Bounded rationality and the choice of jury selection procedures

Martin Van der Linden

January 12, 2017

Abstract

A peremptory challenge procedure allows the parties to a jury trial to dismiss some prospective jurors without justification. Complex challenge procedures offer unfair advantage to parties who are better able to strategize. I introduce a new measure of strategic complexity based on "level-k" thinking and use this measure to compare challenge procedures commonly used in practice. In applying this measure, I overturn some commonly held beliefs about which jury selection procedures are "strategically simple".

1 Introduction

It is customary to let the parties involved in a jury trial dismiss some of the potential jurors without justification. The procedures for dismissal are known as peremptory challenge procedures, and are the status-quo in many common-law countries, including the United States.[1] Side-stepping the lively controversy on the value of such procedures, it is natural to ask whether particular challenge procedures are more desirable than others.[2]

Fairness is an important issue in jury selection. One feature of procedures which impacts fairness is strategic complexity. If a procedure is complex, parties with better strategic skills are likely to secure more favorable juries. This

[1] In *Swain v. Alabama*, the Supreme Court affirmed that "the [peremptory] challenge is one of the most important of the rights secured to the accused." (LaFave et al., 2009). Following *Batson v. Kentucky*, a party can disqualify a peremptory challenge by an opponent if she can prove that the challenge was based on a set of characteristics including race or gender. However, *Batson v. Kentucky* is notoriously hard to implement and judges rarely rule in favor of Batson challenges (Marder, 2012; Daly, 2016).

[2] See, e.g., Horwitz (1992) and Keene (2009) for a defense and Broderick (1992) and Smith (2014) for an attack on peremptory challenges.

is particularly relevant in jury selection where the parties put significant resources into adopting an effective strategy and jury selection consultancy has become a well-established industry.[3] Using strategically simple procedures limits the impact on the selected jury of differences in the parties' ability to strategize, or in their financial means to hire jury consultants.

Experienced judges share this concern and have developed procedures which attempt to limit the parties' ability to strategize. In a report on judges' practices regarding peremptory challenges, Shapard and Johnson (1994, p. 6) write:

> "Some judges require that peremptories be exercised [following procedure X] [...]. This approach [...] makes it more difficult to pursue a strategy prohibited by *Batson* (or any other strategy). [...] A more extreme approach to the same end is [procedure Y] [...]. This approach imposes maximum limits on counsel's ability to employ peremptories in a strategic manner."[4]

In this paper, I develop a new measure of strategic complexity based on *level-k* thinking (see the survey in Crawford et al., 2013) and use this measure to compare the complexity of some challenge procedures commonly used by judges.

The comparison of jury selection procedures presents two challenges. First, jury selection procedures are indirect mechanisms in which the parties' actions do not consist in revealing their preferences.[5] Second, in some procedures commonly used in practice, the parties submit their challenges simultaneously, which induces games of incomplete information. These two difficulties make it impossible to apply measures of strategic complexity previously developed in the literature (Pathak and Sönmez, 2013; de Clippel et al., 2014; Arribillaga and Massó, 2015).

I overcome these difficulties by introducing the concept of a *rationality threshold*. Given a model of the behavior of her opponent, a party has a *straightforward* strategy if one of her strategies is a best response to *any*

[3] As evidenced by the existence of the American Society of Trial Consultants, and its publication "The Jury Expert: The Art and Science of Litigation Advocacy". Jury consultant Roy Futterman writes: "Caditz argues that [...] jury selectors pay [...] little to no attention to the strategic use of strikes. [...] it is a bit of a reach to say that strategy is barely utilized. In my experience, [...] [jury selection] comes closer to a long battle of stealth, counter-punches, misdirection, and hand-to-hand combat than a lofty academic experience." (Excerpt from Roy Futterman's answer to David Caditz's August 20, 2014 post on `http://www.thejuryexpert.com/`).

[4] See footnote 1 regarding *Batson v. Kentucky*.

[5] A procedure based on direct preference revelation would go against the idea of allowing the parties to challenge jurors *without justification*.

strategy of her opponent that is consistent with this model. The rationality threshold measures the complexity of the model of her opponent that a party needs in order to have a straightforward strategy. For example, a rationality threshold of 1 corresponds to the parties having dominant strategies. In this case, a party has a straightforward strategy given *any* model of her opponent. When the rationality threshold is 2, a party only needs to know that her opponent is a best responder in order to have a straightforward strategy.[6]

Using the rationality threshold as a measure of strategic complexity provides new insights and overturns some commonly held beliefs about jury selection procedures. Shapard and Johnson (1994, p. 6) write:

> "Other judges, for the same purposes [limiting the parties' ability to strategize], allow all peremptories to be exercised after all challenges for cause, but with the parties making their choices 'blind' to the choices made by opposing parties (in contrast to alternating "strikes" from a list of names of panel members)."[7]

I show that, contrary to the judges' intuition, procedures in which challenges are sequential tend to be strategically simpler than procedures in which challenges are simultaneous: By generating incomplete information games, simultaneous procedures increase the amount of guesswork needed to determine optimal strategies. I identify theoretical conditions under which this is true and show that these conditions are relevant for jury selection using simulations as well as field data.

I also study the design of "maximally simple" jury selection procedures. I show that it is *im*possible to construct a "reasonable" procedure that allows the parties to challenge jurors and always has a dominant strategy. Hence, the lowest achievable rationality threshold is 2. Such a lowest rationality threshold is attained by a procedure that I call *sequential one-shot* in which the parties sequentially submit a single list of jurors that they want to challenge.

Although the focus of this paper is the study of jury selection procedures, the rationality threshold is a general measure of strategic complexity. Unlike previous measures in the literature, the rationality threshold can be used to compare any type of game, including indirect games and games of incomplete

[6] Ho et al. (1998) previously introduced the term "rationality threshold" to refer to a similar concept. They used their concept of a rationality threshold to explain some aspects of behavior in beauty contest games. The novelty here is to use the rationality threshold as a measure of strategic complexity.

[7] Unlike peremptory challenges, challenges *for cause* must be based on biases recognized by law, such as being a direct relative of one of the parties.

information. The rationality threshold is the first such measure of strategic complexity proposed in the literature.

Related Literature.— This paper differs from the previous game theoretic literature on jury selection procedures in at least two ways (see Flanagan (2015) for a recent review). First, the literature has focused on subgame perfect equilibrium as a solution concept.[8] Subgame perfection requires a high level of strategic sophistication, especially in complex procedures. By relying on the concept of a rationality threshold, this paper accounts for the possibility of boundedly rational parties. I show how the rationality threshold, which measures the "amount" of common knowledge and rationality needed to reach an equilibrium, can be used to measure the strategic complexity of a procedure.

Second, here, jury selection is studied from the point of view of market and mechanism design.[9] Most of the literature focuses on the characterization and properties of equilibria in different procedures.[10] When the performance of procedures is compared, it is typically in terms of their effects on the composition of the jury. This approach has yielded few policy recommendations.[11] By contrast, this paper adopts a traditional mechanism design approach and compares procedures with respect to the standard objective of limiting the parties' ability to strategize. This later approach enables a clear comparison of some of the procedures used by judges.

The paper is organized as follows. Section 2 introduces the model as well as several examples of jury selection procedures and a general class thereof. In Section 3, I show that most "reasonable" jury selection procedures do not have dominant strategies. Section 4 formally introduces the concept of a rationality threshold. The rationality threshold is then applied to comparing the strategic complexity of jury selection procedures in Sections 5 and 6. Omitted proofs can be found in the Appendix.

2 Model and procedures

I focus on *struck* procedures. In addition to peremptory challenges which require no justification, the parties can raise challenges *for cause* which must be based on some bias recognized by law, such as being a direct relative of

[8] Two exception are Bermant (1982) and Caditz (2015).

[9] In this respect, the closest paper is de Clippel et al. (2014), which takes a mechanism design perspective but studies the selection of a *single* arbitrator.

[10] See Roth et al. (1977), Brams and Davis (1978), DeGroot and Kadane (1980), Kadane et al. (1999), Alpern and Gal (2009), and Alpern et al. (2010).

[11] See, however, Bermant (1982) and Flanagan (2015, Section 4.2).

one of the parties. As explained by Bermant and Shapard (1981, p. 92), the defining feature of a struck procedure "is that the judge rules on all challenges for cause before the parties claim any peremptories. Enough potential jurors are examined to allow for the size of the jury plus the number of peremptory challenges allotted to both sides. In a federal felony trial, for example, the jury size is twelve; the prosecution has six peremptories, and the defense has ten. Under the struck jury method, therefore, 28 potential jurors are cleared through challenges for cause before the exercise of peremptories."

This contrasts with *strike and replace* procedures in which challenges for cause and peremptory challenges are intertwined. In a strike and replace procedure, prospective jurors who are challenged (either for cause or peremptorily) are replaced by new jurors from the pool and "to one degree or another, counsel exercise their challenges without knowing the characteristics of the next potential juror to be interviewed" (Bermant and Shapard, 1981, p. 93).

Struck procedures are commonly used in federal courts. In a 1977 survey of judges' practices regarding the exercise of peremptory challenges, 55% of federal district judges reported using a struck procedure (Bermant and Shapard, 1981), as opposed to a strike and replace procedure. Today, the use of a struck procedure is, for example, recommended by law as the preferred method for criminal cases other than first-degree murder in Minnesota.[12] Struck procedures are used in other court circuits as well.[13] However, the details of a procedure are often left to the discretion of the judge and may vary inside a single court circuit or state (Bermant and Shapard, 1981; Cohen and Cohen, 2003), which makes it hard to determine the procedures used in practice without surveying judges.

2.1 The model

The set of prospective jurors left after all challenges for cause have been raised is $N = \{1, \ldots, n\}$. The defendant is d and the plaintiff is p. The defendant and the plaintiff are allowed c_d and c_p peremptory challenges, respectively. Out of N, a jury J of b jurors must be selected. The jurors in J are the **impaneled** jurors. As explained above, when struck procedures are considered, $n = b + c_d + c_p$ in order to allow the parties to challenge up to c_d and c_p jurors.

[12] Minnesota Court Rules, Criminal Procedure, Rule 26.02, Subd.4.(3)b). Minnesota Court Rules, however, recommend using a strike and replace procedure for first-degree murders (Minnesota Court Rules, Criminal Procedure, Rule 26.02, Subd.4.(3)d)).

[13] New Jersey Court Rule 1:8-3(e) suggests the use of a struck procedure.

Let $\mathcal{J}$ be the set of juries containing b jurors and $\Delta\mathcal{J}$ be the set of lotteries on $\mathcal{J}$. The parties have expected utility preferences R_d and R_p on $\Delta\mathcal{J}$, respectively, with corresponding Bernoulli utility functions u_d and u_p on $\mathcal{J}$. A pair of preferences (R_p, R_d) is called a **(preference) profile** and a quintuple (R_d, R_p, c_d, c_p, b) a **(jury selection) problem**.

Example 1. If $\pi(J)$ is the probability that jury J convicts the defendant, then party i's preference could be represented by a utility function of the form $u_i = v_i(\pi(J))$, where v_p is increasing and v_d decreasing.

Throughout, I assume that preferences on juries are **separable**, i.e., if replacing juror h by juror j in jury J is an improvement according to u_i, then the same is true when h is replaced by j in any other jury J'. Formally, for any $i \in \{d, p\}$ and any $J, J' \in \mathcal{J}$, $h \in J \cap J'$ and $j \in N \backslash (J \cup J')$, $u_i(J \cup \{j\} \backslash \{h\}) \geq u_i(J)$ implies that $u_i(J' \cup \{j\} \backslash \{h\}) \geq u_i(J')$.

Separability is mostly assumed for simplicity. Separability eases the exposition because it implies that the preferences R_d and R_p induce well-defined preferences *on the jurors* in N. Below, I explain how some of the results in this paper can be extended when the separability assumption is relaxed. For simplicity, it is also assumed that the preferences on jurors induced by R_d and R_p are *strict*. Slightly abusing the notation, let R_i serve to denote i's preference on jurors.

An extreme kind of profiles are **juror inverse** profiles. A profile is juror inverse if R_d and R_p induce inverse preferences on jurors (i.e., for all $j, h \in N$, $j \; R_p \; h$ if and only if $h \; R_d \; j$). Unlike separability which is assumed throughout the paper, juror inverse profiles are only considered as a special case.

2.2 Procedures

As attested to by Bermant and Shapard (1981), a wide variety of struck procedures are used by judges. One common type of struck procedure are procedures that I call **one-shot**. In a one-shot procedure, each party $i \in \{d, p\}$ submits a list of up to c_i jurors in N that i wants to challenge. Depending on the procedure, the parties submit their lists simultaneously (**one-shot**$_M$) or sequentially (**one-shot**$_Q$). The impaneled jurors are the jurors in N who have not been challenged. If more than b jurors are left unchallenged, the b impaneled jurors are drawn (uniformly) at random from the unchallenged jurors.

The use of one-shot$_M$ is documented by Bermant (1982, Step. 5, Comments by Judges Feikens and Voorhees). I have not found direct evidence of the use of one-shot$_Q$, although Bermant (1982, Step. 5, Comments by Judge Enright) shows that a procedure in which the parties alternate challenges *twice* has been used in practice, with each party allowed to challenge up to $\frac{c_i}{2}$ jurors in each round.[14]

Another common type of struck procedure are the procedures that I call **alternating**. Alternating procedures proceed through a succession of rounds in which the parties can challenge as many jurors in N as they have challenges left. Again, an alternating procedure can be either simultaneous (**alternating**$_M$) or sequential (**alternating**$_Q$) depending on whether challenges are submitted simultaneously or sequentially. In alternating$_M$, if both parties challenge the same jurors in a given round, both parties are charged with the challenge and can challenge one less juror.

Alternating procedures stop when neither of the parties has challenges left, or when both parties abstain from challenging a juror in a single round. The impaneled jurors are the jurors left unchallenged in N, or a random draw of b of these jurors if more than b jurors are left unchallenged.

Alternating$_Q$ is, for example, recommended by law as the preferred method for criminal cases other than first-degree murder in Minnesota (see footnote 12). I have not found direct evidence of the use of alternating$_M$. However, the use of simultaneous challenges in sequential procedures is not unheard of. Simultaneous challenges are, for example, used in the alternating *strike and replace* procedure for civil cases in Tennessee.[15]

One-shot and alternating procedures are part of the class of ***N*-struck procedures** in which parties take turns challenging jurors from N for a number of rounds.[16] Formally, every N-struck procedure consists of a maximum of $f \geq 1$ rounds, where f differs between procedures. Each round $r \in \{1, \dots, f\}$ is characterized by a maximum number of challenges $x_i^r \geq 1$ for each party, with $\sum_{r=1}^{f} x_i^r \geq c_i$. The number of challenges party i has left in round r is ℓ_i^r, with $\ell_i^1 = c_i$. In each round r :

(a) The parties can challenge up to $min\{x_i^r, \ell_i^r\}$ jurors among the jurors in

[14] In the prior game theoretic literature on jury selection procedures, one-shot procedures have sometimes been misleadingly called *struck* procedures. In line with the law literature (see the quote from Bermant and Shapard, 1981, at the beginning of this section), it is more appropriate to use the term "struck procedures" to denote the *class* of struck procedures that includes one-shot procedures.

[15] Tennessee Court Rules, Rules of Civil Procedure, Rule 47.03. The imposition of a uniform procedure for peremptory challenges in Tennessee came with a reform started in late 1997 and described in detail in Cohen and Cohen (2003). The adoption of simultaneous challenges was part of the reform's effort to "permit peremptory challenges to be exercised in a way that does not disclose to potential jurors which lawyer exercised the challenge." (Cohen and Cohen, 2003, pp. 25–26)

[16] The name "N-struck procedure" emphasizes the fact that, in each round, the parties can challenge any juror in N that has not been challenged yet. This is not the case in every struck procedure. See, e.g., the procedure described in footnote 19.

N who have not yet been challenged. Challenges are sequential if the procedure is sequential, and simultaneous if the procedure is simultaneous.

(b) For each party $i \in \{d, p\}$, the number of challenges left is decreased by the number of jurors that the party challenged in (a) (i.e., ℓ_i^{r+1} equals ℓ_i^r minus the number of jurors that the party challenged in (a)).

The procedure terminates when no party has challenges left, when round f is reached, or when both parties abstain from challenging any juror in a single round. The jurors left unchallenged when the procedure terminates are the impaneled jurors. If more than b jurors are left unchallenged when the procedure terminates, the b impaneled jurors are drawn at random from the unchallenged jurors.

One-shot procedures are N-struck procedures with $f = 1$ and $x_i^1 = c_i$ for both parties $i \in \{d, p\}$. Alternating procedures are N-struck procedures with $x_i^r = c_i$ for both $i \in \{d, p\}$ and all $r \in \{1, \dots, f\}$, and $f = 2 \max_{i \in \{d,p\}} c_i$. Besides one-shot and alternating procedures, the class of N-struck procedures includes, for example, the two-round procedure described above.

From a game theoretic point of view, a **(jury selection) procedure** is an extensive game form $\Gamma \colon \mathcal{S}_d \times \mathcal{S}_p \to \Delta\mathcal{J}$ that associates any pair of strategies (s_d, s_p) in some strategy space $\mathcal{S}_d \times \mathcal{S}_p$ with a lottery on juries in $\mathcal{J}$.

3 Impossibility results

Given preference R_i, a **best response** for party i to some strategy s_{-i} of her opponent is a strategy $t_i(s_{-i})$ such that

$$\Gamma\big(t_i(s_{-i}), s_{-i}\big) \ R_i \ \Gamma\big(s_i', s_{-i}\big) \qquad \text{for all } s_i' \in \mathcal{S}_i. \tag{1}$$

When $-i$ plays s_{-i} and i plays $t_i(s_{-i})$, party i **best responds** to $-i$. Given some model $S_{-i} \subseteq \mathcal{S}_{-i}$ of her opponent, a strategy $s_i \in \mathcal{S}_i$ is **straightforward for *i*** if s_i is a best response to *every* strategy $s_{-i} \in S_{-i}$. A **dominant strategy** is a strategy $s_i^* \in \mathcal{S}_i$ that is a best response for i to *any* strategy $s_{-i} \in \mathcal{S}_{-i}$. In other words, a dominant strategy is a strategy that is straightforward for i given *any model* of her opponent.[17]

[17] Conversely, i has a straightforward strategy given model $S_{-i} \subseteq \mathcal{S}_{-i}$ if i has a dominant strategy in the reduced game $\Gamma' : \mathcal{S}_i \times S_{-i} \to \Delta\mathcal{J}$ defined by $\Gamma'(s_i, s_{-i}) = \Gamma(s_i, s_{-i})$ for all $(s_i, s_{-i}) \in \mathcal{S}_i \times S_{-i}$. In this sense, the concept of a straightforward strategy generalizes that of a dominant strategy.

Given some domain of preferences, a **dominant strategy procedure** is a procedure in which both parties have a dominant strategy for every profile in the domain. Dominant strategy procedures are strategically simple because each party can determine an optimal strategy independently of any guess about the strategy of her opponent.[18] Dominant strategy procedures guarantee a form of equality among equals: Two parties having the same preference but different abilities to guess how their opponent will play should be able to secure similar outcomes.

It is useful to relate dominant strategies with level-k thinking (see the survey in Crawford et al., 2013). In the level-k terminology, an L^0 party is a non-strategic party who could potentially play any strategy. An L^1 party assumes that her opponent is L^0, makes a guess about the L^0 strategy s^0 that her opponent will employ, and best responds to s^0. Similarly, an L^k party assumes that her opponent is L^{k-1}, makes a guess about the L^{k-1} strategy s^{k-1} that her opponent will employ, and best responds to s^{k-1}.

Observe that, because an L^0 strategy can be any strategy, i has a dominant strategy if and only if i has an L^1 strategy that is a best response to *every* L^0 strategy of her opponent. In the language of level-k thinking, a dominant strategy procedure limits the impact of differences in strategic skills because i can determine an optimal strategy independently of her belief about her opponent's level of rationality k_{-i}, or her guess about which $L^{k_{-i}}$ strategy her opponent will employ.

Unfortunately, most reasonable procedures that permits challenges do not have a dominant strategy. Consider one-shot$_M$. In one-shot$_M$, i's only best response to any strategy s_{-i} is to challenge her c_i worst jurors among the jurors that $-i$ does not challenge in s_{-i}. As illustrated in Example 2, such a best response is very dependent on the challenges chosen by $-i$. Hence, one-shot$_M$ is not a dominant strategy procedure.

Example 2. Suppose that each juror has four challenges ($c_d = c_p = 4$) and one juror has to be selected ($b = 1$). A set of nine prospective jurors $N = \{1, \ldots, 9\}$ will therefore remain after all challenges for cause have been raised. Let d's preference on these nine jurors be $1 \; R_d \; 2 \; R_d \; \ldots \; R_d \; 9$. If p challenges the circled jurors in (2), then d's best response is to challenge the squared jurors in (2).

$$R_d : \quad 1 \quad \textcircled{2} \quad \textcircled{3} \quad \boxed{4} \quad \textcircled{5} \quad \textcircled{6} \quad \boxed{7} \quad \boxed{8} \quad \boxed{9} \qquad (2)$$

[18] Requiring the existence of dominant strategies in a procedure is somewhat similar to requiring strategy-proofness in direct mechanisms. One difference is that s_i^* need not be a simple function of i's true preference.

On the other hand, if p challenges the circled jurors in (3), then d's best response is to challenge the squared jurors in (3).

$$R_d: \quad 1 \quad \textcircled{2} \quad \textcircled{3} \quad \textcircled{4} \quad \textcircled{5} \quad \boxed{6} \quad \boxed{7} \quad \boxed{8} \quad \boxed{9} \tag{3}$$

Clearly, the challenge of the squared jurors in (3) is not a best response for d to p challenging the circled jurors in (2), which shows that one-shot$_M$ is not a dominant strategy procedure.

As shown in Proposition 1, the preceding example generalizes to the whole class of N-struck procedures and to any problem. Intuitively, in any N-struck procedure, if $-i$ does not challenge any jurors, then i's best response is to challenge her c_i worst jurors. On the other hand, if $-i$ challenges one of the c_i worst jurors of i, say w, then i is better off not challenging w and challenging her other c_i worst jurors. Recall that a (jury selection) *problem* is a quintuple (R_d, R_p, c_d, c_p, b).

Proposition 1. *For any problem, (i) one of the parties does not have a dominant strategy in one-shot$_Q$ and (ii) neither party has a dominant strategy in any N-struck procedure different from one-shot$_Q$.*

Note that one-shot$_M$ is an N-struck procedure different from one-shot$_Q$. Hence, Proposition 1 shows that, for every problem, neither party has a dominant strategy in one-shot$_M$. One-shot$_Q$ is the exception among N-struck procedures: It is the only N-struck procedure in which one of the parties — the second party to challenge — has a dominant strategy, although the other party does not (see the proof of the proposition).

Of course, N-struck procedures are only a small subset of all possible jury selection procedures. Other procedures used in practice include the strike and replace procedures, as well as other struck procedures.[19] It is therefore natural to ask whether there exists dominant strategy procedures for jury selection outside of the class of N-struck procedures. The next proposition shows that if such procedures exist, then they must either deprive a party from her right to challenge at least one juror in N, or be so intricate that they are unlikely to be used in practice.

A procedure satisfies **finiteness** if the set of decision nodes is finite for both parties and for Nature. A procedure satisfies **minimal challenge** if

[19] For example, Bermant (1982, Step. 5, Comments by Judge Atkins) describes a struck procedure in which a first set of b jurors is selected from N and the parties take turns challenging jurors in that set, with every challenged juror being replaced by another juror from N that has not been selected yet. This procedure is not an N-struck procedure because juror from N cannot be challenged before they have been picked to replace one of the b jurors initially selected.

for every prospective juror $j \in N$, both parties $i \in \{d, p\}$ have a strategy $s_i^j \in \mathcal{S}_i$ such that j is never part of the chosen jury when i plays s_i^j.[20] Every N-struck procedure satisfies both finiteness and minimal challenge (strategy s_i^j can, for example, involve challenging juror j and only juror j in the first round).

Proposition 2. *On the domain of separable preferences, no dominant strategy procedure satisfies both finiteness and minimal challenge.*

In the Appendix, I show that Proposition 2 is true even for smaller domains of profiles, including the domain of additive profiles. The proof of Proposition 2 makes use of Theorem 2 in Van der Linden (2015). That theorem implies that, in order to be a dominant strategy procedure, a procedure satisfying minimal challenge must generate a direct mechanism that has a range of possible outcomes which is in some sense "dense". This, in turn, implies that the parties sometimes have a continuum of actions which conflicts with finiteness.

4 A measure of strategic complexity

Propositions 1 and 2 show that most procedures are not strategically simple in the sense that both parties cannot always follow the simple recommendation of playing a dominant strategy. This does not mean, however, that judges should give up on the idea of using procedures that are *as simple as possible*. This section and the next show that, although procedures generally fail to feature dominant strategies, not all procedures are equal in terms of strategic complexity.

4.1 Motivating example

Brams and Davis (1978, p. 969) have argued that, when the parties have juror inverse preferences, one-shot procedures raise "no strategic questions of timing: given that each side can determine those veniremen it believes least favorably disposed to its cause, it should challenge these up to the limit of its peremptory challenges.". This may come as a surprise given Example 2 and Proposition 1. Certainly, one-shot$_M$ is not a dominant strategy procedure. How can we then make sense of Brams and Davis' claim? The next examples hints at one possible answer.

[20] That is, the probability that j is chosen given that i plays s_i^j is zero for all s_{-i}.

Example 3. Consider one-shot$_M$ with $c_d = c_p = 2$ and $b = 5$. Let d have preference $1 \ R_d \ \ldots \ R_d \ 9$. Also, suppose that the parties have juror inverse preferences.

If d believes that p is best responding to some of her strategies, then d knows that p will challenge two of the circled jurors in (4).

$$\begin{array}{lccccccccc} R_d: & \textcircled{1} & \textcircled{2} & \textcircled{3} & \textcircled{4} & 5 & 6 & 7 & \boxed{8} & \boxed{9} \\ R_p: & 9 & 8 & 7 & 6 & 5 & \textcircled{4} & \textcircled{3} & \textcircled{2} & \textcircled{1} \end{array} \tag{4}$$

Indeed, a best response by p involves challenging her two worst jurors among the seven jurors that she believes d will not challenge. This can never lead to p challenging a juror in $\{5, \ldots, 9\}$.

Thus, a best response by d to the minimal belief that p is a best responder always consists in challenging her two worst jurors (squared in (4)). By symmetry, the same is true for p.

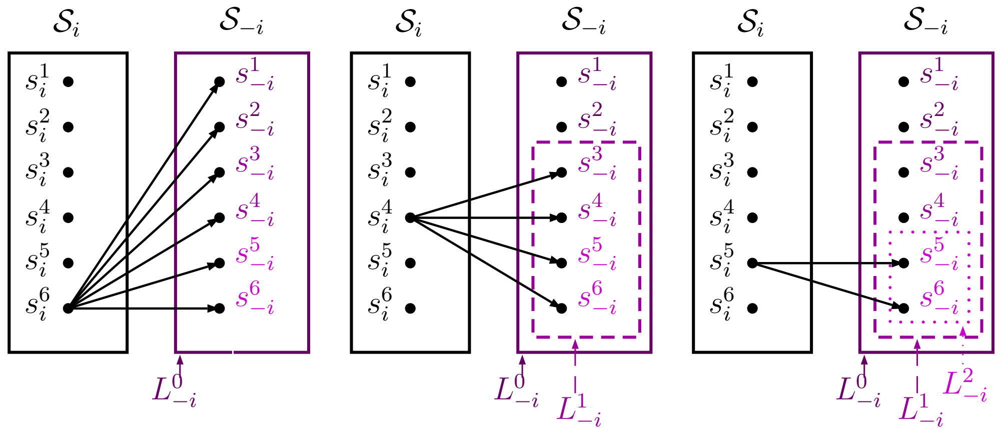
In Example 3, one-shot$_M$ "raises no strategic question" because a party only needs to know that her opponent is a best responder in order to have a straightforward best response. For each party i, challenging her c_i worst jurors is a best response to *any* strategy of party $-i$ that is itself a best response to some of i's strategies. In this sense, each party has a straightforward strategy given a *minimal* model of the strategic behavior of her opponent.

In the rest of this section, I generalize this logic to obtain a measure of strategic complexity. This measure is then applied in the next two sections to compare struck procedures for different assumptions on the problem (R_d, R_p, c_d, c_p, b).

4.2 The rationality threshold

As argued above, first-best procedures are dominant strategy procedures in which the parties have a straightforward strategy *whatever their model* of their opponent. It is then natural for second-best procedures to be procedures in which the parties have a straightforward strategy *given a minimal model* of their opponent. As suggested by Example 3, a meaningful concept of *minimal model* is for a party to assume that her opponent will play a best response to some of her strategies.

In the language of level-k thinking, a procedure is second-best if each party i has an L^2 strategy that is a best responds to *every* L^1 strategy of her opponent. Such second-best procedures limit the impact of differences in strategic skills because i's optimal strategy depends *minimally* on her model



(a) First-best procedure (b) Second-best procedure (c) Third-best procedure

Figure 1: Representation of first-, second-, and third-best procedures. An arrow from strategy s_i to strategy s_{-i} means that s_i is a best response to s_{-i}. In first-best procedures represented in (a), each party i has a strategy — s_i^6 in the figure — that is a best response to every strategy of her opponent (i.e., to every L_{-i}^0 strategy). In second-best procedures represented in (b), each party i has a strategy — s_i^4 in the figure — that is a best response to every strategy of her opponent that is itself a best response (i.e., to every L_{-i}^1 strategy). However, s_i^4 does not need to be a best response to *every* L_{-i}^0 strategy. For example, in the figure, s_i^4 is not a best response to s_{-i}^1. The same logic extends to third-best procedures represented in (c).

of $-i$: i only needs to assume that $-i$ is L^1 to have a straightforward strategy. The difference between first-best and second-best procedures is illustrated in Figure 1(a) and (b).

A second-best procedure guarantees a form of second-best equality among equal parties. Consider two defendants with the same preference who both believe that the plaintiff is L^1 and best responds to one of their strategies. The two defendants might differ in other strategic aspects, such as their ability to guess which of their strategies the plaintiff best responds to. In a second-best procedure, these differences have no impact: the two defendants play equivalent strategies and secure the same outcome.

Similarly, third-best procedures feature straightforward best responses given a model that is *minimally stronger* than in second-best procedures. A natural candidate for such a minimally stronger model is for i to assume that $-i$ is L^2 (see Figure 1(c)). This logic extends to higher level reasoning.

In procedures with multiple rounds, it is important to ensure that best responses be enforced throughout the game tree. Therefore, the measure of strategic complexity defined below relies on iterated *conditional* best responses, rather than iterated best responses. Iterated conditional best responses are obtained by iterative elimination of strategies that are never best responses *in every subgame* of an extensive game.[21]

Definition 1 (Iterated conditional best response)**.** For any procedure Γ and any profile (R_d, R_p), the process of *iterated conditional best response* is defined as follows:

Round 0. For each $i \in \{d, p\}$, the set of $\boldsymbol{CL_i^0}$ **(conditionally level 0) strategies** is $\mathcal{S}_i$.

Round 1. For each $i \in \{d, p\}$, eliminate from CL_i^0 the strategies s_i for which there exists a subgame γ of Γ such that the restriction $s_i|_\gamma$ of s_i to γ is not a best response to any $s_{-i}|\gamma$ in γ.

The remaining set of strategies is denoted by CL_i^1. Any $s_i \in CL_i^1$ is called a $\boldsymbol{CL_i^1}$ **(conditionally level 1) strategy** for i.

⋮

Round k. For each $i \in \{d, p\}$, eliminate from CL_i^{k-1} the strategies s_i for which there exists a subgame γ of Γ such that the restriction $s_i|_\gamma$ of s_i to γ is not a best response to $s_{-i}|_\gamma$ for any $s_{-i} \in CL_{-i}^{k-1}$.

The remaining set of strategies is denoted CL_i^k. Any $s_i \in CL_i^k$ is called a $\boldsymbol{CL_i^k}$ **(conditionally level k) strategy** for i.

Observe that the sets of conditionally level k strategies are nested ($CL_i^0 \supseteq CL_i^1 \supseteq \dots$). Observe also that, for every procedure Γ that satisfies finiteness, the set of conditionally level k strategies is non-empty for every k.[22]

The argument at the beginning of this section suggests using the following concept of a rationality threshold as a measure of strategic complexity.[23]

---

[21] The use of "conditional" to qualify a process of iterated elimination of strategies which applies to every subgame is taken from Fudenberg and Tirole (1991).

[22] When Γ satisfies finiteness, for any $i \in \{d, p\}$ and any strategy s_{-i}, the set $\{\Gamma(s_i, s_{-i}) \mid s_i \in \mathcal{S}_i\}$ is finite because $\mathcal{S}_i$ is finite. Hence, there must exist a strategy $t_i(s_{-i})$ such that $\Gamma\big(t_i(s_{-i}), s_{-i}\big)\ R_i\ \Gamma\big(s_i', s_{-i}\big)$ for all $s_i' \in \mathcal{S}_i$ and CL_i^1 is non-empty. The non-emptiness of CL_i^k then follows by induction.

[23] The term "rationality threshold" is often used in relation to the process of iterated elimination of *undominated* strategies (e.g., Mas-Colell et al., 1995; Ho et al., 1998). Here I use the same term to refer to the iterated elimination of strategies that are never *best responses*. In general, the two processes are *not* equivalent, see the Conclusion.

Definition 2 (Rationality threshold)**.** For any procedure Γ and any profile (R_d, R_p), the *rationality threshold* is the smallest integer r^* such that, for each $i \in \{d, p\}$, there exists a $CL_i^{r^*}$ strategy s_i^* that is a best response to *every* $CL_{-i}^{r^*-1}$ strategy.

If there exists no such integer, then the rationality threshold of Γ is ∞, i.e., the procedure cannot be solved by iterated conditional best response. Note that if the rationality threshold r^* is finite, then there exists a pair of strategies $(s_i, s_{-i}) \in CL_i^{r^*} \times CL_{-i}^{r^*}$ is a subgame perfect equilibrium. Any such equilibrium is, in fact, a *straightforward* subgame perfect equilibrium in the sense that each player i can be viewed as playing a straightforward strategy given some model $S_{-i} \subseteq CL_{-i}^{(r^*-1)}$ of the other player. [24]

5 One-shot procedures

In this section, I show that one-shot$_Q$ is strategically simpler than one-shot$_M$ in the following sense.

Proposition 3. *(i) For* every *problem, the rationality threshold of one-shot$_M$ is* no smaller than *the rationality threshold of one-shot$_Q$. (ii) For* some *problems, the rationality threshold of one-shot$_M$ is* larger than *the rationality threshold of one-shot$_Q$.*

In the rest of this section, I prove and illustrate Proposition 3. The proof of Proposition 3(ii) relies on problems the profiles of which are not juror inverse. I show that such problems are common using simulations and field data.

5.1 One-shot$_Q$ is always maximally simple

The next example illustrates how to compute the rationality threshold of one-shot$_Q$ for a particular problem.

Example 4. The example is illustrated in Figure 2. Suppose that $c_d = c_p = b = 1$ and d is the first party to challenge. Also suppose that the parties have aligned preferences $1\ R_d\ 2\ R_d\ 3$ and $1\ R_p\ 2\ R_p\ 3$.

Because preferences on jurors are strict, p has a unique dominant strategy s_p^* which consists in (a) challenging juror 3 if d did *not* challenge 3 and (b) challenging juror 2 if d *did* challenge juror 3 (dashed blue branches in the

[24] The integer r^* is the smallest integer for which such an equilibrium exists.

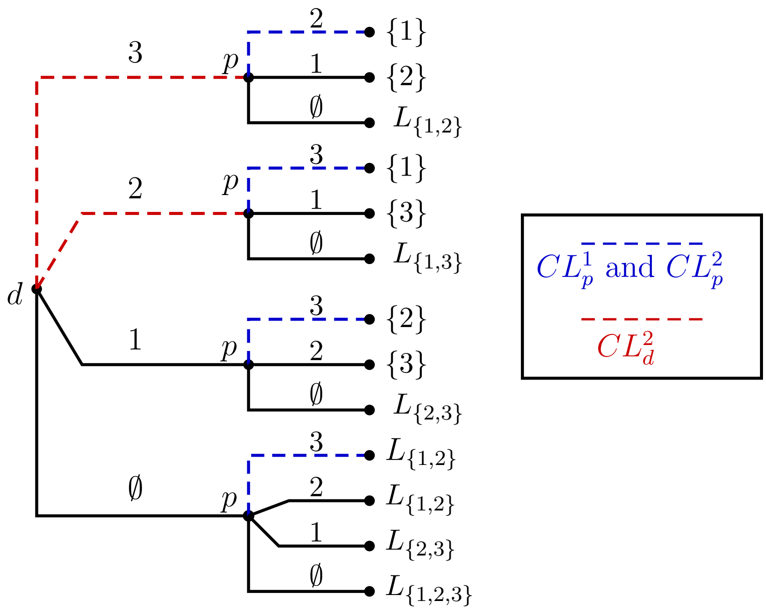


Figure 2: Computing the rationality threshold in Example 4. In the figure, L_T for $T \subset N$ represents a lottery in which two jurors are drawn (uniformly) at random from T. The labels on the branches of the tree indicate the juror who is challenged in the corresponding action.

figure). Strategy s_p^* is, therefore, the unique CL_p^1 strategy. It directly follows from uniqueness that s_p^* is a best response to all CL_d^1 strategies.

Because there is a unique CL_p^1 strategy s_p^*, any CL_d^2 strategy that best responds to s_p^* (either of the red dashed branches in the figure) is a best response to *all* CL_p^1 strategies. Hence, the rationality threshold of one-shot$_Q$ is at most 2 for this problem. But by Proposition 1, because one-shot$_M$ is an N-struck procedure, the rationality threshold of one-shot$_Q$ is at least 2 for every problem. Thus, the rationality threshold of one-shot$_Q$ is 2 for this problem.

It is not hard to see how the argument in Example 4 generalizes to any problem. In general, the party $-i$ who challenges second in one-shot$_Q$ has a unique dominant strategy s_{-i}^*. Then, any best response to s_{-i}^* by i is a best response to every CL_{-i}^1 strategy.

Proposition 4. *For any problem, the rationality threshold of one-shot$_Q$ is 2.*

As can be seen from the proof in Appendix, Proposition 4 does not rely on the separability assumption. Rather, the proof relies on the fact that preferences on the outcomes of the procedure are strict. The proposition also extends to situations in which perfect information (which is implicit in the definition of a rationality threshold of 2) is relaxed. Consider Example 4. In order to have a straightforward best response, d only needs to know that p will challenge juror 3 if she challenges juror 2. Hence, d only needs to know which juror is p's *worst juror* in order to have a straightforward best response (as opposed to knowing *all* of p's preference on jurors).

By Proposition 1, because one-shot$_M$ is an N-struck procedure, the rationality threshold of one-shot$_M$ is at least 2 for every problem. Together with Proposition 4, this implies that the rationality threshold of one-shot$_M$ is never smaller than the rationality threshold of one-shot$_Q$, which proves Proposition 3(i).

5.2 One-shot$_M$ is often complex: one-common profiles

I now show that one-shot$_M$ is more complex than one-shot$_Q$ when the profile is not juror inverse and preferences on jurors satisfy some "commonality at the bottom". One such profile is presented in Example 5. In the example, both parties agree that juror 3 is the worst juror. Therefore, any best responding party would challenge juror 3 if her opponent did not. But each party also prefers a situation in which her *opponent* challenges 3, and she challenges her second worst juror. That is, each party would like to make a credible threat *not* to challenge juror 3 and free ride on her opponent's challenge of juror 3. But because the procedure is simultaneous, such a credible threat is impossible. As explained in detail in Example 5, the impossibility for the parties to commit to leaving juror 3 unchallenged makes the rationality threshold of one-shot$_M$ larger than 2 for this problem. Together with Proposition 4, Example 5 therefore proves Proposition 3(ii).

In fact, the rationality threshold in Example 5 is ∞, which shows just how complex one-shot$_M$ can become when the profile is not juror inverse. As the example illustrates, for some profiles that are not juror inverse, one-shot$_M$ cannot be solved by iterated conditional best response. For these profiles, the parties' common knowledge of each others' rationality, preferences, knowledge of each others' preferences, etc., is not sufficient to provide the parties with straightforward strategies and make the game strategically simple. Even for "high levels of common knowledge", the game induced by one-shot$_M$ remains akin to a *game of chicken* in which each party prefer to swerve (i.e., challenge some of her worst jurors) if her opponent stays straight (i.e., does *not* challenge some of her worst jurors), but prefers to stay straight if her opponent swerves.

	CL_p^0 strategies	CL_d^1 strategies	CL_p^2 strategies	CL_d^3 strategies	...
Juror	3	2	3	2	...
challenged	1	3	1	3	...

Figure 3: For the profile of preferences and the values of c_d, c_p and b in Example 5, the rationality threshold of one-shot$_M$ is ∞.

Example 5. The example is illustrated in Figure 3. Suppose that $b = c_d = c_p = 1$. Also suppose that the parties' preferences are $1\ R_d\ 2\ R_d\ 3$ and $2\ R_p\ 1\ R_p\ 3$.

Both challenging juror 3 and challenging juror 1 are CL_p^0 strategies. Challenging juror 2 is d's best response to p challenging juror 3, and challenging juror 3 is d's best response to p challenging juror 1. Hence, both challenging juror 2 and challenging juror 3 are CL_d^1 strategies. But no strategy of p is a best response to *both* of these CL_d^1 strategies. Therefore, the rationality threshold of one-shot$_M$ is at least 3 for this problem.

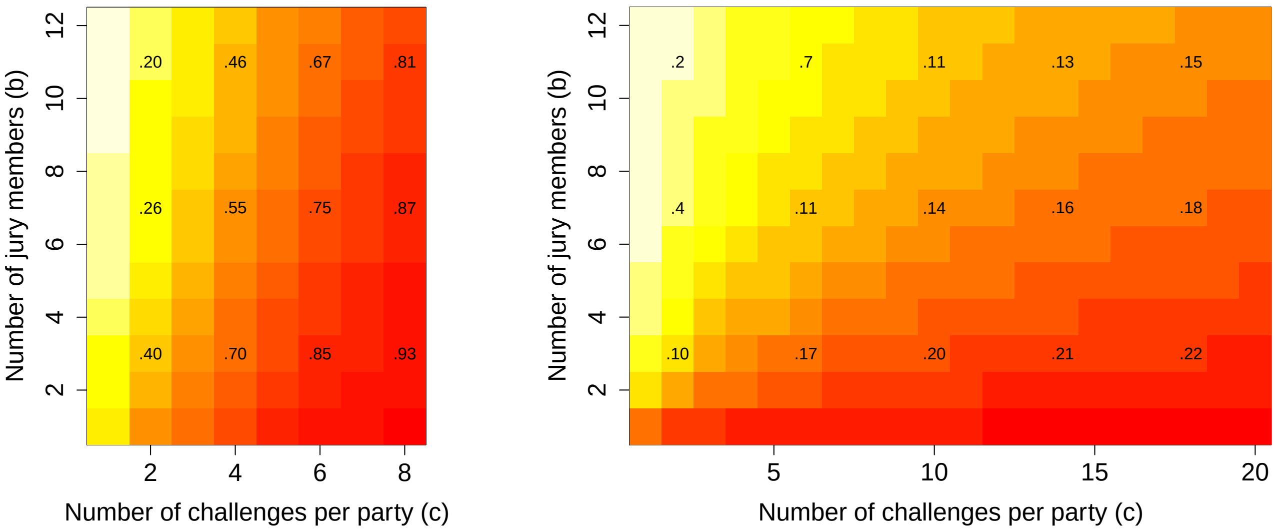
Party p's best responses to these two CL_d^1 strategies are to challenge juror 3 (p's best response to d challenging juror 2) and to challenge juror 1 (p's best response to d challenging juror 3). Thus, both challenging juror 3 and challenging juror 1 are CL_p^2 strategies. But these two CL_p^2 strategies are the two CL_p^0 strategies considered at the beginning of the example. The argument therefore extends by induction to show that the rationality threshold of one-shot$_M$ is ∞ for this problem.

5.2.1 One-common profiles

Profiles for which the rationality threshold of one-shot$_M$ is ∞ are not rare. Given c_d and c_p, a profile is **one-common** if a juror w that is among the c_d worst jurors of d is also among the c_p worst jurors p. Intuitively, the rationality threshold of one-shot$_M$ is ∞ for one-common profiles because the free rider problem described in Example 5 extends to one-common profiles. When the profile is one-common, each party would like to make a credible threat *not* to challenge juror w and free ride on her opponent's challenge of juror w. But if her opponent does not challenge w, each party prefers challenging w herself than leaving w unchallenged.

Proposition 5. *If the profile is one-common, then the rationality threshold of one-shot$_M$ is ∞.*

Although Proposition 5 relies more directly than Proposition 4 on the



(a) (b)

Figure 4: Proportion of one-common profiles relative to the set of (a) all profiles and (b) almost juror inverse profiles. All proportions are proportions of profiles of preferences *on jurors*. Also, the proportions are for the case $c_d = c_p = c$. Alternatively, the proportions are lower bounds for $c = \min\{c_d, c_p\}$. See the Appendix for details on the computation of these proportions.

separability assumption,[25] the core intuition behind Proposition 5 applies when separability is relaxed. Regardless of the assumptions on preferences, if for some juror j, both parties have best responses that include challenging j, then the rationality threshold is larger than 2 in one-shot$_M$. The proposition also extends to situations of imperfect information in which the parties only know that they have a common juror w at the bottom of their ranking of jurors (but do not know each other's complete preference on juror).

One-common profiles arise in a number of natural jury selection situations. For example, both parties might dislike a juror who they view as too unpredictable. Both parties may also dislike "devil advocate" or "irresolute" jurors who are likely to induce hung juries forcing retrials of the case. Finally, d may dislike juror j because of its foreseen position some charges, while p dislike juror j because of its foreseen position a different charges.

As shown in Figure 4(a), the proportion of one-common profiles relative to the set of all profiles is high. (In the figure and henceforth, *proportions of*

[25] One-common profiles are not well-defined without the separability assumption.

profiles refer to proportions of profiles of preferences *on jurors*.) Even among profiles that are close to being juror inverse, the proportion of one-common profiles can be significant when c is high relative to b. Figure 4(b) shows the proportion of one-common profiles among *almost juror inverse* profiles. A profile is **almost juror inverse** if it can be constructed from a juror inverse profiles by changing the ranking of a single juror in the preference of one of the parties.[26] This is illustrated in Example 6.

Example 6. Suppose that $c_d = c_p = 4$ and $b = 1$. The following profile is almost juror inverse because it is constructed from a juror inverse profile by changing the ranking of a single juror — namely juror 7 — in the preference of p.

$$\begin{array}{llllllllll} R_d: & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ R_p: & 9 & 8 & 6 & 5 & 4 & 3 & 2 & 7 & 1 \end{array} \tag{5}$$

The above profile is also one-common because 7 is among the four worst jurors of both d and p, and $c_d = c_p = 4$.

Procedures in which c is high relative to b exist in practice. In the United States, the number of peremptory challenges tends to increase with the gravity of the charges. For example, in federal cases for which the death penalty is sought by the prosecution, $b = 12$ and $c_d = c_p = 20$. In this case, Proposition 5 together with the proportion in Figure 4(b) show that the rationality threshold is ∞ for 15% of almost juror inverse profiles.

5.2.2 One-common profiles in the field

To obtain further evidence of the prevalence of one-common profiles, I consider real-world arbitration cases from the New Jersey Public Employment Relations Commission (Bloom and Cavanagh, 1986; de Clippel et al., 2014). From 1985 to 1996, the Commission used a *veto-rank mechanism* to select an arbitrator in cases involving a union and an employer.

In the veto-rank mechanism used by the Commission, the union and the employer are presented with seven potential arbitrators. The union and the employer *simultaneously* challenge three potential arbitrators and rank order the remaining arbitrators. The chosen arbitrator is the unchallenged arbitrator with the lowest combined rank. Except for the way an arbitrator

[26] Formally, profile (R_d^*, R_p^*) is *almost juror inverse* if there exists a juror inverse profile (R_d, R_p), a party $i \in \{d, p\}$ and a juror $j^* \in N$ such that (a) for all $j, k \in N$, $j \; R_{-i}^* \; k$ if and only if $j \; R_{-i} \; k$, (b) for all $j, k \in N$ with $j, k \neq j^*$, $j \; R_i^* \; k$ if and only if $j \; R_i \; k$, and (c) for some $j \in N$ with $j \neq j^*$, $j \; R_i^* \; j^*$ and $j^* \; R_i \; j$ or $j^* \; R_i^* \; j$ and $j \; R_i \; j^*$.

is selected when challenges overlap, this procedure is equivalent to one-shot$_M$ with $b = 1$ and $c_d = c_p = 3$.[27]

Out of 750 cases, de Clippel et al. (2014) report that the frequency of overlaps in the challenges was as indicated in Table 1.

Number of common challenges	Proportion
0	13%
1	50%
2	34%
3	3%

Table 1: Overlap in challenges in the 750 arbitration cases.

Let a party be **truthful** if she challenges her three worst arbitrators (regardless of her reported ranking over the remaining arbitrators). If both parties were truthful in each of the 750 cases, then the data in Table 1 would imply that the profile of preferences on arbitrators was one-common in 87% of the 750 cases. However, because truthfulness is *not* a dominant strategy in the veto-rank mechanism, this is not an accurate estimate of the proportion of one-common profiles in the 750 cases.

To obtain a more realistic estimate of this proportion, I consider the laboratory experiment on the veto-rank mechanism described in de Clippel et al. (2014). In the experiment, participants play the veto-rank mechanisms with five arbitrators (i.e., $b = 1$ and $c_d = c_p = 2$). The participants are randomly assigned to four different profiles of preferences, denoted Pf1, Pf2, Pf3, and Pf4 (see de Clippel et al., 2014). For each profile, de Clippel et al. (2014) observe 350 instances of the game being played.

Based on the experimental data from de Clippel et al. (2014), I compute for each profile the proportion of plays in which *both* parties were truthful. I also compute this proportion across the four profiles. These proportions are reported in Table 2.

I propose to use the values in Table 2 as estimates of the proportion of the 750 New Jersey cases in which *both* parties were truthful. Whichever estimate x from Table 2 is used, $x - 13\%$ is a lower bound on the proportion of one-common profiles. This lower bound is obtained by assuming that both parties were truthful in *all* 13% of cases in which the challenges did not overlap.

[27] de Clippel et al. (2014) report that after 1996, the Commission started selecting the arbitrator at random from the list of unchallenged arbitrators, which means the Commission effectively used one-shot$_M$ after 1996.

Profile	Proportion of plays in which *both* parties challenged their two worst arbitrators
Pf1	65%
Pf2	45%
Pf3	38%
Pf4	24%
Across the four profiles	43%

Table 2: Proportion of experimental plays of the veto-rank mechanism in de Clippel et al. (2014) in which *both* players challenged their two worst arbitrators.

This lower bounds is illustrated in Figure 5 for different values of the truthfulness estimate. Even under the most conservative truthfulness estimate (i.e., 24%), the lower bound on the proportion of one-common profiles is 11%. Using the truthfulness estimate based on the average across the four profiles, the lower bound on the proportion of one-common profiles is 30%.

Overall, the results in this section contrasts with the judges' intuitions that "blind" (i.e., simultaneous) procedures leave less room for the parties to strategize than sequential ones.[28] Contrary to the judges' intuition, the rationality thresholds suggest that one-shot$_Q$ is strategically simpler than one-shot$_M$: By making past actions observable, one-shot$_Q$ allows the parties to make credible threats about the jurors they challenge, which reduces the amount of guesswork involved in determining an appropriate strategy. The next section shows that similar results hold for other N-struck procedures.

6 Alternating and other N-struck procedures

In general, it is unclear how alternating$_M$ and alternating$_Q$ compare. However, extending the logic of Proposition 5, it is possible to obtain a partial comparison for a significant subset of profiles. For this subset of profiles, the rationality threshold of any *simultaneous* N-struck procedure (including alternating$_M$) is infinite, whereas the rationality threshold of any *sequential* N-struck procedure (including alternating$_Q$) is finite.

If preferences on the outcomes of a sequential N-struck procedure are strict (including preferences on lotteries), then the procedure always has a

[28] See the last quote from Shapard and Johnson (1994) in the Introduction.

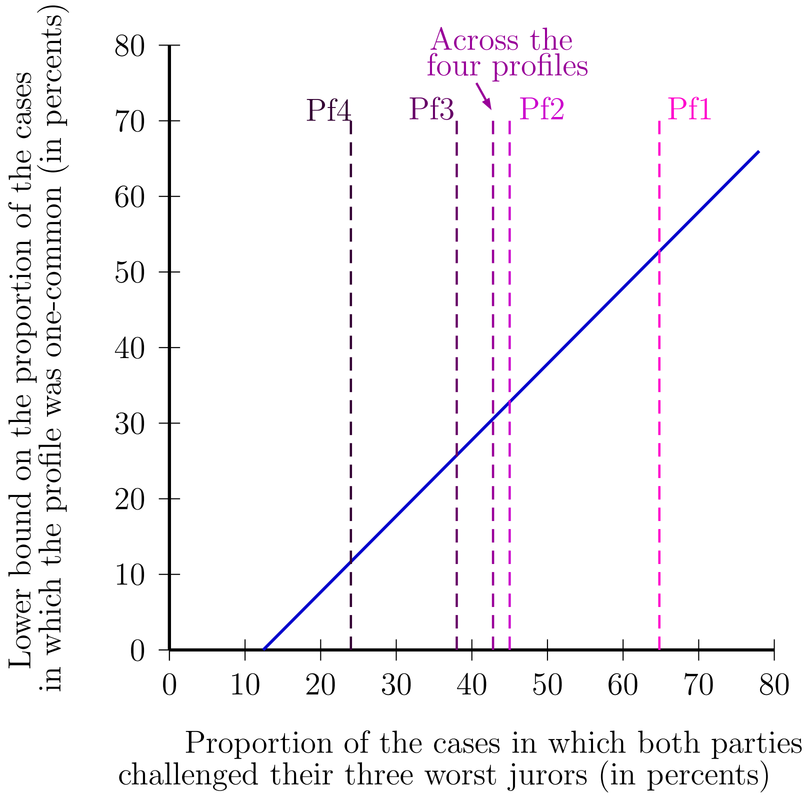


Figure 5: Lower bounds on the proportion of one-common profiles in the 750 arbitration cases (blue line) as a function of the proportion of cases in which both parties were truthful. Estimates for the later proportion using experimental data from de Clippel et al. (2014) are shown by the dashed vertical lines (see Table 2).

finite rationality threshold. This follows from the fact that, with no indifferences on outcomes, sequential N-struck procedures induce games of complete information that can be solved uniquely by backward induction.[29] Then, the number of rounds of backward induction required to solve the game is an upper bound for the rationality threshold.

Proposition 6. *For any sequential N-struck procedures, if preferences on the outcomes of the procedure are strict, then the rationality threshold is finite and smaller than the depth of the game tree.*

Recall that one-shot$_M$ has an infinite rationality threshold when the profile is one-common because each party would like to free ride on her opponent's challenge of one of the jurors they both dislike (see Example 7). This

[29] More precisely, multiple strategy profiles can survive backward induction, but each of these profiles must yield the same outcome.

idea generalizes to the class of simultaneous N-struck procedures as a whole. Below, I identify for each simultaneous N-struck procedure Γ a set of Γ-one-common profiles. In Proposition 7, I show that any Γ-one-common profiles induces an infinite rationality threshold in Γ.

In particular, together with Proposition 6, Proposition 7 implies that, whenever the profile is alternating$_M$-one-common (and preferences on outcomes are strict), the rationality threshold of alternating$_Q$ is lower than the rationality threshold of alternating$_M$. That is, Proposition 3 extends partially to alternating procedures.

Informally, given a simultaneous N-struck procedure Γ, a profile is Γ-*one-common* if, in one of the last subgames of Γ, the set of jurors that remain unchallenged gives rise to the free rider problem described above. Formally, given Γ, a profile is **Γ-one-common** if there exists a subgame γ of Γ such that (a) both parties can still challenge jurors in γ (i.e., $\ell_i^\gamma \geq 1$ for both $i \in \{d, p\}$), (b) the first round of γ is the last round of γ in which both parties can challenge jurors,[30] and (c) among the unchallenged jurors, one of the ℓ_d^γ worst jurors according to R_d is also one of the ℓ_p^γ worst jurors according to R_p.

For example, consider alternating$_M$-one-common profiles. The only subgame of one-shot$_M$ is one-shot$_M$ itself. Hence, in the case of one-shot$_M$, (a), (b) and (c) boil down to requiring that among N, one of the c_d worst jurors according to R_d is also one of the c_p worst jurors according to R_p, which is the definition of a one-common profile. Because the sets of alternating$_M$-one-common and one-common profiles are identical, the next proposition generalizes Proposition 5.

Proposition 7. *For any simultaneous N-struck procedure Γ, if the profile is Γ-one-common, then the rationality threshold of Γ is ∞.*

In general, the Γ-one-common and the one-common conditions are not equivalent. For example, there exist alternating$_M$-one-common profiles that are *not* one-common.

Example 7. Consider alternating$_M$ and any problem in which $c_d = c_p = 2$, $b = 1$, and the preferences on jurors are,

$$\begin{array}{cccccc} R_d: & 1 & 2 & 3 & 4 & 5 \\ R_p: & 2 & 4 & 5 & 1 & 3 \end{array} \tag{6}$$

The profile is not one-common because $\{4, 5\} \cap \{1, 3\} = \emptyset$.

[30] This could arise because the first round of γ is the terminal round of Γ, or because both parties only have one challenge left in γ.

However, consider the subgame γ^* that follows from d challenging juror 4 and p challenging juror 5 in the first round. Subgame γ^* satisfies (a) and (b) in the definition of an alternating$_M$-one-common profile. Also, both players have the same least prefer jurors among $\{1, 2, 3\}$, the set of unchallenged jurors at the beginning of γ^*. Hence, condition (c) in the definition of an alternating$_M$-one-common profile is also satisfied, and profile (6) is alternating$_M$-one-common.

To see why the rationality threshold is infinite in subgame γ^*, observe that in γ^*, each party wants to free-ride on her opponent's challenge of juror 3. This induces an infinite rationality threshold for the same reason that the rationality threshold is infinite in Example 5.

The procedure used to reach subgame γ^* in Example 7 can be generalized. In alternating$_M$, for any set of jurors T containing $b+2$ jurors, there exists a subgame γ satisfying (a) and (b) such that T is the set of unchallenged jurors and each party has one challenge left.[31] Hence, a profile is alternating$_M$-one-common if there exists a set T containing $b + 2$ jurors such that R_p and R_d have the same worst juror among the jurors in T.

This sufficient condition can be used to prove the following result.

Proposition 8. *Every one-common profile is alternating$_M$-one-common.*

Hence, the proportion of alternating$_M$-one-common profiles is no smaller than the proportion of one-common profiles, and the proportions in Figure 4 are therefore lower bounds for the proportions of alternating$_M$-one-common profiles.[32] The lower bounds on the proportions of one-common profiles in the arbitration data in Figure 5 are also lower bounds on the proportion of alternating$_M$-one-common profiles in the same data.

7 Conclusion

This paper shows how jury selection procedures can be compared in terms of their strategic complexity by computing their *rationality thresholds*, i.e., the number of rounds of elimination of strategies that are not best responses required for the parties to have a dominant strategy. The results in this paper notably show that procedures in which challenges are made sequentially tend to be strategically simpler than procedures in which challenges are simultaneous.

[31] Such a subgame is reached, for example, after d alone challenges jurors in $N\backslash T$ for $c_d - 1$, followed by p alone challenging remaining jurors among $N\backslash T$ for $c_p - 1$ rounds.

[32] In fact, together with Example 7, Proposition 8 implies that the proportion of alternating$_M$-one-common profiles is *strictly* larger than that of one-common profiles.

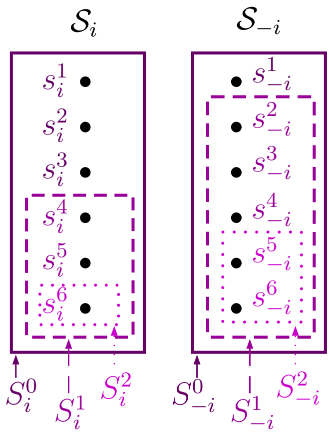


Figure 6: An arbitrary hierarchical model for a given some preference profile (R_d, R_p).

The rationality threshold offers a new method to compare the strategic complexity of mechanisms. Unlike previous methods in the literature (Pathak and Sönmez, 2013; de Clippel et al., 2014; Arribillaga and Massó, 2015), it allows for comparisons even when the mechanisms at stake are indirect or induce games of incomplete information.

More generally, the rationality threshold shows how hierarchical models can be used to compare the strategic complexity of mechanisms. As illustrated in Figure 6, for any profile (R_d, R_p), a **hierarchical model** specifies a pair $(\{S_d^0, \ldots, S_d^m\}, \{S_p^0, \ldots, S_p^m\})$ of nested strategy sets, i.e., $\mathcal{S}_i \subseteq S_i^0 \subseteq \cdots \subseteq S_i^m$ (m could be infinite). As k increases, the sets S_i^k represent increasingly restrictive models of the strategies that i could potentially play.

This paper studies the (conditional) *level-k* hierarchical model $(\{CL_d^0, \ldots, CL_d^m\}, \{CL_p^0, \ldots, CL_p^m\})$ defined in Section 4. Given a profile (R_d, R_p), I define the rationality threshold as the smallest hierarchical level r^* for which each party i has a strategy $s_i^* \in \mathcal{S}_i$ that is a best response to *every* strategy in $CL_{-i}^{r^*-1}$. I then use the rationality threshold as a measure of strategy complexity.

Clearly, this logic is not specific to the *level-k* hierarchical model. A natural alternative would be to use the "*undominated*" hierarchical model UD defined by the process of iterated undominated strategies. One could then define an alternative UD-rationality threshold,[33] and perform analysis similar to those of this paper.

[33] I.e., the smallest hierarchical level r_{UD}^* for which each party i has a strategy $s_i^* \in \mathcal{S}_i$ that is a best response to *every* strategy of her opponent that survives $r_{UD}^* - 1$ rounds of iterated undominated strategies.

Another option is to use a hierarchical model in which the level of sophistication of the parties is fixed, say to level-1, and the parties' *information* about each others' preferences is refined in each iteration of the model. The corresponding I-rationality threshold could, for example, be the smallest r_I^* such that each party i has a strategy $s_i^* \in \mathcal{S}_i$ that is a best response to *every* CL_{-i}^1 strategy for *any* preference $\widetilde{R}_{-i}$ that has the same r_I^* worst jurors as R_{-i}. That is, i only needs to know the r_I^* worst jurors according to R_{-i} and the fact that $-i$ is a best responder in order to have a dominant strategy.[34]

In general, there is no logical relation between the UD-rationality threshold and the level-k rationality threshold of a game, henceforth "CL-rationality threshold". However, some of the results in this paper extend to the UD-rationality threshold.

First, it is not hard to see that the UD-rationality threshold of one-shot$_Q$ is 2. Second, it can be shown that for every problem, the UD-rationality threshold of one-shot$_M$ is at least as large as the CL-rationality threshold of one-shot$_M$.[35] Hence, the results in Sections 5.2 extend to the UD-rationality threshold, and the UD-rationality threshold of one-shot$_M$ is larger than 2 for a significant set of problems.Whether the results of Section 6 also extend to the UD-rationality threshold is left as an open question.

Acknowledgments

I am grateful to Benoit Decerf, Hanna Frank, Greg Leo, Clayton Masterman, Claudia Rei, Roberto Serrano, Yves Sprumont, Rodrigo Velez and John Weymark for helpful discussions and comments. I also want to thank Paul Edelman, Eun Jeong Heo, Myrna Wooders as well as the participants to presentations at Université Catholique de Louvain, Vanderbilt University, the 13th meeting of the Society for Social Choice and Welfare and the fifth world congress of the Game Theory Society for their questions and suggestions. I thank CORE and CEREC for their hospitality during visits. Support from the Kirk Dornbush summer research grant and the National Science Foundation grant IIS-1526860 is gratefully acknowledged.

[34] Observe that $r_I^* = 1$ in one-shot$_Q$ for the problem described in Example 4.

[35] This follows from the fact that, by the assumption that preferences on jurors are strict, best responses are unique in one-shot$_M$. Thus, every best response is also an undominated strategy, and the set of strategies that survive k rounds of iterated conditional best response is a *subset* of the set of strategies that survive k rounds of iterated undominated strategies.

Appendix

A Omitted proofs

For any (separable) preferences (R_d, R_p) and any integer z, let $R(z)$ be the z-th juror according to (R_d, R_p)'s ranking of jurors.

Proof of Proposition 1.

Consider any problem. (i). Let $i \in \{d, p\}$ be the party who challenges jurors first and let w be i's worst juror. Let s^w_{-i} be the strategy in which $-i$ challenges only w (provided i did not already challenge w), and $-i$ challenges no other juror otherwise. Also, let s^0_{-i} be the strategy in which $-i$ does not challenge any jurors. Because i's preference on jurors is strict, i's unique best response $t_i(s^w_{-i})$ consists in challenging her c_i worst jurors among $N \backslash \{w\}$. Differently, party i's unique best response $t_i(s^0_{-i})$ consists in i challenging her c_i worst jurors among N, which includes challenging juror w. Thus, no strategy of i is a best response to both s^w_{-i} and s^0_{-i}, and i does no have a dominant strategy.

(ii). The proof of (ii) is similar to the proof of (i). Let $\tilde{s}^w_{-i}$ be the strategy in which, in each round, $-i$ challenges as many as possible of i's worst jurors among the jurors that remain unchallenged. Also, let $\tilde{s}^0_{-i}$ be the strategy in which $-i$ never challenges any juror.

In any best response $t_i(\tilde{s}^w_{-i})$, i never challenges any of her c_{-i} worst jurors and challenges her c_i other worst jurors over the course of the procedure. Differently, in any best response $t_i(\tilde{s}^0_{-i})$, i challenges her c_i worst jurors among N over the course of the procedure, which includes challenging some of her c_{-i} worst jurors. Thus, no strategy of i is a best response to both $\tilde{s}^w_{-i}$ and $\tilde{s}^0_{-i}$, and i does no have a dominant strategy.

Proof of Proposition 2. Let $\mathcal{D}$ be the set of separable preferences on $\Delta\mathcal{J}$. In order to derive a contradiction, assume that there exists a challenge procedure Γ satisfying finiteness and minimal challenge such that, for every profile $(R_d, R_p) \in \mathcal{D} \times \mathcal{D}$, both parties have a dominant strategy.

Let $s^*_i(R_i)$ be one of i's dominant strategies when i's preference are R_i. Consider the direct mechanism $M^\Gamma : \mathcal{D} \times \mathcal{D} \to \Delta\mathcal{J}$ constructed from Γ by setting

$$M(R_d, R_p) \coloneqq \Gamma(s^*_d(R_d), s^*_p(R_p)), \tag{7}$$

where $\Gamma(s^*_d(R_d), s^*(R_p))$ is the lottery that obtains when $(s^*_d(R_d), s^*(R_p))$ is played in Γ.

Because $s_i^*(R_i)$ is a dominant strategy given R_i, for all $i \in \{d,p\}$ and all $R_i \in \mathcal{D}$,

$$\Gamma(s_i^*(R_i), s_{-i}^*(R_{-i}))\ R_i\ \Gamma(s_i^*(R_i'), s_{-i}^*(R_{-i})) \qquad \text{for all } R_i' \in \mathcal{D} \text{ and } R_{-i} \in \mathcal{D}.$$

But then, by construction of M, for all $i \in \{d,p\}$ and all $R_i \in \mathcal{D}$,

$$M(R_i, R_{-i})\ R_i\ M(R_i', R_{-i}) \qquad \text{for all } R_i' \in \mathcal{D} \text{ and } R_{-i} \in \mathcal{D},$$

and M is strategy-proof on the domain of separable profiles.

Notice that the domain of additive profiles $\mathcal{D}^{add} \times \mathcal{D}^{add}$ is a subset of the domain of separable profiles.[36] By Van der Linden (2015, Example 3), any domain of profiles that contains the domain of additive profiles is a *negative leximin* domain (see Van der Linden, 2015, Domain Property 3). But then, M contradicts Van der Linden (2015, Corollary 3) which states that, on a negative leximin domain, no mechanisms constructed from a procedure satisfying finiteness and minimal challenge as in (7) is strategy-proof.

As stated in the body of the paper, this proof shows that Proposition 2 is true on any *negative leximin* domain of profiles, which includes the domain of additive profiles.

Proof of Proposition 4. Consider any problem. Let i be the first party to challenge. For $-i$, the unique best response s_{-i}^* to *any* strategy by i is to challenge her c_{-i} worst jurors among the jurors that i did not challenge (uniqueness follows again from preferences on jurors being strict). This strategy is, therefore, the unique CL_{-i}^1. It directly follows from uniqueness that s_{-i}^* is a best response to all CL_i^1 strategies.

Because there is a unique CL_{-i}^1 strategy s_{-i}^*, any CL_i^2 strategy that best responds to s_{-i}^* is a best response to *all* CL_{-i}^1 strategies. Hence, the rationality threshold of one-shot$_Q$ is at most 2 for this problem. But by Proposition 1, because one-shot$_M$ is an N-struck procedure, the rationality threshold of one-shot$_Q$ is at least 2 for every problem. Thus, the rationality threshold of one-shot$_Q$ is 2 for this problem.

Proof of Proposition 5. The proof generalizes Example 5 and is illustrated in Figure 7. Consider any problem in which the profile is one-common. Let $w \in N$ be the juror that is among the c_d worst jurors of d and among the c_p worst jurors of p. For any set $\widetilde{N} \subseteq N$ containing at least c_i jurors, let $t_i(\widetilde{N})$ be the set of the c_i worst jurors in $\widetilde{N}$ according to R_i.

Induction basis. For each $i \in \{d,p\}$, both challenging jurors $t_i(N)$ and challenging jurors $t_i(N \backslash \{w\})$ are CL_i^0 strategies of one-shot$_M$. Observe that $w \in t_i(N)$ and $w \notin t_i(N \backslash \{w\})$.

[36] Preferences R_i are additive if there exists $u_i : N \to \mathbb{R}$ such that, for all $L, L' \in \mathcal{J}$, $L\ R_i\ L' \Leftrightarrow \sum_{J \in \mathcal{J}} L(J) \sum_{t \in J} u_i(t) \geq \sum_{J \in \mathcal{J}} L'(J) \sum_{t \in J} u_i(t)$.

	CL^0_{-i} strategies	CL^1_i strategies	CL^2_{-i} strategies	CL^3_i strategies
Set of jurors challenged	$t_{-i}(N)$ $\ni w$	$t_i(t_{-i}(N))$ $\not\ni w$	$t_{-i}(t_i(t_{-i}(N)))$ $\ni w$	$\dots$
	$t_{-i}(N\backslash\{w\})$ $\not\ni w$	$t_i(t_{-i}(N\backslash\{w\}))$ $\ni w$	$t_{-i}(t_i(t_{-i}(N\backslash\{w\})))$ $\not\ni w$	$\dots$

Figure 7: For any one-common profile of preferences, the rationality threshold of one-shot$_M$ is ∞.

Induction step. For any $k \in \mathbb{N}$, suppose that for each $i \in \{d, p\}$, the set of CL_i^k strategies of one-shot$_M$ contains a strategy s_i^k in which w is challenged, and a strategy $\tilde{s}_i^k$ in which w is *not* challenged. Then, for each $i \in \{d, p\}$, we have $w \notin t_i(s_{-i}^k)$ and $w \in t_i(\tilde{s}_{-i}^k)$. Hence, for each $i \in \{d, p\}$, the set of CL_i^{k+1} strategies of one-shot$_M$ contains a strategy $s_i^{k+1} = t_i(\tilde{s}_{-i}^k)$ in which w is challenged, and a strategy $\tilde{s}_i^{k+1} = t_i(s_{-i}^k)$ in which w is *not* challenged.

By the induction step, s_i^k and $\tilde{s}_i^k$ are well-defined for each $i \in \{d, p\}$ and for all $k \in \mathbb{N}$. But observe that, for every $k \in \mathbb{N}$, $w \notin t_i(s_{-i}^k)$ and $w \in t_i(\tilde{s}_{-i}^k)$ which implies that $t_i(s_{-i}^k) \neq t_i(\tilde{s}_{-i}^k)$. Because $t_i(s_{-i}^k)$ and $t_i(\tilde{s}_{-i}^k)$ are unique best responses, no strategy of i is a best response to every CL_{-i}^{k-1} strategy of one-shot$_M$. Hence, the rationality threshold is at least $k+1$. Because this is true for all $k \in \mathbb{N}$, this concludes the proof.

Proof of Proposition 6. Let Γ be any sequential N-struck procedure. Let $\bar{e}$ be the depth of the game tree. It is convenient to locate subgames of Γ in terms of the height of their root node, where the height of a node is the length of the longest path from that node to a leaf node. A subgame of Γ the root node of which has height h is denoted γ^h.

Induction basis. Consider any subgame γ^1 of Γ. Because preferences on outcomes are strict, (a) the outcome of γ^1 is the same under any level-1 profile, and (b) the rationality threshold of γ^1 is 1.

Induction step. For any $h \in \{1, \dots, \bar{e}\}$, suppose that for any subgame γ^{h-1} of Γ, (a') the outcome of γ^{h-1} is the same under any level-$(h-1)$ profile, and (b') the rationality threshold of γ^{h-1} is no larger than $h-1$.

Consider any subgame γ^h of Γ. Let i be party who moves at the root node of Γ. By nestedness, any level-h profile is a level-$(h-1)$ profile. Hence, by (a'), for any subgame γ^{h-1} directly following γ^h, the outcome of γ^{h-1} is the same under any level-h profile. Thus, if the outcomes of γ^h differ under two level-h profiles, it must be that i's action at the root node of γ^h lead to two subgames $\widehat{\gamma}^{h-1}$ and $\widetilde{\gamma}^{h-1}$ with different level-$(h-1)$ outcomes. But then, because preferences on outcomes are strict, i's action at the root node of γ^h cannot be part of a level-h strategy. Therefore, (a") the outcome of γ^h must be the same under any level-h profile.

It follows from (a") that for each $i \in \{d, p\}$, any best response to a CL_{-i}^{h-1} strategy is a best response to all CL_{-i}^{h-1} strategies. Hence, (b") the rationality threshold of γ^h is no larger than h.

Proof of Proposition 7. For any simultaneous N-struck procedure Γ, consider the subgame γ described in the definition of a Γ-one-common profile. Let $j \in N$ be (one of) the unchallenged juror(s) in γ that is among the ℓ_d^γ worst jurors according to R_d and among the ℓ_p^γ worst jurors according to R_p. The proof is then a straightforward adaptation of the proof of Proposition 5 with c_i replaced by ℓ_i^γ and the challenges described in the proof of Proposition 5 occurring in the first round of γ.

Proof of Proposition 8. Consider an arbitrary one-common profile (R_d, R_p). Let $N_i^1 := \{R_i(1), \ldots, R_i(b + c_{-i})\}$ and $N_i^2 := \{R_i(b + c_{-i}+1), \ldots, R_i(n)\}$. By assumption, there exits a juror $w \in N$ such that $w \in N_d^2 \cap N_p^2$. Observe that $\#N_i^1 = b+c_{-i}$ and $\#N_i^2 = c_i$, where for any set T, $\#T$ is the cardinality of T.

Because $w \in N_d^2 \cap N_p^2$, we have $w \notin N_d^1 \cap N_p^1$. Hence, if $\#(N_d^1 \cap N_p^1) \geq b+1$, then $\#(N_d^1 \cap N_p^1) \cup \{j\} \geq b + 2$ and w is the worst juror for both R_d and R_p among $(N_d^1 \cap N_p^1) \cup \{j\}$, making (R_d, R_p) an alternating$_M$ profile by the sufficient condition identified in the text.

Hence, in order to derive a contradiction, suppose that $\#(N_d^1 \cap N_p^1) \leq b$. In words, for some $i \in \{d, p\}$, *at most* b of the jurors in N_i^1 belong to N_{-i}^1. Because N_{-i}^1 and N_{-i}^2 partition N, all the jurors in N_i^1 that do not belong to N_{-i}^1 must belong to N_{-i}^2. Hence, because there are $b + c_{-i}$ jurors in N_i^1 at most b of which belong to N_{-i}^1, *at least* c_{-i} of the jurors in N_i^1 must belong to N_{-i}^2. Recall that $w \in N_i^2$ by assumption. Thus, because N_i^1 and N_i^2 partition N, $w \notin N_i^1$, and w cannot be one of the c_{-i} jurors in N_i^1 that belong to N_{-i}^2. But this implies that there are at least $c_{-i} + 1$ jurors in N_{-i}^2, a contradiction.

B Computing the proportions in Figure 4

Because the one-common property is preserved under relabeling of the jurors, let us, without loss of generality, fix

$$b + 2c \; R_{-i} \; b + 2c - 1 \; R_{-i} \; \ldots \; R_{-i} \; 2 \; R_{-i} \; 1.$$

Also, for any R_i, let the c worst jurors according to R_i be $\{j_1^{R_i}, j_2^{R_i}, \ldots, j_c^{R_i}\}$ with $j_1^{R_i} \; R_i \; \ldots \; R_i \; j_c^{R_i}$.

Figure 4a). The proportion of one-common profiles is then equal to the proportion of R_i that induce a one-common profile given the fix R_{-i}. Clearly, this proportion can be treated as a probability, where the set of outcomes is the set of preferences R_i, and the probability of drawing any particular R_i is $\frac{1}{(b+2c)!}$.

We are interested in the probability that one of the c worst juror of i is among the c worst jurors of $-i$, i.e.,

$$\mathbb{P}\big(j_c^{R_i} \in \{1,\dots,c\} \cup \dots \cup j_1^{R_i} \in \{1,\dots,c\}\big),$$

This probability is equal to

$$\begin{aligned}
&\mathbb{P}\big(j_c^{R_i} \in \{1,\dots,c\}\big)+\\
&\mathbb{P}\big(j_{c-1}^{R_i} \in \{1,\dots,c\} \cap j_c^{R_i} \notin \{1,\dots,c\}\big)+\\
&\vdots\\
&\mathbb{P}\big(j_1^{R_i} \in \{1,\dots,c\} \cap j_c^{R_i} \notin \{1,\dots,c\} \cap \dots \cap j_2^{R_i} \notin \{1,\dots,c\}\big).
\end{aligned} \tag{8}$$

The elements of the sum in (8) can be computed recursively. In step $r=1$, we initiate the recursion by computing $\mathbb{P}\big(j_c^{R_i} \in \{1,\dots,c\}\big)$. To do so, observe that

$$\mathbb{P}\big(j_c^{R_i} \in \{1,\dots,c\}\big) = \mathbb{P}\big(j_c^{R_i} = 1 \cup \dots \cup j_c^{R_i} = c\big).$$

Hence, because $j_c^{R_i} = 1, \dots, j_c^{R_i} = c$ are mutually exclusive, we have

$$\mathbb{P}\big(j_c^{R_i} \in \{1,\dots,c\}\big) = \sum_{i=1}^{c} \mathbb{P}\big(j_c^{R_i} = i\big) = \frac{c}{b+2c}.$$

Then, in each step $r \in \{2,\dots,c\}$, Bayes' rule implies that

$$\begin{aligned}
&\mathbb{P}\big(j_r^{R_i} \in \{1,\dots,c\} \cap j_c^{R_i} \notin \{1,\dots,c\} \cap \dots \cap j_{r-1}^{R_i} \notin \{1,\dots,c\}\big)\\
&= \mathbb{P}\big(j_r^{R_i} \in \{1,\dots,c\} \big| j_c^{R_i} \notin \{1,\dots,c\} \cap \dots \cap j_{r-1}^{R_i} \notin \{1,\dots,c\}\big)\\
&\qquad \mathbb{P}\big(j_c^{R_i} \notin \{1,\dots,c\} \cap \dots \cap j_{r-1}^{R_i} \notin \{1,\dots,c\}\big).
\end{aligned} \tag{9}$$

For all $r \in \{2,\dots,c\}$, the first term on the right-hand side of (9)

$$\begin{aligned}
&\mathbb{P}\big(j_r^{R_i} \in \{1,\dots,c\} \big| j_c^{R_i} \notin \{1,\dots,c\} \cap \dots \cap j_{r-1}^{R_i} \notin \{1,\dots,c\}\big)\\
&\qquad = \frac{c}{b+2c-(r-1)}.
\end{aligned} \tag{10}$$

The second term on the right-hand side of (9)

$$\begin{aligned}
&\mathbb{P}\big(j_c^{R_i} \notin \{1,\dots,c\} \cap \dots \cap j_{r-1}^{R_i} \notin \{1,\dots,c\}\big)\\
&= 1 - \mathbb{P}\big(j_c^{R_i} \in \{1,\dots,c\} \cup \dots \cup j_{r-1}^{R_i} \in \{1,\dots,c\}\big),
\end{aligned}$$

and

$$\begin{aligned}&\mathbb{P}\big(j_c^{R_i} \in \{1,\ldots,c\} \cup \cdots \cup j_{r-1}^{R_i} \in \{1,\ldots,c\}\big)\\&= \sum_{i=1}^{r-1} \mathbb{P}\big(j_i^{R_i} \in \{1,\ldots,c\} \cap j_c^{R_i} \notin \{1,\ldots,c\} \cap \cdots \cap j_i^{R_i} \notin \{1,\ldots,c\}\big),\end{aligned} \tag{11}$$

where the terms inside the sum on the right-hand side of (11) have been obtained in steps $\{1,\ldots,r-1\}$.

Figure 4b). Again, computing the proportion in Figure 4b) is equivalent to computing the probability that (R_i, R_{-i}) is one-common when R_i is drawn uniformly among the preferences that make (R_i, R_{-i}) almost juror inverse. That is, R_i is drawn uniformly at random among the preferences that differs from R_i^* with $1\ R_i^*\ \ldots\ R_i^*\ b+2c$ by the ranking of a single juror, say $\bar{j}^{R_i}$. For (R_i, R_{-i}) to be one-common, $\bar{j}^{R_i}$ must be one of the c worst jurors in R_{-i} *and* in R_i. That is, the relevant probability is

$$\mathbb{P}\big(\bar{j}^{R_i} \in \{1,\ldots,c\} \cap \bar{j}^{R_i} \in \{j_1^{R_i},\ldots,j_c^{R_i}\}\big).$$

Applying Bayes' rule, this is equal to

$$\mathbb{P}\big(\bar{j}^{R_i} \in \{j_1^{R_i},\ldots,j_c^{R_i}\} \big| \bar{j}^{R_i} \in \{1,\ldots,c\}\big) P\big(\bar{j}^{R_i} \in \{1,\ldots,c\}\big).$$

Because R_i is drawn at random among the preferences that differ from R_i^* by the ranking of a single juror, $P(\bar{j}^{R_i} \in \{1,\ldots,c\}) = \frac{c}{b+2c}$. Also,

$$\mathbb{P}(\bar{j}^{R_i} \in \{j_1^{R_i},\ldots,j_c^{R_i}\} \big| \bar{j}^{R_i} \in \{1,\ldots,c\}) = \frac{c}{b+2c-1},$$

where $b+2c-1$ is the number of (equally likely) positions in which $\bar{j}$ could be ranked,[37] and c is the number of these positions that induce a one-common ranking given $\bar{j}^{R_i} \in \{1,\ldots,c\}$). [38] Hence,

$$\mathbb{P}\big(\bar{j}^{R_i} \in \{1,\ldots,c\} \cap \bar{j}^{R_i} \in \{j_1^{R_i},\ldots,j_c^{R_i}\}\big) = \frac{c^2}{(b+2c)(b+2c-1)}.$$

[37] By definition of an almost juror inverse profile, $R_i \neq R_i^*$, Because $\bar{j}$ is the only juror the ranking of which changes between R_i and R_i^*, this implies that $\bar{j}$ must be ranked differently in R_i and R_i^*.

[38] Indeed, observe that given $\bar{j}^{R_i} \in \{1,\ldots,c\}$, because (R_i^*, R_{-i}) is juror inverse, $\bar{j}^{R_i} \notin \{j_1^{R_i^*},\ldots,j_c^{R_i^*}\}$.

References

Alpern, S., Gal, S., 2009. Analysis and design of selection committees: A game theoretic secretary problem. International Journal of Game Theory 38, 377–394.

Alpern, S., Gal, S., Solan, E., 2010. A sequential selection game with vetoes. Games and Economic Behavior 68, 1–14.

Arribillaga, R.P., Massó, J., 2015. Comparing generalized median voter schemes according to their manipulability. Theoretical Economics 11, 547–586.

Bermant, G., 1982. Jury Selection Procedures in United States District Courts. volume 25. Federal Judicial Center, Washington D.C.

Bermant, G., Shapard, J., 1981. The voir dire examination, juror challenges, and adversary advocacy, in: Sales, B.D. (Ed.), The Trial Process. Springer, Berlin, pp. 69–114.

Bloom, D., Cavanagh, C.L., 1986. An analysis of the selection of arbitrators. American Economic Review 76, 408–422.

Brams, S.J., Davis, M.D., 1978. Optimal jury selection: A game-theoretic model for the exercise of peremptory challenges. Operations Research 26, 966–991.

Broderick, R.J., 1992. Why the peremptory challenge should be abolished. Temple Law Review 65, 369.

Caditz, D.M., 2015. Selection under veto with limited foresight. Mimeo .

Cohen, N.P., Cohen, D.R., 2003. Jury reform in Tennessee. University of Memphis Law Review 34, 1–71.

Crawford, V.P., Costa-Gomes, M.A., Iriberri, N., 2013. Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications. Journal of Economic Literature 51, 5–62.

Daly, M., 2016. Foster v. Chatman: Clarifying the Batson Test for discriminatory peremptory strikes. Duke Journal of Constitutional Law and Public Policy Sidebar 11, 148–162.

de Clippel, G., Eliaz, K., Knight, B., 2014. On the selection of arbitrators. American Economic Review 104, 3434–3458.

DeGroot, M.H., Kadane, J.B., 1980. Optimal challenges for selection. Operations Research 28, 952–968.

Flanagan, F.X., 2015. Peremptory challenges and jury selection. Journal of Law and Economics 58, 385–416.

Fudenberg, D., Tirole, J., 1991. Game Theory. MIT Press, Cambridge MA.

Ho, T.H., Camerer, C., Weigelt, K., 1998. Iterated dominance and iterated best response in experimental “p-beauty contests”. American Economic Review , 947–969.

Horwitz, B.L., 1992. Extinction of the peremptory challenge: what will the jury system lose by its demise. U. Cin. L. Rev. 61, 1391–1440.
Kadane, J.B., Stone, C.A., Wallstrom, G., 1999. The donation paradox for peremptory challenges. Theory and Decision 47, 139–155.
Keene, D.L., 2009. Fairness, justice and true understanding: The benefits of peremptory strikes. The Jury Expert 2, 24–25.
LaFave, W., Israel, J., King, N., Kerr, O., 2009. Criminal Procedure. West Academic Publishing, St. Paul, MN. 5th edition.
Marder, N.S., 2012. Batson Revisited - Batson Symposium. SSRN Scholarly Paper ID 2165561. Social Science Research Network. Rochester, NY.
Mas-Colell, A., Green, J.R., Whinston, M.D., 1995. Microeconomic Theory. Oxford University Press, Oxford.
Pathak, P.A., Sönmez, T., 2013. School admissions reform in Chicago and England : Comparing mechanisms by their vulnerability to manipulation. American Economic Review 103, 80–106.
Roth, A., Kadane, J.B., Degroot, M.H., 1977. Optimal peremptory challenges in trials by juries: A bilateral sequential process. Operations Research 25, 901–919.
Shapard, J., Johnson, M., 1994. Memorandom on a Survey of Active Jurdges Regarding their Voir Dire Practices. Technical Report. Federal Judicial Center, Research Division. Washington D.C.
Smith, A., 2014. A call to abolish peremptory challenges by prosecutors. Georgetown Journal of Legal Ethics 27, 1163–1186.
Van der Linden, M., 2015. Impossibilities for strategy-proof veto in selection problems. Working paper, Department of Economics, Vanderbilt University .

Cases

Batson v. Kentucky, 476 U.S. 79 (1986).
Swain v. Alabama, 380 U.S. 202 (1965).