



Swiss Federal Institute of Technology Zurich

Seminar for
Statistics

Department of Mathematics

Master Thesis

Fall 2018

Christopher Salahub

Seen to be done

A graphical investigation of peremptory challenge

Submission Date: March 3 2019

Co-Adviser:

Adviser: Prof. Dr. Marloes Maathuis

Preface

This work would be nowhere near as polished or complete without the effort of Prof. Dr. Marloes Maathuis to ensure I was performing analysis with a clear direction and purpose. I would like to thank her for finding time in her busy schedule to allow for weekly meetings. The group meetings organized by her Ph.D. student Marco Eigenmann were also critical in the development of more nuanced analysis and intuitive visualizations. I thank Marco Eigenmann for organizing them, and Jinzhou Li, Armin Fingerle, Sanzio Monti, and Qikun Xiang for attending my presentations and listening attentively. A special thanks is extended to Cédric Bleutler and Leonard Henckel, both of whom were especially engaged and participated in lengthy discussions both during and outside of the group meetings.

I would like to acknowledge in particular Prof. Dr. Tilman Altwicker for his detailed suggestions on where to look for more legal context on the topic and Prof. Dr. Samuel Baumgartner for his research suggestions. They were very important at providing the necessary information to begin a first investigation of the topic. Of course, without the cooperation of Dr. Ronald Wright, Dr. George Woodworth, Dr. Barbara O'Brien, and Dr. Catherine Grosso for generously providing me with the data from their investigations on the subject of preemptory challenge. Without this data, the visualizations and modelling presented here simply would not have been possible, and so I am exceptionally grateful that they were so enthusiastic to share the fruits of their labour to help cultivate mine.

Abstract

Short summary of my thesis.

Contents

0.1	Terms	xi
0.2	Variables	xi
	Notation	xi
1	Introduction	1
2	Peremptory Challenges	3
2.1	Jury Selection Procedures	3
2.2	The Role of the Jury	5
2.3	Modern Peremptory Challenge Controversy	5
2.4	The Role of the Peremptory Challenge	7
2.5	History	7
2.5.1	Pre-English History	8
2.5.2	In English Law (1066–1988)	8
2.5.3	In American Law (ca. 1700–1986)	9
2.5.4	In Canadian Law (ca 1800–2018)	10
2.6	Summary	11
3	Data	13
3.1	Jury Sunshine Project	13
3.1.1	Methodology	13
3.1.2	Cleaning	14
3.1.3	Variable Synthesis	18
3.2	Stubborn Legacy Data	18
3.2.1	Methodology	18
3.2.2	Cleaning	19
3.3	Philadelphia Data	19
3.3.1	Methodology	19
3.3.2	Cleaning	20
4	Empirical Analysis	21
4.1	Extremes of Partiality	21
4.2	The Impact of Race	23
4.2.1	In the Stubborn and Philadelphia Data	26
4.3	Other Factors	27
4.3.1	Political Affiliation in the Sunshine Data	28
4.3.2	Gender in the Sunshine Data	30
4.3.3	The Stubborn and Philadelphia Data Sets	32
4.4	Modelling	33
4.4.1	Multinomial Logistic Regression	33
4.5	Case Level Summary	34
4.6	Useful Quoted Bits	35
5	Summary	37
5.1	Future Work	37
	Bibliography	39

A	Developing an Effective Visualization of Conditional Probability	43
A.1	The Mobile Plot	45
B	Complementary information	47
B.1	Prosecution and Judge Homogeneity	47
B.2	Jury Sunshine Irregularities	48
B.3	Jury Sunshine Charge Classification	48
C	Mathematical Results	51
C.1	Conditional Distribution of a Poisson Expectation Given Marginal Counts .	51
D	Code	53
D.1	Data Processing Code	53
D.2	Analysis Code	53
	Epilogue	55

List of Figures

3.1	Simple Charge Tree Example	17
4.1	The “Mobile Plot” of Racial Combination and Strikes (Sunshine)	24
4.2	Racial Combination and Strikes with Confidence Intervals (Sunshine)	25
4.3	Strikes by Race and Defendant Race (All Capital Trial Data)	27
4.4	Political Affiliation by Race and Gender (Sunshine)	28
4.5	Strikes by Political Affiliation, Race, and Defendant Race (Sunshine)	29
4.6	Strike Source by Race and Gender (Sunshine)	30
4.7	Strike Source by Race and Gender (Sunshine)	31
4.8	Strikes by Race and Gender of Venire Member (All Capital Trial Data)	32
A.1	Mosaic Plot of Defendant and Venire Member Race	44
A.2	First Parallel Coordinate Attempt	45
B.1	Lawyer Experience (Sunshine)	47
B.2	Regular expression charge tree visualized	49

List of Tables

4.1	Implied Rejection Boundaries	22
4.2	Strike Rate by Race	23
4.3	Model Comparisons to Determine the Importance of Race	34
B.1	Jury Sunshine Irregularities	48

Notation and Terms

0.1 Terms

In order to facilitate clarity despite brevity, a list of terms used in this paper is presented here.

Prosecution/State The legal representation which argues for conviction

Defence The legal representation which argues against conviction

Court Reference to the judge, prosecution, and defence

Venire The population sample from which a jury is selected (according to [Mirriam-Webster \(2019a\)](#) derived from the latin *venire facias*: “may you cause to come”)

Jury The final group of (usually) twelve chosen venire members which judge the guilt or innocence of the accused/defendant

Accused/Defendant The individual on trial for a crime

Voir dire From old French “to speak the truth” (see [Mirriam-Webster \(2019b\)](#)), this is the questioning process used by the court to assess the suitability of a venire member to sit on the jury

Struck In the context of a venire member being rejected from the jury, struck indicates removal by peremptory challenge or challenge with cause

Litigants The accuser and the accused

Disposition The outcome of a venire member in the jury selection process: either kept, struck with cause, struck by prosecution, or struck by defence

0.2 Variables

Across data sets and analyses, the variable names and mathematical notation will be as follows. Note that the use of a capital letter indicates a random variable and a lowercase letter a particular realization of a random variable.

- $\mathbf{x}_i = (r_i, e_i, p_i, g_i, s_i)^T$: the variable combination for a particular venire member
- $d \in \{1, 2, 3, 4\}$: indicator of disposition, with the levels corresponding to kept, struck with cause, struck by defence, and struck by prosecution respectively

- $r \in \{1, 2, 3\}$: indicator of venire member race, with the levels corresponding to black, other, and white respectively
- $e \in \{1, 2, 3\}$: indicator of defendant race, with levels as for the venire member race
- $p \in \{1, 2, 3, 4\}$: indicator of venire member political affiliation, with levels Democrat, Independent, Libertarian, and Republican respectively
- $g \in \{1, 2\}$: indicator of venire member gender, with levels female and male respectively
- $s \in \{1, 2\}$: indicator of defendant gender, with levels as for the venire member
- $\pi_{i|jklmn} \in [0, 1]$: the probability of disposition i given factor levels $jklmn$, may be written as π_i for convenience
- $y_{ijklmn} \in \mathbb{N}$: the count of venire members with $\mathbf{x} = (i, j, k, l, m, n)^T$

This work also uses hat notation for estimates (i.e. the estimate for π is $\hat{\pi}$ and the estimator for π is $\tilde{\pi}$).

Chapter 1

Introduction

The Gerald Stanley murder trial was noteworthy for all of the wrong reasons. The first reason was the crime itself. The rural region around Biggar, Saskatchewan ([Quenneville \(2018\)](#)) is not known for crime, indeed, the crime statistics collected by Statistics Canada suggest it is one of the safest in the province ([Statistics Canada \(2018\)](#)). Any murder at all would be worthy of attention and subject to plenty of drama. But beyond the damage this trial has done to the community, this trial is noteworthy because it led to a significant re-examination of the legal jurisprudence surrounding the jury selection process culminating in the proposition of Bill C-75 by the Canadian government in March of 2018 ([42nd Parliament of Canada \(2018a\)](#)), less than two months after the trial’s verdict ([Quenneville and Warick \(2018\)](#)).

Bill C-75, in part, aims to ameliorate one of the critical points of contention about the Gerald Stanley case: the use of peremptory challenges in jury selection. The outsized impact of the case was due, in large part, to its racial aspect. Gerald Stanley, a white man, was accused of second degree murder in the killing of Colten Boushie, a First Nations man. Given Canada’s troubled history with First Nations groups, this alone would have been enough to make the trial a flash point for race issues, but that was not the worst aspect of the trial. Rather, it was the alleged use of peremptory challenges to strike five potential jurors who “appeared” to be First Nations, resulting in an all-white jury, that proved to be the most controversial and influential facet of the entire affair ([Harris \(2018\)](#), [MacLean \(2018\)](#)).

With Bill C-75 currently moving through the Canadian parliamentary system, having completed its second reading in June 2018 ([42nd Parliament of Canada \(2018b\)](#)), a close re-examination of the practice of peremptory challenge is warranted. A great deal of ink has already been spilled on both sides of the debate (see [Hasan \(2018\)](#), [Zinchuk \(2018\)](#), and [Roach \(2018\)](#)), but startlingly little of this discussion has been based on any hard evidence on the impact of peremptory challenge in jury selection. This paper aims to provide analysis and evidence to illuminate the topic further by analyzing three separate peremptory challenge data sets collected in the United States, namely [Wright, Chavis, and Parks \(2018\)](#), [Grosso and O’Brien \(2012\)](#), and [Baldus, Woodworth, Zuckerman, and Weiner \(2001\)](#). While this data cannot tell us if challenges were racially motivated in the Stanley trial, stepping back from this fraught legal episode to take a wider view of the practice of peremptory challenge provides a more sober place to start the discussion of its place in modern jury trials.

This paper will proceed in five parts. Chapter 2 provides a brief history of the practice of peremptory challenges in jury trials, in particular explaining their original motivation, past implementations, and how they have developed in the United States, the United Kingdom, and Canada. Chapter 3 proceeds to discuss the three data sets obtained, explaining the sources and collection methods before detailing the cleaning and preprocessing. Chapter 4 then provides the details and results of the analysis performed on the different data sets. It begins discussing the Jury Sunshine data set, which was used as a 'test' set of sorts, where analysis could be flexibly performed before the final analysis methods were turned to the other two data sets. The results of this analysis are compared to previous works in Chapter ???. Finally, the results and findings are summarized in ??, and recommendations based on the observations obtained here are provided.

Chapter 2

Peremptory Challenges

The focus of this text is the practice of peremptory challenges in a jury trial system, a highly specific practice in a particular context which may not be known in detail to the reader. As a consequence, a brief exploration of their history, motivation, and current use is presented here. It is not meant to be exhaustive, but rather to provide context and references for an interested and motivated reader to learn more. Indeed, many details have been omitted from the summary of the history in particular.

2.1 Jury Selection Procedures

Before reviewing the history, it is best to give some context and an explanation for readers unfamiliar with the jury system and general courtroom procedures. While the process of jury selection varies by jurisdiction and crime severity, the general steps shared by jury trials are outlined below. More detail and a discussion of the diversity of jury selection procedures can be found in [Ford \(2010\)](#), [Hans and Vidmar \(1986\)](#), and [Van Dyke \(1977\)](#). To select a jury:

- i.) Eligible individuals are selected at random from the population (using a list known as the *jury roll*) of the region surrounding the location of the crime, the sampled individuals are called the *venire*
- ii.) The venire is presented to the court, either as a group or sequentially (borrowing the names of [Ford \(2010\)](#): the “struck-jury” system and the “sequential-selection” system, respectively)
- iii.) The presented venire member(s) are questioned in a process called *voir dire*, which can result in three possible outcomes for each venire member:
 - (a) The venire member is removed with cause, the cause provided by either the prosecutor or defence lawyer and admitted by the judge
 - (b) The venire member is removed by a *peremptory challenge* by the prosecutor or defence lawyer, where no reason need be provided to the court; such privileged rejections of a venire member are limited in number for both lawyers (in Canada a maximum of 20 such challenges per side per defendant are allowed [[Government of Canada \(1985\)](#)])

- (c) The venire member is accepted into the jury, and so becomes a juror
- iv.) Steps i-iii are repeated until the prosecution and defence fail to reject the desired number of jurors.

As mentioned above, the details in this process can vary greatly by region. One of the greatest sources of this variation is the creation of jury rolls. In the United States the method is somewhat homogeneous: they are typically selected using lists of registered voters (see [Van Dyke \(1977\)](#) chapter two and [Hans and Vidmar \(1986\)](#) page 53), but in Canada their creation is far more varied. Ontario uses a combination of municipal voter lists and First Nations band lists (see [Ministry of the Attorney General of Ontario \(2018\)](#)), while in Saskatchewan - the province of the Gerald Stanley trial - the jury roll is created from the data in the central government health insurance agency in accordance with [Government of Saskatchewan \(1998\)](#).

Clearly, the variation in these methods will create differences in the universe of the sampled jury rolls relative to the population they are meant to reflect. Such differences are no doubt important to the coverage of the population which is present in the jury selection process (see [Iacobucci \(2013\)](#)), but these differences are not of primary interest to this paper. Rather, the steps presented afterwards are those to be investigated.

This leads to the two presentation methods presented in step ii, [Ford \(2010\)](#) and [Van Dyke \(1977\)](#) both note that the predominant method in the United States and Canada is the sequential-selection system. This is perhaps due to the relative efficiency of the method, as it is clear that in the sequential system voir dire need not be performed on the entire venire, only a subset. Contrast this with the struck-jury system, where the entire venire must be reviewed in every trial.

Finally, the scope of voir dire is radically different in the United States and much of the British Commonwealth. [Van Dyke \(1977\)](#) notes on page 143 that Canada and the United Kingdom do not allow questions in areas of “non-specific” bias, or bias which is not directly related to the case before the court. That is to say, while it would be perfectly valid to ask a venire member for a murder case about their work history in the United States, such a question would only be allowed in Canada or the United Kingdom if occupation was specifically related to the crime.

This difference in procedure creates a far greater emphasis on the voir dire process in the United States, as noted by [Hans and Vidmar \(1986\)](#). [Hans and Vidmar](#) go further than this, and surmise that the key reason for this marked departure in procedure is a difference in philosophy. To borrow a quote from page 63 of [Hans and Vidmar \(1986\)](#):

In Canada... the courts have said that we must start with an initial presumption that “a juror will perform his duties in accordance with his oath”

This doctrine places a responsibility on the jurors themselves to overcome their biases and accept arguments in spite of them. This stands in stark contrast to the American attitude implied by the emphasis on expansive voir dire: that certain prejudice cannot be overcome by jurors themselves. The public statements of the Stanley trial critics indicate that they subscribe to this viewpoint more than to the Canadian philosophy.

2.2 The Role of the Jury

Such a difference in viewpoint is especially relevant given the purpose of the jury. The central function of a jury in a jury trial system is to judge the innocence or guilt of an accused in light of the presented evidence, a function which has had drastically different forms throughout history. In the distant past, [von Moschzisker \(1921\)](#) and [Hoffman \(1997\)](#) report that the central function of the jury was to collect evidence, essentially assuming the role commonly performed today by police detectives. Such a role justified the archaic practice of selecting only the most “trustworthy” individuals of some renown.

This is contrasted by the modern jury, which performs no collection of evidence. It is, ideally, a panel of peers or “equals” of the accused sampled at random from the population, an idea which did not develop until 19th century Britain (see page 28 of [Hans and Vidmar \(1986\)](#)) and was not applied using random sampling until some time later (see [Hoffman \(1997\)](#), page 29 of [Hans and Vidmar \(1986\)](#), and page 16 of [Van Dyke \(1977\)](#)). The modern jury is meant to apply the law, as told to them by the judge¹, to the case at hand. Evidence of the guilt of the accused is presented to the jury by the prosecutor, while evidence meant to exonerate is presented by the defence.

The jury listens to the evidence, considers the law as presented by the judge, and must (typically) reach a unanimous decision of guilt or acquittal. Such a decision cannot be overturned by the judge of the court, and the judge must then determine sentencing based on the decision of the jury and the letter of the law¹. It should be clear that the jury therefore has tremendous power in the judgement of any case. The philosophical and ethical justification for such power is well explained by [Woolley \(2018\)](#), and best summarized by a quote from [Supreme Court of Canada \(1991\)](#):

The jury, through its collective decision making, is an excellent fact finder; due to its representative character, it acts as the conscience of the community; the jury can act as the final bulwark against oppressive laws or their enforcement; it provides a means whereby the public increases its knowledge of the criminal justice system and it increases, through the involvement of the public, societal trust in the system as a whole.

While such enthusiastic support for juries has not been expressed by all countries which practice them, the justification is entirely consistent with the histories and discussions presented by [Hoffman \(1997\)](#), [von Moschzisker \(1921\)](#), [Hans and Vidmar \(1986\)](#), [Van Dyke \(1977\)](#), and others. This suggests that the [Supreme Court of Canada \(1991\)](#) lionization of the jury system is a fair representation of the perceived role of the jury throughout those countries which use them.

2.3 Modern Peremptory Challenge Controversy

If the general utility and importance of the jury is clear, the same cannot be said for peremptory challenges. The privileged removal of a venire member² without any justification has seen persistent allegations of abuse, often around the use of these challenges by

¹[Hans and Vidmar \(1986\)](#) note that this system actually varies throughout the US, though the jury and judge powers described here are consistent across Canada.

²To be replaced by another, *randomly selected* venire member

state prosecutors.

In the United States, the criticism has focused on racial discrimination, and has led to significant changes in their allowed use, through cases such as *Swain v. Alabama* ([Supreme Court of the United States \(1965\)](#)) and *Batson v. Kentucky* ([Supreme Court of the United States \(1986\)](#)). The first of these cases, *Swain v. Alabama*, established in 1965 that the systematic exclusion of venire members of a particular race would be unconstitutional discrimination under the Fourteenth Amendment, but argued that a “*prima facie*” (or “based on first impression”) argument of discrimination was not adequate to prove this³. This placed a significant burden on the side taking issue with a particular peremptory challenge to demonstrate that the choice had been discriminatory.

However, this ruling was overturned only 21 years later in the 1986 case *Batson v. Kentucky*, which allowed the party objecting to a challenge to use a *prima facie* argument which must be countered by a race-neutral reason that satisfies the judge. If no such reason can be supplied, the challenge would not be allowed. This created a new challenge which could be used to keep a venire member despite the use of a peremptory challenge: the so-called “Batson Challenge”. While the effectiveness of this system of additional challenges is questionable both practically and in abstract (see [Page \(2005\)](#) and [Morehead \(1994\)](#), and a particularly strong response in [Hoffman \(1997\)](#)), it has only been extended to allow Batson challenges for both the sex and race of venire members⁴.

In Canada, the controversy has also had a racial component. Racial bias in Manitoba against First Nations venire members was alleged in 1991 in a report produced after an inquiry by the provincial government (see [Roach \(2018\)](#)). More damning still was the [Iacobucci Report](#) on First Nations representation in juries. This report proposed an explicit restriction to the practice when it recommended:

an amendment to the Criminal Code that would prevent the use of peremptory challenges to discriminate against First Nations people serving on juries.

Despite these recommendations and allegations, there had not been a significant political effort to reform the peremptory challenge system until the Gerald Stanley trial culminated in the tabling of Bill C75 [42nd Parliament of Canada \(2018b\)](#), which would abolish the peremptory challenge outright. As of the writing of this paper, the bill has not been approved by the Government of Canada, but it seems likely to become law in the near future.

In doing so Canada would join the United Kingdom. Significant controversy around the use peremptory challenges in the United Kingdom has already resulted in the abolition of the practice by the Criminal Justice Act of 1988. The specific controversy was the result of the Cyprus spy case in the late 1970s, which led to a “sustained campaign in Parliament and in the press alleging that defence counsel were systematically abusing it” (see [Hoffman \(1997\)](#))⁵.

³In the actual case, not a single black juror had sat in Kentucky in the previous 15 years, despite composing 26% of the jury-eligible population. In Swain’s trial, six of the eight black venire members were rejected by state prosecutor peremptory challenges, and the other two removed for cause, leaving not a single black juror to judge Swain, a black man. This was the *prima facie* argument presented by Swain’s defence team against the state prosecutors of Alabama, and it was rejected as insufficient to prove discrimination

⁴The use of Batson Challenges for sex was established in [Supreme Court of the United States \(1993\)](#)

⁵It should be noted that this did not abolish the use of “standing-aside” by the Crown, although the practice was restricted to national security trials and heavily curtailed, with strict guidelines to its use

2.4 The Role of the Peremptory Challenge

Despite these legal changes, recommendations, and a great deal of articles providing analysis against the practice (see, for example, [Hoffman \(1997\)](#)), the topic of the peremptory challenge remains controversial. The modern motivation and justification for the practice in spite of all of the controversy was perhaps best described by Justice Byron R. White in [Supreme Court of the United States \(1965\)](#):

The function of the challenge is not only to eliminate extremes of partiality on both sides, but to assure the parties that the jurors before whom they try the case will decide on the basis of the evidence placed before them, and not otherwise. In this way, the peremptory satisfies the rule that, “to perform its high function in the best way, justice must satisfy the appearance of justice.”

Such a justification is reminiscent of the now famous words of Lord Chief Justice Hewart in *R. v. Sussex Justices* in 1924: “Justice should not only be done, but should manifestly and undoubtedly be seen to be done” (as reported in [Richardson Oakes and Davies \(2016\)](#)). While these words originally only referred to the pecuniary interest of court staff involved in the case, they have since come to express the idealized expectation that both the defence and prosecution find the judge and jury acceptable, as explored by [Richardson Oakes and Davies \(2016\)](#)⁶.

This defence suggests two modern justifications for the peremptory challenge. The first is that of removing venire members with “extreme” bias, and the second is the creation of a jury which is composed of jurors mutually acceptable to both the defense and the prosecution. Those who defended the practice of peremptory challenges in Canada after the Gerald Stanley trial, including [Hasan \(2018\)](#) and [Macnab \(2018\)](#), seem to use this defence or some variant of it to argue in favour of keeping the practice. However philosophically appealing these two claims are, in light of all of the controversy surrounding the peremptory challenge, perhaps a critical and empirical examination of these assertions is warranted.

2.5 History

Such an analysis most appropriately begins with a historical explanation of the peremptory challenge. Roughly, the presentation of the history of jury trials here follows the comprehensive and exhaustively referenced description provided by [Hoffman \(1997\)](#). Two of the references [Hoffman](#) uses extensively, [Hans and Vidmar \(1986\)](#) and [Van Dyke \(1977\)](#), provided useful context while specific details provided by [von Moschzisker \(1921\)](#), [Forsyth \(1994\)](#), [Brown, McGuire, and Winters \(1978\)](#), and [Brown \(2000\)](#) helped to create a clearer picture of particular periods of jury history. Information regarding the history of the Canadian system was provided by [Brown \(2000\)](#) and [Petersen \(1993\)](#). For an excellent exploration of the nineteenth century, a formative time for the development of challenge law, see [Brown \(2000\)](#).

outlined by [Attorney General’s Office of the United Kingdom \(2012\)](#).

⁶Such grand generalizations and myth-making can also be seen in the common belief that the right to a trial by jury was originally established in the Magna Carta, an idea which is not supported by the relevant historical evidence (see [Hoffman \(1997\)](#) and [Van Dyke \(1977\)](#) for a detailed discussion and more accurate history).

It must be noted that certain important trials in the development of the peremptory challenge system have been excluded from the summary provided here. This was done deliberately, as the history presented here is only meant to explore the practice of peremptory challenges throughout history in broad terms. All of the sources listed above are much more thorough, by merit of their singular focus on the analysis of the practice from a legal and historical perspective, while this work devotes more to empirical and statistical analysis.

2.5.1 Pre-English History

Although precise timelines are hard to establish, there is evidence that jury trials have occurred in some form or another since antiquity. The concept, that of judgement by a group of peers, is so ancient that it is prevalent not only in historical records, but in myth. As [Hoffman \(1997\)](#) indicates, both Norse and Greek mythology feature groups of individuals assessing the guilt or collecting evidence about the actions of a peer.

Outside of the realm of myth, [Hoffman \(1997\)](#) reports that there is evidence of the use of juries in Ancient Egypt, Mycenae, Druid England, Greece, Rome, Viking Scandanavia, the Holy Roman Empire, and Saracen Jerusalem. It should be noted that in none of these areas was the jury trial the primary form of conflict resolution practiced. Nonetheless, it is clear the jury trial has a broad and long history of use.

Something similar to the modern peremptory challenge does not appear until Rome, however. The Roman *Judices* were groups of senators selected to judge the guilt of the accused in a legal case. According to [Hoffman \(1997\)](#), 81 Senators would be chosen to sit on one of these *Judices*, after which the litigants were permitted to remove fifteen of these Senators each. This egalitarian reduction of the jury size seems analogous to the modern peremptory challenge system, as it places the power of removal with the litigant and suggests no justification is necessary for their removal.

2.5.2 In English Law (1066–1988)

Peremptory challenge did not reach its modern form, as outlined in [2.1](#), until it was established in the English legal system. It should be noted that despite some previous debate on the topic, the most modern historical evidence suggests that the basis of the English practice was not related to the system used in the selection of *Judices* in Rome. The English system appears to be its own beast entirely.

The dominant historical interpretation is presented by [von Moschzisker \(1921\)](#) and [Hoffman \(1997\)](#): that the jury system was introduced to England during the Norman conquest of 1066 by William the Conqueror. The practice, however, was not made official until the Assize of Clarendon in 1166 by Henry II, and it was not until the outlaw of trials by ordeal (the most common method of trial at that time) in 1215, that peremptory challenges began to appear in England in the late thirteenth century. The challenges were officially recognized in 1305 when Parliament outlawed their use by the Crown, only to replace them with an analogous system of so-called “standing-aside”⁷.

⁷For a detailed explanation of this system see [Hoffman \(1997\)](#) and [Brown \(2000\)](#)

It should be noted here that although the challenges issued between the Assize of Clarendon and this 1305 act are called “peremptory,” they may not have served the same purpose, nor shared the same justification, as the modern challenges. Indeed, as [Hoffman \(1997\)](#) argues convincingly, these challenges may have been closer to modern challenges with cause. The argument hinges on the paradigm of royal infallibility and absolutism which was present in the late medieval period when the peremptory challenge first appeared (see [Burgess \(1992\)](#)).

Under royal absolutism and infallibility the argument for peremptory challenges is quite simple. If the king cannot be wrong in his judgement and he has some reason to feel that a venire member cannot serve on the jury, then he need not say why he thinks that is so, as his judgement is correct in any case. Indeed, asking for an explanation would be disrespectful and providing one undignified. The Crown prosecutors, as representatives of the king, would be similarly shielded from criticism.

Such an argument is further supported by the abolition of their royal use in 1305, the language of which suggests that peremptory challenges were originally the privilege of the Crown (see [Hoffman \(1997\)](#) and [Van Dyke \(1977\)](#)), with none being granted to the defence. [Hoffman \(1997\)](#) suggests that as royal infallibility grew out of favour, peremptory challenges were granted to the defence rather than being removed entirely.

Whatever the logic of the expansion of these challenges to the defence, their legal limits are recorded more precisely⁸. From a maximum of 35 challenges allowed at their peak in the fourteenth century, the number of challenges allowed only decreased over time until their abolition in 1988 (discussed in [2.4](#)).

2.5.3 In American Law (ca. 1700–1986)

[von Moschzisker \(1921\)](#), [Hoffman \(1997\)](#), and [Van Dyke \(1977\)](#) all agree that the early English colonists that came to North America accepted the jury system with peremptory challenges as common law well before the establishment of the United States of America. [Hans and Vidmar \(1986\)](#) note, however, that the difficulty of ocean travel and the overall indifference of appointed Crown representatives in the colonies led to an increased importance of the jury trial and the role of challenges to these early colonists as a way to exercise some degree of community control in the face of laws drafted in a distant country and implemented by unsympathetic authorities⁹.

It is somewhat interesting then, that the United States constitution makes no mention of the practice of peremptory challenges. The Sixth and Seventh Amendments specify a great deal of the jury system, including the right to public defense and an impartial jury drawn from the district of the crime, but make no mention of a right to the exercise of peremptory challenges, or any challenges whatsoever (see [Constitution of the United States \(1788\)](#)).

⁸see [Brown \(2000\)](#) for a detailed examination of the case law developing around challenges in the nineteenth century

⁹For more detail on this development among the early colonists, it is instructive to read about the Zenger trial of 1734 (described on pages 33-35 of [Hans and Vidmar \(1986\)](#)). Not only does this trial say a great deal about the attitudes of the colonists at the time, but it also presents the idea of a jury assessing guilt and “wrongness” using their own conscience rather than just settling fact. The precept of the modern jury trial in Canada (see [Woolley \(2018\)](#)) is based on this very idea

As [Hans and Vidmar \(1986\)](#) report on page 37, an original draft of the Sixth Amendment expressly included challenges for cause, but the debate around their inclusion resulted in the removal of their mention. They continue to say that at the time, even some proponents of the challenge considered the reference unnecessary, as the practice was implied by the text which remained, referring to a trial by an “impartial” jury. Another result of these debates was the adoption of the extensive voir dire process which allows questions of general bias¹⁰.

Critically, there appears to have been no discussion around the inclusion of peremptory challenges (see [Hans and Vidmar \(1986\)](#) and [Hoffman \(1997\)](#)). Despite the clear importance of the jury trial to the drafters of these amendments, it would seem the peremptory challenge was not considered to have anywhere near the same significance as judgement by an impartial jury of local peers¹¹.

Regardless of this, as [Brown \(2000\)](#) notes, the importance and use of challenges increased in the United States in the nineteenth century following American independence due to a desire to prevent the tyranny of the state. This desire also led to the adoption of a limited number of peremptory challenges for the prosecution, rather than the possibly unlimited stand-asides that were allowed under British law to prosecutors (see [Van Dyke \(1977\)](#), page 150).

While the specific numbers of peremptory challenges allowed to both sides and the required motivation of challenges for cause have varied over time (see [Hoffman \(1997\)](#) and [Brown \(2000\)](#)), they have remained a feature of the American legal system, and numerous Supreme court cases (detailed by [Hoffman \(1997\)](#)) have merely served to make the use of challenges more specific and codified. It was not until *Batson v. Kentucky* in 1986 that this system of challenges was drastically changed with the introduction of Batson challenges (described in 2.3).

2.5.4 In Canadian Law (ca 1800–2018)

Canadian law, inspired by a close relationship to both the British Crown and the United States, seems to have adopted elements of both legal systems in its development of peremptory challenges in the nineteenth century. As discussed by [Brown \(2000\)](#), Canada adopted the American practice of replacing prosecutorial stand-asides in favour of a more egalitarian limited number of peremptory challenges to both sides. Despite this, the Canadian voir dire process remains limited and much more similar to the British one, as does the system of challenges for cause (see page 48 of [Hans and Vidmar \(1986\)](#)).

One perfect demonstration of this departure is the Canadian constitution. As in the United States, the Canadian constitution fails to mention challenges. The British North America Act of 1867 (see [Constitution of Canada \(1982\)](#)), which established Canada’s independence from England, makes no mention of legal rights of the accused, indicating a deference to legal precedent in England. It is not until the Charter of Rights and Freedoms

¹⁰This is described on page 37-38 of [Hans and Vidmar \(1986\)](#), though [Brown \(2000\)](#) notes that 1807 Burr trial was also highly significant in the development of general voir dire in the United States

¹¹Indeed, as *Batson v. Kentucky* and *Swain v. Alabama* have both shown ([Supreme Court of the United States \(1986\)](#) and [Supreme Court of the United States \(1965\)](#)), the modern interpretation of “impartial” may preclude the use of peremptory challenges altogether

in 1982¹² that such rights were guaranteed in a legal Canadian document. Notably, its language is considerably more vague than the United States Sixth and Seventh Amendments, guaranteeing only “the benefit of trial by jury” (see [Constitution of Canada \(1982\)](#)).

This “eclectic” incorporation of both American and English case law, to borrow the term used by [Brown \(2000\)](#), led to a system somewhere between the English and American systems, but decidedly closer in operation to the English system. It should be noted, however, that as Canada grew more populous in the twentieth century and developed a greater legal precedent and more experienced judges of its own, this reliance upon its former colonial master and its more powerful southern neighbour seems to have diminished in importance. Viewing Supreme court rulings from recent decades reveals a great deal of reference to Canadian legal precedent rather than to the precedent of the other two countries.

2.6 Summary

The peremptory challenge, a practice of much controversy in the English-speaking world, seems to have started in its modern form as a privilege of the King of England in the thirteenth century. After its conception, it spread with English conquest and colonization, with new colonies and local governments accepting the practice based primarily on the adoption of English legal precedent. Though it was abolished in England in 1988, it remains a fixture of American jury trials, and is accompanied there by a thorough and invasive voir dire process which is not seen in Canada nor the United Kingdom.

Though the practice has historical longevity, it is not guaranteed by the constitutions of Canada or the United States, and has been a practice of considerable legal debate and significant change throughout its history. In England this culminated in the Cyprus spy trial, in the United States in *Batson v. Kentucky* and *Swain v. Alabama*, and in Canada in *R. v. Stanley*: the Gerald Stanley murder trial. As a consequence, the broad agreement of the importance and propriety of a jury has conferred little consensus on the place peremptory challenges in the selection of juries.

Indeed, it seems increasingly impossible for the jury to function in a way consistent with its demanding ideals with the peremptory challenge still present. Its spotted history and use to exclude certain minorities may undermine its purported use as a tool to ensure the acceptance of a trial’s outcome by both litigants. The three court cases mentioned above are a demonstration how the peremptory challenge can be used to create a jury which is actually unacceptable to one litigant in the case.

¹²This was the year of patriation of the Canadian constitution. As independence was granted by the British Parliament, the British North America Act outlining Canada’s laws was a British law and changing it was the prerogative of the British Parliament rather than the Canadian one. It was not until the Constitution Act of 1982 that the Canadian constitution became a Canadian law. For a more detailed history see [Sheppard \(2018\)](#)

Chapter 3

Data

Without data, performing an analysis that incorporated more than the history and legal argumentation presented in Chapter 2 is impossible. This proved problematic. While the motivation of this text was a Canadian case, no comprehensive Canadian data sets which examined jury selection in Canada could be found. The increased prominence of the jury selection process in the United States garnered a more fruitful search.

The author is heavily indebted to [Wright et al.](#); [Grosso and O'Brien](#); and [Baldus et al.](#). These authors shared their data freely with the author, providing him with a wealth of data to analyse empirically. As a consequence of the multiple separate data sets, however, care must be taken to describe each of the data sets separately in order to capture adequately the different methodologies and sources they represent. Critically, it should be noted that each of these papers represents effort on the part of the authors. As [Wright et al. \(2018\)](#) notes:

limited public access to court data reinforces the single-case focus of the legal doctrines related to jury selection. Poor access to records is the single largest reason why jury selection cannot break out of the litigato's framework to become a normal topic for political debate

Currently, the collection of jury data is difficult, as many courtrooms have not digitized past records and concerns over privacy limit the release of those records, which are stored as paper documents in the case file (see [Wright et al. \(2018\)](#)). This limits the ability of an individual to ask for summaries across numerous trials or to view the jury selection process on a scale beyond the basis of one case. Thus, to gather aggregate data the authors of these papers necessarily used different collection techniques dictated by the scope of collection desired and the procedures of the court systems from which data was collected.

3.1 Jury Sunshine Project

3.1.1 Methodology

The Jury Sunshine Project ([Wright et al. \(2018\)](#)), so named as it was carried out in order to shed light on the jury selection process, is the most extensive data set which was provided to the author. It endeavoured to collect jury data for all felony trial cases in

North Carolina in the year 2011, which ultimately resulted in a data set that detailed the simple demographic characteristics and trial information of 29,624 individuals summoned for jury duty in 1,306 trials. Note that not all entries were complete.

Due to the scope of the project, there are a number of problems which had to be solved by the authors. The first of these was simply identifying which court cases went to trial in 2011, in order to direct resources effectively. This was accomplished by downloading publicly available case data from the North Carolina Administrative Office of the Courts (NCAOC)¹ and determining the case numbers and counties of cases which went to trial. Wright et al. state that this likely missed some cases which went to trial, but that they were confident that a “strong majority” of trials was collected, which did not systematically differ from those excluded.

This list was then used to perform a pilot study to refine recording practices before undertaking a more general survey where “law students, law librarians, and undergraduate students” (called *collectors* for convenience) visited court clerk offices to collect the relevant case data, including the presiding judge, prosecutor, defence lawyer, defendant, venire members, charges, verdict, and sentence. The case files also included data about whether a venire member was removed by cause or peremptorily, and the party which challenged in the peremptory case. Using public voter databases, bar admission records, and judge appointment records, these collectors were able to determine demographic (race, gender, and date of birth) and political affiliation data for the venire members, lawyers, defendants, and judges. This data set was stored in a relational database provided to the author by Dr. Ronald Wright.

The analysis of the data provided in Wright et al. (2018) was limited to aggregate summaries of the trends at the venire member level. That is to say, they examined the strike trends for both the defence and the prosecution, conditioning on some additional variables. There was also spatial analysis performed, where different urban counties were directly compared. These analyses were not statistical in nature, and were displayed using contingency tables. Regardless the stark differences between prosecution and defence with regards to race were a key finding when the aggregate data was analyzed.

3.1.2 Cleaning

Flattening the Data

For greater expediency of analysis, the relational database of the Jury Sunshine Data was first flattened. The relational database was read into Microsoft Excel and the `readxl` package (Wickham and Bryan (2018)) was used to read the excel file into the programming language R. A wrapper for the `merge` function was developed which provided simple output detailing failed matches in an outer join in order to ensure that the flattening of the data into a matrix did not miss important data due to partial incompleteness. The code for this wrapper can be seen in D.1.

¹The link provided in the Jury Sunshine Paper to the specific source (http://www.nccourts.org/Citizens/SRPlanning/Statistics/CARports_fy16-17.asp) does not appear to be working as of January 2019, however the NCAOC seems to provide an API functionality at <https://data.nccourts.gov/api/v1/console/datasets/1.0/search/>

This wrapper revealed only a small number of irregularities in the data, which are detailed in [B.2](#):

- i.) Twenty-nine charges missing trial information such as the presiding judge (all of trials with IDs of the form 710-0XX)
- ii.) Twenty-six prosecutors not associated with any trials and missing demographic data
- iii.) One trial missing charge information

Ultimately, the jurors for trial ID 710-01, the trial missing a charge from above, were included in the data as their records were complete otherwise. The prosecutors and charges which could not be joined were excluded from any future analysis, as they could have easily been included by collectors by accident. Due to the small relative size of these inconsistencies relative to the size of the data set, they did not cause concern.

Uninformative Columns

Of course there were other irregularities in the data than the obvious ones that arose in the flattening process. There were a handful of likely sources for these errors. The first of these is the anonymization of the data for public use. The private data includes a wealth of privileged data such as juror name and address, and these were removed in the data given to the author.

As a consequence of this anonymization as well as the inclusion of rarely used columns such as those for additional notes, some columns of the data contained only NA values. Most baffling of these was the `BirthDate` variable in the `Jurors` table, as there was no clear reason for this data to be missing. Thankfully, none of the missing columns were relevant to the joins performed in flattening, and they would have been only secondary in data analysis. As a consequence, these uninformative columns were simply removed from the data.

Coding Inconsistencies

Related to this problem was the issue of inconsistently coded variable levels. An example of these inconsistencies would be levels recorded as both lower and upper case letters, or the presence of “?” instead of “U” for unknown values. It is very likely this inconsistency was a direct result of the data collection method which used many data collectors working independently in different places at different times. Thankfully, [Wright et al.](#) provided the codebook used by data collectors, which served as the authoritative reference for the admissible factor levels of demographic variables. Solving this problem was as simple as setting all demographic variable levels to be uppercase and replacing obviously mis-specified levels.

One specific inconsistency which should be noted is that of the outcome, which had a handful of entries recorded as “HC”, an inadmissible level not defined by the codebook. It is quite possible that this level represented a typo, as the “H” and “G” keys are adjacent on the American QWERTY keyboard layout, and “GC” was the code for ‘guilty as charged’, but out of a desire to be conservative with the data the occurrence of this outcome was replaced with U instead. Additionally, the inadmissible level “G” was replaced by GC.

Swaps

A more difficult problem of level misspecification was the presence of what appeared to be columns with swapped values, frequently occurring with the gender column (the admissible levels of which are “M”, “F”, and “U”) and the political affiliation column (the admissible levels of which are “D”, “R”, “I”, or “U”). The aforementioned “swaps” appeared as records in which, for example, the gender was recorded as “R” and political affiliation as “M”. More complicated swaps of three columns also occurred. To address this problem, the `IdentifySwap` function was written (see line 108 in [D.1](#)).

The `IdentifySwap` function accepts two arguments: a data frame with named columns and a named list of vectors of the acceptable levels for some of the column names. It then performs vectorized checks of the specified column names and presents any rows which may have swaps or errors interactively to the user, along with a suggested reorder to “un-swap” the row. The user can press enter to accept the suggested reordering, enter some other reordering, or enter 0 to indicate that the row was not a true swap, but simply an error. The un-swapped entries are then returned to the data, and the rows with errors have the erroneous values replaced by “U”, the universal code for “Unknown” within all data variables².

The source of these swaps is also most likely the data collection method. The codebook provided specifically notes that the data collection was meant to record the race (R), gender (G), and political affiliation (P) data in the form RGP, but it is not inconceivable that it would occasionally have been recorded or entered in some other ordering in the tedium of data entry. In any case, this problem affected only 431 records of the nearly 30,000, suggesting that the recorded error rate was not unacceptably large.

Charge Classification

Perhaps the least regular data in this data set was that of the charge text. Due to the lack of any codebook guidance about the standard way of recording a charge in a trial, identical charges were recorded in numerous ways. The first method used to combat this was removing non-alphanumeric characters, extra spaces, and converting all charges to lower case. This still left a considerable variation, however. Consider the charge of breaking and entering, for example, even with this simple preprocessing performed the entries varied significantly (e.g. “break or enter”, “breaking andor entering”, “breaking and or entering”, etc.).

As a consequence, the processing was more involved. First the most common versions of the charge text for the charges were all regularized to be identical (see `StringReg` in [D.1](#)). Next, a regular expression classification tree was developed, which would also account for specific features of a charge. When identifying murder, for example, it seemed important to ensure attempted murder was separated from murder itself, and separating first and second degree was also desired. This tree would, when presented with a charge, apply the regular expressions at each node to the charge. If the charge matched the expression at a node, the regular expressions of that node’s children were applied to the charge until it was classified to some leaf node, each of which had a standardized value which replaced

²One notable exception to this insertion of “Unknown” was the case of the judge Arnold O Jones II, whose gender was not recorded in the data, but who was identifiable as a man using a quick Google search of his unique name

the charge. A small example of this structure is displayed in Figure 3.1, and the full tree is visualized in B in Figure B.2.

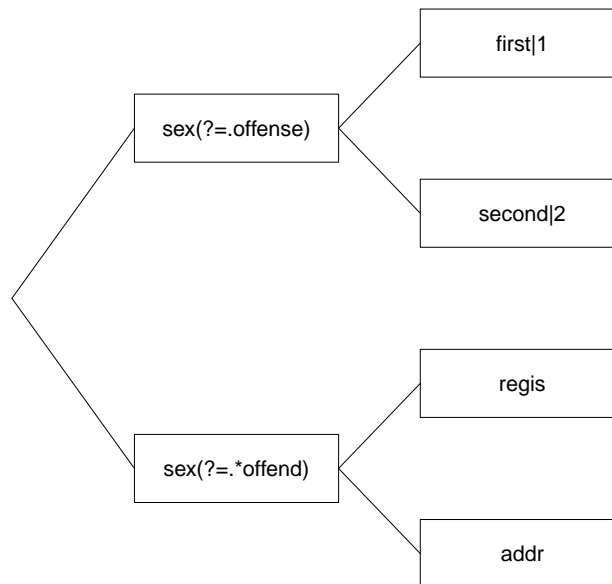


Figure 3.1: A example of a simple charge classification tree to separate the sexual offenses from charges leveled against previously known sex offenders. A charge would be classified from most general on the left to most specific on the right.

By performing regularization using this charge tree, regularized charges were guaranteed. The cost of this regularization was the inability to classify all crimes, however. Of the 1407 charges present in the data, the tree provides regularization for 1209. With additional time and inspection of the failed matches, the tree could conceivably be expanded to regularize all charges. As this was not the primary investigation of the report, however, such effort was not expended.

Instead, a number of helpful aggregation and extraction functions were developed to allow for the charges to be further simplified, as seen in D.1. In this report, they were aggregated by intuitive classes: sex-based offenses, thefts, murders, drug charges, violent offenses not otherwise classified, and driving charges. However, other classes, such as the North Carolina felony classes themselves (as provided by [North Carolina Sentencing and Policy Advisory Commission \(2017\)](#)), may provide a more informative classification rationale.

Variable Level Renaming

The final step of the data cleaning process was to convert the uninformative codes used to indicate variable values to more intuitive and clear names (for example to convert “I” in the political affiliation variable to “Ind”, a clearer indication of independent). Certain variables which were already clear, such as gender (codes “M”, “F”, “U”), were not renamed due to the clarity of the one letter representations.

3.1.3 Variable Synthesis

In order to expand the possible analysis and visualization potential, a number of variables were synthesized from the Jury Sunshine data set. They are detailed below.

Race Match A logical variable which is true for a venire member if they are the same race as the defendant, and false otherwise. This variable was motivated in particular by the Gerald Stanley trial, in which the implicit contention of those who have taken issue with peremptory challenge is that the First Nations venire members were removed by the defence as their race did not match that of Stanley in the racially charged trial.

Guilty Logical indicator indicating whether the trial outcome was guilty or not

Visible Minority Logical indicator of non-white venire member race

Race of Striking Party Factor variable which gives the race of the prosecution if the venire member was struck by the prosecution, the race of the defence if the venire member was struck by the defence

Simplified Race Due to the scarcity of the other minority races, this variable simplified the race provided to white, black, or other for the venire member

Simplified Defendant Race The same as the simplified race for the defendant races

Simplified Disposition This variable combined the categories “Foreman” and “Kept” in the original disposition variable into a variable “Kept”

3.2 Stubborn Legacy Data

3.2.1 Methodology

[Grosso and O’Brien \(2012\)](#) also provided data to the author, albeit a more limited set. This study, also based in North Carolina, focused on the trials of inmates on death row as of July 1, 2010, yielding a total of 173 cases. In each proceeding, the study examined only those venire members not excluded for cause, and critically the analysis of the study focused only on prosecutorial peremptory challenges. Besides collecting demographic data as in the Jury Sunshine Case, this study also collected attitudinal data for the venire members. This attitudinal data for the venire members is somewhat more detailed than the political affiliation data provided in the Sunshine data, including attitudes about the death penalty, employment information, and opinions on the trustworthiness of law enforcement.

Staff attorneys from the Michigan State University College of Law were responsible for the data collection in this study. The work was performed similarly to the Jury Sunshine Data, using case files to collect information about the court proceedings such as the peremptory challenges used, presiding judge, prosecutor, and defence lawyer. Detailed verdict and charge information was not collected, as the preselection criteria of death row inmates made the verdict clear, and the death penalty can only be applied for certain crimes.

To collect demographic and attitudinal data, the juror questionnaire sheets were con-

sulted³. These sheets are typically used as a component of voir dire, in order to make the process more efficient and determine venire members categorically ineligible for jury duty. As a result, they inquire about opinions on the death penalty, for example, as well as demographic questions. As not all jury questionnaires were available, additional information was collected from jury roll lists to determine the races of the final jury members. It should be noted that this collection was done blind and to high standards of proof, and a reliability study carried out in [Grosso and O'Brien \(2012\)](#) indicated that under this system the race coding was 97.9% accurate when the standards were met. Those for whom the standards were not met were marked as "Unknown."

The lack of an examination of political affiliation by this study, instead choosing to input far more detailed data on venire member viewpoints, serves as a barrier to the comparison of this data to the Sunshine data on an identical basis. However, the racial data for the two is recorded in a very similar way, so this effect can, at least, be compared. A comparison of the estimates of racial impacts controlling for political affiliation in the Sunshine data to those generated controlling for much more specific may be also of interest.

3.2.2 Cleaning

The data provided to the author was already exceptionally clean, and so no cleaning was required. There was no variable synthesis performed on the data, rather variables were used and combined to generate analogous measures to those recorded by the Sunshine data. These combinations included using indicators of strike status to make a disposition variable analogous to that provided in the Sunshine data, and combining certain racial indicators into the more universal white/black/other coding.

3.3 Philadelphia Data

3.3.1 Methodology

[Baldus et al. \(2001\)](#) presents a similar data set collected using similar means. Court files such as the juror questionnaire, voter registration, and census data were all used to complete juror demographic information for 317 venires consisting of 14,532 venire members in Philadelphia capital murder cases between 1981 and 1997⁴. It should be noted that this data included only those jurors kept or peremptorily struck, venire members struck for cause were not included in the data. The procedure used to determine race using the census and voter registration polls was quite complicated, but was rigorously performed using accepted census methods to a standard of 98% reliability⁵.

³As [Grosso and O'Brien \(2012\)](#) observe, self-identified race may be the most accurate source of racial group identification.

⁴This study took into account the sampling error by reweighting venires based on the year of the trial and the defendant race, as court records showed that the sample coverage varied over these factors.

⁵Additionally, imputation was only performed in a small minority of cases.

3.3.2 Cleaning

One interesting quirk of the Philadelphia data set was missing values. While the codebook describing the data explicitly stated a number of variables should be recorded as binary values. In the provided data files, however, these variables were missing for a majority of the observations. In the case of the “FINLJURY” variable for example, an indicator of whether the jury member was included in the final jury, there were 4626 records with a value of 1, 3 with a value of 0, and 12890 missing values. These missing values were assumed to be zero, as using this assumption created a data set which was consistent with that reported in [Baldus et al. \(2001\)](#). As with the Stubborn data, the only variable synthesis performed was completed to create analogous variables to the Sunshine data.

Chapter 4

Empirical Analysis

With this data cleaned and processed, questions can now be posed and addressed through analysis. A few obvious questions come to mind, considering the previous work done on this subject and the modern controversy surrounding it. First, there is the obvious question of not only the possible racial imbalance of peremptory challenge use, but how this imbalance changes with the race of the defendant. In the Gerald Stanley trial, for example, the critical aspect of the trial was not the use of peremptory challenges in abstract, but how their use interacted with the race of Stanley.

Aside from these investigations, we may wonder whether the most common arguments posed in favour of peremptory challenge are satisfied in this data. As discussed in 2.4, there are two primary arguments. The first is the argument that the peremptory challenge is necessary to remove the “extremes of partiality” present in the venire for both sides, that is to remove the most extremely biased jurors. This goal is complemented by the ability of the judge to remove jurors with cause, which is also designed to remove those jurors with extreme bias. The second argument is the creation of a jury which is mutually acceptable to both parties in the trial.

4.1 Extremes of Partiality

While creating a quantitative judgement on the acceptability of a jury is somewhat difficult, measuring the extremality or abnormality of observations is a critical function of statistics. With this in mind, a very simple calculation was performed. The central claim of the advocates of the use of peremptory challenge is that it is only used to remove extreme cases of bias. If that is so, then the proportion of venire members removed by peremptory challenge should reflect this concept.

Of course, this cannot be rigorously tested, as there is no way of knowing the true distribution of bias among jurors. That does not mean nothing can be said, however. As [Nisbett and Kunda \(1985\)](#) notes, there is a tendency of people to guess that a distribution is normal when asked to guess the distribution of social attitudes¹. Additionally, math-

¹This problem is not helped by the notoriety of the normal distribution, as it is commonly the distribution used when performing tests (likely due to the utility of the Central Limit Theorem) and generating visualizations of a general distribution

ematical constraints such as the Chebyshev inequality (see [Weisstein \(2018\)](#)) provide an upper limit to the dispersion of any distribution.

This study suggests that it would not be unreasonable to view the overall causal and peremptory challenge rates as the tails of a normal distribution, and the Chebyshev limit gives an estimate of the extremality of rejections given a maximally dispersed distribution of opinions. Table 4.1 provides a summary of the rejection rates of the different data sets and the implied standard distance from the centre that these imply for symmetric rejection. Note that two rows are provided for the Sunshine data, the first for the entire data set of jury trials and the second for first degree murder². This was done to facilitate comparison, as both the Stubborn Legacy and Philadelphia dataset only addressed capital cases³.

Table 4.1: The implied statistical extremity bound for symmetric rejection in the datasets given different distributional assumptions

Data	Rejection Rate	Normal	Chebyshev Limit
Sunshine	0.434	0.781	1.517
Sunshine Capital	0.639	0.470	1.251
Stubborn	0.659	0.442	1.232
Philadelphia	0.736	0.337	1.166

Obviously, it is not possible to comment with authority on the presence of partiality in the population. Indeed, given the large divide that appears to be present politically in the United States and the rest of the Western world today, it may be easier to argue for a maximally spread distribution than a centralized one. Regardless of this difficulty, it is difficult to justify that more than 43% of a dataset is “extreme” statistically. In the normal case, this suggests that the rejection boundary is less than one standard deviation from the mean, i.e. that the typically sampled point will be too extreme. The Chebyshev limit is not much better, suggesting that the rejection boundary is at most 1.5 standard deviations from the mean in either direction.

This observation is, of course, complicated by the increased use of peremptory challenges for the capital cases, as seen by the large difference in the rejection rates between the full Sunshine data and the capital trials alone. There are many possible competing explanations for this difference, chiefly the high stakes of capital cases and the greater publicity these cases often generate before going to trial. In either case, obvious knowledge and strong opinion of a case before the trial constitutes a reason for a rejection with cause, and so the language used in defence of the peremptory challenge lacks adequate precision. Calling 43-73% of the venire “extreme” is not consistent with the vernacular connotation of the word.

Such high rejection rates given any distribution suggests that the peremptory challenge is not simply being used to remove “extremes of partiality.” Rather, it seems that the argument used to support and justify the practice cannot be reconciled well with the data, suggesting systemic over-use relative to its supported use. This leads naturally to the question of how exactly this legal instrument is over-used, and why.

²The only sentence with the potential of a death penalty, as informed by the sentencing guidelines provided in [North Carolina Sentencing and Policy Advisory Commission \(2017\)](#).

³It should be noted that there is still some lack of ability to compare, as the previous studies looked at individuals sentenced to death, while no individuals in this data set were sentenced to death.

4.2 The Impact of Race

The racially-motivated controversy surrounding peremptory provide one hypothesis in the pattern of peremptory strike over-use which may be investigated. To begin, a simple marginal investigation was performed to explore the impact of the simplified race on the peremptory strike probability. The result of this investigation is displayed in Table 4.2. Of particular interest is whether any race is far more likely to be struck by peremptory challenge than the others, as this might suggest that race is the target of the over-use of strikes.

Table 4.2: The conditional probability of a venire member being struck peremptorily by the simplified venire member race across data sets. These values are smaller than the values presented in the extremity analysis as only the individuals which were identifiably removed by peremptory challenge are counted in this table. Regardless, the comparisons remain similar if the unattributed removals are included. Note that the Philadelphia trial data only indicated black and non-black venire members and so only two numbers can be reported.

Data	Black	Other	White
Sunshine	0.23	0.24	0.25
Sunshine Capital	0.22	0.27	0.27
Stubborn	0.65	0.36	0.66
Philadelphia	0.67	0.68	

These probabilities are different, but not greatly so. Indeed, the trend of higher probabilities for the removal of white jurors across all data sets is perhaps counter-intuitive given the history of peremptory challenge controversy in the United States. In any case, the small magnitude of these differences seems to suggest that there is no strong racial bias at the aggregate level, whether or not the results are statistically significant⁴. Perhaps there is a more interesting relationship present at the racial level. Taking inspiration from *Swain v. Alabama*, *Batson v. Kentucky*, and *R. v. Stanley*, perhaps viewing the relationship between venire member race and defendant race would be informative. This relationship is displayed in Figure 4.1.

A detailed description of this plot and its development which includes a discussion of the principles of graphics and perception which were used to devise its form is presented in A⁵. The most interesting patterns visible in the plot will be discussed here.

First, a small explanation of the plot. The plot displays the relationship between three categorical variables: venire member race, defendant race, and disposition (whether a venire member is struck and by whom). The vertical axis corresponds to the conditional probability of the disposition given a race and defendant race combination. Racial combinations are placed along the horizontal axis, and each combination corresponds to one horizontal black line in the plotting area. The length of these lines is proportional to the number of venire members in the data with the corresponding racial combination, and their vertical positions are the mean conditional probability of a venire member being removed by a challenge for that particular combination. The dashed vertical lines, coloured by disposition, start at this mean line and extend to the observed conditional probability

⁴Consider that the impact of a statistically significant difference of a few percent when the jury size is 12 for each trial would be minor.

⁵Here it suffices to mention that much of its design was motivated by the philosophy of Tufte (2001) and the results of Cleveland and McGill (1987) on the accuracy of visual perception.

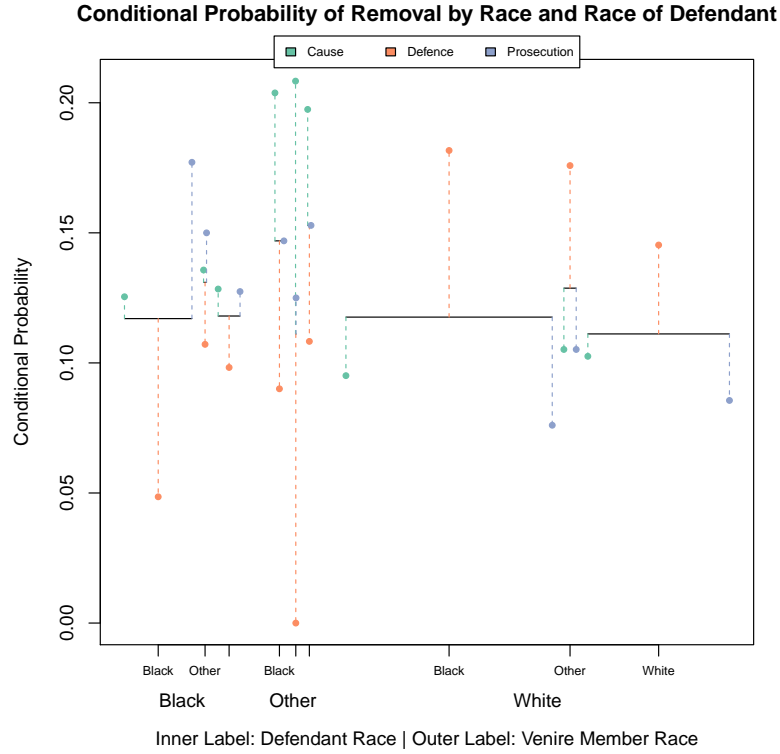


Figure 4.1: The conditional probability of successful challenges given the venire member and defendant race, with the expected value represented by the horizontal black line, and the observed values represented by the point at the end of the dotted line. Each horizontal black line corresponds to a particular venire member and defendant race combination, with a length proportional to the number of venire members with that combination. The dashed vertical lines, coloured by challenge source, start at these horizontal lines and end at points which show the observed probability of a challenge by that source for the given racial combination.

of the corresponding disposition for the relevant racial combination. As a consequence, this plot can be viewed as a visualization of the test of a specific hypothesis:

$$D|D \in \{2, 3, 4\}, R, E \sim \text{Unif}(\{2, 3, 4\}) \quad (4.2.0.1)$$

Where D, R, E are random variables representing the disposition, venire member race, and defendant race respectively as outlined in 0.2. In words: the conditional distribution of the disposition given the racial combination is uniform. This implies that all three strikes occur with the same probability for each racial combination, though they may differ between racial combinations. Such a hypothesis allows for certain racial combinations to experience a higher strike rate generally, but constrains the strike rate to be the same for all parties, which would imply that all parties in the court pursue an identical strike strategy across all venire member and defendant race combinations.

Clearly, Figure 4.1 casts some doubt on this hypothesis. While the horizontal black lines tell a very similar story to Table 4.2, with little variation between them, a number of other striking patterns are visible. The first, and most obvious of these, is the main effect of venire member race. While the aggregate removal rates do not seem to depend on the race of the venire member, it is clear that the defence and prosecution pursue radically different

strategies. The defence seems biased towards a jury with more venire members from visible minorities. All orange points are below the horizontal lines for the black and other venire members, indicating these groups are less likely to be struck by the defence than expected, while the points are above the lines for the white venire members, indicating a higher than expected probability of defence removal for white venire members. The prosecution seems to mirror this tendency, striking the white venire members at a lower rate than expected and the black venire members more often than expected. The prosecution seems to show no deviation from expectation for the venire members of other races.

The addition of defendant races shows another interesting trend. It would seem that the aforementioned tendencies of the prosecution and defence are strongest for black defendants, which have the strongest departure of the conditional probabilities from the expectation. The defense and prosecution seem to have slightly more similar habits when the defendant is white, despite their opposite tendencies in all cases. Finally, it would seem that the removals with cause have tendencies similar to the prosecution, as the points representing the conditional probability of a venire member being removed with cause are always on the same side of the expected line, an event which would occur with probability $2^{-9} \approx 0.002$ under the hypothesis of independent uniform strike rates. Further exploration of the agreement of these two strike tendencies is explored in B.1.

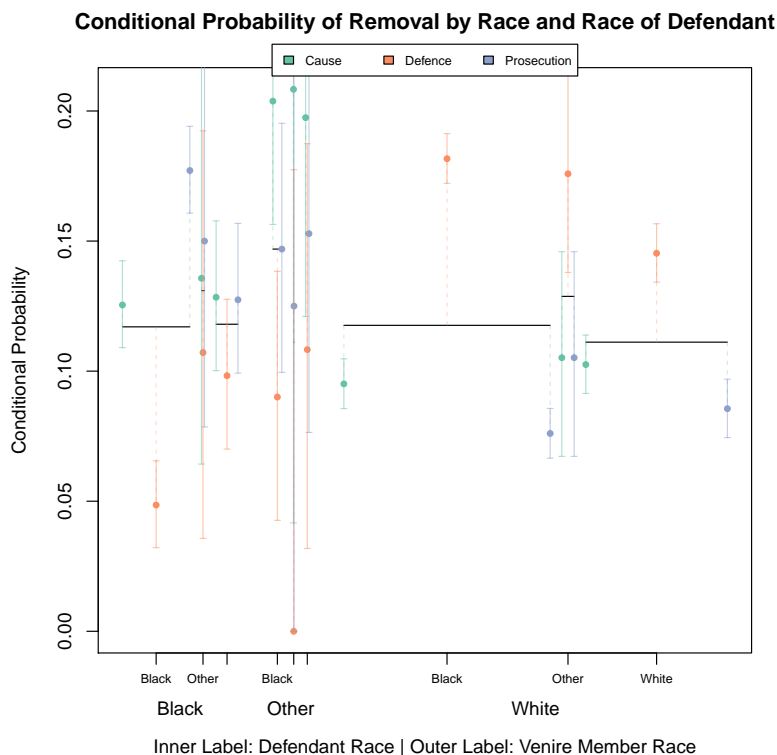


Figure 4.2: The plot of conditional strike probability by racial combination from above with confidence intervals added. Note that many of the seemingly striking departures seen are insignificant when these confidence intervals are applied.

While Figure 4.1 is quite suggestive, the widths of certain horizontal black lines, in particular those for venire members with a race other than white or black, suggest that perhaps some of the more extreme tendencies are simply a result of the well-known higher variation of samples with small sizes. In order to see the true nature of the noted departures some

incorporation of the variation one expects from each observed value is required. This is accomplished by the addition of approximate 95% simultaneous multinomial confidence intervals using the `MultinomialCI` package in R, which implements simultaneous confidence intervals for multinomial proportions following the method presented in [Sison and Glaz \(1995\)](#). These confidence intervals can be seen in Figure 4.2.

As suspected, some of the results for the smaller sample sizes do not seem to be significant. The results for the larger groups, in particular for white venire members or black defendants, are significant, however. It should be noted that these simultaneous confidence intervals do not constitute a proper statistical test of the impact of race, they are rather a way of visually providing a viewer some sense of the expected variability in the data over repeated sampling. More rigorous testing is performed alongside the model building in 4.4.

4.2.1 In the Stubborn and Philadelphia Data

With the patterns for the entire Sunshine data set examined, it is important to see how general this pattern is. It is entirely possible that the Sunshine data simply shows a particularly bad year in North Carolina, and so seeing whether this pattern is present over time and in different regions is important to motivate its generality. To that end, Figure 4.1 was reproduced for the Sunshine capital cases, the Stubborn data, and the Philadelphia data. These plots can be seen in Figure 4.3.

Three important differences between these plots and the Figure 4.2 plot must be noted when interpreting these visualizations. The first of these is that data differences have limited the scope of the comparison to simply the prosecution and defence strike rates, as the Stubborn data does not include any information on strikes with cause. Second, the Sunshine data used to generate the mobile plot in Figure 4.3 is filtered to only the first degree murder trials, as the other two data sets only addressed capital trials (i.e. those with the death penalty), while the Sunshine data had a broader scope. Finally, the race and defendant race were further simplified to logical indicators of whether the race variable was black or not.

Despite the very different study sample universes of these three data sets, all display identical patterns, with only the magnitudes of the different strike rates differing. The similar level of all black lines within each plot shows that in each data set, the aggregate probability of removal is similar across racial minority combinations. However, these aggregate similarities hide the vastly different strategies of the defence and prosecution. The defence has a tendency to retain visible minority venire members, striking them at a lower rate than the white venire members, while the prosecution shows a pattern which mirrors that of the defence.

It should be noted that the Sunshine data set looks most unique of the three, and this may be a result of the sample size. While the Philadelphia data and Stubborn data both collected data which included multiple years, the Sunshine data was restricted to trials which occurred in 2012. This small sample is the reason for the large confidence intervals present in Figure 4.3c. This small sample size does not explain the overall lower strike rate observed in the capital trials in this data, which is also visible Table 4.2. Such a departure may be of interest for further investigation.

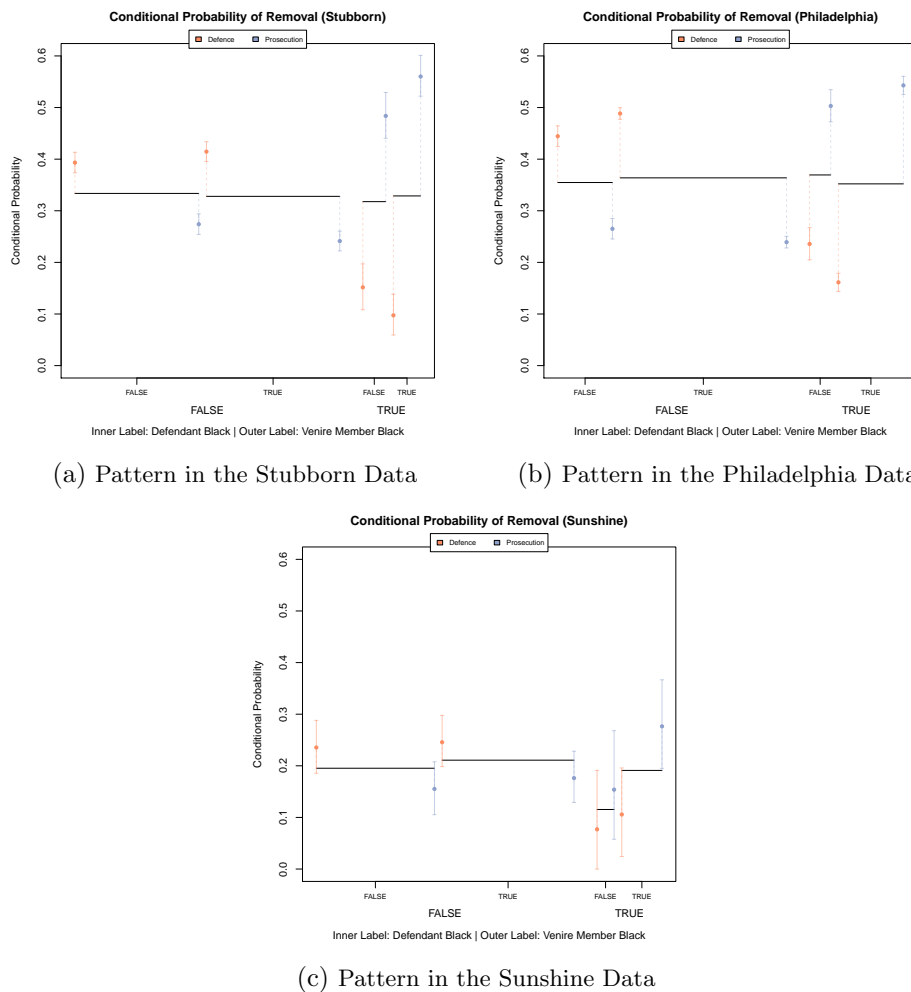


Figure 4.3: The conditional probability of defence and prosecution peremptory challenge by racial minority status across all capital trials in all data sets. The pattern, though sometimes different in magnitude, is quite consistent across the three examined data sets, despite the significant differences in the respective study sample universes.

4.3 Other Factors

Of course, it would be incorrect to conclude immediately that the cause of the racial patterns observed across these data sets is race itself. There may be a plethora of attitudes associated with race that could serve as legitimate cause for a peremptory challenge. As noted by Justice Byron R. White in the majority opinion in [Supreme Court of the United States \(1965\)](#)

[The peremptory challenge] is no less frequently exercised on grounds normally thought irrelevant to legal proceedings or official action, namely, the race, religion, nationality, occupation or affiliations of people summoned for jury duty. For the question a prosecutor or defense counsel must decide is not whether a juror of a particular race or nationality is in fact partial, but whether one from a different group is less likely to be.

This quote leads directly to the heart of the problem. Without detailed transcripts in-

dicating how the venire members were questioned, it cannot be known if the aggregate pattern of removal is the result of racially based strikes, or whether the lawyers determined valid reasons for a peremptory challenge during the voir dire process. For example, if defence attorneys reasonably assumed that trust in and deference to authority and law enforcement would make a venire member predisposed to reject arguments provided about possible mishandling of evidence without proper consideration, this would be reasonable grounds for peremptory challenges of individuals with that opinion. If such opinions are distributed heterogeneously by race, the aggregate pattern may appear to reflect racially-based decision making by the defence attorneys.

4.3.1 Political Affiliation in the Sunshine Data

Revesz (2016) provides, inadvertently, data which might support the above reasoning in defence of peremptory challenges in the United States. He notes that the distribution of political affiliation in the United States is not consistent across races, with black voters far more likely to vote for the Democratic Party and far less likely to vote for the Republican Party. If political affiliation is used as a surrogate for ideology and point of view, this suggests that the observed pattern could be the result of defence lawyers removing conservative venire members and prosecution lawyers attempting to remove liberal ones. As the Sunshine data has political affiliation, these aggregate results can be examined for the data used to generate Figure 4.1. Figure 4.4 displays the conditional probability of political affiliation across races and genders.

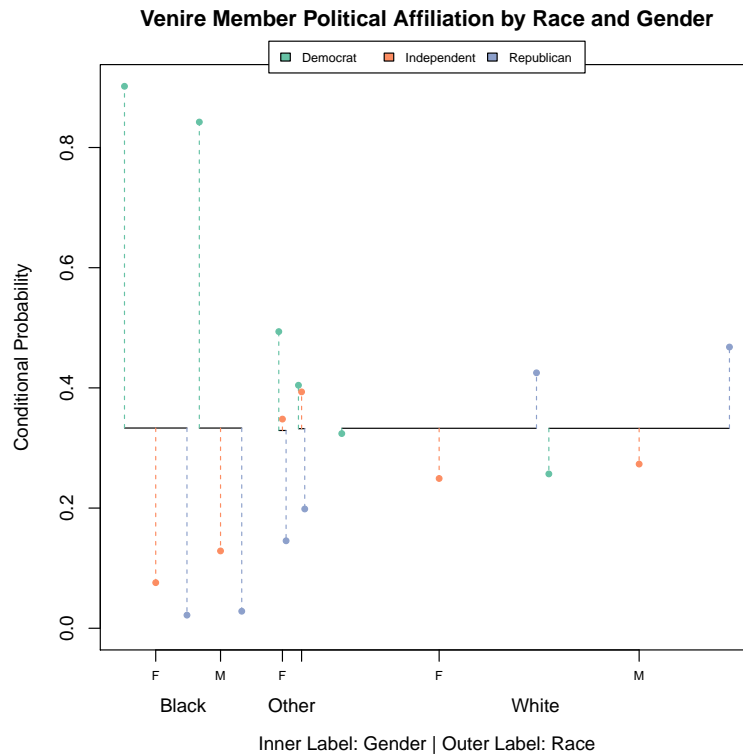


Figure 4.4: Conditional probabilities of political affiliation by race and gender. Note how the black venire members are far more homogeneous than the white venire members for both genders.

What is immediately apparent viewing this plot and the data in [Revesz \(2016\)](#) is how closely the two data sets agree. This “ideological imbalance”, as [Revesz](#) aptly calls it, is a clear confounding factor and a possible source of a legitimate cause for an initially suspect overall trend. As such, it was investigated using the mobile plot.

To control for the defendant race as well, which already appears to be important, the venire members were split into the simplified racial groups black, white, and other. Then mobile plots of the conditional strike probabilities for the different venire races given the defendant race and political affiliation were generated. Figures 4.5a, 4.5c, and 4.5b display these mobile plots.

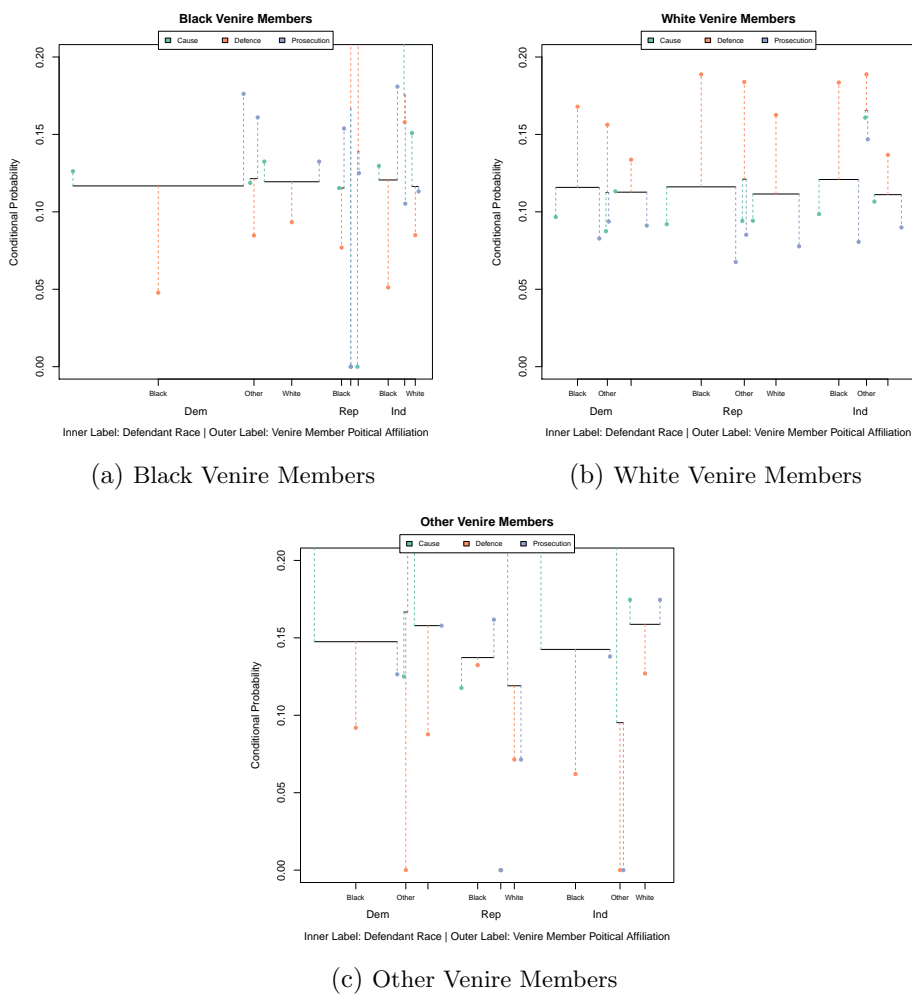


Figure 4.5: Conditional probability of venire member strike by defendant race and political affiliation, split by race. Note how the pattern of conditional probabilities is the same across political affiliations for the same defendant race within a particular venire member race, and this pattern is the aggregate pattern seen in [Figure 4.1](#)

This sequence of three plots immediately suggests that a political argument is insufficient for this data. For each venire member race and defendant race, the political affiliation of the venire member does not radically change the pattern of strikes for any party in the court. Rather, the court tendencies for each political affiliation, venire member race, and defendant race seems to follow the pattern seen in [4.1](#) for all political affiliations with the

exception of some very small subgroups of black venire members⁶. It should be noted that many of the other, less noticeable subgroups represented by this particular set of plots will be too small to provide statistically significant results, the consistency of the pattern of strikes across political affiliations provides rather stark evidence against a possible claim of political affiliation being the reason for the strikes in the data.

4.3.2 Gender in the Sunshine Data

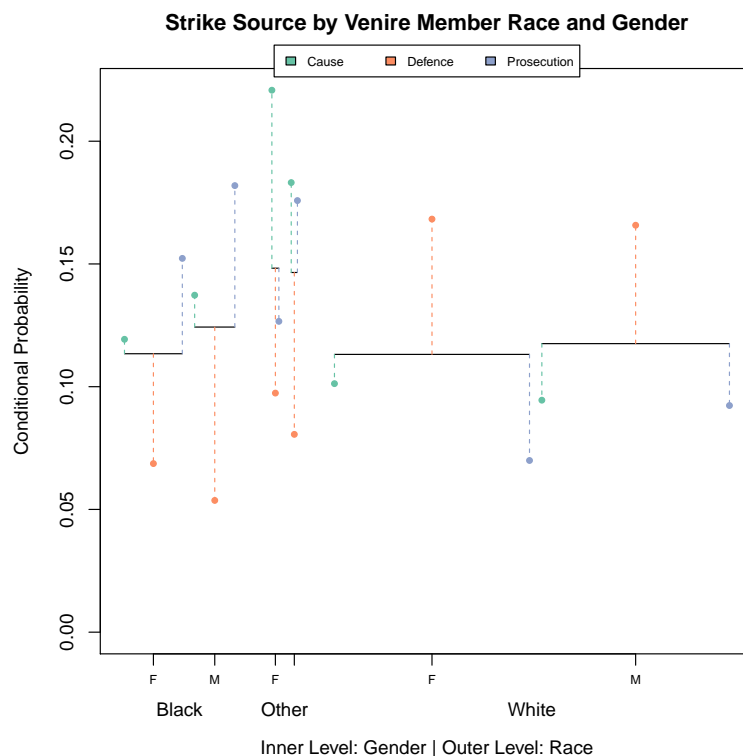


Figure 4.6: Conditional probabilities of strike source by race and gender. Note that the pattern is nearly identical for the genders within each racial group.

Another factor which may have an impact is that of gender. While *J.E.B. v. Alabama* (Supreme Court of the United States (1993)) has also ruled peremptory challenges for this reason alone unconstitutional, it is noted in Van Dyke (1977) on page 152-153 that prosecutor's guidelines have, in the past, recommended using peremptory challenges to remove female venire members⁷, and so perhaps it is a relevant factor. Additionally, there is a relationship between gender and race in the data, as noted in Wright et al. (2018)

⁶The subgroups, black republican venire members for white or other race defendants, have sizes 1 and 10, respectively. The study examines a total of 29636 venire members.

⁷The guidelines for sex written by Jon Sparling, an assistant district attorney in Dallas, read specifically:

- i.) I don't like women jurors because I don't trust them.
- ii.) They do, however, make the best jurors in cases involving crimes against children.
- iii.) It is possible that their "women's intuition"[sic] can help you if you can't win your case with the facts.
- iv.) Young women too often sympathize with the Defendant; old women wearing too much make-up are usually unstable, and therefore are bad State's jurors.

black males are highly under-represented relative to black females. Luckily, the relationship between race, gender, and peremptory challenges can be visualized quite neatly without the need to split the plots as was done in Figure 4.5. Figure 4.6 displays the mobile plot for race, gender, and strike source.

The same pattern is seen here for the most part, that of a dearth of defence strikes against visible minorities and a surplus against white venire members, with a mirrored tendency for both the prosecution and challenges with cause. One may ask the question if such a switch is just a quirk of the data, and if conditioning on gender alone, for example, would create a similar pattern of switching tendencies of the court. Figure 4.7 shows this plot.

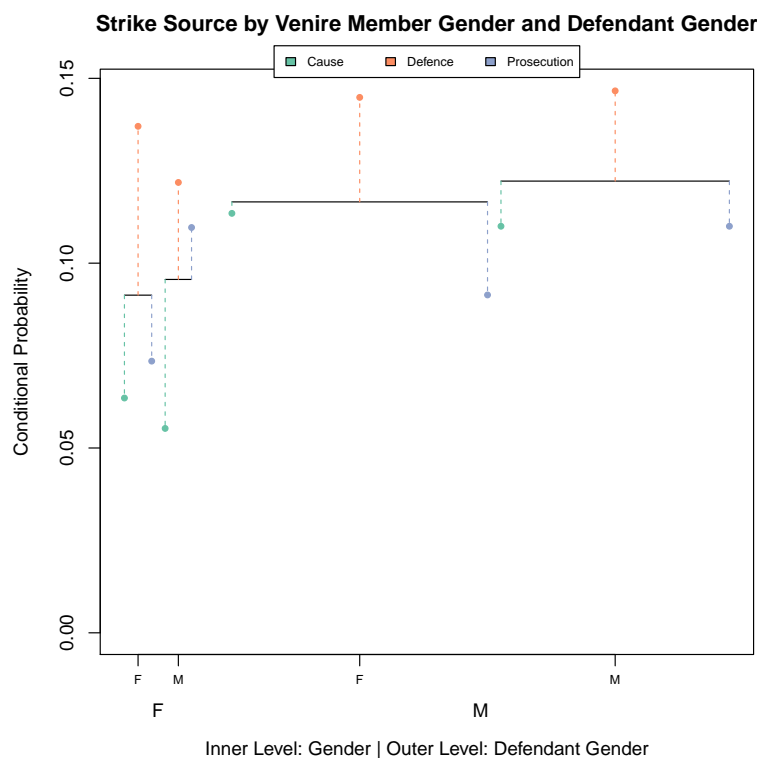


Figure 4.7: Conditional probabilities of strike source by venire member gender and defendant gender. The characteristic swap of the oppositional preferences cannot be seen here.

The characteristic switch is not present. At this point the message from these additional mobile plots should be clear. The dominant determinant of the strike probability for a venire member is their race. The race of the defendant does impact this somewhat, but plots across gender and political affiliation for both venire members and defendants in numerous combinations suggest that race is the dominant factor in determining the probability a venire member will be struck by the prosecution, defense, or removed by cause. This hypothesis, suggested heavily by these plots, is examined more rigorously through the construction and testing of specific models in 4.4.

If data on make-up use and jury outcome is ever collected perhaps Mr. Sparling's bold claim can be tested, but until then it is better treated as prejudice.

4.3.3 The Stubborn and Philadelphia Data Sets

Grosso and O'Brien (2012) and Baldus et al. (2001) did not collect data on the political affiliation of the venire members⁸, and Grosso and O'Brien (2012) did not record the gender of the defendant. Consequently, of the above plots, only Figure 4.6 can be compared to these two data sets. Figure 4.8 displays the plots. Note that as before, the data has been restricted to ensure a fair comparison across all data sets.

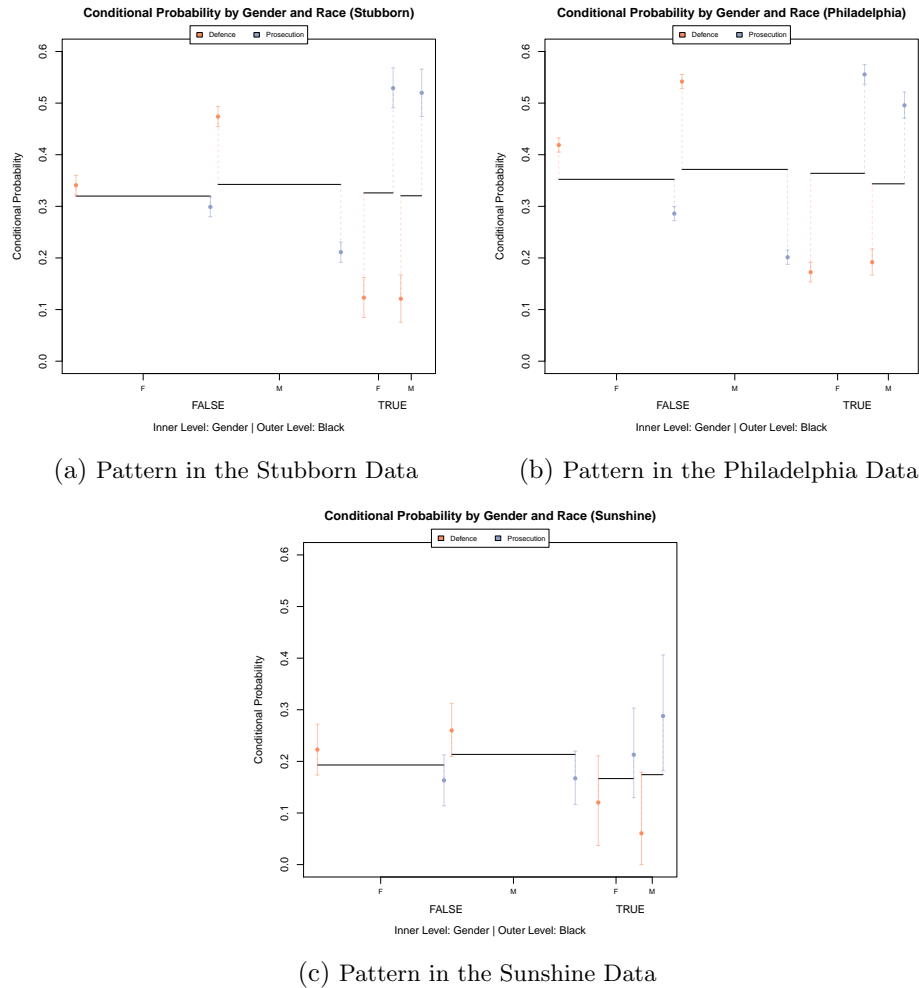


Figure 4.8: The conditional probability of defence and prosecution peremptory challenge by venire member race and gender for all capital trials in all data sets. The pattern, though sometimes different in magnitude, is quite consistent across the three examined data sets, despite the significant differences in the respective study sample universes.

This, again, shows a pattern very similar to the overall Sunshine data, and the differences in pattern between these data sets are minor. When displayed visually as above, the differences and similarities are immediately obvious.

⁸While it is perhaps possible that the different combinations of the attitudinal data collected by these groups could be used with data on the voter profiles most typically associated with the political affiliations of different groups in the United States, such an investigation is beyond the scope of this paper. Lacking such a predictor, it is also unclear which attitudinal variables would be most comparable to the political affiliation, and so attempting to compare this pattern would be futile.

4.4 Modelling

4.4.1 Multinomial Logistic Regression

Controlling for the myriad of possible confounding factors motivated the fitting of multiple models to test the significance and magnitude of the impact of certain variables on the probability of venire member rejection. Here, the use of the mobile plot in Figure 4.1 and others is suggestive of a model by design. The creation of a plot with conditional probabilities contingent on other factors is suggestive of a conditional multinomial distribution, and so fitting multinomial log-linear regression seemed a natural choice.

It should be noted that early modelling instead used Poisson regression, but as the conditional distribution of a particular outcome given some marginal count is multinomial (as shown in C), these two models are essentially equivalent. There are known transforms (Baker (1994)) to move between the models and their equivalencies are well documented (Lang (1996)). For greater interpretability, however, the fitting performed used multinomial models as implemented in the `nnet` package in R, which implements a method of fitting multinomial models discussed in Venables and Ripley (2002). While in Poisson models the interaction terms are the only quantities of interest, and the main effects merely serve as constraints, the multinomial fit leads to the more intuitive main effect importance for the different factors.

For all models, the pivot level chosen to fit odds ratios was the probability that a black female venire member with Democrat voting tendencies was kept in the final jury in a case with a black female defendant (using the notation from 0.2: the reference venire member combination is $\mathbf{x} = (1, 1, 1, 1, 1, 1)^T$). While the choice of race, gender, and political affiliation was not deliberate, the choice to use the probability of a venire member serving on the jury was motivated by the linear relationship between this quantity and the black line displayed in the mobile plots earlier. As the mobile plot horizontal lines have been placed at $p_r = \frac{\pi_2 + \pi_3 + \pi_4}{3}$, and $\pi_1 = 1 - \pi_2 + \pi_3 + \pi_4$, there is a linear relationship between π_1 , chosen as a reference, and the p_r displayed in the mobile plot.

The mobile plots created in 4.3 and 4.2 suggest, primarily, that an interaction between race and defendant race is relevant in modelling the conditional probability of venire member rejection. These plots do not suggest that any other interactions are likely to be significant. This led to a model:

$$\log \frac{\pi_d^{(i)}}{\pi_1^{(i)}} = a_d + \mathbf{x}_i^T \beta + (re)_i \beta_{re} = a_i + r_i \beta_r + e_i \beta_e + p_i \beta_p + g_i \beta_g + s_i \beta_s + (re)_i \beta_{re} \quad (4.4.1.1)$$

The fundamental purpose of fitting this model is to test two effects. The first is whether the interaction between defendant race and venire member race is significant, as has been suggested by the controversial trials named in 2, and the second is whether the venire member race is a significant factor in the presence of the other controlled factors. These two assumptions can be testing by fitting models nested in 4.4.1.1:

$$\log \frac{\pi_d^{(i)}}{\pi_1^{(i)}} = a_d + \mathbf{x}_i^T \beta \quad (4.4.1.2)$$

$$\log \frac{\pi_d^{(i)}}{\pi_1^{(i)}} = a_d + \mathbf{x}_i^T \beta' \quad (4.4.1.3)$$

Where $\beta'_r = 0$. Comparing 4.4.1.2 and 4.4.1.3 to 4.4.1.1 gives striking results. These can be seen in Table ??, which compares the deviances of the different models sequentially.

Table 4.3: Comparison of models 4.4.1.3 and 4.4.1.2 to 4.4.1.1, displaying the residual deviance, residual degrees of freedom, differences, and p-value of these differences for adjacent models.

Model	Residual df	Residual Deviance	Difference	$P(\chi^2)$
4.4.1.3	55527	39496		
4.4.1.2	55521	39087	405	0
4.4.1.1	55509	39023	67	1.4e-9

Even when controlling for defendant characteristics and the venire member’s political affiliation and sex, the race of the venire member and its interaction with the defendant race are both highly significant at the $\alpha = 0.05$ level. This suggests that the rejection of venire members is, at least in part, based on their racial characteristics. Whether this is due to opinions associated with race or not is impossible to say from this data, but the control of political affiliation provides a suggestion that commonly held opinions within racial groups are not the cause.

Such analysis is consistent with the modelling found in [Grosso and O’Brien \(2012\)](#) and [Baldus et al. \(2001\)](#). Both of these studies performed logistic regression analysis of the prosecution strike patterns, controlling for attitudinal variables. In both cases, the venire member race remained a significant predictor of the venire member removal. Additionally, [Baldus et al.](#) fit analogous models for the defence strike use and found the race remained significant even when the attitudinal variables were controlled. Together, these three data sets and analyses provide a strong case that the race of a venire member is utilized by both the defence and prosecution to make peremptory challenge decisions both in different regions and throughout time.

4.5 Case Level Summary

While [Wright et al. \(2018\)](#) reported a great deal of aggregate statistics about the venire members themselves, one piece of investigation which was lacking was an analysis which aggregated and viewed the trends for the cases, rather than simply for individual venire members. As we cannot know why a potential venire member is struck individually, and viewing their aggregate statistics tells us nothing about how different strikes relate to each other, it is possible we are viewing some effect which is not a result of persistent bias across trials, but is rather the result of some other effect.

By aggregating the venire members by trial and viewing the demographic trends in strikes and behaviour at this level, we gain a more detailed insight into the impact of challenges at a more relevant scale. Additionally, such aggregation allows for the synthesis of certain measures, such as a disitributional difference via the Kullback-Leibler divergence ([Kullback and Leibler \(1951\)](#)), which would otherwise not be well defined. This particular perspective of the data has also not been explored by any other studies known to the author.

4.6 Useful Quoted Bits

[Van Dyke \(1977\)](#) spends much of chapters four and five exploring the causes for the underrepresentation of certain groups in jury venires, and his analysis suggests that underrepresentation starts at the jury selection stage due to fewer non-white individuals registering to vote generally (page 89), and the process of applying to be excused from jury duty, in which economic hardship, which impacts disadvantaged economic groups to a greater extent (pages 113-120), is a common reason for excusal from jury duty.

Such explanations provide a plausible reason why black males would be most underrepresented in venires, and why the majority of the venire is white in this data despite the majority of defendants being black. Such issues with the jury selection process will not, and cannot, be solved by simply removed the peremptory challenge. They have much more to do with the relationship between certain groups and wider society, and so require more comprehensive and complex solutions.

Chapter 5

Summary

Summarize the presented work. Why is it useful to the research field or institute?

5.1 Future Work

Possible ways to extend the work.

Bibliography

- 42nd Parliament of Canada (2018a, March). Bill C-75: An Act to Amend the Criminal Code, Youth Criminal Justice Act and other Acts and to make consequential amendments to other Acts. <http://www.justice.gc.ca/eng/csj-sjc/pl/charte-charte/c75.html>.
- 42nd Parliament of Canada (2018b, November). Bill C75. LEGISinfo. <http://www.parl.ca/LegisInfo>.
- Attorney General's Office of the United Kingdom (2012, November). Jury vetting right of stand-by guidelines. <https://www.gov.uk/guidance/jury-vetting-right-of-stand-by-guidelines-2>.
- Baker, S. G. (1994). The multinomial-poisson transformation. *Journal of the Royal Statistical Society, Series D (The Statistician)* 43(4), 495–504.
- Baldus, D. C., G. Woodworth, D. Zuckerman, and N. A. Weiner (2001). The Use of Peremptory Challenges in Capital Murder Trials: A Legal and Empirical Analysis. *University of Pennsylvania Journal of Constitutional Law* 3(1).
- Brown, F. L., F. T. McGuire, and M. S. Winters (1978). The peremptory challenge as a manipulative device in criminal trials: Traditional use or abuse. *New England Law Review* 14, 192.
- Brown, R. B. (2000). Challenges for cause, stand-asides, and peremptory challenges in the nineteenth century. *Osgoode Hall Law Journal* 38(3), 453–494.
- Burgess, G. (1992). The divine right of kings reconsidered. *The English Historical Review* 107(425), 837–861.
- Cleveland, W. S. and R. McGill (1987). Graphical perception: The visual decoding of quantitative information on graphical displays of data. *Journal of the Royal Statistical Society, Series A (General)* 150(3), 192–229.
- Constitution of Canada (1982). Constitution of Canada. Accessed: <https://laws-lois.justice.gc.ca/eng/Const/index.html>.
- Constitution of the United States (1788). Constitution of the United States. Accessed: https://www.senate.gov/civics/constitution_item/constitution.htm.
- Ford, R. (2010). Modeling the effects of peremptory challenges on jury selection and jury verdicts. *George Mason Law Review* 17, 377.
- Forsyth, W. (1994). *History of Trial by Jury* (2 ed.). Lawbook Exchange.
- Friendly, M. (1994). Mosaic plots for multiway contingency tables. *Journal of the American Statistical Association* 89, 190–200.

- Government of Canada (1985). Criminal code section 634. <https://laws-lois.justice.gc.ca/eng/acts/C-46/section-634.html>.
- Government of Saskatchewan (1998). Jury act, 1998. Accessed: <http://www.qp.gov.sk.ca/documents/English/Statutes/Statutes/J4-2.pdf>.
- Grosso, C. M. and B. O'Brien (2012). A Stubborn Legacy: The Overwhelming Importance of Race in Jury Selection in 173 Post-Batson North Carolina Capital Trials. *Iowa Law Review* 97, 1531.
- Hans, V. P. and N. Vidmar (1986). *Judging the Jury* (1 ed.). Plenum Press.
- Harris, K. (2018, February). Liberals review jury selection process after Boushie case uproar. CBC News. <https://www.cbc.ca/news/politics/jury-selection-diversity-indigenous-1.4531792>.
- Hasan, N. R. (2018, April). Eliminating peremptory challenges makes trials less fair. The Star. <https://www.thestar.com/opinion/contributors/2018/04/10/eliminating-peremptory-challenges-make-trials-less-fair.html>.
- Hoffman, M. B. (1997). Peremptory Challenges Should Be Abolished: A Trial Judge's Perspective. *The University of Chicago Law Review* 64(3), 809.
- Iacobucci, F. (2013). First Nations Representation on Ontario Juries: Report of the Independent Review Conducted by the Honourable Frank Iacobucci. Ministry of the Attorney General. Accessed: https://www.attorneygeneral.jus.gov.on.ca/english/about/pubs/iacobucci/First_Nations_Representation
- Kullback, S. and R. Leibler (1951). On information and sufficiency. *The Annals of Mathematical Statistics* 22(1), 79–86.
- Lang, J. B. (1996). On the comparison of multinomial and poisson log-linear models. *Journal of the Royal Statistical Society, Series B (Methodological)* 58(1), 253–266.
- MacLean, C. (2018, February). Gerald Stanley acquittal renews calls for justice reform 27 years after Manitoba inquiry. CBC News. <https://www.cbc.ca/news/canada/manitoba/aboriginal-justice-inquiry-colgen-boushie-gerald-stanley-jury-1.4532394>.
- Macnab, A. (2018, February). Stanley acquittal should not lead to scrapping peremptory challenges, say criminal lawyers. Canadian Lawyer. <https://www.canadianlawyermag.com/legalfeeds/author/aidan-macnab/stanley-acquittal-should-not-lead-to-scrapping-peremptory-challenges-say-criminal-lawyers-15332/>.
- Ministry of the Attorney General of Ontario (2018). The annual jury selection process. Queen's Printer for Ontario. Accessed: https://www.attorneygeneral.jus.gov.on.ca/english/courts/jury/jury_selection_process.php.
- Miriam-Webster (2019a). Miriam-Webster Dictionary Online. Accessed: <https://www.merriam-webster.com/dictionary/venire>.
- Miriam-Webster (2019b). Miriam-Webster Dictionary Online. Accessed: <https://www.merriam-webster.com/dictionary/voir%20dire>.

- Morehead, J. W. (1994). When a peremptory challenge is no longer peremptory: Batson's unfortunate failure to eradicate invidious discrimination from jury selection. *DePaul Law Review* 43, 625.
- Nisbett, R. E. and Z. Kunda (1985). Perception of social distributions. *Journal of Personality and Social Psychology* 48(2), 297–311.
- North Carolina Sentencing and Policy Advisory Commission (2017). *Classification of a Sample of Offenses*. North Carolina Judicial Branch. Accessed: <https://www.nccourts.gov/assets/documents/publications/Sample-list-2017.pdf?MAZluXuS0FWod4nFt7zXecJ8Ifu0qEZx>.
- Page, A. (2005). Batson's blind spot: Unconscious stereotyping and the peremptory challenge. *Boston University Law Review* 85, 155.
- Petersen, C. (1993). Insitutionalized racism: The need for reform of the criminal jury selection process. *McGill Law Journal* 38(1).
- Quenneville, G. (2018, February). What happened on Gerald Stanley's farm the day Colten Boushie was shot, as told by witnesses. CBC News. <https://www.cbc.ca/news/canada/saskatoon/what-happened-stanley-farm-boushie-shot-witnesses-colgen-gerald-1.4520214>.
- Quenneville, G. and J. Warick (2018, February). Shouts of 'murderer' in courtroom after Gerald Stanley acquitted in Colten Boushie shooting. CBC News. <https://www.cbc.ca/news/canada/saskatoon/gerald-stanley-colgen-boushie-verdict-1.4526313>.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Revesz, J. (2016). Ideological imbalance and the peremptory challenge. *Yale Law Journal* 125.
- Richardson Oakes, A. and H. Davies (2016). Justice must be seen to be done: a contextual reappraisal. *Adelaide Law Review* 37(2), 461–494.
- Roach, K. (2018, April). Ending peremptory challenges in jury selection is a good first step. The Ottawa Citizen. <https://ottawacitizen.com/opinion/columnists/roach-ending-peremptory-challenges-in-jury-selection-is-a-good-first-step>.
- Sheppard, R. (2018). Patriation of the constitution. The Canadian Encyclopedia: Historical Canada. Accessed: <https://www.thecanadianencyclopedia.ca/en/article/patriation-of-the-constitution>.
- Sison, C. P. and J. Glaz (1995). Simultaneous confidence intervals and sample size determination for multinomial proportions. *Journal of the American Statistical Association* 90(429), 366–369.
- Statistics Canada (2018, November). Table 35-10-0061-01: Crime severity index and weighted clearance rates, police services in Saskatchewan.
- Supreme Court of Canada (1991). *R. v. sherratt*. Supreme Court Judgments. SCC Case Number: 21501; Accessed: <https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/734/index.do?q=21501>.

- Supreme Court of the United States (1965). *Swain v. Alabama*. Accessed: <https://supreme.justia.com/cases/federal/us/380/202/>.
- Supreme Court of the United States (1986). *Batson v. Kentucky*. Accessed: <https://www.law.cornell.edu/supremecourt/text/476/79>.
- Supreme Court of the United States (1993). *J.E.B. v. Alabama*. Accessed: <https://supreme.justia.com/cases/federal/us/511/127/>.
- Tufte, E. R. (2001). *The Visual Display of Quantitative Information* (2 ed.). Graphics Press.
- Van Dyke, J. M. (1977). *Jury Selection Procedures: Our Uncertain Commitment to Representative Panels* (1 ed.). Ballinger Publishing.
- Venables, W. N. and B. D. Ripley (2002). *Modern Applied Statistics with S* (Fourth ed.). New York: Springer. ISBN 0-387-95457-0.
- von Moschzisker, R. (1921). The historic origin of trial by jury. *University of Pennsylvania Law Review* 70(1).
- Wegman, E. J. (1990). Hyperdimensional analysis using parallel coordinates. *Journal of the American Statistical Association* 85(411), 664–675.
- Weisstein, E. W. (2018). Chebyshev inequality. MathWorld - A Wolfram Web Resource. Accessed: <http://mathworld.wolfram.com/ChebyshevInequality.html>.
- Wickham, H. and J. Bryan (2018). *readxl: Read Excel Files*. R package version 1.1.0.
- Woolley, A. (2018). An Ethical Jury? Reflections on the Acquittal of Gerald Stanley for the Murder/Manslaughter of Colten Boushie. *Slaw: Canada's online legal magazine*. Accessed: <http://www.slaw.ca/2018/02/20/an-ethical-jury-reflections-on-the-acquittal-of-gerald-stanley-for-the-murder-manslaughter-of-colten-boushie/>.
- Wright, R. F., K. Chavis, and G. S. Parks (2018, October). The Jury Sunshine Project: Jury Selection Data as a Political Issue. *University of Illinois Law Review* 2018(4), 1407.
- Zinchuk, B. (2018, March). Both sides wrong about Stanley trial. Prince George Citizen. <https://www.princegeorgecitizen.com/opinion/editorial/both-sides-wrong-about-stanley-trial-1.23199321>.

Appendix A

Developing an Effective Visualization of Conditional Probability

One deficiency of the results of the previous investigations was a failure to generate compelling and effective visualizations of the trends of peremptory challenges for different racial groups. While such visualizations are not necessarily critical to analysis, they can often be incredibly useful to not only communicate data, but to motivate further investigations and models in a way which is clearer and more intuitive than a simple table of values.

The first attempt at such a visualization was the mosaic plot (as discussed by [Friendly \(1994\)](#)) using the `mosaicplot` function in the `graphics` package in R ([R Core Team \(2018\)](#)). Figure [A.1](#) displays this first approach with disposition related to the simplified races of both the defendant and the venire member.

This visualization suffers from a number of limitations, some of which are obvious, and others of which are best explained by the hierarchy of accuracy of visual perception provided in [Cleveland and McGill \(1987\)](#). The obvious limitations are the lack of ability to perceive the differences for the smallest groups, which are compressed enough that their error is nearly imperceptible. Additionally, the ordering of the axes is incredibly important in how the different areas appear visually, and comparing the different areas is unclear if any specific comparisons are to be made.

This may be somewhat unsurprising. [Cleveland and McGill \(1987\)](#), in their ranking of visual displays by accuracy of perception place area low in the hierarchy, below angles, lengths, and positions along common scales. In *The Visual Display of Quantitative Information*, [Tufte](#) gives two more sources of possible criticism of the mosaic plot as displayed in Figure [A.1](#): the concept of data-ink and the dimensionality of visualization.

Of the mosaic plot, one may ask how much of the “ink”, or structure, on the page is necessary to communicate the information present. If one has a desire to “above all else show the data” as Tufte does, then these large shaded rectangles, which are likely not perceived accurately according to [Cleveland and McGill](#), seem unnecessary compared to a simpler visualization. This is the concept of “data-ink,” to reduce the complexity of the structures and chart used to display the data.

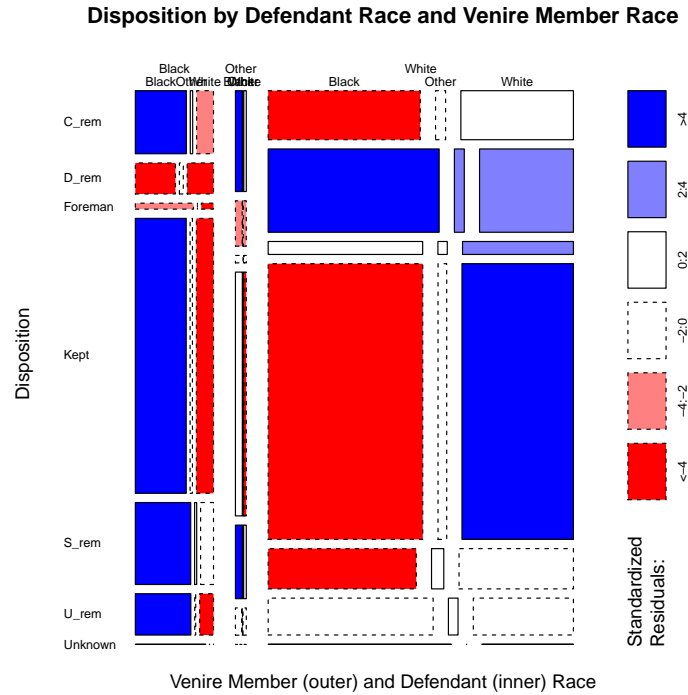


Figure A.1: A mosaic plot of the simplified defendant and venire member race and their relation to the disposition of the venire member.

Hand-in-hand with this concept for this plot is [Tufte's](#) rule that the dimensionality of the visualization should not be larger than the data. In the case of the mosaic plot this is not strictly violated, as the marginal lengths used to create the areas reflect a measurement of the data. Nonetheless, the areas of each rectangle correspond to a simple count in a contingency table, and perhaps an area is not the best way to represent such a singular value.

Motivated by these concepts, parallel coordinates (as in [Wegman \(1990\)](#)) were used to visualize the data next, as can be seen in [Figure A.2](#). This attempted visualization is arguably more difficult to interpret than the mosaic plot. It is cluttered by the parallel coordinate lines, the bars emanating from each point obscure the fact that the end point of the bar is the only feature of interest, and the meaning of the black reference line is entirely unclear without extensive explanation. Finally, by viewing the distribution of each disposition, the wrong conditional density is being examined, $P(Race, Race_{Defendant} | Disposition)$. Multiple edits and re-conceptualizations of the concept eventually resulted in [Figure ??](#), which will be called the “mobile plot” due to its passing resemblance to the mobiles hung above babies’ cribs.

An example of this plot can be seen in [4.1](#). Note that this plot is less cluttered than either the mosaic plot or the first parallel coordinate plot, despite displaying more information. It is also more efficient with data-ink, avoids displaying data with higher dimensions than the data itself, and uses redundant encoding of information in visual cues which are high in the hierarchy presented by [Cleveland and McGill \(1987\)](#).

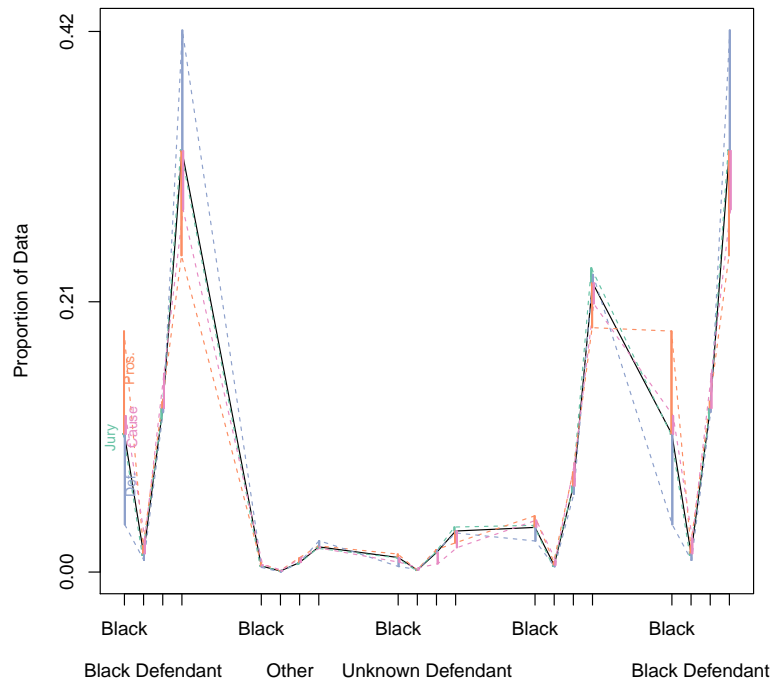


Figure A.2: The first attempt at a parallel coordinate plot attempted. Note that the cramped display and unclear definition of the axis make interpretation even less intuitive than the mosaic plot, suggesting that this first attempt was a decided failure.

An explanation of the features and encoding used in the mobile plot is presented in [A.1](#).

A.1 The Mobile Plot

The mobile plot consists of multiple grouped vertical lines anchored at one end to horizontal black lines, and at the other to points. Information is encoded using length, colour, and position relative to a common scale. The vertical axis is meant to show the value of a continuous variable, while the horizontal axis shows the value of a, possibly hierarchical, categorical variable. It can be used to display the relationship between three categorical variables and a continuous variable in a meaningful two-dimensional plot.

To show the grouping of categories on the horizontal axis, position is used. Those categorical levels which are grouped by some separate categorical variable are placed closer to each other than those which are not in the same group. Each categorical variable combination corresponds to a single horizontal black line, the length of which is proportional to the count of the associated combination in the data being plotted. The vertical position of this line corresponds to the value of the continuous variable expected for that particular combination.

Each of the vertical lines which extend from this horizontal line corresponds to a particular value of a third categorical variable, coloured to show the specific level across the

different horizontal lines. The end points of these lines represent the observed value of the continuous variable for the three way combination of categorical variables represented by the vertical and horizontal line combination. The lengths of these lines correspond to the deviation of the observation from the expectation. If a different expectation is expected for the different values of the third categorical variable, the horizontal lines can be split evenly and placed vertically at this expectation, to the detriment of grouping clarity.

In the case that such a split is not used and the continuous variable is the probability of a particular value of the third categorical variable given the first two, the plot serves as a visual test of a very specific hypothesis: that of a uniform distribution of the third categorical variable with respect to the two variables represented horizontally. Such a plot is powerful because it allows for the simple detection of main effects and interaction effects over the three categorical variables against this hypothesis.

Appendix B

Complementary information

B.1 Prosecution and Judge Homogeneity

As noted in 4.2, the prosecution and judge seem to match in their strike tendencies for every combination. This suggests that both challenges with cause and the prosecution tend to have the same effect on the jury composition, though the magnitudes can differ greatly for these two strikes. An immediate explanation to this is offered by [Hans and Vidmar \(1986\)](#), who outline, on pages 69-70, the skill and tact required to effectively propose challenges with cause. In order to determine an individual's bias, it is frequently the case that a direct question will fail to garner an honest response due to social pressures. As a consequence, the questions asked of venire members must be carefully presented.

Using this as a motivation, an obvious possible explanation for the challenges with cause is that the prosecution is simply more experienced on average than the defence. To determine the veracity of this claim, the licensing year of each lawyer was subtracted from the outcome date of each trial. The resulting distribution of years of experience was then plotted in back-to-back histograms as shown in Figure B.1.

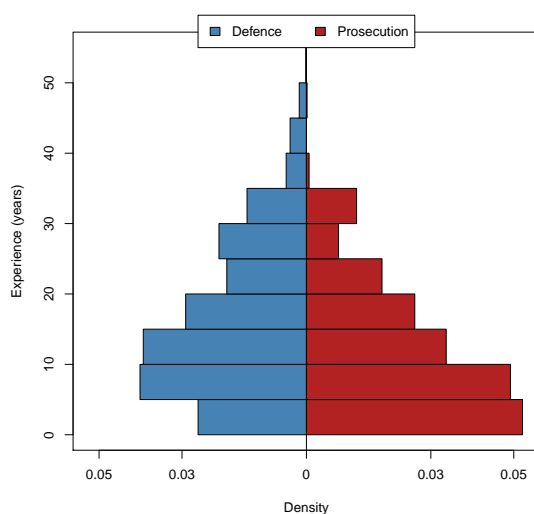


Figure B.1: Distributions of lawyer experience for prosecutors and defence attorneys

Clearly, this hypothesis is not supported by the data. It seems the typical defence lawyer is more experienced than the typical prosecutor, not less. Indeed, the prosecutors seem to be much more likely to be inexperienced than the defence lawyers. This suggests that the cause of this matching departure from the expected independent hypothesis has some different cause. Future investigation may endeavour to answer this question.

B.2 Jury Sunshine Irregularities

Table B.1: Jury sunshine data irregularities noted in data flattening

Charges without trial (ACISID)	08CRS50940, 09CRS1106, 10CRS051975, 10CRS51388, 11CRS051642, 11CRS1745, 11CRS51895, 08CRS50113	08CRS52888, 09CRS50752, 10CRS1215, 10CRS51610, 11CRS051795, 11CRS1783, 11CRS52470,	09CRS000305, 10CRS52031, 10CRS397, 10CRS52410, 11CRS1577, 11CRS51204, 08CRS54836,
Prosecutors without trials (IDs)	1-000, 11B-000, 12-000, 14-000, 15B-000, 16A-000, 16B-000, 17A-000, 17B-000, 19A-000, 19B-000, 20A- 000, 20B-000, 21-000, 22A-000, 22B-000, 24-000, 25- 000, 27A-000, 27B-000, 28-000, 29A-000, 29B-000, 30-000, 6-000, 9-000		
Trial missing charge (ID)	710-01		

B.3 Jury Sunshine Charge Classification

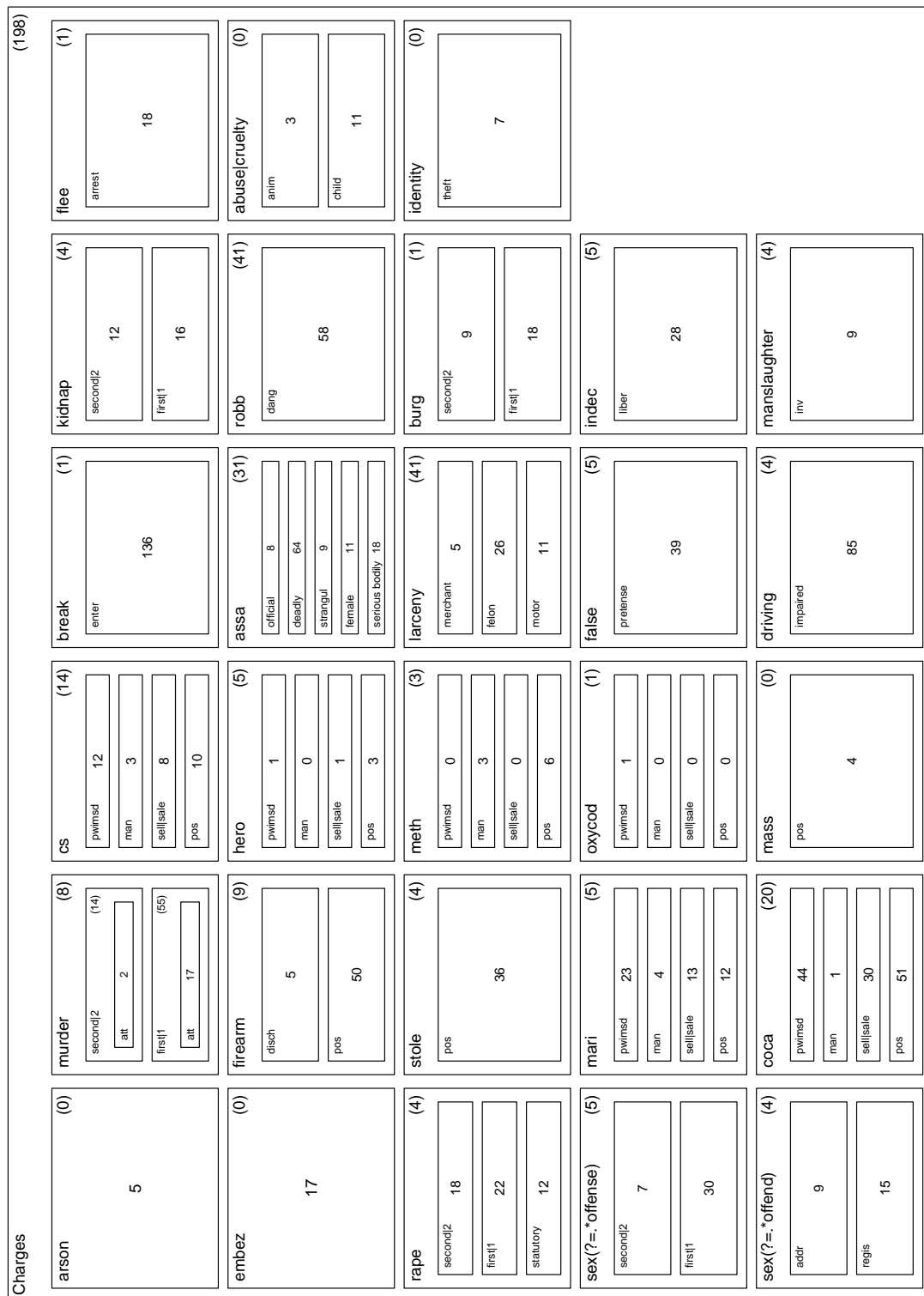


Figure B.2: The regular expression charge tree arranged by hierarchy with counts provided. The counts in brackets indicate the counts of charges which could not be classified to a lower level of the hierarchy

Appendix C

Mathematical Results

C.1 Conditional Distribution of a Poisson Expectation Given Marginal Counts

For simplicity, consider a Poisson random variable Y with a rate dependent only on one discrete random variable: $X \in \{1, \dots, m\}$. Let the count of values observed for $X = i$ be Y_i and denote $E[Y_i|X = i] = \lambda_i$. Additionally denote the sum of all counts as $N = \sum_{i=1}^m Y_i$. The feature of interest is then the distribution of $Y_1, \dots, Y_m|N = n$. Or by the definition of the conditional distribution:

$$P(Y_1 = y_1, \dots, Y_m = y_m|N = n) = \frac{P(Y_1 = y_1, \dots, Y_m = y_m, \sum_{i=1}^m Y_i = n)}{P(\sum_{i=1}^m Y_i = n)} \quad (\text{C.1.0.1})$$

Clearly this density is zero if $\sum_{i=1}^m y_i \neq n$, but consider its value with for $\sum_{i=1}^m y_i = n$. Start with the distribution of $N = \sum_{i=1}^m Y_i$. Note that for $A \sim \text{Pois}(\lambda_A)$ and $B \sim \text{Pois}(\lambda_B)$, where A and B are independent, the distribution of $A + B$ can be derived quite easily using the characteristic function $\varphi_{A+B}(t)$:

$$\varphi_{A+B}(t) = E[e^{it(A+B)}] = E[e^{itA}]E[e^{itB}] = e^{(\lambda_A + \lambda_B)(e^{it} - 1)}$$

This is the characteristic function of a $\text{Pois}(\lambda_A + \lambda_B)$ variable, and so the sum of two Poisson random variables is a Poisson random variable with a rate corresponding to the sum of the two variables. Iterating this, then, one obtains $N \sim \text{Pois}(\sum_{i=1}^m \mu_i)$, and so the denominator of C.1.0.1 is:

$$P\left(\sum_{i=1}^m Y_i = n\right) = \frac{e^{-\sum_{i=1}^m \mu_i} (\sum_{i=1}^m \mu_i)^n}{n!}$$

Additionally, recognizing that the Y_{ij} are independent, and considering only the case where $\sum_{i=1}^m y_i = n$, as the density is zero otherwise, this can be further simplified, as this joint density can be split into a product of marginal densities:

$$P\left(Y_1 = y_1, \dots, Y_m = y_m, \sum_{i=1}^m y_i = n\right) = P(Y_1 = y_1)P(Y_2 = y_2) \dots P(Y_m = y_m)$$

Now each independent marginal is Poisson distributed, so the product of all of these marginals is:

$$P(Y_1 = y_1)P(Y_2 = y_2) \dots P(Y_m = y_m) = \frac{e^{\sum_{i=1}^m \mu_i} \mu_1^{y_1} \mu_2^{y_2} \dots \mu_m^{y_m}}{y_1! y_2! \dots y_m!}$$

And so C.1.0.1 simplifies to

$$\begin{aligned} & \frac{e^{\sum_{i=1}^m \mu_i} \mu_1^{y_1} \mu_2^{y_2} \dots \mu_m^{y_m}}{y_1! y_2! \dots y_m!} \cdot \frac{n!}{e^{\sum_{i=1}^m \mu_i} \left(\sum_{i=1}^m \mu_i\right)^n} \\ &= \frac{m!}{y_1! y_2! \dots y_m!} \left(\frac{\mu_1}{\sum_{i=1}^m \mu_i}\right)^{y_1} \left(\frac{\mu_2}{\sum_{i=1}^m \mu_i}\right)^{y_2} \dots \left(\frac{\mu_m}{\sum_{i=1}^m \mu_i}\right)^{y_m} \end{aligned} \quad (\text{C.1.0.2})$$

C.1.0.2 is recognizably the multinomial distribution, where the probability of a particular class i is given by the ratio of μ_i to the sum over all μ_j . ■

Appendix D

Code

D.1 Data Processing Code

The data processing functions and code are available through the author's GitHub (<https://github.com/Salahub/peremp>), where the processed data is also provided.

D.2 Analysis Code

As for the data processing code, this code is posted on the author's GitHub (<https://github.com/Salahub/peremp>).

Epilogue

A few final words.

Declaration of Originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor .

Title of work (in block letters):

SEEN TO BE DONE
A GRAPHICAL EXPLORATION OF PEREMPTORY CHALLENGE

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):

SALAHUB

First name(s):

CHRISTOPHER

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the Citation etiquette information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work .
- I am aware that the work may be screened electronically for plagiarism.
- I have understood and followed the guidelines in the document *Scientific Works in Mathematics*.

Place, date:

Zürich, Feb. 21

Signature(s):

Christopher Salahub

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.