j jurors from group-b) are selected, whereas if $m = n-k$, no jurors from group-b (and j jurors from group-a) are selected. Second, suppose that $k \in \{p+1,\dots,\lfloor n/2 \rfloor\}$. Then, if $m = k$, $k-p = k-d$ jurors from group-a (and $j-(k-p) = j-(k-d)$ jurors from group-a) are selected, whereas if $m = n-k$, $k-d = k-p$ jurors from group-b (and $j-(k-d) = j-(k-p)$ jurors from group-b) are selected.

Part (b). The proof is similar to the proof of Part (a). Consider the set of panel configurations $\{a, b\}^n$ where, for example, vector $(a, b, a, \ldots, b, b, b) \in \{a, b\}^n$ indicates that the juror with the lowest conviction probability in the panel is a group-a juror, the juror with second-lowest conviction probability is a group-b juror, the juror with the third-lowest conviction probability is a group-a juror, ..., and the jurors with the three highest conviction probabilities are all group-b jurors. To explain the structure of the proof, suppose that n is even (we explain below how the argument generalizes to any n). We first construct a partition of $\{a, b\}^n$ into two subsets S^a and S^b of equal size and construct a bijection q between S^a and S^b. We then show that for every panel configuration $l \in S^a$ which results in m^l group-a jurors being selected, (a) the panel configuration $q[l]$ result $j - m^l$ group-a jurors being selected, and (b) panel configurations l and $q[l]$ are equally likely. As in the proof of Part (b), the result then follows directly.

Similar to the proof of Part (b), the bijection $q[l]$ is obtained by (i) mirroring l around the $\lfloor n/2 \rfloor$ position, and (ii) inverting the group of each juror in the resulting panel configuration. For example, panel configuration $q[(a, a, b, a)]$ is obtained by mirroring (a, a, b, a) around position $\lfloor n/2 \rfloor$, which results in (a, b, a, a), and then inverting the group of each jurors in (a, b, a, a), which results in (b, a, b, b). Formally, if $inv[l]$ denotes the configuration that results from turning all the a's in l into b's and all the b's in l into a's, then $q[(l_1, l_2, \ldots, l_{n-1}, l_n)] = inv[(l_n, l_{n-1}, \ldots, l_2, l_1)]$.

Let S^a and S^b be two sets that together contain all l for which $l \neq q[l]$ and are such that $l \in S^i$ implies $q[l] \notin S^i$. Since $q\big[q[l]\big] = l$, the sets S^a and S^b have equal sizes. Also let S^* contain all l for which $l = q[l]$, if any ($S^* \neq \emptyset$ if and only if n is even). Note that $\{S^a, S^b, S^*\}$ forms of partition of $\{a, b\}^n$. Therefore, if we let $(\#m|l)$ denote the number of group-a juror that are selected conditional on configuration l and $\mathbb{P}(l)$ the probability of configuration l, we have

$$r_{STR} = \sum_{l \in S^a} \mathbb{P}(l) * (\#m|l) + \mathbb{P}(q[l]) * (\#m|q[l]) + \sum_{l \in S^*} \mathbb{P}(l) * (\#m|l).$$

Part (b) then follows from the fact that (A) $\mathbb{P}(l) = \mathbb{P}(q[l])$ for all $l \in S^a$, (B) $(\#m|l) = n - (\#m|q[l])$ for all $l \in S^a$, and (C) $(\#m|l) = j/2$ for all $l \in S^*$.

Properties (B) and (C) follow directly from the construction of q and the fact that $d = p$. Property (A), on the other hand, follows from Lemma 4 which establishes the symmetry of order statistics for symmetric distributions. A formal proof of (A) using Lemma 4 requires heavy and tedious notation. Instead, we show how (A) follows from Lemma 4 in a simple example that clarifies how the argument generalizes to other cases.

Consider the case of (a, a, b) for which $q[(a, a, b)] = (a, b, b)$. We can obtain the probability of any configuration by integrating the p.d.f. of the appropriate order statistics from the bottom to the top of $[0, 1]$. For example, using the notation for order statistics introduced before Lemma 4, we have

$$\begin{aligned}\mathbb{P}[(a, a, b)] &= \mathbb{P}[m = 2] * P[(a, a, b)|m = 2] \\ &= \mathbb{P}[Bi(3, 0.5) = 2] * \int_a^1 f_a^{1,2}(x) \left[\int_x^1 f_a^{2,2}(y) \left(\int_y^1 f_b^{1,1}(w)\ dw\right)\ dy\right]\ dx. \quad (10)\end{aligned}$$

We can also obtain the probability of any configuration by reverting the list of order statistics and integrating from the top to the bottom of $[0, 1]$. For example,

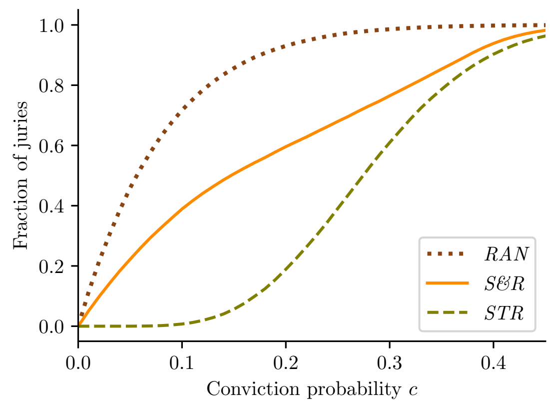
$$\begin{aligned}&\mathbb{P}[(a, b, b)] \\ &= \mathbb{P}[m = 1] * P[(a, b, b)|m = 1] \\ &= \mathbb{P}[Bi(3, a.5) = 1] * \int_a^1 f_b^{2,2}(1-x) \left[\int_x^1 f_b^{1,2}(1-y) \left(\int_y^1 f_a^{1,1}(1-w)\ dw\right)\ dy\right]\ dx. \quad (11)\end{aligned}$$

Finally, by Lemma 4, $f_a^{1,2}(x) = f_b^{2,2}(1-x)$, $f_a^{2,2}(y) = f_b^{1,2}(1-y)$, and $f_b^{1,1}(w) = f_a^{1,1}(1-w)$, which together with symmetry of the binomial with 0.5 probability of success implies that the expressions in (10) and (11) are equal.

B External Appendix: Additional simulations

B.1 Excluding extremes, uniform distribution of conviction probabilities

Figure B.1: Fraction of juries with at least one extreme juror



Note: Results from 50,000 simulations of jury selections with parameters $j = 12$, $d = p = 6$, and $C \sim U[0, 1]$

B.2 Minority representation when minorities favor conviction

Table B.1: Representation of Group-a jurors in the effective jury when Group-a is a minority of the jury pool

Polarization	Extreme		Moderate		Mild		(All)
Procedure	*S&R*	*STR*	*S&R*	*STR*	*S&R*	*STR*	*RAN*
Average fraction of minorities	0.12	0.08	0.18	0.16	0.23	0.23	0.25
Standard deviation	0.11	0.11	0.12	0.12	0.12	0.12	0.12
Fraction of juries with at least 1	0.76	0.45	0.89	0.85	0.96	0.95	0.97

(a) Group-a represents 25% of the jury pool

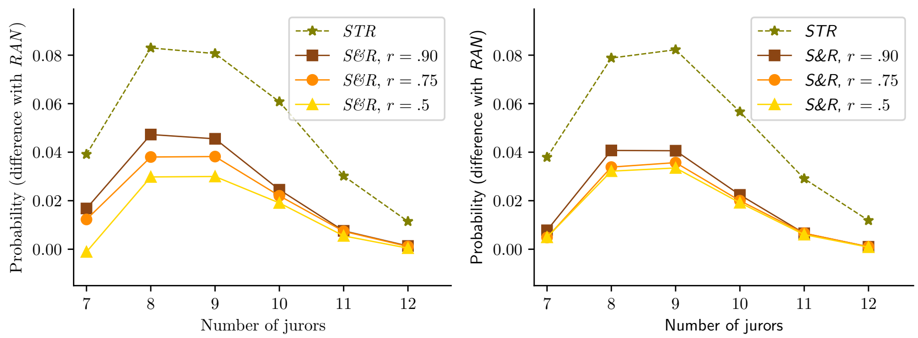
Polarization	Extreme		Moderate		Mild		(All)
Procedure	*S&R*	*STR*	*S&R*	*STR*	*S&R*	*STR*	*RAN*
Average fraction of minorities	0.01	0.00	0.05	0.04	0.09	0.08	0.10
Standard deviation	0.03	0.02	0.06	0.06	0.08	0.08	0.09
Fraction of juries with at least 1	0.09	0.02	0.44	0.38	0.66	0.64	0.72

(b) Group-a represents 10% of the jury pool

Note: The rows report the average number and standard deviation of group-a jury members, and the percent of juries with at least one group-a jurors, out of 50,000 simulations of jury selection with parameters $j = 12$ and $d = p = 6$. Conviction probabilities are drawn for from $Beta(1,5)$, $Beta(5,1)$, respectively for Group-a, Group-b jurors (Extreme), from $Beta(2,4)$, $Beta(4,2)$ (Moderate), and from $Beta(3,4)$, $Beta(4,3)$ (Mild); see Figure 3 for the shape of these distributions.

B.3 Excluding unbalanced juries, simulations from mild and moderate polarization

Figure B.2: Probability of selecting jurors below the median, difference with *RAN*



(a) Moderate polarization **(b) Mild polarization**

Note: The chart displays the probability of selecting a number of jurors with c_i below the median under STR (green dashed line) and $S\&R$ (orange lines) relative to the same probability under RAN, i.e. $\mathbb{T}_M(x; med[C]) - \mathbb{T}_{RAN}(x; med[C])$. The model parameters are $j = 12$, $d = p = 6$ and $C \sim r \cdot Beta(2,4) + (1-r) \cdot Beta(4,2)$ for Panel (a) and $C \sim r \cdot Beta(3,4) + (1-r) \cdot Beta(4,3)$, for $r = \{0.1, 0.25, 0.5\}$ Values for $S\&R$ are the results from 50,000 simulations of jury selection, whereas values for RAN and STR are computed analytically and are independent of r (see Footnote 36).

References

Anwar, S., P. Bayer, and R. Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." *The Quarterly Journal of Economics*, 127(2): 1017–1055.

Beck, Coburn R. 1998. "The Current State of the Peremptory Challenge." *William & Mary Law Review*, 39(3): , p. 42.

Bermant, Gordon, and John Shapard. 1981. "The Voir Dire Examination, Juror Challenges, and Adversary Advocacy." In *The Trial Process.* ed. by Sales, Bruce Dennis, Berlin: Springer, 69–114.

Biedenbender, Alice. 1991. "Holland v. Illinois: A Sixth Amendment Attack on the Use of Discriminatrory Peremptory Challenges." *Catholic University Law Review*, 40(3): , p. 31.

Bonebrake, James G. 1988. "Sixth and Fourteenth Amendments–The Lost Role of the Peremptory Challenge in Securing an Accused's Right to an Impartial Jury." *The Journal of Criminal Law and Criminology*, 79(3): , p. 23.

Brams, Steven J., and Morton D. Davis. 1978. "Optimal Jury Selection: A Game-Theoretic Model for the Exercise of Peremptory Challenges." *Operations Research*, 26(6): 966–991.

Broderick, Raymond J. 1992. "Why the Peremptory Challenge Should Be Abolished." *Temple Law Review*, 65, p. 369.

Cohen, Neil P., and Daniel R. Cohen. 2003. "Jury Reform in Tennessee." *University of Memphis Law Review*, 34 1–71.

Craft, Will. 2018. "Peremptory Strikes in Mississippi's Fifth Circuit Court District." *APM Reports.*

Daly, Meghan. 2016. "Foster v. Chatman: Clarifying the Batson Test for Discriminatory Peremptory Strikes." *Duke Journal of Constitutional Law and Public Policy Sidebar*, 11 148–162.

Diamond, Shari Seidman, Destiny Peery, Francis J. Dolan, and Emily Dolan. 2009. "Achieving Diversity on the Jury: Jury Size and the Peremptory Challenge." *Journal of Empirical Legal Studies*, 6(3): 425–449.

Flanagan, Francis X. 2015. "Peremptory Challenges and Jury Selection." *Journal of Law and Economics*, 58(2): 385–416.

Flanagan, Francis X. 2018. "Race, Gender, and Juries: Evidence from North Carolina."

The Journal of Law and Economics, 61(2): 189–214.

Hochman, Rodger L. 1993. "Abolishing the Peremptory Challenge: The Verdict of Emerging Caselaw." *Nova Law Review*, 17, p. 1367.

Horwitz, Barbara L. 1992. "Extinction of the Peremptory Challenge: What Will the Jury System Lose by Its Demise." *University of Cincinnati Law Review*, 61 1391–1440.

Keene, Douglas L. 2009. "Fairness, Justice and True Understanding: The Benefits of Peremptory Strikes." *The Jury Expert*, 2(21): 24–25.

LaFave, Wayne, Jerold Israel, Nancy King, and Orin Kerr. 2009. *Criminal Procedure*. St. Paul, MN: West Academic Publishing, , 5th edition.

Marder, Nancy S. 1994. "Beyond Gender: Peremptory Challenges and the Roles of the Jury." *Texas Law Review*, 73 1041–1138.

Marder, Nancy S. 2012. "Batson Revisited - Batson Symposium." SSRN Scholarly Paper ID 2165561, Social Science Research Network, Rochester, NY.

Raphael, Michael J., and Edward J. Ungvarsky. 1993. "Excuses, Excuses: Neutral Explanations under Batson v. Kentucky." *University of Michigan Journal of Law Reform*, 27 229–276.

Rose, Mary R. 1999. "The Peremptory Challenge Accused of Race or Gender Discrimination? Some Data from One County.." *Law and Human Behavior*, 23(6): 695–702.

Sacks, Patricia E. 1989. "Challenging the Peremptory Challenge: Sixth Amendment Implications of the Discriminatory Use of Peremptory Challenges." *Washington University Law Quarterly*, 67(2): , p. 29.

Shapard, John, and Molly Johnson. 1994. "Memorandom on a Survey of Active Judges Regarding Their Voir Dire Practices." *Federal Judicial Center, Research Division*.

Small, Mario L., and Devah Pager. 2020. "Sociological Perspectives on Racial Discrimination." *Journal of Economic Perspectives*, 34(2): 49–67.

Smith, Abbe. 2014. "A Call to Abolish Peremptory Challenges by Prosecutors." *Georgetown Journal of Legal Ethics*, 27 1163–1186.

Turner, Billy M., Rickie D. Lovell, John C. Young, and William F. Denny. 1986. "Race and Peremptory Challenges during Voir Dire: Do Prosecution and Defense Agree?" *Journal of Criminal Justice*, 14(1): 61–69.

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