BOSTON UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

THE ROLE OF RATIONAL CHOICE IN EXPLAINING
ELECTRICITY CONSUMPTION AND
JURY COMPOSITION

by

JEREMY BLAIR SMITH
Bac.Soc.Sc. (Hon.), University of Ottawa, 2002
M.A., Queen’s University, 2003

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
2013
I would like to express my deepest gratitude to Kevin Lang for his advice, encouragement, kindness, and, not least of all, his patience. I would also like to express my sincere gratitude to Marc Rysman and Johannes Schmieder for their advice and support. And I would like to express my gracious thanks to Michael Manove, Claudia Olivetti, and Hsueh-Ling Huynh for their guidance and kindness.

Thank you to my co-authors, Jee, Dave, and Katrina, for their patience with me and the inspiration and support they have given me.

Finally, thank you to all of my family and friends for their generous support and encouragement.
THE ROLE OF RATIONAL CHOICE IN EXPLAINING
ELECTRICITY CONSUMPTION AND
JURY COMPOSITION

JEREMY BLAIR SMITH

Boston University Graduate School of Arts and Sciences, 2013

Major Professor: Kevin Lang, Professor of Economics

ABSTRACT

I test for the presence of economically rational behavior in two non-standard settings.

The choice of how much electricity to consume at what times of day – the focus of Chapters 1 and 2 – is complicated by the fact that consumers have limited information on the rates they are charged and the quantities they consume. I study a policy that forced residential customers onto a time-of-use electricity tariff – which charges a higher rate on weekday afternoons and evenings, and a lower rate at all other times – if they breached a monthly usage threshold. I find that the reform led to a lower effective electricity price for some consumers, but that these consumers responded by decreasing their total electricity usage, exhibiting a clear violation of the standard model of utility maximization. I also find that high-intensity users displayed larger responses to the reform than low-intensity users.

The choice of which potential jurors to strike during jury selection – the focus of Chapters 3 and 4 – is complicated by the fact that attorneys have limited
information on the predispositions of the panelists. Using data from felony trials in four large and diverse counties, I find that blacks and jurors with lower incomes are more likely to hold opinions and sympathies favoring the defense, and that juries with a greater proportion of such jurors acquit more often. Giving attorneys more freedom in jury selection can allow them to strike panelists according to these stereotypes more easily, but it can also allow them to identify the panelists’ predispositions more accurately. I find that skilled and empowered attorneys successfully retain jurors predisposed to their side at a greater rate, confirming that they are responding rationally to the constraints they face. However, I find no effect of attorney empowerment or skill on the socio-demographic composition of juries, suggesting that they do not necessarily accomplish this by striking according to stereotypes alone.

This work demonstrates how a better understanding of choice behavior can inform policies in such divergent and significant areas as energy conservation and criminal justice.
Contents

1 Are Residential Electricity Consumers Utility Maximizers? Evidence from a Natural Experiment 1
   1.1 Introduction ................................................. 1
   1.2 Theoretical Framework ................................. 4
   1.3 Program Design ..................................... 6
   1.4 Experimental Setting and Data ......................... 9
   1.5 Treatment Effects for Total Usage and Total Bills .......... 13
      1.5.1 Methods ......................................... 13
      1.5.2 Preliminary Evidence ............................ 15
      1.5.3 Treatment Effects .............................. 17
   1.6 Interpretation ..................................... 23
   1.7 Conclusion ......................................... 29

2 Understanding Residential Response to Time-of-Use Electricity Pricing 31
   2.1 Introduction ............................................. 31
   2.2 Setting and Data .................................... 33
   2.3 Design ............................................... 35
   2.4 Baseline Results .................................... 38
   2.5 Treatment Effects by Household Characteristics ........... 43
   2.6 Responses to a Non-Price Treatment .................... 50
   2.7 Discussion ........................................... 54

3 A Multidimensional Examination of Jury Composition, Trial Outcomes, and Attorney Preferences 56
   3.1 Introduction ............................................ 56
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Electricity Rates, 2008, Cents per kWh</td>
<td>7</td>
</tr>
<tr>
<td>1.2</td>
<td>Summary Statistics, July of the Qualification Period</td>
<td>13</td>
</tr>
<tr>
<td>1.3</td>
<td>Treatment Effects (%), Total Usage and Total Bill</td>
<td>22</td>
</tr>
<tr>
<td>2.1</td>
<td>Electricity Rates, 2010, Cents per kWh</td>
<td>34</td>
</tr>
<tr>
<td>2.2</td>
<td>Treatment Effects (%), Total Usage and Total Bill</td>
<td>42</td>
</tr>
<tr>
<td>3.1</td>
<td>Case Summary Statistics</td>
<td>71</td>
</tr>
<tr>
<td>3.2</td>
<td>Pre-Deliberation Juror Opinions with Case Fixed Effects</td>
<td>75</td>
</tr>
<tr>
<td>3.3</td>
<td>Summary of Judge Opinions</td>
<td>81</td>
</tr>
<tr>
<td>3.4</td>
<td>The Effect of Average Jury Characteristics on Verdicts</td>
<td>84</td>
</tr>
<tr>
<td>3.5</td>
<td>Attorney Voir Dire Satisfaction</td>
<td>94</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary Statistics</td>
<td>142</td>
</tr>
<tr>
<td>4.2</td>
<td>Prediction #1 (Ordered Logistic Regressions)</td>
<td>145</td>
</tr>
<tr>
<td>4.3</td>
<td>Prediction #1 (OLS Regressions)</td>
<td>149</td>
</tr>
<tr>
<td>4.4</td>
<td>Alternative Measures of Attorney Empowerment</td>
<td>150</td>
</tr>
<tr>
<td>4.5</td>
<td>Determinants of Attorney Skill</td>
<td>153</td>
</tr>
<tr>
<td>4.6</td>
<td>Prediction #2 (Logistic Regressions)</td>
<td>156</td>
</tr>
<tr>
<td>4.7</td>
<td>Number of Strikes (OLS Regressions)</td>
<td>159</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Density of the Forcing Variable at the 4000 kWh Threshold . . . . . . 10
1.2 Intent to Treat Effect, Propensity to be Treated, July 2008 . . . . . . 16
1.3 Intent to Treat Effect, Total Bill, July 2008 . . . . . . . . . . . . . . 17
1.4 Intent to Treat Effect, Total Usage, July 2008 . . . . . . . . . . . . 18
1.5 Intent to Treat Effects by Month, Propensity to be Treated . . . . . 20
1.6 Intent to Treat Effects by Month, Total Bill . . . . . . . . . . . . . 20
1.7 Intent to Treat Effects by Month, Total Usage . . . . . . . . . . . . 21
1.8 Budget Lines, Unbundled Rates . . . . . . . . . . . . . . . . . . . . 24
1.9 Budget Lines, Bundled Rates . . . . . . . . . . . . . . . . . . . . . 25

2.1 Intent to Treat Effect, Propensity to be Treated, April 2010 . . . . 36
2.2 Intent to Treat Effects by Month, Propensity to be Treated . . . . . 37
2.3 Density of the Forcing Variable at the 2000 kWh Threshold . . . . . 38
2.4 Intent to Treat Effect, Total Usage, April 2010 . . . . . . . . . . . . 39
2.5 Intent to Treat Effects by Month, Total Usage . . . . . . . . . . . . 40
2.6 Intent to Treat Effects by Month, Total Bill . . . . . . . . . . . . . 40
2.7 Frequency of Households by Energy Intensity . . . . . . . . . . . . . 45
2.8 Frequency of Households by Presence of Pool . . . . . . . . . . . . . 45
2.9 Treatment Effects, Total Usage, High-Intensity Households . . . . . 47
2.10 Treatment Effects, Total Usage, Low-Intensity Households . . . . . 47
2.11 Mean Usage by Energy Intensity . . . . . . . . . . . . . . . . . . . . 49
2.12 Treatment Effects, Total Usage, Households with Pools . . . . . . . 51
2.13 Treatment Effects, Total Usage, Households with No Pool . . . . . . 51
2.14 Warning Letter Treatment Effects, Total Usage . . . . . . . . . . . 53
3.1 Difference in Voir Dire Satisfaction within a Case  

A.1 Sensitivity to Bandwidth at the 4000 kWh Threshold  

A.2 Peak-to-Off-Peak Usage Ratio for Non-TOU Households
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SLS</td>
<td>Two-Stage Least Squares</td>
</tr>
<tr>
<td>EDC</td>
<td>Electric Distribution Company</td>
</tr>
<tr>
<td>FRD</td>
<td>Fuzzy Regression Discontinuity</td>
</tr>
<tr>
<td>ICPSR</td>
<td>Interuniversity Consortium for Political and Social Research</td>
</tr>
<tr>
<td>ITT</td>
<td>Intent to Treat</td>
</tr>
<tr>
<td>kWh</td>
<td>Kilowatt Hour</td>
</tr>
<tr>
<td>NCSC</td>
<td>National Center for State Courts</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary Least Squares</td>
</tr>
<tr>
<td>SJQ</td>
<td>Supplemental Juror Questionnaire</td>
</tr>
<tr>
<td>TOU</td>
<td>Time of Use</td>
</tr>
<tr>
<td>WARP</td>
<td>Weak Axiom of Revealed Preference</td>
</tr>
</tbody>
</table>
CHAPTER ONE

Are Residential Electricity Consumers Utility Maximizers?
Evidence from a Natural Experiment

1.1 Introduction

A major policy goal in coming decades is to reduce greenhouse gas emissions. Most economists, including ourselves, favor market-based approaches due to their ability to achieve an emissions target cost effectively. However, the strength of such instruments relies partly on the assumption that consumers will respond optimally to prices. But a growing body of evidence suggests that within the energy choice setting, consumer behavior does not strictly adhere to the predictions derived from standard models. In some cases, non-monetary incentives such as moral license or pressure to conform to social norms can dominate financial incentives. Altruism and green identity also play important roles, with environmental concerns becoming a relevant aspect of consumer decisions. In other instances, features of the setting such as complex pricing structures or noisy signals about consumption may limit responses to prices. Understanding when and under what conditions households respond to prices as standard theory would predict is crucial to achieve climate change mitigation objectives and efficiency in

---

1This chapter is the product of joint work with Katrina Jessoe, of the University of California, Davis, Department of Agricultural and Resource Economics, and David Rapson, of the University of California, Davis, Department of Economics.

2Voluntary enrollees in a carbon offset program increase their electricity consumption despite also facing higher prices (Harding and Rapson, 2012), and customers informed of their neighbors’ electricity usage respond by using less themselves (Allcott, 2011).

3Kotchen and Moore (2007) provide evidence that environmental concern and altruistic attitudes are correlated with green program enrollment.

4Consumers facing an increasing-block electricity rate structure appear to respond more to average price than marginal price (Ito, 2011), and high frequency information about real-time consumption increases price elasticity of electricity demand (Jessoe and Rapson, 2012).
energy markets. It is thus valuable to document consumer behavior in different settings, and develop and test hypotheses about what drives observed behavior.

In this chapter, we document an instance in which households do not respond to a retail electricity price intervention as standard theory would predict. Our empirical setting offers a unique opportunity to test whether consumers are static utility maximizers, and we find conclusively that in this instance they are not. The results suggest that there may be risk in adhering too ideologically to price interventions in terms of missing policy goals or achieving them only imperfectly or inefficiently; an assertion that price incentives always work can be disproven by the counter-example we provide. This highlights the question of how generally we can expect the utility maximization assumption to hold in this setting, and the potential of price alone to change consumer behavior.

We partnered with an electricity distribution company (EDC) in the northeast US to evaluate a large-scale mandatory residential time-of-use (TOU) program that forced households to switch irrevocably from a flat-rate tariff to a TOU tariff after breaching a monthly usage threshold. The setting gives rise to a regression discontinuity framework in which we compare outcomes of households just above the usage threshold to those of households falling just below the cutoff. Due to customers’ inability to perfectly control monthly usage, in the neighborhood of the usage threshold assignment to the TOU rate is as good as random. The large-scale deployment of the program exhibits a high density around the threshold, creating a large sample of treatment and control households.

\footnote{TOU electricity pricing divides electricity use into two blocks according to the time of day at which electricity is consumed, and applies a higher rate to the block corresponding to historically high-cost times. It is a small step towards aligning retail electricity prices with marginal production costs. It is also the most common corrective measure used by electricity regulators to achieve such an alignment, due largely to the crucial advantage of being easy for consumers to understand and, in principle, respond to.}
with which to test our hypothesis.

In the first summer months of the program in 2008, TOU rates were low relative to the flat rate alternative. Whereas the standard formulation of TOU prices is for the on-peak rate to be substantially higher than the flat rate and the off-peak rate substantially lower, in our setting TOU households faced on-peak rates in June to September of 2008 that were either lower than the relevant flat rate, or only slightly higher. Off-peak rates were correspondingly even lower. The financial incentives for TOU households are clear: total electricity use in those months – regardless of substitution patterns across on-peak and off-peak hours – should increase.

We find the opposite. TOU customers reduced total electricity consumption, as measured by our estimates of the treatment effect at the threshold. These results are inconsistent with static utility maximization. It is clear that households are responding to incentives other than contemporaneous prices, and there are ways to rationalize their choices. We discuss these but do not perform formal tests, which our data cannot support. Nevertheless, the simple documentation of this result is an important step towards understanding energy demand behavior, and how future policy interventions might be improved.

The chapter is organized as follows: in section 1.2 we present a review of the standard theoretical framework for demand in this setting; Section 1.3 describes the program design, which forms the basis for our empirical setting; we explain how the setting can be viewed as a natural experiment in Section 1.4, which also includes a description of our data; treatment effects are reported in

---

Customers may purchase the generation component of their electricity services from either our EDC partner or an alternate supplier. This choice affects the relative on-peak and off-peak prices (the “TOU gradient”). In the discussion below we demonstrate why this does not affect our conclusions.
Section 1.5 and interpreted in Section 1.6; and Section 1.7 concludes.

1.2 Theoretical Framework

Several studies assessing early TOU experiments characterized consumer preferences using a framework that has since become the standard for modeling short-run household electricity demand by time of use. This framework, presented in detail in Aigner and Poirier (1979), was first used by Hausman, Kinnucan and McFadden (1979) and Caves and Christensen, (1980) to estimate the on-peak and off-peak price elasticities corresponding to TOU experiments in Connecticut and Wisconsin, respectively. We will rely on this framework as a baseline to define rational household responses to TOU pricing in a static optimization setting.

Following Hausman, Kinnucan and McFadden (1979), we specify a household’s monthly utility function as

\[ U = U(x^{on}, x^{off}, y), \]

where \( x^{on} \) and \( x^{off} \) are the household’s monthly on-peak and off-peak electricity usage, respectively, and \( y \) is a vector of all other goods.\(^7\) We then make a weak separability assumption so that utility can be characterized as

\[ U = U(f(x^{on}, x^{off}), y), \]

where \( f(x^{on}, x^{off}) \) represents a homogeneous of degree one Hicksian aggregation of on-peak and off-peak electricity consumption, and \( y \) is a Hicksian composite of all the goods in \( y \). Normalizing the price of \( y \) to unity permits \( y \) to be interpreted

\(^7\)An important assumption underlying this utility specification is that the stock of electricity-using appliances is fixed. Therefore, \( x^{on} \) and \( x^{off} \) should be thought of as derived electricity demand based on demand for household services that use these appliances and the times of day that the household prefers to consume such services.
simply as expenditure on all goods besides electricity.

The weak separability condition allows the household’s monthly maximization problem to be decomposed into two levels. One level represents the household’s choice of how much to spend on total electricity usage, where the remainder of its (fixed) income is spent on all other goods. The other level describes the household’s choice of how to allocate electricity consumption across on-peak and off-peak hours, which depends only on electricity rates.

The choice of total electricity usage, \( X \equiv x^{on} + x^{off} \), will depend on an aggregated price of electricity \( p \) given by

\[
p = p^{on} s^{on} + p^{off} s^{off},
\]

(1.3)

where \( s^{on} \) and \( s^{off} \) are the shares of on-peak and off-peak usage in total usage as determined in the allocation level of the maximization problem.\(^8\) These shares sum to unity and depend only on the parameters of \( f() \). Therefore, the aggregated price of electricity is a weighted average of the on-peak rate \( p^{on} \) and the off-peak rate \( p^{off} \).

Thus, conditional on the optimal choices of \( x^{on} \) and \( x^{off} \) from the allocation level of the maximization problem, total electricity consumption can be viewed as being determined by a straightforward utility-maximizing division of total income between good \( X \), with price \( p \), and expenditure on all other goods. Once properties of \( U() \) are specified, deriving predictions concerning total electricity consumption is accomplished as in any two-good setting characterized by these properties. For example, if \( U() \) is such that \( X \) is a normal good, the model clearly predicts that a drop in \( p \) will lead unambiguously to an increase in

\(^8\)The aggregated price in our empirical setting will also include a small adjustment for a fixed monthly charge and for the increasing-block structure of the non-TOU rate. This will be discussed in more detail in the following section.
the quantity demanded of $X$. Further predictions of the model will be discussed in the context of our empirical results in the interpretation section below.

1.3 Program Design

Beginning in 2006, an electric distribution company in the northeastern United States implemented a mandatory time-of-use (TOU) program for residential customers. Prior to the introduction of this program, most residential customers were billed according to a seasonal flat rate, with the price of electricity varying seasonally but remaining constant within a day. Approximately 12% of customers chose to be placed instead on a seasonal TOU rate, with the price of electricity varying seasonally and within a day. In the analysis that follows, we exclude these voluntary adopters.

Under the policy, when a residential customer’s electricity usage in any 30-day billing period exceeded a pre-determined threshold, the customer was automatically placed onto TOU pricing. Beginning November 2006, a household was to be placed on TOU pricing by January of 2008 if usage in any 30-day billing period exceeded 4000 kWh. This threshold applied until December 31, 2007.9

The residential TOU rate plan charges a high per-kWh rate at on-peak times (noon through 8pm on weekdays) and a low per-kWh rate at off-peak times (all other times and days). Table 1.1 shows the TOU rates that were in effect over the period of our analysis, and compares them to the corresponding non-TOU rates. In our study, the summer non-TOU tariff had an increasing-block structure, with the first 500 kWh of usage in a billing month charged at a base “headblock”

---

9The threshold was lowered to 3000 kWh in 2008 and to 2000 kWh in 2009. The present chapter focuses on households that crossed the 4000 kWh threshold. The following chapter provides a detailed analysis of household responses to TOU pricing during the latter phases of the policy.
Table 1.1: Electricity Rates, 2008, Cents per kWh

<table>
<thead>
<tr>
<th></th>
<th>Non-TOU</th>
<th>TOU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Headblock</td>
<td>Tailblock</td>
</tr>
<tr>
<td>unbundled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun.</td>
<td>7.9</td>
<td>11.8</td>
</tr>
<tr>
<td>Jul.</td>
<td>8.6</td>
<td>12.6</td>
</tr>
<tr>
<td>bundled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun.</td>
<td>20.1</td>
<td>24.1</td>
</tr>
<tr>
<td>Jul.</td>
<td>20.6</td>
<td>24.6</td>
</tr>
</tbody>
</table>

Notes: Unbundled rates include distribution, transmission, and delivery charges plus fees only. Bundled rates also include the generation prices that were charged by the utility to those customers opting to keep the utility as both distributor and supplier. About 55% of the customer base opted to pay generation prices charged by alternate suppliers; no alternate suppliers had TOU generation prices, so the unbundled rates represent the relative on-peak/off-peak all-inclusive rates faced by these customers, but not the absolute level. The headblock is the first 500kWh of total usage in the billing month. The July rates stayed in place through September.

per-kWh rate and the remaining usage in that billing month charged at a higher “tailblock” per-kWh rate.

Given this increasing-block structure for the flat rate and the fact that all of the households in our analysis exceed 500 kWh in total electricity consumption in every month, the non-TOU monthly budget constraint can be expressed as

\[ p^t(X - 500) + p^h500 + g = E, \]  \hspace{1cm} (1.4)

where \( E \) is total electricity expenditure, \( p^t \) is the tailblock rate, \( p^h \) is the headblock rate, and \( g \) is a fixed monthly charge. Noting once again that total electricity consumption is simply the sum of on-peak and off-peak consumption, this can be re-written as

\[ p^t x^{on} + p^t x^{off} - (p^t - p^h)500 + g = E, \]  \hspace{1cm} (1.5)

which emphasizes the fact that the marginal rate faced by non-TOU customers in both on-peak and off-peak hours is the tailblock rate. Meanwhile, the TOU
monthly budget constraint is given by

\[ p_{on}^x x_{on} + p_{off}^x x_{off} + g = E, \] (1.6)

where the fixed monthly charge \( g \) is the same as that for non-TOU customers in all months.

Within the theoretical framework presented in the previous section, total electricity expenditure is defined as the product of the aggregated electricity price and total electricity consumption, or \( E \equiv pX \). Inserting this definition into equations (1.5) and (1.6) and dividing by total consumption gives expressions for the aggregated non-TOU and TOU electricity prices,

\[ p_N = p^t + \phi_N \] (1.7)

and

\[ p_T = p_{on}^s s_{on} + p_{off}^s s_{off} + \phi_T, \] (1.8)

where the subscript \( s \in \{N, T\} \) refers to non-TOU and TOU respectively, and the \( \phi_s \) are small constants based on the fixed charge and the headblock adjustment.

We can now link the rates in Table 1.1 – and thus the change in the aggregated electricity price experienced by a household that was switched from the flat rate to the TOU rate – to predictions generated by the theoretical framework. Setting \( \phi_T = \phi_N \) as a convenient approximation for now, it is clear that \( p_T < p_N \) if \( p^t > p_{on}^m > p_{off}^m \), which was the case with the unbundled rates in Table 1.1 throughout the summer of 2008.\(^{10}\) Further, \( p_T < p_N \) as well if

\(^{10}\)The unbundled rates include delivery and distribution charges only. They reflect the on-peak/off-peak gradient faced by all customers that chose to pay the generation rates of alternate suppliers, though the absolute level of the all-inclusive rates depends on the specific
\( p^{\text{on}} > p^{\text{t}} > p^{\text{off}} \) and \( s^{\text{on}} \) is sufficiently small. Therefore, as a first approximation, Table 1.1 indicates that households that were switched to TOU in 2008 experienced a decrease in the aggregated price of electricity that they faced compared to households that remained on the flat rate. As discussed in the previous section, the standard theoretical framework would hence predict that households that were switched to TOU would increase their total electricity consumption. This and other predictions will be stated and tested more formally in the interpretation section below.

1.4 Experimental Setting and Data

In this section we explain in detail how the TOU program we study gives rise to a regression discontinuity design, and discuss some nuances of our empirical setting. We then describe the billing data used to identify the effect of mandatory TOU pricing on total usage and total bills.

The key feature of the regression discontinuity design in general is that assignment to the treatment group is triggered by crossing some threshold. In our setting, this occurs when monthly usage exceeds a pre-determined threshold. For the treatment effect to be valid, it must be the case that within the neighborhood of the threshold, assignment to TOU is effectively random. This will occur if some idiosyncratic factors push some individuals over the threshold but not others, or as described by Lee and Lemieux (2010), households lack precise control over the “forcing variable”, defined in our study, according to the rules of the program design discussed above, to be maximum monthly electricity usage between alternate supplier that a given household was served by, which we do not observe. The bundled rates are the all-inclusive rates that were faced by all customers that chose the EDC as their supplier, which includes about 45% of the EDC’s overall customer base. All customers had the EDC in question as distributor, as there are no alternative distributors in the region.
Figure 1.1: Density of the Forcing Variable at the 4000 kWh Threshold

Note: Data are smoothed into bins of width 20 kWh. Separate quadratic predictions on each side.

November 2006 and December 2007 net of the 4000 kWh threshold.\footnote{This forcing variable will be defined more formally in the following section.}

If “bunching” in the density of households just below the crossing threshold occurred, this might indicate that households could manipulate usage to avoid crossing the TOU threshold. However, it seems reasonable to assume that households have only imprecise control over their exact usage in any billing period, since precise control would likely require sophisticated equipment for monitoring and regulating usage. We also present visual evidence to support the exogeneity assumption: Figure 1.1 indicates no evidence of bunching around the crossing threshold. Having thus provided evidence that crossing the threshold is random, differences in outcomes between households on either side of the threshold can be interpreted as causal effects.
However, one feature of the program – the varying lag across households between crossing a threshold and receiving the TOU treatment – complicates the regression discontinuity design. To address this, we divide the sample into three periods: the pre-experiment period; the qualification period; and the treatment period. The pre-experiment period is defined as the set of months preceding the introduction of the mandatory TOU rule. The qualification period is defined as November 2006 through December 2007, the months during which a household, should it exceed 4000 kWh, would eventually be assigned to TOU pricing. No household was actually assigned to TOU pricing until February 2008.\footnote{There were some households that had previously adopted TOU on a voluntary basis, but again, voluntary adopters have been excluded from the analysis. This was done because such self-selection into treatment would invalidate the experimental design.} Thus, up until this month there is no difference between crossers and non-crossers in the propensity to be treated. However, not all qualifying households were switched at this point, and indeed, some were not switched for several more months.\footnote{The long delays between crossing and switching and the failure to switch some qualifying households altogether mainly occur because of technical and administrative difficulties associated with installing requisite metering equipment. Households suffering from a serious illness or other life threatening situation necessitating the use of specialized electrical devices could apply for exemption from the program. We observe a small number of crossers that were switched to TOU but eventually allowed to revert to a non-TOU, and interpret this to be the result of the granting of a medical exemption. These households have been removed from the analysis. It is possible that some of the crossers that were never switched to TOU were granted a medical exemption pre-emptively, but we cannot observe this.} Therefore, the propensity to be treated does not immediately jump to 100% in February 2008. The treatment period, the focus of our analysis, is comprised of June - September 2008, some months when households that crossed the threshold ("crossers") should have been switched onto TOU. We choose these months to be the treatment period, since households on TOU faced a lower per kWh rate (net generation) than households on the non-TOU rate.

Another nuance in our setting is that customers would also qualify to be
switched to TOU if they breached a lower threshold in any month in 2008. This implies that it is possible for some households who never crossed the 4000kWh threshold to nonetheless be on TOU in the later months of 2008. The joint effect of these two features is that the propensity to be treated will increase over time for both groups, and thus that the difference in this propensity across groups will be substantial for a limited window only.

It follows that, unlike in a canonical “sharp” regression discontinuity setting, in our setting crossing the TOU usage threshold is not a perfect determinant of being in the treatment group in any given month. Instead, the empirical setting should be viewed as having been generated by the Fuzzy Regression Discontinuity (FRD) design, where the “fuzziness” refers to the imperfection of the crosser/non-crosser distinction as a predictor of TOU status in a given experimental month. The FRD design allows us to interpret differences in outcomes between crossers and non-crossers as causal treatment effects, though we must adjust their magnitudes for the propensity for each group to be treated. These treatment effects can be estimated consistently only for treatment months in which a sufficiently high proportion of crossers is on TOU relative to the proportion of non-crossers on TOU (i.e. in which the crosser/non-crosser distinction is a strong instrument for treatment status).

Before turning to a more precise discussion of how we implement the estimation of these treatment effects, we describe the billing data and present summary statistics. Monthly billing data beginning in June 2006 on total usage, total expenditure (net of generation) and rate class were provided for a sample of about 35,000 households. Table 1.2 presents descriptive results at the preferred

\[14\] The sample is comprised of the population of households with usage above 1500kWh in September 2010, and a random sample of households with usage between 1300 and 1500kWh in September 2010. The year 2010 was chosen so that the included households would be most
Table 1.2: Summary Statistics, July of the Qualification Period

<table>
<thead>
<tr>
<th></th>
<th>Total Usage (kWh)</th>
<th>Total Bill ($)</th>
<th>Crossers (%)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000kWh Experiment (2007)</td>
<td>3,309</td>
<td>382</td>
<td>0.339</td>
<td>1,096</td>
</tr>
</tbody>
</table>

Notes: Standard deviations are in square brackets. The households included for each experiment are those within an optimally-chosen bandwidth around the threshold; see the text for details.

The experiment consists of 1,096 households, 34% of which crossed the 4000 kWh threshold at some point in the qualification period. Mean usage and net-of-generation expenditure for this sample amount to 3,309 kWh and $382 in July 2007, though within this bandwidth there is substantial variation in both usage and expenditure.

1.5 Treatment Effects for Total Usage and Total Bills

1.5.1 Methods

We begin by comparing crossers to non-crossers along several dimensions, separately for each month in the entire sample. Specifically, we estimate

\[ Y_i = \beta_0^{Y_t} + \beta_1^{Y_t}C_i + \beta_2^{Y_t}f(\tilde{X}_i) + \beta_3^{Y_t}C_i \times f(\tilde{X}_i) + \varepsilon_i^{Y_t} \]  

(1.9)

individually for each month \(t\) and for various dependent variables \(Y\). The variable \(C_i\) is a dummy variable indicating whether household \(i\) is a crosser. The variable \(\tilde{X}_i\) is the “forcing variable” that determines whether household \(i\) is a crosser. More precisely, \(\tilde{X}_i\) is household \(i\)’s maximum total usage across all billing periods during the qualification period net of the kWh threshold. Under the rules representative of the EDC’s current customer base. September was chosen because it corresponded to the annual system peak that year.

\(^{15}\) For a discussion on the optimal bandwidth choice, please refer to Section A.1.
of the program, if $\tilde{X}_i$ is strictly greater than zero, household $i$ is a crosser and should receive the TOU treatment eventually.$^{16}$

The dependent variables we consider are total usage, total bills, and a dummy variable $TOU_{it}$ indicating whether household $i$ was on TOU pricing (i.e. was treated) in month $t$. Specification (1.9) allows for a flexible relation between the outcome variable of interest and the forcing variable through the function $f(\cdot)$, and allows this relation to differ for crossers and non-crossers.$^{17}$ The parameter $\beta_{1Yt}$ measures the effect of being a crosser on the level of outcome variable $Y$ in month $t$ as the distance from the threshold approaches zero, and is interpreted as the Intent to Treat effect (ITT). These are causal effects by virtue of our assumption that, as the distance from the threshold approaches zero, a household’s crossing status is exogenous.

The fuzzy regression discontinuity treatment effect for outcome $Y$ in any month $t$ in the treatment period for a given experiment is defined as

$$\tau_{Yt}^{FRD} \equiv \frac{\beta_{1Yt}}{\beta_{1TOUt}}. \quad (1.10)$$

That is, the treatment effect for the outcome of interest is the ratio of the ITT for the outcome of interest to the ITT for the propensity to be treated. It can be

$^{16}$Formally, let $X_{it}$ be household $i$’s total electricity usage in month $t$. Further, let usage on a standardized 30-day-billing-period basis be $\bar{X}_{it} \equiv X_{it}/d_{it} \times 30$, where $d_{it}$ is the number of total days actually in the billing period corresponding to household $i$’s bill in month $t$. Then

$$\tilde{X}_i = \left( \max_{t \in Q} \{\bar{X}_{it}\} - \bar{X} \right),$$

where $Q$ is the set of months in the qualification period and $\bar{X}$ is the threshold; and $C_i \equiv 1\{\tilde{X}_i > 0\}$, where $1\{\}$ is the indicator function. The households included in these regressions are only those with a value of the forcing variable $\tilde{X}_i$ within a selected bandwidth around zero, i.e. households “close to” the threshold. When presenting our results, we first use a wide bandwidth to visually examine the data and then use an optimal bandwidth to estimate the treatment effects.

$^{17}$We first define $f(\cdot)$ as a fourth-order polynomial to visually examine the data, then as linear to estimate the treatment effects. Within the optimal bandwidth, we do not find alternatives to the linear form to qualitatively affect our estimated treatment effects.
estimated by applying two-stage least squares to the following system of equations for any outcome variable \( Y \) in a given treatment-period month \( t \):

\[
Y_i = \tau_{Yt} Y_{0i} + \tau_{Yt} Y_{1i} \text{TOU}_i + \tau_{Yt} Y_{2i} f (\tilde{X}_i) + \tau_{Yt} Y_{3i} \times f (\tilde{X}_i) + \omega_{Yt} \tag{1.11}
\]

\[
\text{TOU}_i = \beta_{0\text{TOU}} \text{TOU}_0 + \beta_{1\text{TOU}} C_i + \beta_{2\text{TOU}} f (\tilde{X}_i) + \beta_{3\text{TOU}} C_i \times f (\tilde{X}_i) + \varepsilon_{\text{TOU}} \tag{1.12}
\]

where \( \hat{\tau}_{1,2SLS} \) is numerically equivalent to inserting the ITTs estimated via specification (1.9) into equation (1.10). Note that we apply two-stage least squares as a computational convenience, not to address endogeneity concerns.\(^{18}\)

We calculate standard errors for the treatment effects based on non-parametric bootstrap methods. While the robust 2SLS covariance matrix is asymptotically valid, we opt to report bootstrapped standard errors to ease the interpretation of our results. Bootstrap methods allow us to estimate the model, including the relationship between the forcing variable and the dependent variables, in levels but present the results, for ease of interpretation, in percent. We discuss bootstrap methods in Section A.2.

### 1.5.2 Preliminary Evidence

We begin by visually examining the propensity to be treated, total billed amount, and total usage on each side of the threshold in July 2008. Specifically, we estimate specification (1.9), including households within a very wide range

---

\(^{18}\)A household’s time-of-use status in a given treatment-period month depends on its crossing status in the preceding qualification period and on unobservable factors. However, crossing status is exogenous at the threshold by assumption, and the unobservable factors are ostensibly exogenous issues related to various meter installation and administrative hurdles faced by the utility. We therefore do not consider concerns about endogeneity between TOU status and either total expenditure or total usage to be present.
Notes: Data are smoothed into bins of width 80 kWh.

around the threshold and allowing the relation between the outcome variable and the forcing variable to have a separate quartic form on each side of the threshold. This provides a first look at whether the relation exhibits a discontinuity at the threshold (i.e. an intent to treat effect), and allows us to diagnose any non-linearities that may complicate the identification of a discontinuity.

Figure 1.2 shows the estimated propensity to receive the TOU treatment for crossers (households that exceeded the 4000 kWh threshold) and non-crossers. Crossing the threshold is clearly a strong predictor of having received the TOU treatment by July 2008, as illustrated by the dramatic discontinuity at the threshold. However, it is not a perfect indicator, as some non-crossers just to the left of the threshold – i.e. whose maximum 30-day usage during the 4000 kWh qualification period was very close but did not exceed the 4000 kWh threshold – have a small but positive propensity to be treated. Likewise, a few crossers still
In Figures 1.3 and 1.4, we present the estimated total billed amount and usage, respectively, on each side of the 4000kWh threshold in July 2008. These graphs illustrate a discontinuity both in expenditure and usage at the threshold, suggesting that a crosser had a substantially lower electricity bill than a non-crosser at the threshold (by about $100). While the relation in Figure 1.3 exhibits some non-linearity, particularly for very high levels of the forcing variable, these figures provide fairly clear evidence that the difference in usage and expenditure is indeed the result of a discontinuity.

1.5.3 Treatment Effects

Having provided visual evidence of the discontinuity, we now restrict specification (1.9) to be linear in the forcing variable and in its interaction with crossing status, and include only households within a narrower, optimally-chosen
We use this form to identify ITTs for each dependent variable for several pre-qualification and qualification months, as well as our treatment months of June through September 2008. To present the results as compactly as possible, we graph time series of the set of estimated coefficients for each of the three dependent variables. For dependent variable $Y$, we graph $\hat{\beta}_Y^t$ – the estimate of outcome $Y$ in month $t$ for a non-crosser exactly at the threshold – and $\hat{\beta}_0^Y + \hat{\beta}_1^Y$ – the same for a crosser exactly at the threshold – for every month, also indicating when the difference between the two is statistically significant.

Figure 1.5 graphs the ITT of the probability that a crosser receives the

---

Notes: Data are smoothed into bins of width 80 kWh.

---

19 The method used to determine the optimal bandwidth is described in Section A.1. A larger bandwidth leads to more precise estimates of the discontinuity. However, it also means that households further away from the threshold are being used to identify the discontinuity at the threshold, which can impart a bias; and this bias can be large if the relation is highly non-linear around the threshold. We choose an optimal bandwidth to apply uniformly for the estimation of all ITTs and treatment effects in each month of the treatment period.
TOU treatment for each individual month between June 2006 and January 2009.\textsuperscript{20} The months between the vertical lines delineate the qualification period, and the months further to the left are the pre-experiment period. This figure illustrates that crossing the TOU threshold is a strong predictor of TOU pricing in the treatment period. In the pre-experiment and qualification periods the propensity to be on time-of-use pricing is zero for both crossers and non-crossers by construction, since we restrict our sample to households that did not receive the treatment during the qualification period.\textsuperscript{21}

We present the estimated ITTs on the total bill in Figure 1.6. The large discontinuity illustrated in Figure 1.3 for July 2008 is also present for the other treatment months, with 95 percent confidence. We also observe that the estimated total bill was nearly identical for crossers and non-crossers throughout the pre-experiment and qualification periods. This balance on pre-determined observables is consistent with the intent to treat being randomly assigned at the threshold. It also suggests that the large ITTs observed in the summer 2008 are not spuriously caused by systematic difference in summer usage patterns between crossers and non-crossers.

Figure 1.7 illustrates the estimated ITTs on total electricity usage over time. Total usage was nearly identical between crossers and non-crossers.

\textsuperscript{20}The bandwidth is symmetric, so encompasses households with a value of the forcing variable between -600kWh and 600kWh. Note that the data in Figure 1.2 have been smoothed for ease of presentation, so that each point represents several households. The point for July 2008 in Figure 1.5 is based on straight lines of best fit through the first 7-8 points on each side of the threshold in Figure 1.2.

\textsuperscript{21}Households with a value of the forcing variable substantially higher than the upper bandwidth cut-off of 600kWh are more likely to have crossed the 4000 kWh threshold for the first time early in 2007, and such households were required to have been switched to TOU before the end of 2007. A few of these households were indeed switched in late 2007, but most were not switched until February 2008. The delay in rolling out the program for these larger households (that are not included within the bandwidth we consider in any case) appears to be due to unforeseen technical and administrative issues by the utility.
Figure 1.5: Intent to Treat Effects by Month, Propensity to be Treated

Notes: The qualification period is the set of months between the vertical lines. The estimates for each month are from a separate application of equation (1.9) with a bandwidth of 600 kWh.

Figure 1.6: Intent to Treat Effects by Month, Total Bill

Notes: The qualification period is the set of months between the vertical lines. The estimates for each month are from a separate application of equation (1.9) with a bandwidth of 600 kWh.
Figure 1.7: Intent to Treat Effects by Month, Total Usage

Notes: The qualification period is the set of months between the vertical lines. The estimates for each month are from a separate application of equation (1.9) with a bandwidth of 600 kWh.

throughout the pre-experiment and qualification periods, providing evidence of another observable along which the two groups are balanced. However, during the treatment periods, there is a significant difference in total usage in June and July 2008, when crossers at the threshold exhibited lower usage than non-crossers at the threshold. We also see some visual evidence of lower usage for TOU households in August and September, though we cannot distinguish these from zero with confidence. The absence of significant differences in total usage between crossers and non-crossers during the pre-experiment and qualification periods indicates that the differences in June and July 2008 are not driven by pre-existing differences between the groups. It also indicates that non-crossers were not purposely restraining their usage during the qualification period to avoid crossing the threshold, which would violate the random assignment assumption.

Table 1.3 shows the treatment effects, adjusted for the propensity to be
Table 1.3: Treatment Effects (%), Total Usage and Total Bill

<table>
<thead>
<tr>
<th></th>
<th>Total Usage</th>
<th>Total Bill</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun. 2008</td>
<td>-9.24 ** (4.69)</td>
<td>-21.50 *** (4.19)</td>
<td>1,105</td>
</tr>
<tr>
<td>Jul. 2008</td>
<td>-9.85 *** (3.73)</td>
<td>-30.06 *** (2.96)</td>
<td>1,105</td>
</tr>
<tr>
<td>Aug. 2008</td>
<td>-5.39 (4.04)</td>
<td>-26.15 *** (2.96)</td>
<td>1,107</td>
</tr>
<tr>
<td>Sep. 2008</td>
<td>-2.11 (5.82)</td>
<td>-22.31 *** (4.36)</td>
<td>1,095</td>
</tr>
</tbody>
</table>

Notes: Standard errors (in parentheses) are based on a non-parametric bootstrap with 1,000 replications. Significance at the 1% (***) , 5% (**), and 10% (*) levels is indicated. Each estimate is from a separate application of equations (1.11)-(1.12) with a bandwidth of 600 kWh, and is the estimated TOU treatment effect as a percentage of the estimated non-TOU level at the threshold for the respective dependent variable. Of the 1,105 households included in the regressions for July 2008, 373 are crossers; and the distribution of households is similar in other months.

treated, on total usage and total bills for each month in the treatment period. To give a better sense of magnitudes, treatment effects are reported as a percentage of the level of the respective variable for non-TOU households at the threshold.22 We find that the switch to TOU pricing caused economically and statistically significant reductions in electricity expenditure in all treatment months by at least 21% and as much as 30% in July. This is matched by statistically significant declines in total electricity usage in June and July of 9-10%, and noisy declines of 5 and 2 percent in August and September, respectively.

When interpreting the expenditure estimates, it seems natural that electricity expenditure would decline or remain unchanged since customers on TOU faced lower peak and off-peak rates compared to flat rate households. In contrast, basic intuition suggests that demand for electricity should increase with a reduction in electricity prices; yet we find the opposite to be true. We now

---

22That is, each entry shows $\hat{\tau}Y_t/\hat{\tau}Y_t \times 100$ from a separate two-stage least squares application of equations (1.11)-(1.12).
investigate the revealed choice behavior more directly.

1.6 Interpretation

To assist us in digesting the empirical results, we turn to Figures 1.8 and 1.9, which provide a simple visual way to evaluate the nature of consumer choice. These figures present graphs of budget frontiers and revealed choices as implied by the empirical results described in Table 1.3. Each graph presents the consumption bundle chosen by TOU customers, as well as two budget frontiers. These features of the choice setting are derived directly from prices and estimates of behavior in treatment (TOU) and control (non-TOU) at the threshold. The TOU consumption bundle is revealed arithmetically from the relationship between total consumption \( (\hat{\tau}_0^{\text{kWh}} + \hat{\tau}_1^{\text{kWh}}) \) and the TOU tariff rates. Budget frontiers are determined by the revealed consumption level at the threshold (from the application of the 2SLS estimation) and relative prices.\(^{23}\)

The first frontier is based on non-TOU rates and the level of expenditure of the non-TOU household, and has a slope of -1 to reflect equality of peak and off-peak prices. This frontier represents all combinations of on-peak and off-peak usage that sum to the estimated non-TOU total usage at the threshold. Note that any point on the interior of this frontier is unequivocally a drop in total consumption relative to the non-TOU bundle. The second, analogous, frontier is based on TOU rates and the expenditure of the non-TOU household at the threshold. Were expenditure for treated households to remain at the revealed non-TOU level, their TOU bundle would reside on this second frontier. Each

\(^{23}\)Algebraically, the budget frontiers are expressed by equations (1.4) and (1.6), with values of actual rates and revealed total expenditure inserted where appropriate. The imputation of the TOU consumption bundle is discussed in Section A.3. A technical issue involving adjustment of calendar-month rates for billing cycles is discussed in Section A.4.
Figure 1.8: Budget Lines, Unbundled Rates

- Non-TOU rates and Expenditure
- TOU rates and non-TOU Expenditure
- Imputed TOU bundle

June 2008
- Off-peak usage vs. on-peak usage

July 2008
- Off-peak usage vs. on-peak usage

August 2008
- Off-peak usage vs. on-peak usage

September 2008
- Off-peak usage vs. on-peak usage
Figure 1.9: Budget Lines, Bundled Rates

- June 2008
- July 2008
- August 2008
- September 2008
budget constraint is presented for the months June to September 2008.

We present two different rate types – unbundled in Figure 1.8 and bundled in Figure 1.9 – to reflect differences in the TOU gradient between two customer types in our sample. Unbundled rates are paid by customers who have elected to purchase electricity generation from an “alternate supplier” (i.e. not from the regulated electricity distribution company). During the period of analysis, all alternate supplier generation rates were time-invariant\textsuperscript{24}, implying that the entire TOU gradient was transmitted through the unbundled price for these customers. On the other hand, bundled rates include generation charges that are paid to the electricity distribution company. In our setting, generation charges transmit an additional peak/off-peak price gradient.\textsuperscript{25} As such, the choice setting is different for customers who have elected to purchase generation from an alternate supplier than for those purchasing exclusively from the regulated utility, so we present budget frontiers separately for each.

Recall that the TOU rates/non-TOU expenditure frontier represents the theoretical consumption possibilities available to a household that is switched to TOU pricing and retains the non-TOU level of electricity expenditure. This frontier describes the set of on-peak/off-peak bundles from which a static utility maximizing household would choose. The non-TOU bundle lies somewhere on the interior line, and from these figures it is apparent that the chosen TOU bundle was feasible under the non-TOU budget. Thus the original non-TOU bundle is

\textsuperscript{24}A thorough search of alternate supplier rates by the authors in 2010 confirmed what our utility partners asserted: time-varying generation rates were not offered by alternate suppliers until more recently.

\textsuperscript{25}Readers may note that the unbundled rate budget frontiers appear to be further right (higher) than the bundled frontiers. This is a direct result of excluding generation charges. In reality all customers pay for generation, and frontiers of those on alternate suppliers would exhibit a parallel inward shift were the flat (but unknown to the researchers) generation rates to be included.
revealed preferred to the TOU bundle. Note that this is true irrespective of the presence of crossing of the budget constraints (which we discuss below). In this simplified two-good setting, each graph reveals an outcome in which treated consumer choices appear to violate the Weak Axiom of Revealed Preference (WARP).

A more realistic interpretation of the setting includes an outside consumption good in addition to both electricity goods. Here we will distinguish between instances in which this TOU frontier lies completely above the non-TOU frontier and those in which the TOU and non-TOU budget constraints cross. When we allow for the presence of an outside good, the possibility exists that an electricity price change will induce substitution towards the outside good. In regions where the TOU budget is on the exterior of the non-TOU budget constraint, the outside good has become relatively more expensive in treatment. If the non-TOU bundle resided in this range, substitution away from electricity and towards the outside alternative would imply that the entire bundle of non-electricity household expenditures was a Giffen good. We consider this to be an unsatisfactory explanation with which to reconcile the empirical findings.

In regions where the TOU budget is interior to the non-TOU budget (i.e. in the lower-right region of graphs where the frontiers cross), the story becomes slightly more nuanced. For households with non-TOU bundles in that region, the switch from flat rate to TOU implies a decrease in the relative price of the outside good. In this case substitution away from electricity and towards the outside good does not require the aggregate good to be a Giffen good. However, the likelihood of a household’s chosen peak/off-peak bundle initiating on this region of the non-TOU budget frontier is essentially zero. In months where we observe a cross in the budget frontier, this crossing occurs at an extremely high peak/off-peak
ratio. Using an external data set on the peak-to-off-peak usage ratio for a random sample of customers, we can examine the likelihood that the observed “crossing” ratio falls within the observed range of ratios.\textsuperscript{26} With the exception of bundled rates for June, the crossing ratio is much higher than any peak to off-peak ratio observed in the data. Even for a customer on a bundled rate in June, the “crossing” ratio is in the 99th percentile of the observed distribution and thus an unlikely explanation for our results.

The abundance of evidence does not allow much scope for household choice behavior to be consistent with static utility maximization. A two-good view leads immediately to violations of WARP. When we allow for the consumption of an outside good, for static utility maximization to hold, it must be that the composite bundle of non-electricity expenditures is Giffen, or that the initial non-TOU bundle lay on a region of the budget frontier that is inconsistent with observed data. So where does this leave us?

One potential hypothesis that could explain the observed outcome is that households were not only responding to contemporaneous rates, but also engaging in what is becoming known as “intermittent updating”. Under this hypothesis, consumers are attentive to choices infrequently, and thus may exhibit choices that were optimal at some time in the past but do not correspond to utility maximization in each moment.\textsuperscript{27} Another hypothesis allows for a dynamic and in a sense more sophisticated view of consumer choice. Perhaps consumers are making durable goods investments in response to expected future rates. Finally,\textsuperscript{26}This load profile dataset comprises hourly usage data between January 2006 and October 2011 for a random sample of 1,300 households present for between 2 and 48 months. The distribution of the peak-to-off-peak usage ratio in this dataset is described in more detail in Section A.3.\textsuperscript{27}While we remain agnostic about mechanisms, support for the “intermittent updating” hypothesis resides in the fact that hundreds of households just below the threshold could have saved substantial amounts by volunteering for TOU, but didn’t.
there is a growing body of evidence on the importance of “behavioral” considerations in this choice setting. While our study is limited to identifying an instance of curious behavior that is relevant to current policy debates, testing some of the aforementioned alternative hypotheses would be valuable.

1.7 Conclusion

This study exploits a natural experiment to document a setting in which households do not respond to price incentives in the way that standard theory predicts. Despite growing evidence that non-monetary and inter-temporal factors are important in a wide range of environments, there are few well-identified cases of this behavior in environmental economics. Evaluating customer response to price incentives is a necessary step in understanding consumer choice in this setting.

The randomized nature of assignment into TOU pricing that arises from the structure and implementation of the program provides us with an empirical setting to evaluate consumer behavior. Customers were automatically placed on the TOU rate after exceeding the usage threshold, creating an appropriate setting in which to apply a regression discontinuity design. This differentiates our research design from most studies of time-varying electricity pricing which rely on framed field experiments in which participants are aware of their participation.\textsuperscript{28} Thus, our paper offers a novel estimate of how certain residential consumers behave when exposed to TOU pricing.

The baseline model of consumer behavior employed in this paper is rooted in the traditional framework for modeling consumer electricity choice (Aigner and

\textsuperscript{28}We refer here to the taxonomy of field experiments proposed by Harrison and List (2004). Wolak (2006) and Jessoe and Rapson (2012) are examples of recent studies of the effect of time-varying pricing that are based on framed field experiments.
Poirier, 1979; Hausman, Kinnucan and McFadden, 1979; and Caves and Christensen, 1980). And while this framework should serve as a starting point to frame consumer response to electricity prices, we highlight that in some instances it may not describe customer behavior accurately. In these cases, structural estimates based on the classic theoretical framework may lead to misleading conclusions.

Admittedly, the households in our setting are very large, and not representative of the “average” electricity user. On the other hand, the intensity of electricity use which they exhibit makes them a particularly important target for energy conservation efforts. Questions of external validity, though, are almost beside the point since what we observe in this setting is potentially present in many other settings. An assertion that price incentives always work can be disproven by the counter-example we have provided.

We must achieve a more thorough understanding of what drives behavior to inform planners about how to effectively achieve climate change mitigation. If what we observe here suggests that failure of the utility maximization assumption could arise more generally in some settings, then market-based policies designed on the basis of this assumption may fail to achieve the emissions target and/or fail to achieve a given reduction cost effectively. On the other hand, if we knew the mechanism driving consumer behavior, this information could be leveraged to effectively introduce price-based policies (for example, by coupling prices with real-time feedback) or non-price interventions. Our findings suggest that there may be a risk in adhering too ideologically to price interventions alone, in terms of missing policy goals or achieving them only imperfectly or inefficiently.
CHAPTER TWO

Understanding Residential Response to
Time-of-Use Electricity Pricing

2.1 Introduction

Time-of-use (TOU) electricity pricing divides a household’s total electricity use into two or three blocks according to the time of day at which it is consumed, and applies higher rates to blocks corresponding to times that are historically associated with high production costs. It is a small step towards aligning retail electricity prices with marginal production costs. But its attractiveness to some regulators lies to a greater degree with its crucial purported advantage of being easy for consumers to understand and, in principle, respond to.

In this chapter, we study short-run household responses to a large-scale mandatory residential TOU program, and seek to draw lessons about TOU’s potential as a tool to achieve efficiency in electricity markets. Such potential could be realized in several ways. First, and of most concern to utilities, TOU could induce a reduction in peak consumption. Shedding peak load is desirable for utilities from a cost-containment perspective, and desirable more generally to the extent that it reduces reliance on dirtier generation sources. Second, though this is not an explicit goal of the policy, TOU might induce a general conservation effect depending on how the rates are set, which again would lead to a reduction in both production and environmental costs. Finally, TOU could provide a platform for regulators and rate-setters both to familiarize consumers with time-varying rates and to learn about the dimensions along which consumers respond to rate

---

29 This chapter is the product of joint work with Katrina Jessoe, of the University of California, Davis, Department of Agricultural and Resource Economics, and David Rapson, of the University of California, Davis, Department of Economics.
changes. This would be valuable in terms of guiding future policies aimed at correcting both sources of inefficiency.

We find that the TOU program caused the households we examine to substantially reduce total electricity consumption in non-summer months. However, it failed to achieve one of the primary goals of some utilities and regulators – to reduce demand during seasonal peaks – as the households exhibited no response in summer months. We also find suggestive evidence that at least some households did not respond to the actual TOU rate design, but rather exhibited a psychological response related to information and notification received about the program.

A concrete lesson that policymakers may be able to take away from our results is that interventions such as the TOU program we study could be more cost effectively implemented if they were targeted at households more likely to respond strongly to them. We find that the reductions in total electricity usage in non-summer months were concentrated nearly exclusively amongst energy intense customers that were likely to have had unexploited energy saving opportunities easily available. Learning to pre-emptively identify such customers would obviate the need for overly costly large-scale implementations that end up garnering little notice from broad portions of the customer base.

The larger message, though, is that TOU does not appear as easy for consumers to understand as it might be supposed. Our findings indicate that there is still much work to be done to learn how households respond to TOU and other electricity pricing interventions, and how such policy instruments can be constructively deployed in the pursuit of either environmental or system efficiency goals.
2.2 Setting and Data

Our analysis is based on the same TOU program that was the object of the previous chapter, but here we focus on a later stage of its implementation. Throughout the calendar year of 2009, households that exceeded 2000 kWh in total usage in any 30-day billing period, and that had not previously been switched to TOU, were to be placed on TOU pricing within six months of the first breach.

Table 2.1 shows the non-TOU and TOU electricity rates that were in effect during 2010, when most of the households that crossed the 2000 kWh threshold in 2009 were being switched onto TOU pricing. Comparing with the 2008 rates in Table 1.1, two observations are immediately clear. First, the increasing block structure of the summer rates was abandoned by the utility. The single non-TOU rate is therefore the marginal as well as the average rate faced by non-TOU customers. Second, the non-TOU rate is substantially lower than the TOU on-peak rate. Households could hence face a higher or lower effective price of electricity when being switched from the non-TOU to TOU tariff, depending on the proportion of their consumption that occurs at on-peak times. We thus do not have a simple prediction concerning the direction of change in total usage that we would expect for these households.

Our analysis relies on monthly billing data for the population of the utility’s customers that had usage greater than 1500 kWh in September 2010.\(^{30}\) We observe total usage and total electricity expenditure for each customer and each month in our analysis window. We do not observe the on-peak/off-peak

\(^{30}\)We also have a random sample of customers that had between 1300 and 1500 kWh of usage and were on the non-TOU tariff in September 2010. The latter restriction implies that these households have both low average usage and low month-to-month variation, as they would have been switched to TOU if they had exceeded 2000 kWh in total usage in any of the previous 21 months. This restriction makes the households in this second segment unrepresentative, and we have therefore chosen to exclude them from our analysis.
Table 2.1: Electricity Rates, 2010, Cents per kWh

<table>
<thead>
<tr>
<th></th>
<th>Non-TOU</th>
<th>TOU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Headblock</td>
<td>Tailblock</td>
</tr>
<tr>
<td><strong>unbundled</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan.</td>
<td>9.2</td>
<td>12.6</td>
</tr>
<tr>
<td>Jun.</td>
<td>9.6</td>
<td>14.0</td>
</tr>
<tr>
<td>Oct.</td>
<td>9.2</td>
<td>12.6</td>
</tr>
<tr>
<td><strong>bundled</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan.</td>
<td>20.8</td>
<td>26.7</td>
</tr>
<tr>
<td>Jun.</td>
<td>21.2</td>
<td>28.1</td>
</tr>
<tr>
<td>Oct.</td>
<td>20.8</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Notes: Unbundled rates include distribution, transmission, and delivery charges plus fees only. Bundled rates also include the generation prices that were charged by the utility to those customers opting to keep the utility as both distributor and supplier. About 55% of the customer base opted to pay generation prices charged by alternate suppliers; no alternate suppliers had TOU generation prices, so the unbundled rates represent the relative on-peak/off-peak all-inclusive rates faced by these customers, but not the absolute level. The headblock is the first 500kWh of total usage in the billing month. The October rates stayed in place through December.

breakdown of usage, though we can impute this for TOU household months as discussed in Section A.3.

The billing data also include street addresses for all customers, though our use of such identifying information is restricted due to confidentiality concerns. We have been able to use address information to add data on selected household characteristics from two external sources to the confidential module of our dataset.

The first of these external sources is Zillow.com, which is an online real estate marketplace that allows buyers and sellers to share and exchange information on properties and homes for sale or rent. The website maintains a database covering over 100 million U.S. homes that includes data on number of rooms, number of bathrooms, square footage, lot size, year built, and other characteristics, as well as past sale prices and the company’s own Zestimate®, which is an estimate of the current home value based on recent sales in the area, physical attributes, tax records, and other information. We used Perl script routines to scrape data directly from this database through the company’s website.
portal for a large proportion of the households in our billing data.

The second of these external sources is Acxiom, a private market analytics firm that maintains a database of household characteristics and purchasing habits designed to aid clients in the identification and analysis of their customer base. We purchased data on a few variables covering the demographic characteristics of the owners and physical attributes of the home that had substantial coverage in the Acxiom database for the households in our billing sample. For our present purposes, we focus on one particularly relevant characteristic that is reliably measured with good coverage: whether or not the household has a pool of any kind.

2.3 Design

As with the 4000 kWh experiment that was the subject of the previous chapter, the TOU program design likewise generates a 2000 kWh experiment based on the Fuzzy Regression Discontinuity (FRD) design. We include households that were continuously present from the beginning of 2008 through the end of 2010. In order to focus on the effect of the mandatory switch to TOU, we again remove households that had previously volunteered to be placed on the TOU tariff, and in addition, we remove households that had previously been forced onto TOU by virtue of crossing a higher binding threshold in an earlier year.

We define the qualification period to be May through September 2009. That is, any household with maximum monthly usage greater than 2000 kWh in this period is designated a “crosser” and is, by the rules of the policy, intended to receive the treatment, i.e. be irrevocably switched to TOU pricing within six

---

31 We also include households that were present in the dataset for all months but which have missing usage and expenditure data due to billing or recording anomalies. We hence do not have a perfectly balanced panel.
Figure 2.1: Intent to Treat Effect, Propensity to be Treated, April 2010

Notes: Data are smoothed into bins of width 30 kWh.

months of the breach. We choose to exclude the later months of 2009 from the qualification period because households crossing for the first time in the summer of 2009 began to be switched to TOU as early as October, and we do not wish usage and expenditure in the qualification period to be confounded with the treatment. We choose to exclude the earlier months of 2009 from the qualification period because, for this same reason, we would then have to truncate the qualification period even earlier, and this would exclude a large volume of crossers, as most households crossed for the first time or came close to crossing in the summer months.

Figure 2.1 shows the propensity to be treated around the 2000 kWh threshold for both crossers and non-crossers in April 2010, about six months after crossers first started to be switched to TOU. The figure demonstrates that crossing has strong predictive power for treatment status, which is crucial for our
Figure 2.2: Intent to Treat Effects by Month, Propensity to be Treated

Notes: The qualification period is the set of months between the vertical lines. The estimates for each month are from a separate application of equation (1.9) with a bandwidth of 200 kWh. Ability to consistently estimate treatment effects. Figure 2.2 shows this difference across crossers and non-crossers in their propensity to be treated for every month from 2008 to 2010. By construction, no households in the sample, either crossers or non-crossers, were treated before the end of the qualification period in September 2009. But already by October 2009, some households that crossed early in the summer of 2009 were beginning to be switched to TOU. Crossing status is a strong predictor of treatment status for December 2009 through November 2010, which is the set of months that we hence restrict our treatment period to. By the beginning of 2011, most of the crossers from the summer 2009 qualification period had been switched to TOU, but a high proportion of non-crossers had also been switched, by virtue of crossing the 2000 kWh threshold – which remained in effect from 2009 onwards – sometime in 2010.
Another crucial ingredient to a valid experimental design is a lack of manipulation of the forcing variable. Figure 2.3 indicates no visual evidence of manipulation by households in terms of their maximum monthly usage in the summer of 2009. We therefore proceed under the assumption that assignment to treatment in the neighborhood of the threshold is as good as random, which will permit the interpretation of our estimates as causal treatment effects.

2.4 Baseline Results

Figure 2.4 illustrates the intent to treat effect on total usage in April 2010, using a wide bandwidth and flexible functional form. It provides initial visual evidence that households reduced total usage in response to exceeding the 2000 kWh threshold – and thus becoming eligible for the TOU treatment – the previous
Figure 2.4: Intent to Treat Effect, Total Usage, April 2010

Notes: Data are smoothed into bins of width 30 kWh.

summer.

Figure 2.5, which graphs this total usage intent to treat effect for every month in our analysis window, restricts the bandwidth to our optimally-chosen size of 200 kWh and restricts the functional form to be linear on each side.\textsuperscript{32} It shows that the negative intent to treat effect on total usage in April 2010 from Figure 2.4 was also present in all other treatment months from the autumn through the spring. However, it also shows that crossers were virtually identical to non-crossers in the summer months of 2010 in terms of their total usage.

Figure 2.6 shows the corresponding intent to treat effects on total

\textsuperscript{32}We use a bandwidth of 200 kWh throughout our analysis of the 2000 kWh experiment. Each point in Figure 2.4 and its analogues is an average for households with a value of the forcing variable within 15 kWh on either side. Therefore, a bandwidth of 200 kWh on either side of the threshold corresponds to the first six to seven points on each side of Figure 2.4. As this and similar graphs suggest, non-linearities in the relation between treatment variables and the forcing variable become more severe beyond this range. For an extended discussion of optimal bandwidth, see Section A.1.
Notes: The qualification period is the set of months between the vertical lines. The estimates for each month are from a separate application of equation (1.9) with a bandwidth of 200 kWh.
electricity expenditure for all months in the analysis window. They show an identical pattern for total billed amount as for total usage: negative intent to treat effects in the autumn through spring months, and zero intent to treat effects in the summer months. Interestingly, this is precisely the opposite pattern that we found for the 4000 kWh experiment, in which crossers exhibited lower usage and expenditure in the summer months of 2008 but zero intent to treat effects in other months. Finally, note that Figures 2.5 and 2.6 illustrate balance on pre-determined variables, as crossers and non-crossers at the threshold had indistinguishable levels of total usage and total electricity expenditure throughout the pre-experiment period.

Table 2.2 shows the corresponding treatment effects, estimated via the usual 2SLS implementation of equations (1.11)-(1.12). Recall that these are the intent to treat effects adjusted for the difference between crossers and non-crossers in the propensity to be treated. For ease of interpreting magnitudes, these treatment effects are expressed as a percentage of the non-TOU level of total usage or total expenditure, respectively.

The magnitude of the declines in total usage induced by the switch to TOU range from a very substantial 12% to a still notable 7% in the non-summer months, beginning to wane in size and significance in the early summer of 2010, and becoming imprecisely estimated zero effects in June through October. The declines in total expenditure are of a very similar magnitude.

Though the focus here is not on testing the standard model of static utility maximization, as it was in the previous chapter, an interesting point can be noted. The rates in Table 2.1 suggest that, if the effective aggregate price of electricity was higher under TOU than under non-TOU in the non-summer months in 2010, this would have been the case to an even greater degree in the summer months of
Table 2.2: Treatment Effects (%), Total Usage and Total Bill

<table>
<thead>
<tr>
<th></th>
<th>Total Usage</th>
<th>Total Bill</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(3.36)</td>
<td>(2.92)</td>
<td></td>
</tr>
<tr>
<td>Jan. 2010</td>
<td>-8.12 **</td>
<td>-8.06 **</td>
<td>2,522</td>
</tr>
<tr>
<td></td>
<td>(3.64)</td>
<td>(3.28)</td>
<td></td>
</tr>
<tr>
<td>Feb. 2010</td>
<td>-8.91 **</td>
<td>-11.08 ***</td>
<td>2,523</td>
</tr>
<tr>
<td></td>
<td>(3.85)</td>
<td>(3.33)</td>
<td></td>
</tr>
<tr>
<td>Mar. 2010</td>
<td>-11.66 ***</td>
<td>-13.40 ***</td>
<td>2,515</td>
</tr>
<tr>
<td></td>
<td>(3.55)</td>
<td>(3.06)</td>
<td></td>
</tr>
<tr>
<td>Apr. 2010</td>
<td>-8.15 **</td>
<td>-10.52 ***</td>
<td>2,521</td>
</tr>
<tr>
<td></td>
<td>(3.63)</td>
<td>(3.09)</td>
<td></td>
</tr>
<tr>
<td>May 2010</td>
<td>-7.34 *</td>
<td>-9.79 ***</td>
<td>2,522</td>
</tr>
<tr>
<td></td>
<td>(3.78)</td>
<td>(3.21)</td>
<td></td>
</tr>
<tr>
<td>Jun. 2010</td>
<td>0.48</td>
<td>-1.89</td>
<td>2,521</td>
</tr>
<tr>
<td></td>
<td>(3.11)</td>
<td>(2.84)</td>
<td></td>
</tr>
<tr>
<td>Jul. 2010</td>
<td>0.20</td>
<td>-1.82</td>
<td>2,523</td>
</tr>
<tr>
<td></td>
<td>(2.49)</td>
<td>(2.32)</td>
<td></td>
</tr>
<tr>
<td>Aug. 2010</td>
<td>0.87</td>
<td>-0.74</td>
<td>2,521</td>
</tr>
<tr>
<td></td>
<td>(2.00)</td>
<td>(1.85)</td>
<td></td>
</tr>
<tr>
<td>Sep. 2010</td>
<td>-1.16</td>
<td>-2.42</td>
<td>2,521</td>
</tr>
<tr>
<td></td>
<td>(1.89)</td>
<td>(1.71)</td>
<td></td>
</tr>
<tr>
<td>Oct. 2010</td>
<td>-3.92</td>
<td>-7.01 ***</td>
<td>2,522</td>
</tr>
<tr>
<td></td>
<td>(3.15)</td>
<td>(2.71)</td>
<td></td>
</tr>
<tr>
<td>Nov. 2010</td>
<td>-8.45 **</td>
<td>-10.21 ***</td>
<td>2,516</td>
</tr>
<tr>
<td></td>
<td>(3.87)</td>
<td>(3.29)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Significance at the 1% (***), 5% (**), and 10% (*) levels is indicated. Each estimate is from a separate application of equations (1.11)-(1.12) with a bandwidth of 200 kWh, and is the estimated TOU treatment effect as a percentage of the estimated non-TOU level at the threshold for the respective dependent variable. Of the 2,521 households included in the regressions for April 2010, 1,125 are crossers; and the distribution of households is similar in other months.

2010. We thus see a weaker (indeed, zero) response precisely when households should have had a stronger price incentive to respond.

We will leave a detailed examination of load shifting – that is, the substitution of usage between peak and off-peak hours in response to the switch to TOU – for further work. For now, we mention that preliminary evidence suggests that the declines in total usage indicated by Table 2.2 appear to have occurred
roughly uniformly by time of day.\textsuperscript{33}

2.5 Treatment Effects by Household Characteristics

In order to elucidate the behavior behind our baseline treatment effects further, we attempt to decompose them by household characteristics to shed light on which types of households might be responding more strongly than others. From our external data sources, we have data on age and square footage of the home, number of bedrooms and bathrooms, an estimate of home value, and an indicator for whether a pool is present. We have decided to use data on the structure of the home to estimate an energy intensity index, and focus separately on two characteristics: energy intensity and pool presence.

The general idea behind our energy intensity index is to identify households that use a lot of electricity in relation to households with homes of similar size, construction, etc. – i.e. to identify heavy conditional users. We begin by estimating average electricity usage as a function of age of home, square footage of home, number of bedrooms, and number of bathrooms.\textsuperscript{34} We then form an index

\textsuperscript{33}Section A.3 outlines a method to estimate the extent of load shifting. As discussed there, our inability to either observe or impute the on-peak/off-peak allocation of electricity consumption for non-TOU household months – a result, in the first place, of the utility having no need to meter non-TOU households by time of day, and, in the second place, of the mathematical fact that total expenditure is simply a scalar multiple of total usage under the non-TOU bill structure – requires us to estimate it from an external source. Unfortunately, the load profile sample that we have access to is not of adequate size to support precise estimation of the required parameters. Preliminary work reveals point estimates that suggest marginally larger proportional declines in on-peak usage compared to off-peak usage in the early months of 2010, though the standard errors around these point estimates are very large, and a broad range of load shifting – including zero and some degree of reverse substitution – cannot be ruled out.

\textsuperscript{34}Our preferred specification is linear in age and size, with a set of dummy variables each for number of bedrooms and number of bathrooms. We have experimented with various functional forms and the inclusion of additional explanatory variables, and have not found the resulting index to be very sensitive to these choices. We calculate the household average of total usage over all months in the analysis window, i.e. 2008-2010. We estimate the specification of average usage as a function of characteristics using households within the same bandwidth that we use for the estimation of the treatment effects.
based on the residuals of this estimation: those households with a positive residual use more electricity than those with similar characteristics, so receive a higher value of the index. The simplest representation for such an index – and the one we focus on – is simply a binary indicator for whether a household is a high-intensity user (those with a positive residual) or a low-intensity user (those with a negative residual). The natural interpretation of this index is that high-intensity users have more options to potentially exercise in order to reduce average usage. For example, a high-intensity user might be one with electric heat and poor insulation, while a low-intensity user might be one in an identical house and location that has converted to gas heating and installed energy-saving insulation and windows.

Figure 2.7 shows the distribution of households by their distance from the 2000 kWh threshold separately for high-intensity and low-intensity households. Within the chosen bandwidth of 200 kWh on either side of the threshold, the frequency of high-intensity users is slightly increasing in usage in the summer of 2009, while that of low-intensity users is slightly decreasing; but each distribution is close to uniform over this range. There is no visual evidence of manipulation of the forcing variable within either group of households.

Our data on presence of pools is already in binary form, simply indicating whether zero or a positive number of pools is present, with no additional information on number, size, or type of pool if present. Figure 2.8 shows the distribution of households by their distance from the 2000 kWh threshold separately for those with a pool and those without a pool. The number of households with a pool is much lower than the number of households without a pool, but the distribution is fairly uniform for both groups within the 200 kWh symmetric bandwidth around the threshold.

These histograms establish the validity of dividing the 2000 kWh
Figure 2.7: Frequency of Households by Energy Intensity

Note: Data are smoothed into bins of width 10 kWh.

Figure 2.8: Frequency of Households by Presence of Pool
experiment into independent sub-experiments in which participants are first divided by characteristic – either energy intensity or presence of pool – and then into control and intent-to-treat sub-groups. We therefore run separate FRD specifications and estimate treatment effects separately for high-intensity households and low-intensity households, and separately again for households with pools and households with no pools.\footnote{We retain the 200 kWh bandwidth throughout each of these separate specifications, even though the size of the groups varies widely. We have not found a great deal of sensitivity to bandwidth, and have therefore opted for the consistency afforded by a uniform bandwidth.}

Figure 2.9 shows the treatment effects on total usage for high-intensity households, in level form and adjusted for the difference in propensity to be treated between crossers and non-crossers, for each month in the treatment period (December 2009 through November 2010). It illustrates the sharp distinction between non-summer and summer months that was evident in the baseline results for the overall 2000 kWh experiment. The decreased total usage compared to non-TOU households at the threshold throughout the spring is about 200 kWh per month, or upwards of 15\%.

In contrast, Figure 2.10, which presents the analogous treatment effects on total usage for low-intensity households, shows negligible treatment effects in magnitude and significance throughout the entire treatment period. The two sets of treatment effects thus suggest that all of the response indicated by the baseline results is accounted for by high-intensity households. This is perhaps a sensible and predictable finding in relation to our definition of energy intensity, which, as discussed above, is consistent with the highest-intensity households being those that have access to the lowest-hanging fruit in terms of their options to rein in electricity consumption. However, the complete lack of a response by low-intensity households in all months and high-intensity households in summer months raises
Figure 2.9: Treatment Effects, Total Usage, High-Intensity Households

Figure 2.10: Treatment Effects, Total Usage, Low-Intensity Households

Notes: The estimates for each month are from a separate 2SLS application of equations (1.11)-(1.12) with a bandwidth of 200 kWh.
the question of what the precise nature was of the means to save energy that high-intensity households were able to take advantage of in non-summer months, and why these means were apparently unavailable, even to a limited degree, more widely.

To begin to address this question, we examine the seasonal usage patterns of high-intensity and low-intensity households. Figure 2.11 shows the sample mean of total usage, by month over the entire 2008-2010 analysis window, for each group. It illustrates a dramatic difference in non-summer usage, with high-intensity households consuming much more electricity, particularly in winter months. However, it also indicates that high-intensity and low-intensity households have nearly identical summer usage. The most compelling conclusion suggested by Figure 2.11 is thus that high-intensity households differ from low-intensity households primarily in their reliance on electricity for winter heating. This in turn implies that heating is the only margin along which households respond to the TOU treatment – or, in other words, that high-intensity households do not respond in the summer because they do not heat their homes in the summer, and that low-intensity households do not respond at all because their heating is independent of their electricity use.\footnote{An alternative hypothesis is that high-intensity households responded to the switch to TOU by immediately reducing their heating demand, but then forgot about the switch and simply reverted to past behavior after a few months. However, this alternative hypothesis is not supported by the reappearance of the negative treatment effects in late 2010, which coincide with the beginning of a new heating season. Unfortunately, the high propensity for non-crossers in this experiment to have received the TOU treatment by the end of 2010 precludes us from examining whether the high-intensity households continued to exhibit negative treatment effects throughout the subsequent winter and spring. According to the U.S. \textit{Historical Census of Housing}, in 2000, about 10\% of households in states in the northeast relied on electricity for heating, while over 80\% relied on gas or oil for heating.}

This provides some optimism from a policy perspective, as it suggests that initiatives aimed at inducing energy savings could be both successful and cost
Figure 2.11: Mean Usage by Energy Intensity

Note: Sample means are calculated separately by period and group over households within the 200 kWh bandwidth.

effective if targeted at high-intensity users, or, even more specifically, deployed only in non-summer months to households with electric heat. But it leaves unanswered the question of why households were unable or unwilling to respond in the summer, which is especially salient in the policy context, since summer months correspond to system-wide peak demand and hence periods in which demand reductions would be most beneficial in terms of cost containment for utilities and lower reliance on pollution-intensive peak generation sources more broadly. High-intensity households appear to have been motivated to adjust their indoor temperature in non-summer months by turning their heat down, but neither high-intensity nor low-intensity households appear to have been motivated to do so in summer months by turning their air conditioning down. Since, as mentioned above, the price incentive to reduce was slightly stronger in the summer of 2010 than in the spring, this may suggest that households’ indoor climate preferences
are very sensitive to season or other factors, and hence that much sharper incentives – in terms of prices or otherwise – might be necessary to induce adjustments to indoor temperature in the summer compared to other seasons.

Figures 2.12 and 2.13 show treatment effects for households with pools and households with no pool, respectively. Neither pool owners nor households with no pool exhibit large treatment effects, but TOU households without a pool do have lower usage than their non-TOU counterparts in the non-summer months, mirroring the pattern of the overall 2000 kWh treatment effects in direction. TOU pool owners show only one non-summer month of lower usage than their non-TOU counterparts, and actually exhibit a positive and statistically significant treatment effect in both June and September. These positive summer treatment effects may indicate that pool users take advantage of the low off-peak rate under the TOU tariff by running their pool pumps at night rather than during the day, but also run them for longer than they otherwise would.\(^{37}\)

2.6 Responses to a Non-Price Treatment

To attempt to get a better sense of the role of prices in inducing the responses we have estimated, we take advantage of a type of placebo treatment that the utility administered in the form of a warning letter. In 2008, when the 3000 kWh threshold was binding, the utility – which had already informed its customer base of the TOU policy at its inception in late 2006 – sent a bill insert to non-TOU households immediately following the first billing period in which they exceeded 1750 kWh in total usage.

\(^{37}\)Since presence of a pool was not included in the regression underlying our energy intensity index, we might expect pool ownership to be positively correlated with energy intensity. However, the treatment effects for households with pools bear little resemblance to those for high-intensity households, further supporting our deduction that the principal trait distinguishing high-intensity households is their reliance on electricity for heating.
Figure 2.12: Treatment Effects, Total Usage, Households with Pools

Figure 2.13: Treatment Effects, Total Usage, Households with No Pool

Notes: The estimates for each month are from a separate 2SLS application of equations (1.11)-(1.12) with a bandwidth of 200 kWh.
The purpose of the letter was to remind households that continuing to consume in that approximate range would, starting the following year, result in triggering the switch to TOU. However, the letter could be read as painting TOU in a somewhat negative or foreboding hue. It states that “[Regulator] guidelines require [us] to move customers who use more than 2000 kWh in any billing cycle to a different rate,” and that “once you are switched . . . your account cannot switch back.” It even suggests that a customer might prefer to remain on a flat rate, in which case “you should monitor your energy use and take energy reducing measures to remain below the 2000 kWh threshold.”

Given the tone of this letter, we might expect two things: that households would indeed attempt to rein in their usage to avoid the implied unpleasantness of TOU; and, for those that failed in this endeavor and were eventually switched to TOU by virtue of crossing the 2000 kWh threshold at some point thereafter, that they might try to avoid the implied unpleasantness by reining in usage even more after the fact. In other words, a response to the simple receipt of this letter in late 2008 or early 2009 might suggest that part of the TOU treatment effect we find in late 2009 and beyond is a delayed or continued response to the letter and unrelated to the actual price change implied by the switch to TOU.

To explore these issues, we estimate treatment effects on total usage, where the treatment in question is receipt of the letter. We cannot observe if any household actually received the letter, but we do know that it was supposed to have been sent to all households that exceeded 1750 kWh in any billing month in 2008. We therefore apply a sharp regression discontinuity framework, taking an indicator for whether a household breached the 1750 kWh threshold as the treatment indicator and examining households immediately on either side of the threshold.
Figure 2.14: Warning Letter Treatment Effects, Total Usage

Figure 2.14 illustrates these treatment effects for late 2008 through late 2009. We find negative treatment effects throughout this period, suggesting that households indeed appear to have responded to receiving the letter by reducing electricity use. This may have been driven by an attempt to avoid the switch to TOU, but it may also have been related to some households interpreting the letter as a warning of being in violation of a social norm or some other broader notion of acceptable behavior. In any case, while not trivial in magnitude, the treatment effects are estimated somewhat imprecisely. Further, any attempted manipulation of total usage to avoid the switch to TOU appears to have been unsuccessful, as it was not of sufficient magnitude or targeted accurately enough to leave a visual trace in the density of the 2000 kWh forcing variable in Figure 2.3.

These estimated treatment effects are the result of preliminary work. We are still in the process of exploring issues such as the mechanism and timing of the letter’s delivery and hence the appropriate definition of the qualification period, optimal bandwidth, balance on predetermined variables and other tests for a valid
experimental design, and other communication between the utility and segments of its customer base concerning the TOU program. Nonetheless, they provide suggestive evidence that some households do respond to non-price interventions – enough, at least, to raise the intriguing question of whether the 2000 kWh treatment effects discussed above are driven more by such behavioral effects than by the actual rate change induced by the switch to TOU.

2.7 Discussion

From a policy perspective, perhaps the most important message to take away from our findings is that there may be great value in targeting interventions rather than rolling them out to the entire customer base. Our identification of high-intensity users has been performed using ostensibly publicly available data. Utilities should be able to estimate even more sophisticated and accurate indicators of the intensity of a given customer’s use conditional on observable characteristics. Programs could then be designed that specifically target such households, with concomitant cost savings associated with rolling out the rate change and any requisite equipment on a more limited basis.\footnote{While no direct estimates are available, informal evidence suggests that, for the TOU program we study, the utility faced substantial costs associated with providing and installing new meters capable of tracking household usage by time of day.}

Why neither high-intensity nor low-intensity users responded to the TOU program in summer months in the particular instance that we study remains less than clear. While high-intensity users – who, as suggested by seasonal usage patterns, tend to be those with electric heat – appear to have treated the adjustment of indoor temperatures as an easy way to reduce electricity consumption in non-summer months, neither group of households seemed willing to contemplate such an adjustment in the summer. Unfortunately, features of the
experimental design do not allow us to test whether this lack of a summer response was peculiar to the unusually long and hot summer of 2010. Regardless, our present results suggest at least some sensitivity of indoor climate preferences to season or other factors, which would represent an important dimension to consider in terms of policy design and targeting.

Ultimately, our work indicates that there is still much to be done to learn how households respond to TOU and other electricity pricing interventions, and how such policy instruments can be constructively deployed in the pursuit of either environmental or system efficiency goals. Of course, if equivalent responses can be induced simply through various forms of communication with customers, the targeting and fine-tuning of price interventions may become of less interest – though the channels through which such non-price interventions operate remain perhaps the least understood of all.
CHAPTER THREE

A Multidimensional Examination of Jury Composition, Trial Outcomes, and Attorney Preferences\textsuperscript{39}

3.1 Introduction

Legal scholars and the public have frequently voiced concerns that the American legal system fails to provide defendants with their constitutional right to an impartial jury in criminal prosecutions.\textsuperscript{40} The focal point of such concerns is typically the jury selection process, which is regarded as giving attorneys a means to stack juries with biased jurors.\textsuperscript{41} But underlying this view is a supposition that attorneys can identify such biased jurors in the first place. This identification is straightforward if juror biases are correlated with observable juror characteristics such as race, which they are commonly assumed to be. Thus, the trial of the police officers charged with the beating of Rodney King – in which an entirely non-black jury failed to convict four white men of excessive force against a black victim despite visceral video evidence – is seen as epitomizing a modern manifestation of the old maxim attributed to Clarence Darrow that a case is won or lost by the time the jury is sworn in.\textsuperscript{42}

\textsuperscript{39}This chapter is the product of joint work with Jee-Yeon K. Lehmann, of the University of Houston, Department of Economics.

\textsuperscript{40}The Sixth Amendment of the U.S. Constitution states that “The accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed.”

\textsuperscript{41}For example, Rothwax (1996) suggests that attorneys “can mold a jury in the hope that it will be swayed by emotion and innuendo, not fact.”

\textsuperscript{42}In the initial state trial of the officers, the jury was comprised of ten Caucasians, one white Hispanic, and one Asian. Three of the officers were acquitted of all charges, while the jury was hung on one charge and acquitted on others for the fourth officer. The infamous videotape of the beating was, of course, not the only evidence presented at trial, and it remains difficult to predict if the trial outcome would have been different if a jury of a different racial composition had heard all of the same evidence, testimony, and judge instructions. See Linder (2007) for further details and discussion.
In this chapter, we assess the degree to which such concerns are justified along several dimensions. First, we ask if certain demographic and socioeconomic characteristics are indeed related to the sentiments and biases that individual jurors take with them into the deliberation room or to the verdicts at which juries collectively arrive. In examining this question, we do not focus solely on race, but also jointly consider other juror characteristics – such as sex, age, religiousness, education, and income – that existing literature has largely neglected. Second, we place the size of the jury composition effects on verdicts in the context of other important aspects of the trial, especially the strength of the evidence presented in the courtroom. Finally, we assess whether attorneys are aware of the effects of juror characteristics on pre-deliberation biases and trial outcomes. In addressing these questions, we take advantage of a uniquely rich dataset on non-capital felony jury trials held in four major U.S. state trial courts during 2000 to 2001 that includes detailed survey information on various aspects of the case from all trial participants, including the jurors, the judge, and the attorneys.

We find strong evidence of significant relations between individual jurors’ demographic and socioeconomic characteristics and their pre-deliberation leanings. In particular, we find that jurors with higher levels of education and religiousness and, most consistently, those with higher income are significantly more likely to hold biases and sentiments favoring the prosecution compared to other jurors who heard the same case. Female jurors hold some pre-deliberation opinions weakly favoring the prosecution that depend to some degree on the gender of the victim. Black jurors show greater favoritism for the defense in general, and especially tend

---

43 Los Angeles (CA), Maricopa (AZ), Bronx (NY), and Washington D.C.
44 The particular dimensions along which we examine juror bias include sympathy for the victim, trust in the police, subjective evaluation of each side’s overall case, and interpretation of the trial evidence.
to express pro-defense sentiments when the defendant is also black.

We then examine whether the juror characteristics that are associated with significant individual pre-deliberation biases also have strong effects on the verdicts arrived at post deliberation. Using a rich set of case-level controls, including attorney demographics and experience measures as well as the judge’s assessment of attorney skills and trial evidence, we find that juries with higher average income and religiousness hand out significantly fewer acquittals, closely mirroring our findings regarding jurors’ pre-deliberation biases. We also find that juries with a greater proportion of blacks convict on fewer counts when both the defendant and victim are black. Finally, we find that a greater proportion of females and greater average age and years of education amongst seated jurors are weakly associated with trial outcomes favoring the defendant. These jury composition effects are small compared to the effect of evidentiary strength, and account for less of the explained variation in verdicts than do the factual and legal aspects of the trial. Specifically, reducing acquittals by the same amount as would be induced by increasing the strength of the evidence by one standard deviation would require, other things equal, replacing four black jurors with non-black jurors on a twelve-person jury or increasing the income of each juror by about two fifths.

Having established the significance of jury characteristics in shaping juror predispositions and trial outcomes, we examine whether attorneys correctly anticipate these associations by asking if they hold preferences for juror characteristics that tend to favor their side. Capitalizing on the availability of a self-reported measure of attorney satisfaction regarding the seated jury, we ask whether the prosecutor’s and the defense attorney’s levels of satisfaction can be explained by the characteristics of the seated jury, once again conditioning our
analysis on a rich set of controls.\textsuperscript{45} Robust to all specifications, we find that seated juries with higher average income are associated with significantly higher satisfaction levels for the prosecution and lower satisfaction levels for the defense. This is fully consistent with our findings that higher income is strongly related to pro-prosecution juror sentiments and fewer acquittals. On the other hand, in contrast to the weak association between gender and verdicts and to the tendency for female jurors to hold some biases favoring the prosecution, we find strong evidence that defense attorneys prefer having more women on the jury.

Our results regarding prosecuting and defense attorney preferences for blacks on the seated jury are inconclusive. While the effects of race on attorney satisfaction are sometimes large, they are often very imprecisely estimated. Further, they suggest that prosecuting attorneys may sometimes prefer juries with a greater proportion of blacks, in contrast to our findings that black jurors tend to hold strong predispositions in favor of the defense in general. We speculate on some factors that may be confounding the estimation and interpretation of these race effects on attorney satisfaction, including the fact that it is ostensibly illegal to remove potential jurors on the sole basis of race during jury selection.\textsuperscript{46}

Our analysis improves upon existing studies in several ways. First, while past studies of the impacts of jury composition on trial outcomes have focused predominantly on race (e.g. Anwar et al., 2012b; Lee, 2010; Bowers et al., 2001), our results demonstrate that other demographic and socioeconomic characteristics also play an important role in shaping the biases of juries. Furthermore,

\textsuperscript{45}We also limit our analysis to cases in which attorneys reported their satisfaction level with the seated jury before learning the verdict, so that their evaluations should not be influenced by the eventual trial outcome.

\textsuperscript{46}The Supreme Court ruling in \textit{Batson v. Kentucky} (1986) and other subsequent rulings forbid attorneys to remove jurors solely on the basis of race or gender. However, the effectiveness of these rulings has been widely questioned. This will be discussed in more detail in the following section.
accounting jointly for race and other characteristics correlated with race allows us to disentangle effects that may have been misattributed solely to race in past studies that omit these other characteristics.\textsuperscript{47} Second, our dataset contains trials from multiple courtrooms across several jurisdictions with varying racial diversities and histories of racism. As such, results from this paper should be helpful in assessing the nature and the extent of jury composition effects on trial outcomes, as well as the degree to which attorneys have preferences over certain juror characteristics, outside the deep southern states and across counties with varying racial mixes.\textsuperscript{48} Third, we take a stated preference approach to assessing which juror characteristics are preferred by attorneys on both sides. Past studies have instead taken a revealed preference approach by exploring the types of potential jurors removed by each attorney, which may be more susceptible to confounding effects of constraints faced by attorneys during jury selection.\textsuperscript{49}

Finally, due to the richness of our dataset, we are able to incorporate into our

\textsuperscript{47}Sommers and Ellsworth (2003) review evidence from the legal literature on the effects of race in criminal jury trials. Also see Pfeifer (1990) for a critical appraisal of some of the earliest of this evidence. Lieberman and Sales (2007) and Baldus et al. (2001) also review evidence on the effects of gender and other socio-demographic characteristics. The studies reviewed by these authors are often based on small samples constructed from archival material, on mock trials, or on public opinion polls, and consider the selected juror characteristics in isolation rather than jointly. In contrast, Anwar et al. (2012b) account for age and gender jointly with race, while Anwar et al. (2012a) undertake a similar joint analysis focusing on age, both employing a large dataset on actual trials. Shayo and Zussman (2011), Alesina and La Ferrara (2011), Abrams et al. (2011), and Iyengar (2011) provide evidence that the race of judges, victims, and defendants can affect trial outcomes in other settings.

\textsuperscript{48}Anwar et al. (2012a,b) use data from Lake County and Sarasota County in Florida, which are predominantly white. Rose (1999) uses data from a county in North Carolina described as “largely biracial”. Bowers et al. (2001) use data from the Capital Jury Project, which encompasses trials from 14 states, though 11 of these are in the south. Diamond et al. (2009) use data from just a single courtroom, though in the more racially diverse Cook County, Illinois. Baldus et al. (2001) use data from Philadelphia.

\textsuperscript{49}This method has been implemented by Anwar et al. (2012a) to identify attorney preferences over juror age; by Baldus et al. (2001) to identify attorney preferences over juror race, gender, and age; by Grosso and O’Brien (2012) and Rose (1999) to identify attorney preferences over juror race; and, for civil trials, by Diamond et al. (2009) to identify attorney preferences over juror race, gender, age, and income. The jury selection process and the various restrictions that can be imposed upon attorneys during the process will be discussed in the following section.
analysis many aspects of trials and participants that are unobserved in most previous studies, including, most notably, information on the evidence presented at trial and jurors’ interpretation of that evidence.

Overall, our results give some justification for concerns about a lack of impartiality and trustworthiness in U.S. criminal jury trials. One implication is that defendants face the prospect of an unfair trial simply by virtue of being unlucky in the characteristics of the jurors drawn from the jury pool on the day of the trial. But an even more disconcerting implication is that attorneys, with at least some apparent ability to correctly anticipate relations between juror characteristics and biases, might attempt to use this knowledge to manipulate trial outcomes. However, our results also provide some important context for such concerns. First, our multidimensional approach reveals that the sources of jury bias are more nuanced than an analysis based on race alone would suggest. Second, our results on attorney preferences indicate that attorneys have only a partial understanding of these sources of jury bias, which may limit their opportunities to attempt to leverage them to their advantage. Finally, our comparison of the effects of jury composition with that of evidentiary strength suggest that the fundamental issues of fact and law remain the primary determinants of the outcome of any given trial. Investigating these issues further and exploring how our results generalize to regions and types of cases not covered by our dataset are crucial next steps for evaluating the fairness of jury trials and the validity of the verdicts delivered by the U.S. justice system.

The rest of the chapter is organized as follows. Section 3.2 provides a description of the jury selection process and other relevant institutional details. Section 3.3 describes the data set and the restricted sample of cases used in our empirical analysis. Section 3.4 examines the link between juror characteristics and
both individual biases and trial verdicts. In Section 3.5, we present our investigation of attorney preferences in jury selection. Finally, we conclude and discuss some additional policy and research questions that our findings raise.

### 3.2 Institutional Background

Throughout our analysis, we will be using data on the characteristics of seated jurors. It is thus important to understand the many forces that can influence which jurors come to be seated on the jury for a given criminal trial.

The broad goal of the overall jury selection process is to randomly draw a panel – also called a venire – of potential jurors from the population of the county in which the alleged crime was committed; then, through a procedure known as voir dire, to remove any potential jurors with an inability to be impartial from the possibility of serving on the seated jury. Of interest here are the details of how the process is actually implemented, and how this broad goal can fail to be met.\(^{50}\)

There are a number of reasons why the initial panel may not be a random draw from the county population. Juror rolls, from which the names of those to be summonsed on a particular day are randomly drawn, are usually constructed from drivers’ license, tax, voter registration, or other administrative databases, and these can systematically under-represent certain groups. Summonses that are undeliverable, never responded to, or received by those with automatic exemptions can likewise over-represent certain groups.\(^{51}\)

There is also the question of which cases actually make it to a jury trial in

\(^{50}\)Gobert et al. (2009) and Starr and McCormick (2001) each provide a very comprehensive description of jury selection, from both a theoretical and practical perspective. This section draws primarily from these two sources and from a number of interviews we have conducted with legal professionals who have experience with the process.

\(^{51}\)See Mize et al. (2007), Sommers (2008), and American Civil Liberties Union of Northern California (2010) for further details.
a certain county. Defense attorneys are occasionally successful at requesting a change of venue into or out of a given county on the grounds that it would be too difficult to form an impartial jury in the original trial location. Defendants sometimes request a bench (judge-only) rather than a jury trial. And the set of cases that reach a conclusion prior to a jury rendering a verdict or are never prosecuted is determined by a series of decisions and negotiations made by the court and attorneys on both sides. All of these issues raise the possibility that the types of cases heard by a jury in a given courtroom can depend to some extent on attorney expectations of the types of jurors who will eventually comprise the jury.

We will not attempt to diagnose or account for potential non-randomness in the venire or in the trials represented in our dataset. As with past studies, we have no satisfactory means to accomplish this. At the same time, we note that the states covered by our dataset have been above-average or amongst the leaders in jury reform efforts related to jury pool diversity (Mize et al., 2007), and that defense motions for changes of venue are rarely successful outside of very high-profile cases.\footnote{For example, litigation consultant Gary Moran has stated that such requests are often motivated primarily by delay and diversion. He concludes: “The judge is going to sit there and sit there and after four hours . . . he will finally utter the magic word – ‘Denied.’ And then we all move along” (quoted in Kressel and Kressel, 2002, 59).} We, therefore, do not expect these issues to substantially affect our results. On the other hand, voir dire presents a serious potential source of endogeneity, which we make a concerted effort to redress in our analysis. Voir dire can be leveraged by attorneys for far more than simply identifying potential jurors with an inability to be impartial. In fact, lawyers are taught to identify potential jurors who are likely to hold unfavorable predispositions to their side and to attempt to remove them. This often takes the form of targeting specific socio-demographic groups that community surveys or mock trials have suggested
are more likely than others to hold certain opinions on various aspects of the case. However, attorneys face a number of institutional constraints in this endeavor and must also consider the strategies of their opposing counsel.

During voir dire, potential jurors are questioned, and on the basis of their responses, can be deselected from (rather than selected to) the seated jury. The sole legal rationale for voir dire is to ensure that the seated jury is impartial. Individual judges have wide latitude in deciding how this is to be accomplished: there are no constitutionally-mandated procedures, and while most states and some local courts have guidelines or a set of common practices in place, these are rarely binding for any given trial.

The examination of potential jurors can be carried out primarily by the judge, by the judge with suggestions or a greater degree of participation from the attorneys, or primarily by the attorneys. It can also be carried out in the open courtroom or, less often, privately in the judge’s chambers or at the sidebar. The primary concerns that are examined in any voir dire are the existence of personal relationships between panelists and any other parties involved in the trial, and the capacity to understand and contemplate the relevant legal issues and evidence in a dispassionate and impartial manner. But many other subjects are often explored as well, especially if attorneys have a high degree of participation. Attorneys may also have access to the results of questionnaires filled in by potential jurors: in almost all cases, they observe the information on the potential jurors that the court collects, which includes basic demographic and occupational data at the least; and they are sometimes also permitted to design and administer

\footnote{In addition to Gobert et al. (2009) and Starr and McCormick (2001), which include much practical instruction for attorneys, see also Hoffman (2006) and Kressel and Kressel (2002) on attorney strategies, and Lieberman and Sales (2007) and Posey and Wrightsman (2005) on the professional practice of what has come to be known as scientific jury selection.}
“Supplementary Juror Questionnaires” (SJQs) to elicit more detailed information. The judge has discretion over each of these facets of the examination phase; individual judges tend to have strong views on how these decisions should be made, and to hold adamantly to them.

Attorneys can remove panelists via two instruments: strikes for cause and peremptory strikes. An attorney on a given side can challenge a potential juror for cause by arguing that an inability to be impartial has been demonstrated, and the challenge can then be debated amongst the attorneys and the judge. If the judge rules in favor of the strike, the potential juror is dismissed. Attorneys can strike an unlimited number of potential jurors for cause, as long as the judge can be sufficiently convinced that a basis for disqualification has been demonstrated. Each attorney can also remove a limited number of potential jurors without stating any reason by exercising peremptory strikes. The total number of peremptory strikes available is generally small relative to the size of the panel and is usually divided equally between the defense and prosecution, though again, these matters are ultimately at the discretion of the judge.54

Although attorneys ostensibly have full discretion over how to exercise their peremptory strikes, they are in fact legally obliged to do so in a non-discriminatory manner. If an attorney suspects that the opposing counsel has exercised a peremptory strike solely on the basis of a potential juror’s race or gender, a “Batson objection” can be raised.55 When a Batson objection is raised

54The order in which strikes are made – across attorneys and types of strikes – depends to some extent on whether the struck jury method or some version of the strike-and-replace method is being employed. These are further details that can vary widely by trial and courtroom, though they are less relevant to the present discussion. See Gobert et al. (2009) and the related discussion in the following chapter for more on variations in the sequence in which voir dire can be conducted.

55The seminal ruling in Batson v. Kentucky (1986) established that the prosecuting attorney must be able to state a race-neutral justification if the defense counsel can make a prima facie case that a peremptory strike has been exercised against an African American potential juror.
and sustained by the judge, the opposing counsel must provide a race- or
gender-neutral reason for the strike in order for it to be allowed. If such a reason
is not provided, the strike will be disallowed, and other remedies can occasionally
be imposed at the judge’s discretion as well, even leading to a new voir dire or a
mistrial in extreme cases. Many legal commentators regard Batson as only a minor
impediment to how attorneys exercise peremptory strikes, since attorneys can and
often do comply with its requirements by stating neutral but extremely arbitrary
and improbable reasons for the strikes.56 Less direct ways in which Batson might
influence attorney behavior, such as heightening concerns over reputation or the
likelihood of verdicts being overturned on appeal, have not been studied.

This description indicates that attorneys have some control over the
composition of juries, with the degree of control increasing in the number of
peremptory strikes available and in the freedoms afforded to attorneys in eliciting
information about the potential jurors.57 Attorneys can thus influence trial
outcomes through multiple channels: by affecting which jurors from the pool of
potential jurors with heterogeneous biases will hear the case; and through the
usual processes of argument and presentation of evidence during the trial itself.
One would thus not be surprised to find correlations between verdicts and jury
composition, since both can be affected by common elements of attorney
strategies, which themselves can depend on other aspects of the case. In order to
isolate the causal effect of jury composition on trial outcomes, we will thus need to
address this endogeneity. We intend to do so by controlling for a rich set of

56See, for example, Equal Justice Initiative (2010) and Norton et al. (2007).
57We provide some theoretical and empirical support for this intuition in the following
chapter, and also confirm the importance of relative attorney skill. See also Neilson and Winter
covariates reflecting the freedoms and constraints faced by attorneys in voir dire, the attorneys’ skills in designing and implementing effective strategies, and the aspects of each trial that attorneys might condition their strategies on. We will return to this in Section 3.4, but will first introduce our dataset and the wealth of information it contains.

3.3 Data Description

3.3.1 The NCSC/ICPSR Hung Juries Dataset

Our empirical analysis relies on a detailed examination of 351 felony jury trials in four major county courts across the U.S. collected by Hannaford-Agor et al. (2002, 2003) for the National Center for State Courts (NCSC) during 2000 and 2001 and disseminated by the Interuniversity Consortium for Political and Social Research (ICPSR). Although the NCSC’s main goal in the study was to provide an empirical evaluation of hung juries, the dataset is not limited to trials ending in a hung jury, and indeed includes all felony trials held in these courts during the specified period of data collection.\textsuperscript{58} The four courts – Los Angeles County Superior Court in California, Maricopa County Superior Court in Arizona, Bronx County Supreme Court in New York, and District of Columbia Superior Court in Washington, D.C. – were chosen based on their high volume of felony jury trials and their willingness to cooperate with the data collection process and guidelines. Each court is fairly evenly represented in the dataset, with a single location accounting for about 85 trials each.\textsuperscript{59}

The NCSC study provides a comprehensive look at the trial and all the parties involved: the defendant(s) and the victim(s) (if any), the judge, the

\textsuperscript{58}The period of data collection varied by county, ranging from 4 to 11 months.
\textsuperscript{59}D.C. has a slightly larger representation, with about 97 cases.
attorneys, and, most importantly for our analysis, the seated jurors. We do not have information about the initial jury pool from which the seated jury was formed, except for the initial panel size.\textsuperscript{60}

The case data in the NCSC study provide researchers with an extensive look at all the charges, the race and sex of the defendant(s) and victim(s), and the voir dire process that led to the seated jury. The NCSC questionnaires also asked the presiding judge of each trial for his/her evaluation of the evidence, case complexities, and attorney skills. The main variables of interest for our empirical analysis come from the attorney and the jury questionnaires. The attorneys in each trial were asked about their satisfaction with the voir dire process, legal experience, and basic demographic information, including sex and age. In addition, the juror questionnaire provides a rich set of demographic information about each seated juror, including age, sex, race, education, income, and religious beliefs. The NCSC data also contain a description of the dynamics of each juror’s opinion formation, and of pre- and post-deliberation perceptions and opinions regarding the defendant(s) and victim(s). To our knowledge, the richness of the jury demographic and opinion information available in the NCSC study is unmatched by any other readily available dataset.

Regrettably, although the NCSC dataset contains information on 351 cases and 3,497 jurors in total, not every juror in our data answered every survey question. Since much of our analysis relies on the average characteristics of the seated juries, we restrict our sample of interest to cases in which these average characteristics are well-measured. Specifically, we focus on trials for which six or

\textsuperscript{60}We can, however, infer the average characteristics of the panel from those of the county as a whole, since, as was discussed previously, a panel is essentially a random sample of the county population. We include county fixed effects in all of our empirical specifications, in part to account for the composition of the panel as much as possible.
more jurors responded to questions about education, income, age, race, gender, and religion to reduce any bias that may result in our calculation of average jury characteristics due to systematic non-reporting. Small changes to this cutoff point do not alter our main findings. Additionally, we further restrict our sample to those cases in which attorneys reported their voir dire satisfaction measures before learning the verdict, in order to minimize the possibility that our measure of voir dire satisfaction is confounded by satisfaction with broader aspects of the trial, specifically outcomes.

3.3.2 Summary Statistics

Table 3.1 provides some descriptive statistics of our restricted sample. Panel A shows that the majority of cases in our restricted sample fall into the categories of property or drug-related crimes. Consequently, a possible criticism of our dataset is that most of our cases are restricted to non-capital felonies that are not as “serious” as homicides (actual or attempted) or sexual crimes. In turn, attorneys might not pursue the same strategies, and jurors might not act as strongly on their personal prejudices as would be the case with more serious charges. Therefore, the results that we find might be underestimates of the level of bias present in more serious cases. While a valid concern, we believe that our results remain interesting for the following reasons. First, many of the previous studies on bias and discrimination in jury trials have focused on capital crimes. However, we believe that it is important to assess potential impediments to fairness across various types of trials/crimes that are much more prevalent and frequent in courtrooms. Second, in all of our regression specifications, we control for the type of crime. Moreover, in alternative specifications, we interact our main variables of interest with an indicator of whether the crime is a murder or a sex
crime and find no significant differences in our main results across case types.

Panel B shows the case-level means of the various sets of controls on which we rely in our empirical analysis. The case descriptive variables show that the majority of the cases have black defendants and black victims. Over half of the cases are represented by a public defender, which generally speaks to the low income levels of the defendants. There is wide variation in the voir dire process, from who conducted the voir dire\textsuperscript{61} to the length of time it took to shape the final jury. A Batson objection was raised in about 14% of the cases. A typical case in our sample involves close to three different charges against the defendant(s). Of these charges, about 56% result in a conviction, 38% in an acquittal, and the remaining 6% in a hung jury.

The next set of variables summarizes the average characteristics of the seated jury. These are our main variables of interest. In the jury supplemental survey, jurors only provide categorical responses to questions about their education, income, and age. In order to facilitate a more useful interpretation of our coefficients, we translate the categorial variables into appropriate levels for each variable by taking the mid-point of each category. However, using the categorical variable directly does not change our main results. In our restricted sample, a typical seated jury has jurors with about 15 years of education who earn $58,000 in annual income and are about 42 years old. On average, roughly 66% of the jurors seated for a particular case describe themselves as religious.\textsuperscript{62} A typical jury in our restricted sample is over 50% women and about a quarter black, and the sex and race of the foreperson generally reflect the proportions in the overall

\textsuperscript{61}“Who conducted” is a variable ranging from 1 to 4, increasing in attorney involvement, where 1 is “Judge with little or no attorney involvement” and 4 is “Attorney with little or no judge involvement.”

\textsuperscript{62}Answered 1 or 2 on a scale of 1 to 5 where 1 is very religious, 2 is religious, and 5 is very nonreligious.
### Table 3.1: Case Summary Statistics

#### Panel A:

<table>
<thead>
<tr>
<th>Case Type</th>
<th>% of Cases</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murder (1st, 2nd, attempted) and Manslaughter</td>
<td>16.18</td>
<td>22</td>
</tr>
<tr>
<td>Sexual crime</td>
<td>2.94</td>
<td>4</td>
</tr>
<tr>
<td>Robbery, Burglary, Larceny, Theft, Assault, Arson</td>
<td>39.71</td>
<td>54</td>
</tr>
<tr>
<td>Others, including drug-related crimes</td>
<td>41.18</td>
<td>56</td>
</tr>
</tbody>
</table>

#### Panel B:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victim</td>
<td>0.594</td>
<td>0.493</td>
<td>138</td>
</tr>
<tr>
<td>Female Victim</td>
<td>0.231</td>
<td>0.423</td>
<td>134</td>
</tr>
<tr>
<td>Non-black Victim</td>
<td>0.333</td>
<td>0.473</td>
<td>135</td>
</tr>
<tr>
<td>Female Defendant</td>
<td>0.098</td>
<td>0.299</td>
<td>132</td>
</tr>
<tr>
<td>Black Defendant</td>
<td>0.555</td>
<td>0.499</td>
<td>137</td>
</tr>
<tr>
<td>Public Defense</td>
<td>0.519</td>
<td>0.502</td>
<td>133</td>
</tr>
<tr>
<td>Total Number of Charges</td>
<td>2.713</td>
<td>2.277</td>
<td>136</td>
</tr>
<tr>
<td>Number of Charges Convicted</td>
<td>1.415</td>
<td>1.975</td>
<td>135</td>
</tr>
<tr>
<td>Voir Dire Characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who Conducted (1 “Judge” to 4 “Attorney”)</td>
<td>2.309</td>
<td>0.923</td>
<td>136</td>
</tr>
<tr>
<td>Questionnaire Used</td>
<td>0.223</td>
<td>0.418</td>
<td>139</td>
</tr>
<tr>
<td>Voir Dire Time (hrs)</td>
<td>5.343</td>
<td>6.278</td>
<td>137</td>
</tr>
<tr>
<td>Anonymous Jury</td>
<td>0.273</td>
<td>0.447</td>
<td>139</td>
</tr>
<tr>
<td>Batson Objection Raised</td>
<td>0.137</td>
<td>0.345</td>
<td>139</td>
</tr>
<tr>
<td>Jury Characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education (years)</td>
<td>15.101</td>
<td>1.090</td>
<td>139</td>
</tr>
<tr>
<td>Income ($, thousands)</td>
<td>57.603</td>
<td>12.462</td>
<td>139</td>
</tr>
<tr>
<td>Age</td>
<td>41.523</td>
<td>3.982</td>
<td>139</td>
</tr>
<tr>
<td>Religious</td>
<td>0.663</td>
<td>0.160</td>
<td>139</td>
</tr>
<tr>
<td>Female</td>
<td>0.569</td>
<td>0.179</td>
<td>139</td>
</tr>
<tr>
<td>Black</td>
<td>0.261</td>
<td>0.235</td>
<td>139</td>
</tr>
<tr>
<td>Woman Foreperson</td>
<td>0.508</td>
<td>0.502</td>
<td>130</td>
</tr>
<tr>
<td>Black Foreperson</td>
<td>0.248</td>
<td>0.434</td>
<td>133</td>
</tr>
<tr>
<td>Attorney Characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defense:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voir Dire Satisfactory (1 “Low” to 7 “High”)</td>
<td>5.403</td>
<td>1.483</td>
<td>139</td>
</tr>
<tr>
<td>Legal Practice (yrs)</td>
<td>12.903</td>
<td>8.459</td>
<td>134</td>
</tr>
<tr>
<td>Previous Criminal Trials</td>
<td>62.323</td>
<td>73.089</td>
<td>133</td>
</tr>
<tr>
<td>Age</td>
<td>41.880</td>
<td>10.239</td>
<td>133</td>
</tr>
<tr>
<td>Female</td>
<td>0.313</td>
<td>0.465</td>
<td>131</td>
</tr>
<tr>
<td>Prosecution:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voir Dire Satisfactory</td>
<td>5.396</td>
<td>1.377</td>
<td>111</td>
</tr>
<tr>
<td>Legal Practice (yrs)</td>
<td>8.696</td>
<td>6.569</td>
<td>112</td>
</tr>
<tr>
<td>Previous Criminal Trials</td>
<td>39.584</td>
<td>57.515</td>
<td>113</td>
</tr>
<tr>
<td>Age</td>
<td>36.372</td>
<td>7.913</td>
<td>113</td>
</tr>
<tr>
<td>Female</td>
<td>0.495</td>
<td>0.502</td>
<td>105</td>
</tr>
</tbody>
</table>

**Notes:** Sample limited to cases with more than 5 jurors responding to school, income, age, race, gender, and religion questions, and non-missing site information. Also restricted to cases in which both the defense and the prosecution answered voir dire satisfaction question before learning verdict.
Throughout our empirical analysis, we estimate specifications in which we control for attorney characteristics, because the attorney’s satisfaction with the voir dire is a self-reported variable and his/her experience or sex may influence his/her evaluation of the seated jury. The mean level of satisfaction is the same for the defense as for the prosecution, and the dispersion of the satisfaction measure is similar for the two sides as well. A typical defense attorney in our sample is older, has much more legal experience, and is more likely to be male compared to a typical prosecutor. Unfortunately, our data do not include the race of the attorneys.

3.4 Juror Characteristics and Biases

In this section, we first explore the link between individual juror characteristics and the pre-deliberation opinions held by jurors. Next, we examine the role of jury composition in shaping trial outcomes by assessing the relationship between average jury characteristics and verdicts. We then put these jury composition effects on verdicts in the context of the effect of evidentiary strength, treating the judge’s assessment of the trial evidence as an objective evaluation of the case against the defendant.

63 Comparing our restricted sample with the county population demographics, we find that females are overrepresented in our sample in Bronx (62.1% in NCSC sample versus 53.4% in the Census) and D.C. (64.3% in the NCSC sample versus 52.9% in the Census) while blacks are overrepresented in L.A. (15.8% versus 9.4%) and Bronx (40.4% versus 31.9%) and underrepresented in D.C. (44.9% versus 59.1%) The median household income of the seated jurors in our sample are significantly higher than in the county population estimates from the Census by about $15,000 to $20,000. These differences can be partially attributed to the selective pools from which initial jury summonses are drawn. County demographic estimates are from U.S. Census Bureau Population Estimates, April 1, 2000 to July 1, 2010. Retrieved March 8, 2011 from http://www.census.gov/popest/estimates.html. Household median income data from U.S. Census Bureau Small Area Income and Poverty Estimates. Retrieved March 15, 2011 from http://www.census.gov/did/www/saipe/data/index.html.

64 We will discuss this variable in more detail in Section 3.5.
3.4.1 Pre-Deliberative Juror Biases

The survey of jurors in our dataset asks a series of questions about opinions concerning the defendant, the victim, the attorneys, and other aspects of the case, as well as crime more generally. We ask whether differences in these stated opinions across jurors within each case can be explained by the jurors’ demographic and socioeconomic characteristics. More specifically, we estimate the case fixed effects model

\[ y_{jc} = \beta X_{jc} + \alpha_c + \varepsilon_{jc}, \]

where \( y_{jc} \) is the level of agreement of juror \( j \) with various statements about case \( c \) and the parties involved (on a scale ranging from 1 “Strongly Disagree” to 7 “Strongly Agree”), \( X_{jc} \) is a vector of juror characteristics, and \( \alpha_c \) captures case-level characteristics.\(^{65}\) Introducing case fixed effects ensures that variations in jurors’ opinions cannot be attributed to differences in unobservable case-level characteristics.

Table 3.2 reports OLS estimates of equation (3.1) for six separate questions that jurors were asked to state an opinion on. We present results from a linear model in lieu of ordered logit results because the fixed effects ordered logit model has been shown to encounter the incidental parameters problem, which renders the maximum likelihood estimator inconsistent even when the model is properly specified (Greene and Hensher, 2009). However, utilizing the fixed effects ordered logit estimators proposed by Baetschmann et al. (2011) to address these concerns does not change our main results.

\(^{65}\)We cannot observe whether the same jurors appear in more than one case. However, given the short period of data collection in each county, it is highly unlikely that one individual would have been summoned for jury duty more than once within that period, let alone making it on to the seated jury more than once. We therefore treat each observation in our dataset as representing a distinct juror.
Table 3.2 provides evidence of significant relations between the education, income, and religiousness of individual jurors—characteristics that have been largely ignored in past studies on jury discrimination—and their pre-deliberation leanings. Compared to others who heard the same case, jurors with higher education levels are less likely to express sympathy for the victim (column 1). More educated jurors are also significantly more likely to interpret the trial evidence as strongly favoring the prosecution relative to the defense (column 6). Specifically, an extra year of education is associated with a 0.03 point decrease on the 7-point scale that measures juror interpretation of evidentiary strength.66

Most consistently across the six predispositions, jurors with higher income hold pre-deliberation leanings that are more favorable for the prosecution. While individuals with higher earnings are more critical of attorney skills for both the prosecution and the defense, we find that higher income is significantly associated with less sympathy for the victim (column 1), greater trust in police in the community (column 2), and a greater likelihood of evaluating the trial evidence in favor of the prosecution (column 6). Specifically, a $10,000 increase in juror income has the same effect as an additional year of education (i.e. a 0.03 point decrease on the 7-point scale) on the interpretation of the trial evidence. Similarly, jurors who self-identify as “Religious” or “Very Religious” express significantly greater trust in the police, a higher rating of the prosecution’s case (column 4), and an interpretation of the evidence inclined towards the prosecution. Thus, higher education, income, and religiousness appear to be associated with a general bias towards the prosecution.

Table 3.2 also reveals important sex and race effects on individual juror

---

66The dependent variable in column 6 is the juror’s response to the question “All things considered, how close was the case?” to which the juror can answer on a scale of 1 to 7 where 1 is “Evidence strongly favored the prosecution” and 7 is “Evidence strongly favored the defense.”
Table 3.2: Pre-Deliberation Juror Opinions with Case Fixed Effects: OLS Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>education (years)</td>
<td>-0.113**</td>
<td>-0.001</td>
<td>0.008</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.018)</td>
<td>(0.020)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>income ($, thousands)</td>
<td>-0.013***</td>
<td>0.005***</td>
<td>-0.004**</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>age</td>
<td>-0.000</td>
<td>0.002</td>
<td>0.003</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>religious</td>
<td>0.027</td>
<td>0.251***</td>
<td>0.050</td>
<td>0.211***</td>
</tr>
<tr>
<td></td>
<td>(0.225)</td>
<td>(0.084)</td>
<td>(0.090)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>female</td>
<td>-0.560</td>
<td>0.076</td>
<td>0.307**</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>(0.443)</td>
<td>(0.153)</td>
<td>(0.133)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>female × vfemale</td>
<td>0.626</td>
<td>-0.064</td>
<td>0.063</td>
<td>-0.218</td>
</tr>
<tr>
<td></td>
<td>(0.405)</td>
<td>(0.198)</td>
<td>(0.175)</td>
<td>(0.202)</td>
</tr>
<tr>
<td>black</td>
<td>0.593</td>
<td>-0.434*</td>
<td>-0.480**</td>
<td>-0.079</td>
</tr>
<tr>
<td></td>
<td>(0.655)</td>
<td>(0.256)</td>
<td>(0.188)</td>
<td>(0.207)</td>
</tr>
<tr>
<td>black × dbblack</td>
<td>-0.155</td>
<td>0.051</td>
<td>0.342</td>
<td>-0.243</td>
</tr>
<tr>
<td></td>
<td>(0.680)</td>
<td>(0.276)</td>
<td>(0.221)</td>
<td>(0.226)</td>
</tr>
<tr>
<td>black × vnblack</td>
<td>-0.272</td>
<td>-0.336</td>
<td>0.051</td>
<td>-0.126</td>
</tr>
<tr>
<td></td>
<td>(0.666)</td>
<td>(0.300)</td>
<td>(0.245)</td>
<td>(0.235)</td>
</tr>
</tbody>
</table>

Female effect when
vfemale = 0: -0.560, 0.076, 0.307*, 0.066, -0.319*, -0.089
vfemale = 1: 0.066, 0.012, 0.371*, -0.153, -0.094, 0.062

Black effect when
dblack = 0, vnblack = 0: 0.593, -0.434*, -0.480**, -0.079, 0.032, 0.380
dblack = 1, vnblack = 0: 0.438, -0.383**, -0.138, -0.322**, 0.227*, 0.351**
dblack = 0, vnblack = 1: 0.320, -0.770**, -0.429*, -0.206, 0.382*, 0.271
dblack = 1, vnblack = 1: 0.166, -0.719*, -0.087, -0.439**, 0.577*, 0.241

Notes: Robust standard errors clustered at the case level are reported in parentheses. Sample limited to those cases with more than 5 jurors responding to school, income, age, race, gender, and religious questions and to those cases in which both the defense attorney and the prosecutor answered the voir dire satisfaction question before learning verdict. “How close” asks “All things considered, how close was the case?” and the juror can answer from 1 to 7 where 1 is “Evidence strongly favored prosecution” and 7 “Evidence strongly favored defense.” All regressions also control for the juror’s employment status, previous jury experience, previous criminal trial experience, and his/her race and sex interacted with the race and sex of the foreperson. ∗=10%, ∗∗=5%, ∗∗∗=1%.
biases before group discussion. For clarity and ease of interpretation, the middle panel in Table 3.2 reports the full sex and race effects under different sex/race composition identities of the victim and defendants in the case.\textsuperscript{67} Female jurors tend to rate the skill of the prosecuting attorney more highly (column 3) in all cases, but appear to have some other biases that depend to some degree on the gender of the victim. For example, female jurors are weakly less likely than other jurors to report sympathy for the victim or a high rating of the defense’s overall case when the victim is male, but display a smaller difference in these regards compared to other jurors when the victim is female. However, the coefficients are imprecisely estimated, and equality of the gender effects by sex of the victim cannot be ruled out. While firm conclusions are more difficult to draw for the case of juror gender, the greater tendency to rate the prosecutor’s skill highly and to regard the defense’s case as weak indicate that women may have a moderate general bias towards the prosecution.

Examining the relation between juror race and pre-deliberative biases, we find large effects consistently across the six predispositions and most combinations of defendant and victim race. Black jurors report significantly lower trust in the police, lower ratings of the prosecutor’s skill and case, higher ratings of the defense’s case, and an interpretation of the evidence more in favor of the defense. The divergence in evaluations of the prosecutor’s case and the defense’s case by black jurors relative to other jurors is most prominent when the defendant is black. On the other hand, the divergence in the evaluation of the prosecutor’s skill by black jurors relative to other jurors is most prominent when the defendant is not black. Finally, although the race effect on the interpretation of the evidence is

\textsuperscript{67}For example, the full black effect when the defendant and the victim are black (i.e. $\text{dblack} = 1$ and $\text{vunblack} = 0$) is the sum of coefficients on black and black $\times$ dblack.
only statistically significant when both the defendant and the victim are black, the magnitudes are similar for cases involving all defendant/victim race combinations. Specifically, black jurors report an interpretation of the evidence that is roughly 0.25 to 0.35 points higher on the 7-point scale than other jurors. To put the magnitude of this race effect in context, an equivalent change in the interpretation of evidence would require a drop in income of about $100,000 or a decline in education of about 10 years. Thus, black jurors appear to have a strong general bias towards the defense.

3.4.2 Average Jury Characteristics and Verdicts

Having demonstrated that there are significant relations between individual juror characteristics and pre-deliberation leanings, we now explore whether these individual biases feed through into biases in the determinations of guilt that juries make collectively. More precisely, we will assess the relation between the average characteristics of the seated jury and the trial verdicts. Before proceeding with the analysis, we first address some methodological issues.

While our dataset includes information on how individual jurors voted at various stages of the deliberations, we have chosen to focus on actual trial verdicts rather than individual juror voting patterns. This choice was motivated by several considerations. First, as noted by some legal scholars, the human dynamics involved in jury deliberation provide a potential channel through which individual prejudices can be attenuated or amplified within a group-decision setting.68 We

\footnote{68This idea is related to the notion of “group polarization” found in the psychology literature, which states that, if a group is like-minded, discussion strengthens its prevailing opinions. In an influential study, Myers and Bishop (1970) found that talking about racial issues increased prejudice in a high-prejudice group of high school students and decreased it in a low-prejudice group. Bringing this idea into the legal field, Sunstein (2000) writes that, in small deliberative groups such as juries, there may be a “robust pattern” of polarization whereby the initial inclinations of individuals before deliberation become more severe during deliberation.}
therefore consider the question of how jury composition relates to the verdicts arrived at when deliberations have concluded to be of much more policy relevance than the question of how individual juror characteristics relate to individual juror votes at earlier stages of deliberation or to how individual jurors would have decided the case on their own. Second, while estimating a case fixed effects model relating the voting patterns of individual jurors to their characteristics might mitigate bias due to unobserved heterogeneity, it would necessarily entail dropping all cases in which all jurors voted the same way. Since hung juries occur for only 8 percent of the counts in our estimation sample, this restriction would drastically reduce our sample size and statistical power. Finally, our dataset only contains information on individual jurors’ votes on the “most serious charge” against the defendant. The wording of the survey questions appears to have led to differing interpretations across respondents and, more generally, a low response rate, which limits our confidence in the quality of this information.

Having thus decided to focus on verdicts rather than voting patterns, we must still confront the issue of how to precisely define and specify our dependent variable. Past studies – most notably Anwar et al. (2012a,b) – have focused on a binary dependent variable indicating whether there was a conviction on any count for a given trial. However, these studies report that few of the trials being examined had multiple charges. In contrast, Table 3.1 shows that the average number of counts per trial in our dataset is 2.7, with a wide variance around this mean. We consider the loss of information that would result from the collapse of the outcomes on such a large number of charges into a binary conviction indicator to be undesirable. The alternative that we have adopted is to represent trial outcomes with a pair of count variables, indicating the absolute number of convictions and acquittals for a given trial respectively. While we do not attempt
to account for varying seriousness of charges or to otherwise weight individual
counts, our approach allows us to distinguish outcomes more finely – and perhaps
also more in line with the objectives of attorneys on each side – than an approach
based on a binary indicator. We model trial outcomes thus defined by way of an
independent poisson regression for each dependent variable.

Finally, we must return to the endogeneity concerns raised by the
discussion of the voir dire process in Section 3.2. The unit of observation for our
analysis of the impact of jury composition on trial outcomes will necessarily be
individual cases, so that it will no longer be possible to employ case fixed effects.
This raises the possibility that our results will be biased by unobservable factors
related to both verdicts and jury composition. For example, attorneys might
choose jury selection strategies based on observable juror characteristics jointly
with strategies for other aspects of the trial, thereby inducing some simultaneity

69Of the 2.7 counts per trial, 1.4 result in a conviction on average, while 1.1 result in an
acquittal, and the remaining 0.2 result in a hung jury. In contrast, headline conviction rates for
American criminal trials in general are typically reported as two thirds or higher (see, for
example, the Sourcebook of Criminal Justice Statistics Online, Table 5.57). However, such
headline conviction rates refer to convictions on any charge for a given trial, including secondary
counts and pleas to lesser charges. The relatively low conviction rate suggested by the average
number of acquittals and hung juries per count in our data is precisely the result of our decision
to examine counts within a case individually rather than compressing this information into a
simple any conviction/no conviction assessment for each trial.

70Our modeling choice has primarily been driven by efficiency concerns. The results that we
present below remain qualitatively similar if we instead use a binary dependent variable, either
estimated by way of a linear probability model as by Anwar et al. (2012b) or a logistic model;
and are likewise very similar to those that are obtained from a linear or negative binomial model
applied to the count dependent variables. However, the results from the poisson model are
estimated the most precisely, and are therefore the results that we focus on. This decision also
receives some support from various specification tests, in which the null hypothesis that the
poisson model is the correct one cannot be rejected in our data. While the negative binomial
model would allow us to relax the assumption implicit in the poisson model that the mean and
variance of the dependent variable are the same (which, in any case, is not obviously false in our
data), we instead address this potential concern by using robust standard errors within the
poisson model, which has the additional attractive feature of being robust to misspecification.
Unfortunately, our relatively small sample size also raises precision concerns when attempting to
account for cross-equation correlation in our estimation, for example by way of the
seemingly-unrelated poisson approach suggested by King (1989).
in the determination of jury composition and verdicts. We address this endogeneity issue by controlling for factors affecting attorney strategy and as many other aspects of the trial as possible. Our dataset contains a rich set of such proxy variables for many potential sources of unobserved heterogeneity.

Specifically, we control for all of the case, voir dire, and attorney characteristics that are summarized in Table 3.1. The attorney experience variables, in conjunction with the voir dire characteristics, are meant to capture both the means and the ability of attorneys to design and implement jury selection strategies that could affect the composition of the seated jury. The remaining case-level variables are meant to capture both the direct effects of these and other characteristics with which they are correlated, as well as any indirect effects that may result from attorneys attempting to condition their strategies on these variables.

In addition, we include a set of variables describing the trial judge’s opinion of various aspects of the case, including strength of evidence, attorney performance for each side, and the complexity of the evidence and the law. Table 3.3 summarizes the seven variables accounting for these judge opinions, which we treat as objective assessments of these aspects of the trials. The attorney performance variables provide additional indicators of attorney ability to shape jury composition and trial outcomes, while the other aspects, in addition to their direct interest, may also guide or constrain attorney strategies. Moreover, the

---

71 The voir dire satisfaction variables are excluded. We treat voir dire satisfaction as a separate outcome variable, and analyze its determinants in detail in the following section.

72 We do not observe identifying or demographic information about the trial judges. It is therefore possible, in the absence of other proxies for judge effects, that these judge assessments – as well as the voir dire characteristics, since, as discussed above, judges have a high degree of discretion over the voir dire process – will also capture the effects of otherwise unobservable judge characteristics. Given the size of the jurisdictions covered by our dataset, we regard the number of trials presided over by a single judge to be small, though we ultimately have no way to verify this.
Table 3.3: Summary of Judge Opinions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence presented at trial complex</td>
<td>2.328</td>
<td>1.526</td>
<td>134</td>
</tr>
<tr>
<td>(1 = “Not at all” to 7 = “Very complex”)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How complex was the law</td>
<td>2.597</td>
<td>1.576</td>
<td>134</td>
</tr>
<tr>
<td>(1 = “Not at all” to 7 = “Very complex”)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evidence favored which side</td>
<td>3.120</td>
<td>1.360</td>
<td>133</td>
</tr>
<tr>
<td>(1 = “Prosecution strongly” to 7 = “Defense strongly”)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attorneys presented all relevant evidence</td>
<td>5.561</td>
<td>1.540</td>
<td>132</td>
</tr>
<tr>
<td>(1 = “Completely disagree” to 7 = “Completely agree”)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skillful prosecutor during trial</td>
<td>5.000</td>
<td>1.535</td>
<td>136</td>
</tr>
<tr>
<td>(1 = “Not at all” to 7 = “Very skillful”)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skillful defense attorney during trial</td>
<td>5.199</td>
<td>1.444</td>
<td>136</td>
</tr>
<tr>
<td>(1 = “Not at all” to 7 = “Very skillful”)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How important was police testimony</td>
<td>5.052</td>
<td>1.669</td>
<td>134</td>
</tr>
<tr>
<td>(1 = “Not at all” to 7 = “Very important”)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Sample limited to cases with more than 5 jurors responding to school, income, age, race, gender, and religious questions. Sample also limited to those cases in which both attorneys answered voir dire satisfaction question before learning verdict.

evidence variables provide us with a unique opportunity to assess the relative prominence of jury composition effects as drivers of trial outcomes.\(^{73}\)

Table 3.4 reports our poisson results for the number of convictions and the number of acquittals. To avoid over-fitting our data and to preserve as many observations as possible, we include some specifications that limit the control variables included. Columns 1 and 4 and columns 2 and 5 control for measures of the presiding judge’s opinion of the case and attorney characteristics, respectively, while columns 3 and 6 include both sets of judge and attorney controls. Moving from left to right, the number of trials for which data on all control variables in a given specification are available decreases. Restricting the sample to only those cases included in the specifications in columns 3 and 6 does not change the results in the other columns.\(^{74}\)

\(^{73}\)Though measured on a seven-point Likert scale, we include these variables linearly in our main specifications. Alternative categorical specifications do not alter our results.

\(^{74}\)Many control variables are suppressed in Table 3.4 to save space, but are included in the
Our results reveal that some of the pre-deliberation biases we found to be associated with various juror characteristics do indeed appear to feed through into the verdicts that juries with concentrations of these juror characteristics arrive at. Juries made up of higher income individuals are strongly associated with a decreased number of acquittals. In the specification controlling for both attorney characteristics and judge opinions (column 6), a $10,000 increase in the jury’s average income is approximately associated with a 64 percent decrease in the expected number of acquittals. Likewise, a higher proportion of religious individuals on the jury is associated with a large increase in the expected number of convictions, and a large decrease in the expected number of acquittals, with a 10-percentage point increase in the proportion of religious jurors approximately associated with a 19 percent decrease in the expected number of acquittals.

These results directly correspond with the generally pro-prosecution bias we regressions. In addition to the variables already discussed, all equations also include controls for county, case type, and number of counts. Constraining the coefficient on the natural log of the total number of counts in a given trial to be one would impose the restriction that the rate of conviction or acquittal stays constant as the number of counts rises. Imposing this restriction makes little difference to the rest of our results, though the null hypothesis that it holds can be rejected at conventional significance levels. Specifically, the unconstrained coefficient estimate on the natural log of counts is greater than one in the acquittals equations and, correspondingly, less than one in the convictions equations, indicating that the rate of acquittal is actually increasing in the total number of counts. This may indicate that prosecutors add less serious charges when the main charge lacks evidence in an attempt to secure at least one conviction, or that defense attorneys focus their efforts on a large number of secondary charges when conviction on the main charge seems assured; either would lend further support to our assertion that a binary conviction indicator would seem to miss some potentially important information.

Poison coefficients are semi-elasticities, i.e. proportional changes in the dependent variable for an infinitessimal level change in the independent variable of interest. Of course, these are only approximate when the contemplated change in the independent variable is large. Poisson coefficients can alternatively be exponentiated to calculate what are known as the associated incidence-rate ratios, which give the exact ratio of the level of the dependent variable corresponding to a unit increase in the independent variable of interest to the level of the dependent variable corresponding to some baseline set of values of the regressors. We discuss our results in terms of semi-elasticities in order to retain a transparent link to the estimated coefficients reported in Table 3.4.

Increasing the proportion of the jury with a given characteristic by ten percentage points (i.e. by 0.1) is approximately equivalent to adding one person with that characteristic to a twelve-person jury.
identified previously with religious and high-income individual jurors.

In contrast, some of the effects of jury composition on verdicts that we uncover do not correspond as directly with relations between the associated individual juror characteristics and leanings. Whereas jurors with more years of education are more likely to interpret evidence in favor of the prosecution, juries with a higher average level of education convict at a weakly lower rate and acquit at a weakly higher rate, though the coefficients are not estimated precisely. On the other hand, whereas juror age was neither statistically nor substantially associated with any individual juror leanings, the results in Table 3.4 indicate that juries with a higher average age convict on significantly fewer counts. This latter finding is notable because it stands in contrast to the findings of Anwar et al. (2012a), which suggest that older jurors are more likely to convict than younger jurors. One possible explanation for these differing conclusions about the effect of age is that Anwar et al. (2012a) cannot control for juror characteristics beyond age, race, and gender. Since older individuals are more likely to earn higher incomes, the age effect that Anwar et al. (2012a) find may be biased by the negative impact of income on acquittals, as they are forced to omit income from their analysis due to data limitations.\footnote{It should be mentioned, however, that the research design of Anwar et al. (2012a) allows them to isolate the effect of random variation in the age composition of juries on conviction rates. So while they are able to deal with endogeneity in return for potentially suffering from a separate omitted variable issue, our ability to include additional juror characteristics comes with a potential endogeneity issue that our proxy variable approach may not fully address. Our analysis of individual juror biases, which includes case fixed effects in addition to a full suite of juror characteristics, does not produce such directly contrasting results to the age effect identified by Anwar et al. (2012a), though neither does it provide corroborating evidence.}

As with our analysis of individual juror inclinations, it is difficult to draw firm conclusions concerning effects of the gender composition of juries on trial outcomes. There is some evidence that juries with a greater proportion of women
Table 3.4: The Effect of Average Jury Characteristics on Verdicts: Poisson Results

|                          | Convictions | Acquittals |  |  |  |  |
|--------------------------|-------------|------------|  |  |  |  |
| Dep. Var. = Number of counts that resulted in |             |            |  |  |  |  |
|                          | (1)         | (2)        | (3) | (4) | (5) | (6) |
| education (years)        | -0.187*     | -0.172     | -0.085 | 0.211 | 0.204 | 0.000 |
| (0.100)                  | (0.158)     | (0.173)    | (0.129) | (0.204) | (0.169) |
| income ($, thousands)    | 0.010       | 0.002      | 0.002 | -0.014 | -0.035* | -0.064*** |
| (0.009)                  | (0.015)     | (0.014)    | (0.012) | (0.019) | (0.019) |
| age                      | -0.021      | -0.064***  | -0.091*** | 0.039 | 0.048 | -0.005 |
| (0.017)                  | (0.034)     | (0.030)    | (0.025) | (0.040) | (0.026) |
| religious                | 0.446       | 1.036*     | 3.218** | -0.401 | -1.990** | -1.893* |
| (0.447)                  | (0.616)     | (1.328)    | (0.640) | (0.888) | (0.986) |
| female                   | -0.241      | -0.646     | -1.036 | 0.057 | 0.864 | 4.333*** |
| (0.624)                  | (0.806)     | (1.002)    | (0.713) | (0.816) | (1.067) |
| female × vfemale         | -0.979      | 1.003      | 1.908 | 0.409 | -1.174 | 2.278 |
| (1.205)                  | (1.741)     | (3.068)    | (1.310) | (3.196) | (3.295) |
| black                    | -1.029      | -0.263     | -2.370 | 2.294* | 2.933* | 7.163*** |
| (1.079)                  | (1.595)     | (2.206)    | (1.206) | (1.591) | (1.769) |
| black × dblack           | -0.330      | -3.293***  | -2.000 | -2.241** | -1.422 | -6.210*** |
| (0.912)                  | (1.360)     | (1.512)    | (0.967) | (1.336) | (1.533) |
| black × vnblack          | 1.507*      | 1.814      | 3.859*** | -1.307 | -1.820 | -7.544*** |
| (0.798)                  | (1.307)     | (1.314)    | (1.106) | (1.618) | (1.584) |
| Evidence favored         | -0.297***   | -0.442***  | 0.374*** | 1.082*** |
| (1 “Pros.” to 7 “Def.”)  | (0.065)     | (0.100)    | (0.077) | (0.149) |
| Female effect when       |             |            |  |  |  |  |
| vfemale = 0              | -0.241      | -0.646     | -1.036 | 0.057 | 0.864 | 4.333*** |
| (0.624)                  | (0.806)     | (1.002)    | (0.713) | (0.816) | (1.067) |
| vfemale = 1              | -1.221      | 0.357      | 0.872 | 0.466 | -0.311 | 6.011** |
| (0.981)                  | (1.435)     | (2.745)    | (1.234) | (3.148) | (3.202) |
| Black effect when        |             |            |  |  |  |  |
| dblack = 0, vnblack = 0  | -1.029      | -0.263     | -2.370 | 2.294* | 2.933* | 7.163*** |
| (1.079)                  | (1.595)     | (2.206)    | (1.206) | (1.591) | (1.769) |
| dblack = 1, vnblack = 0  | -1.359*     | -3.556***  | -4.369*** | 0.052 | 1.511 | 0.953 |
| (0.708)                  | (1.114)     | (1.273)    | (0.841) | (1.416) | (0.888) |
| dblack = 0, vblack = 1   | 0.477       | 1.551      | 1.489 | 0.927 | 1.113 | -0.382 |
| (1.079)                  | (1.065)     | (2.086)    | (1.087) | (1.486) | (1.522) |
| dblack = 1, vblack = 1   | 0.148       | -1.742     | -0.511 | -1.314 | -0.309 | -6.591*** |
| (0.935)                  | (1.341)     | (1.499)    | (1.153) | (1.482) | (1.614) |

| N                        | 165         | 98         | 93         | 165       | 98         | 93         |

| Judge’s Opinion?         | ✓           | ✓           | ✓          | ✓         | ✓          | ✓          |
| Attorney Controls?       | ✓           | ✓           | ✓          | ✓         | ✓          | ✓          |

Notes: Robust standard errors are reported in parentheses. Sample limited to cases with more than 5 jurors responding to school, income, age, race, gender, and religious questions. ∗=10%, ∗∗=5%, ∗∗∗=1%. All columns also include number of counts; case, voir dire, and foreperson characteristics; and county and case type fixed effects.
are more lenient towards defendants in terms of convicting on fewer counts and acquitting on more counts. This in contrast to the weak general bias in favor of the prosecution exhibited by individual female jurors that we discussed above. However, the effect on trial outcomes is only statistically significant in one specification. There are also signs that such leniency towards defendants is attenuated when the victim is female, but again, the effects are weak and imprecisely estimated.

Finally, we find racial composition effects on trial outcomes to depend primarily on victim race. In cases with a black victim, we find that a greater proportion of blacks on the jury is associated with significant increases in the number of acquittals and decreases in the number of convictions. The positive effect on the number of acquittals is most prominent when the defendant is not black, while the negative impact on convictions is strongest when the defendant is black. The effects are quite large, with a 10-percentage point increase in the proportion of blacks associated with a decrease of between 10 and 44 percent in convictions depending on the specification and the race of the defendant. These results are consistent with the general pro-defense biases we noted amongst individual black jurors. When the victim is not black, the effects are imprecisely estimated, and vary widely in sign and magnitude across specifications.

Note that, whereas the variable ‘black’ referred to an indicator of the individual juror’s race in Table 3.2, it refers to the proportion of the jury that is black in Table 3.4. Perhaps surprisingly, in a case with a black defendant and a non-black victim, our results from the specification including all controls suggest that a greater proportion of blacks on the jury is associated with a substantial and significant decrease in the number of acquittals. One possible explanation for this somewhat surprising result may be that the victim not black indicator has a composite excluded category capturing all cases with a black victim as well as cases without a victim, and that black jurors are only more likely to acquit in cases without a victim. We do, however, control directly for whether there is a victim in the case, which mitigates this concern; and specifications with alternative victim variable definitions and interactions do not produce qualitatively different findings. A more plausible explanation is that the specification in column 6 is over-fitting the data, which is supported by the small number of included trials with non-black defendants and/or non-black victims and the corresponding large jumps in the magnitude and
Although we find that many aspects of jury makeup are significant predictors of trial outcomes, we find that more substantive aspects of the trial are also strong predictors of verdicts. For example, the judge’s opinion regarding the evidence presented at trial as favoring the defense versus the prosecution has a large impact on the number of convictions and acquittals, in the expected direction.\textsuperscript{80}

Even though we cannot know what the “true” or “correct” outcome should have been for any of the trials in our dataset, we can use this measure of the strength of evidence to put the magnitude of the jury composition effects in the context of factors that are closer to core notions of justice. Suppose that there could never be evidence stronger than one standard deviation below the mean strength of evidence against any defendant who is truly innocent, and, symmetrically, that there could never be evidence weaker than one standard deviation above the mean against a defendant who is truly guilty. Then, for a truly guilty defendant to be acquitted, factors besides the strength of evidence would have to have an effect equivalent to at least a two-standard-deviation shift in the strength of evidence. As shown in Table 3.3, the sample standard deviation of the judge’s assessment of the strength of evidence is 1.36, so that the estimates in Table 3.4 imply that a two-standard-deviation shift in the strength of evidence in favor of the defendant would be associated with an 81-120 percent decrease in the number of convictions per trial and a 102-294 percent increase in the number of acquittals. To achieve the same effect through jury composition would require, other things equal, a decrease in the average income of the jury by between

\textsuperscript{80}While not shown in Table 3.4 in consideration of space, we also find that the judge’s assessments of other characteristics of the trial and evidence are significant correlates of verdicts.
$16,000 and $46,000, or an increase in the number of blacks on the jury of between 3 and 6. Or equivalently, to make a defendant confronted with evidence of average strength appear truly innocent or truly guilty, the race of roughly one third of the seated twelve-person jury would have to be changed, or the average income of the jury would have to be changed by roughly two fifths of the sample average jury income level.

An alternative way to place the jury composition effects in perspective is to employ the regression-based technique proposed by Fields (2004) that decomposes the explanatory power of of an overall model into the contributions of the constituent independent variables. The basic idea is as follows. Let $s(X^k)$ be the share of the variation in the dependent variable $Y$ that can be attributed to variation in the $k$-th independent variable. Fields defines $s(X^k)$ as

$$ s(X^k) \equiv \frac{cov(X^k \hat{\beta}^k, Y)}{var(Y)}, $$

where $\hat{\beta}^k$ is the estimated coefficient on $X^k$ from our regression of choice. Then $\sum_{k=1}^{K} s(X^k) = R^2$ where $K$ is the total number of independent variables in the model. We can then express the $s(X^k)$ in terms of their percentage contribution to the $R^2$ of the regression model:

$$ p(X^k) \equiv \frac{s(X^k)}{R^2}. $$

Performing the Fields (2004) regression-based decomposition thus allows us to compare the relative explanatory power of jury composition with that of evidentiary strength.\textsuperscript{81} For convictions, variation in jury composition – in terms of

\textsuperscript{81}In the poisson model, the traditional $R^2$ is not well defined, so the decomposition must be performed using the pseudo-$R^2$. An alternative would be to perform the decomposition for the analogous linear model. This alternative is suggested by Fields (2004), though he states that the practical differences tend to be minor for count data applications. We have likewise found little
education, income, age, religiousness, sex, and race – accounts for about 39% of the explained variation in our model in column 3, about half of which is from the effect of race. However, roughly 43% of the explained variation in convictions can be attributed to the set of variables associated with the judge’s opinion of the evidentiary complexity and strength. For acquittals, 33% of the explained variation can be attributed to jury composition, and the judge’s evaluation of the case’s legal and evidentiary complexity and strength account for about 44% of the explained variation. Therefore, while the proportion of variation in trial outcomes explained by jury composition is not trivial, it is nonetheless exceeded by the proportion of variation explained by the more fundamental aspects of fact and legal context.

3.4.3 Summary

We have demonstrated in this section that there is some basis for concerns that jury composition can affect whether justice is achieved in criminal trials. Opinions, sympathies, and inclinations that would lead a juror to be biased towards one side are indeed correlated with observable juror characteristics. And these individual biases do indeed appear to feed through into verdicts when juries are composed of a larger number of individual jurors with the associated observable characteristics.

However, it is important to note that the relevant observable characteristics are not restricted to race, which has been the primary focus of several past studies. Most notably, we also find income to be an important predictor of individual biases and determinant of trial outcomes. While this is a novel and potentially important finding on its own, it also raises the possibility that the race qualitative difference in results when applying the decomposition to a linear model.
effects identified in past studies are confounded with effects that are properly attributed to separate characteristics excluded from the analysis. Further, these jury composition effects are smaller and have less explanatory power than effects associated with the more fundamental aspects of the trial, i.e. those related to fact and the relevant legal doctrine.

Nonetheless, concerns that jury composition effects – regardless of size or source – can impede fairness in criminal trials remain. Furthermore, they raise a potential further concern, which we now proceed to evaluate: whether attorneys are aware of these jury composition effects, and might therefore attempt to make use of them to manipulate trial outcomes.

3.5 Attorney Preferences Over Juror Characteristics

Motivated by our findings that there exist significant and robust relations between juror characteristics – especially income, religiousness, and race – and both individual juror biases and verdict patterns, we now turn to assessing whether attorneys anticipate these relations. If attorneys are unaware of these effects, this may mitigate concerns of their impact on justice (and would certainly mitigate concerns of endogeneity in our trial outcomes specifications).

If, however, attorneys are aware of these effects, we would expect them to prefer seated juries with greater proportions of jurors exhibiting characteristics that we have found to be amenable to their side. In order to analyze attorney preferences over jury composition and assess if this is true, we capitalize on a self-reported measure of satisfaction concerning the jury selection process that attorneys were requested to state before learning the verdict. In the supplementary attorney survey, attorneys responded to the question, “How adequate was the voir dire in this trial?” on a scale of 1 to 7, where 1 is “Very
inadequate” and 7 is “Very adequate.” We interpret higher responses on this scale as indicating that the responding attorney felt better able to achieve a subjectively desirable jury composition.\textsuperscript{82} We then ask whether this proxy measure for attorney satisfaction with the seated jury itself can be explained by variation in the average characteristics of the seated jury, controlling for case characteristics, various aspects of the voir dire process, attorney characteristics, and county fixed effects.\textsuperscript{83}

Before proceeding with the analysis, we take a closer look at the voir dire satisfaction variables. Figure 3.1 illustrates the distribution across cases of the difference between the defense attorney’s level of satisfaction and that of the prosecuting attorney. In about a third of the cases, the defense and the prosecution report the same level of satisfaction, and in another third of the cases, there is a one-point difference between the prosecution and the defense. While one might expect to see a greater proportion of cases with more polarized satisfaction, the shape of the distribution in Figure 3.1 could be driven by a few factors. First, one strategy that both attorneys might follow is to exercise their strikes against the potential jurors whom they perceive as having the most extreme prejudices against their side. After all (or most of) the peremptory strikes have been exercised, most of those who are left on the seated jury should be jurors who are perceived as harboring only moderate inclinations to one side or the other, and are therefore not associated with an extreme preference held by either attorney.

\textsuperscript{82}In all of our analyses, we restrict our sample to trials for which the attorneys explicitly stated that they had responded to the voir dire satisfaction question before learning the verdict. This supports our interpretation of these satisfaction ratings by making it less likely that they simply capture post hoc rationalization of the broader trial outcome.

\textsuperscript{83}It is important to note that we do not attempt to examine the degree to which attorneys actually manage to transform the composition of the jury in relation to that of the jury pool, nor do we attempt to identify the precise actions taken or strategies employed by attorneys in voir dire. While these questions are interesting in their own right, their answers depend both on attorney preferences and on a number of institutional and trial-specific constraints, and it is the former that we wish to focus on exclusively here. In the following chapter, we develop a model of attorney behavior in voir dire to examine jury selection strategies and outcomes in detail.
Figure 3.1: Difference in Voir Dire Satisfaction within a Case: Defense Minus Prosecution

Notes: Sample limited to cases with more than 5 jurors responding to school, income, age, race, gender, and religious questions and those cases in which both defense and prosecution answer voir dire satisfaction question before learning the verdict.

Second, if attorneys’ preferences in voir dire were simply based on a single characteristic of the jurors – for example, race – then one might expect a greater divergence in satisfaction with the seated jury, since one side’s gain would necessarily be the other side’s loss. However, if preferences are not so unidimensional, then it is not clear that the prosecution and the defense should express such antithetical satisfaction with the overall makeup of the jury. For example, although the racial makeup may benefit the prosecution, the religiousness and income levels of the jurors might be simultaneously more favorable to the defense. Finally, we do not know the benchmark against which any given attorney might express his or her satisfaction. So, for example, in a trial
in which the jury pool has an unusually large proportion of blacks, the defense attorney might express a high level of satisfaction for having a greater proportion of blacks on the seated jury than in previous trials, while the prosecuting attorney might also express a high level of satisfaction for successfully avoiding a seated jury with as many blacks as might have resulted from the particular jury pool.

The empirical approach we take is to estimate separate ordered logistic models for defense attorneys and prosecuting attorneys.\textsuperscript{84} Besides the jury characteristics of primary interest, we control for our full set of case-level covariates, including voir dire characteristics and the characteristics of the attorney of the given side.

The primary value of considering the two sides separately is that it allows us to retain as many observations as possible.\textsuperscript{85} However, the disadvantage is that our ability to address endogeneity concerns becomes limited to the proxy variable approach we employed in our trial outcomes specifications. But even with our rich sets of controls, our ordered logistic estimates may be biased by remaining unobserved heterogeneity across trials that simultaneously affects both attorney satisfaction and jury composition.

Voir dire satisfaction is only observable at the trial level, but in principle we observe it for both sides, implying that we should be able to form a panel out of our dataset and leverage this panel structure to eliminate unobserved heterogeneity. However, the problem with this approach is the aforementioned limited number of trials for which we have a balanced panel, which leads to

\textsuperscript{84}The ordered logistic functional form allows us to account for the ordinal nature of the satisfaction ratings. Treating the dependent variables as cardinal and estimating linear models does not substantially alter our results.

\textsuperscript{85}There are unfortunately only 80-90 cases in our dataset for which both attorneys responded to the voir dire satisfaction question before learning the verdict and for which various other control variables are also available.
imprecise estimates and a susceptibility to over-fitting the models. We have nonetheless attempted various panel and pseudo-panel methods to diagnose and correct for endogeneity concerns. While the magnitude and significance of some effects vary widely across these specifications, the main results that we present and focus on below are relatively stable. We have therefore opted to report the separate ordered logistic results for the clear interpretation they permit and for their advantage in terms of precision of the estimates.  

These results are presented in Table 3.5. The first three columns examine the determinants of the defense attorney’s satisfaction, and the last three columns report the correlates of the prosecutor’s satisfaction. Columns 1 and 4 control for all variables summarized in Table 3.1 except for attorney and foreperson characteristics, which are excluded in order to retain as many observations as possible. Specifications with these additional controls are reported in the remaining columns, and demonstrate the robustness of our main results to these different sets of controls.

The most striking finding in Table 3.5 is that the defense and prosecuting attorneys’ levels of satisfaction are significantly correlated with the average income

---

86 The conditional fixed effects models – either ordered logistic or binomial logistic based on collapsing the satisfaction ratings from a seven-point to a two-point scale in various ways – not only require a balanced panel, but also a different value of the dependent variable for the two sides, leaving us with as few as 40 trials. A more promising set of methods involves using the satisfaction rating of one side to attempt to capture case fixed effects in specifications explaining the other side’s satisfaction. One way in which we implemented this idea was to treat the satisfaction ratings as cardinal and use their difference as the dependent variable. Alternatively, we included one side’s satisfaction directly in the opposing side’s specification, entering as a set of dummy variables or simply linearly. We also made some attempts to increase efficiency, by pooling observations across attorneys and constraining some coefficients to be the same for both sides on one hand, and by accounting for cross-equation correlation in a bivariate probit model – again, with collapsed binary satisfaction ratings – on the other. Each of these specifications is adversely affected by the small sample sizes, whether through convergence problems, imprecise estimates, or nearly exactly determined coefficients for some explanatory variables. However, the two main results that we focus on are present and and stable in sign and significance across all specifications.
Table 3.5: Attorney Voir Dire Satisfaction: Ordered Logit Results

<table>
<thead>
<tr>
<th>Dep. Var. = Attorney Satisfaction with Voir Dire (1 to 7 where 1=Low and 7=High)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>education (years)</td>
<td>0.239</td>
<td>0.213</td>
<td>0.152</td>
<td>-0.432*</td>
<td>-0.360</td>
<td>-0.344</td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.218)</td>
<td>(0.242)</td>
<td>(0.252)</td>
<td>(0.260)</td>
<td>(0.299)</td>
</tr>
<tr>
<td>income ($ thousands)</td>
<td>-0.046**</td>
<td>-0.054***</td>
<td>-0.060**</td>
<td>0.067***</td>
<td>0.076***</td>
<td>0.078**</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.020)</td>
<td>(0.023)</td>
<td>(0.025)</td>
<td>(0.028)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>age</td>
<td>-0.017</td>
<td>-0.018</td>
<td>-0.006</td>
<td>-0.001</td>
<td>-0.000</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.063)</td>
<td>(0.067)</td>
<td>(0.051)</td>
<td>(0.064)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>religious</td>
<td>1.548</td>
<td>1.851</td>
<td>0.703</td>
<td>0.555</td>
<td>0.955</td>
<td>0.465</td>
</tr>
<tr>
<td></td>
<td>(1.207)</td>
<td>(1.290)</td>
<td>(1.497)</td>
<td>(1.040)</td>
<td>(1.158)</td>
<td>(1.202)</td>
</tr>
<tr>
<td>female</td>
<td>3.678***</td>
<td>4.183***</td>
<td>4.684***</td>
<td>0.784</td>
<td>-1.427</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(1.111)</td>
<td>(1.209)</td>
<td>(1.438)</td>
<td>(1.972)</td>
<td>(2.004)</td>
<td>(2.017)</td>
</tr>
<tr>
<td>female × vfemale</td>
<td>-0.203</td>
<td>-0.062</td>
<td>1.022</td>
<td>-3.376</td>
<td>-3.064</td>
<td>-7.122</td>
</tr>
<tr>
<td></td>
<td>(3.480)</td>
<td>(3.138)</td>
<td>(3.499)</td>
<td>(3.195)</td>
<td>(3.777)</td>
<td>(4.818)</td>
</tr>
<tr>
<td>black</td>
<td>1.026</td>
<td>2.085</td>
<td>1.641</td>
<td>5.432*</td>
<td>6.096*</td>
<td>6.331*</td>
</tr>
<tr>
<td></td>
<td>(2.457)</td>
<td>(2.708)</td>
<td>(3.120)</td>
<td>(3.254)</td>
<td>(3.161)</td>
<td>(3.720)</td>
</tr>
<tr>
<td>black × dblack</td>
<td>-2.634</td>
<td>-4.182</td>
<td>-4.401</td>
<td>-5.640**</td>
<td>-6.371**</td>
<td>-5.862*</td>
</tr>
<tr>
<td></td>
<td>(2.480)</td>
<td>(2.780)</td>
<td>(2.906)</td>
<td>(2.857)</td>
<td>(2.770)</td>
<td>(3.232)</td>
</tr>
<tr>
<td>black × vnblack</td>
<td>0.883</td>
<td>0.368</td>
<td>1.353</td>
<td>-0.895</td>
<td>-1.227</td>
<td>-1.451</td>
</tr>
<tr>
<td></td>
<td>(2.491)</td>
<td>(2.868)</td>
<td>(2.790)</td>
<td>(2.077)</td>
<td>(2.077)</td>
<td>(2.772)</td>
</tr>
</tbody>
</table>

**Female effect when vfemale = 0**

| 3.678*** | 4.183*** | 4.684*** | 0.784 | -1.427 | -0.016 |
| (1.111) | (1.209) | (1.438) | (1.972) | (2.004) | (2.017) |

**Female effect when vfemale = 1**

| 3.475 | 4.121 | 5.706* | -2.592 | -5.090 | -7.139* |
| (3.268) | (2.892) | (3.351) | (2.874) | (3.387) | (4.137) |

**Black effect when dblack = 0, vnblack = 0**

| 1.026 | 2.085 | 1.641 | 5.432* | 6.096* | 6.331* |
| (2.457) | (2.708) | (3.120) | (3.254) | (3.161) | (3.720) |

**Black effect when dblack = 1, vnblack = 0**

| -1.608 | -2.097 | -2.760* | -0.208 | -0.276 | 0.469 |
| (1.314) | (1.390) | (1.479) | (1.490) | (1.598) | (1.749) |

**Black effect when dblack = 0, vnblack = 1**

| 1.909 | 2.453 | 2.993 | 4.537* | 4.868* | 4.880* |
| (2.097) | (2.409) | (2.855) | (2.608) | (2.559) | (2.947) |

**Black effect when dblack = 1, vnblack = 1**

| -0.725 | -1.729 | -1.408 | -1.103 | -1.503 | -0.983 |
| (2.981) | (3.342) | (3.046) | (2.108) | (2.091) | (2.403) |

N: 142 132 124 133 122 115
Log-Likelihood: -212.285 -186.467 -175.082 -203.166 -178.198 -167.010

Attorney Controls ✓ ✓ ✓ ✓ ✓ ✓
Foreperson Controls ✓ ✓ ✓ ✓ ✓ ✓

Notes: Robust standard errors are reported in parentheses. Sample limited to cases with more than 5 jurors responding to school, income, age, race, gender, and religion questions and to cases in which corresponding attorneys answered the voir dire satisfaction question before learning the verdict. ∗=10%, ∗∗=5%, ∗∗∗=1%. All models also control for county and case type fixed effects, and case and voir dire characteristics.
level of the seated jury, in opposite directions and in line with our findings on individual biases and trial outcomes. Seated juries with lower average income are associated with higher levels of satisfaction for the defense, while the prosecution’s satisfaction level rises with a richer jury. This corresponds directly with our findings of a general pro-prosecution bias amongst individual jurors with higher incomes and the tendency for juries with higher average incomes to acquit on fewer counts. The significance and robustness of the estimates provide a strong indication that attorneys are well aware of these effects.

A second striking finding is that defense attorneys have a strong preference for seated juries with a high proportion of women. This preference is most prominent when there is no victim or a male victim, though the effect is still positive, if much less precisely estimated, when there is a female victim. This result is puzzling in relation to the effects of juror gender on individual biases and trial outcomes. While, as discussed above, juries with a greater proportion of women do appear to be weakly more lenient towards defendants, the effects are imprecisely estimated and somewhat sensitive to specification. Moreover, individual female jurors appear to have at least a weak general bias towards the prosecution. In contrast, the estimates in Table 3.5 indicate that defense attorneys have a very strong preference for female jurors.

One potential explanation for this apparent divergence between our previous results and attorney preferences is that attorneys use gender as an indicator of juror characteristics that they have trouble observing directly. For example, attorneys might expect that women tend to have lower incomes and fewer years of schooling or to be less religious, and – consistent with the results in Table 3.2 showing that these traits are associated with a greater tendency to interpret evidence in favor of the defense – prefer juries with a greater proportion
of women on that basis. However, it appears that there would be little justification for attorneys to form such beliefs. When we estimate the specifications in Tables 3.2 and 3.4 excluding juror characteristics that attorneys may not be able to observe or approximate accurately, the effects of juror gender are qualitatively unaffected.\textsuperscript{87} Thus, none of these alternative specifications provides evidence that a greater proportion of women on seated juries, when used as a proxy for other characteristics, would indicate any substantial advantage for the defense.

A more straightforward interpretation of this divergence between our previous results and the strong preference of defense attorneys for female jurors may therefore be that attorneys lack awareness of true relations between gender and individual biases or trial outcomes. This might suggest that attorneys form preferences for the gender of seated jurors based on anecdotal notions – even if they often turn out to be incorrect – that women are markedly sympathetic to defendants or are very likely to possess other traits that are correlated with such sympathies.\textsuperscript{88}

Attorney difficulties with observing some juror characteristics may nonetheless be affecting our results through a different channel. In contrast to our findings of a strong association between juror religiousness and pro-prosecution trial outcomes and individual juror biases, attorneys do not appear to hold strong preferences for the average religiousness of the seated jury. A possible explanation for the absence of a clear link between religiousness and attorney preferences is

\textsuperscript{87}Most notably, the strong tendency for female jurors to rate the prosecutor’s skill highly, from column 3 of Table 3.2, is unchanged in size and significance across every alternative specification. Similarly, the female effect on the rating of the defense’s case becomes slightly smaller in absolute magnitude and loses significance in some alternative specifications, but is always negative.

\textsuperscript{88}Table 3.5 also provides weak evidence that prosecuting attorneys prefer juries with a lower proportion of females when there is a female victim. This might suggest that attorneys also rely to some extent on anecdotal and frequently mistaken impressions that women are harsher judges of female victims.
that the religiousness of a potential juror is less likely to be observable to the attorneys than other characteristics. Hence, while attorneys may be aware of an effect of religiousness on verdicts in general, they may have only a very limited notion either of the religiousness of the seated jurors or the expected magnitude of the effect of their religiousness in any particular case.\textsuperscript{89}

Table 3.5 also suggests that prosecuting attorneys have a weak preference for juries with fewer years of education on average, while defense attorneys may have a very weak preference for the opposite. This corresponds with the direction of the weak effects we found of education on trial outcomes, though not with the association we found between individual juror education and a greater tendency to interpret trial evidence in favor of the prosecution. Unfortunately, the imprecision of the estimates makes attorney preferences over juror education difficult to analyze in more depth.

Finally, our results on attorney preferences for the proportion of blacks on the jury are inconclusive and sometimes in conflict with what would be expected given our findings on the impact of racial composition on verdicts. As discussed in Section 3.4, individual black jurors exhibit strong pro-defense biases that are only slightly more prominent when the defendant is black; while, in cases with a black

\textsuperscript{89}One hypothesis arising from these conjectures is that religiousness may be more strongly correlated with attorney satisfaction in cases in which attorneys were able to gather greater individual information about the potential jurors through the use of questionnaires. We test this hypothesis by interacting the average level of religiousness of the jury with a dummy indicating whether questionnaires were used in voir dire. In some specifications, we find that, when a questionnaire was used, a higher proportion of religious jurors is weakly associated with lower relative satisfaction for the defense attorney compared to the prosecution, broadly in line with our findings on individual juror biases and trial outcomes. As mentioned in Section 3.2, regardless of whether attorneys have access to results from supplementary questionnaires, they do know the basic occupational and biographical details that courts routinely collect from all members of the jury pool upon or prior to arrival at the courthouse. We therefore consider all of the other juror characteristics we observe to also have been at least approximately known by the attorneys from their own observation and separate information sources. Correspondingly, we find little impact of juror questionnaires on the effects of the other juror characteristics on attorney preferences in alternative specifications with extended interactions.
victim, a higher proportion of blacks on the jury is associated with more favorable trial outcomes for the defendant regardless of defendant race. However, results in Table 3.5 suggest that the prosecutor holds some preference for a greater proportion of black jurors when the defendant is not black, and also appear to suggest that the defense may hold a weak preference for a less black-dominated jury when the defendant is black, regardless of victim race. These results could perhaps be interpreted, similar to those regarding juror gender, as indicating some degree of reliance on folk wisdom amongst attorneys: prosecutors may believe that black jurors are unconditionally more likely to be sympathetic towards black defendants, but simply not pay much attention to juror race when the defendant is not black. However, inferences of this nature are difficult to support, as the estimates of these race effects are imprecise, and their magnitude and significance are sensitive to alternative specifications.

Another potential interpretation of these race results is that attorneys are expressing their satisfaction relative to some unobserved benchmark: for example, as mentioned earlier, a pool of mostly black potential jurors is likely to result in most of the seated jurors also being black regardless of what the attorneys accomplish in jury selection, in which case the prosecutor might be relieved and the defense attorney disappointed with any situation except an all-black seated jury. It is also possible that, in response to increasing scrutiny of race-based removal of potential jurors during jury selection, attorneys on both sides may have moved away from treating racial composition as a predominant goal in jury selection and instead formed stronger preferences for other juror characteristics.\(^{90}\)

\(^{90}\)As discussed in Section 3.2, striking jurors based solely on race is illegal, although there are practical difficulties with enforcing this. We control for whether any Batson objections were raised in all specifications in Table 3.5, but cannot observe which side raised such objections or what, if any, remedy was provided as a result. We have attempted to test for the sensitivity of attorney preferences over race to the degree of scrutiny concerning race-based strikes by
Our findings on the strong and robust effects of the jury’s average income and female representation on attorney satisfaction certainly support the hypothesis that attorneys have multi-dimensional preferences across characteristics besides race. It is thus possible that any estimated race effects on attorney satisfaction are driven by a few isolated trials in our dataset, implying that the best interpretation might be that these apparent race effects are just unreliably estimated and masking a true zero effect.

In summary, we find evidence that attorneys are aware of some relations between juror characteristics and trial outcomes. Of particular note are the strong and opposing preferences over the average income of the jury held by attorneys on both sides, which indicate that the ability to correctly anticipate the role of juror income in predicting juror biases and verdicts is widespread in the legal profession. This serves to underscore the concerns that arise from the existence of these relations in the first place, as it raises the possibility – or at least the suspicion – that attorneys could use this knowledge to manipulate trial outcomes. At the same time, our findings suggest that any attorney efforts in this regard may be hampered by an imperfect understanding of the relation between some other juror characteristics and trial outcomes and by an inability to observe some characteristics.

3.6 Conclusion

Our results give some justification for concerns about a lack of impartiality and trustworthiness in U.S. criminal jury trials. Defendants whose trials are heard by juries with a higher average income and fewer blacks can expect the individual interacting with the Batson indicator, but unfortunately the small sample size prevents these effects from being estimated precisely enough to draw any conclusions.
jurors to interpret the case and evidence less in their favor and can expect a greater likelihood of conviction. Moreover, as attorneys appear to have some ability to correctly anticipate these relations, defendants must also worry that the prosecution will attempt to push the composition of the jury in these specific directions. These findings hardly inspire confidence in the ability of the legal system to deliver fair outcomes.

However, our findings also suggest some potential silver linings. First, our results do not point to race as the primary source of bias in criminal trials. There is a popular perception that attorneys predominantly target jurors for removal on the basis of race during jury selection, resulting in widespread discriminatory exclusion of minorities from jury service and a severe bias in verdicts against minority defendants. This perception may be primarily driven by certain high-profile cases (such as the O.J. Simpson criminal trial or the first trial of the officers who beat Rodney King), but it has also received some support from various notable studies (especially Anwar et al., 2012b, Baldus et al., 2001, and Bowers et al., 2001). Our results certainly confirm that black jurors tend to hold sentiments favorable to defendants. On the other hand, our estimated effects of juror race on verdicts are less pronounced than those on individual predispositions, and less pronounced as well than our estimated effects of certain other juror characteristics on verdicts. Furthermore, our results on attorney preferences in jury selection, while difficult to interpret, do not appear to indicate that attorneys consider juror race to be associated with a substantial advantage to either side. The difference in degree between our findings and those of other studies can likely be explained by two important factors: the high and varying racial diversity in and across the counties covered by our dataset; and the fact that we are able to jointly control for such a breadth of juror characteristics. The
robust income effects we find throughout our analysis suggest that this may be an especially important omitted variable in other studies, leading to an upward bias in the magnitude of their estimated race effects due to the generally greater population-wide propensity for minorities to have low incomes.

A second potential reason for optimism revealed by our results is that the substantive evidentiary and legal aspects of trials remain the predominant determinants of verdicts. The jury composition effects that we identify are small compared to the effect of evidentiary strength, and account for less of the explained variation in verdicts than do the variables related to the facts and legal doctrines central to the trial. Of course, the presence of any significant jury composition effect could be interpreted as undermining the fairness and validity of jury trials in general; and in a specific trial in which a defendant is truly innocent but nonetheless faces moderately strong circumstantial evidence, even a small effect could induce a false conviction if the prosecutor manages to alter the jury composition enough.

This question of how well attorneys can leverage jury composition effects to affect verdicts is one that our present work only partially illuminates. To be able to do so at all, attorneys must of course be aware of the jury composition effects in the first place, and this points to a third potential silver lining of our results. While attorneys on both sides seem to be well aware of the effect of juror income on trial outcomes, there are also signs that they are at least partially mistaken about the effects of other characteristics. The defense’s strong preference for women on seated juries, in particular, seems misguided in comparison to the effects of juror gender on verdicts and individual biases that we estimate. And whether attorneys are aware of the strong effects of religiousness or not, it appears that they have trouble observing this trait in jurors.
But the ability of attorneys to affect trial outcomes through jury composition effects also hinges on how much they can actually manage to manipulate jury composition in voir dire, and our present work is silent on this issue. No matter how strong some individual juror biases might be and how accurately attorneys anticipate them, the success of an attorney of a given side in stacking the jury will be tempered by the actions and strategies of the opposing counsel and by the institutional constraints imposed within the jury selection process. In other words, the question of whether attorneys are aware of jury composition effects is moot if they have no freedom to make use of such awareness in voir dire or if the actions of one attorney can be perfectly offset by the other.

This might suggest that a potentially attractive approach to addressing some of the concerns arising from the effects of jury composition on trial outcomes would be to simply limit attorney freedoms in voir dire. However, the work that we will turn to in the following chapter, in which we model attorney behavior in jury selection to evaluate these questions in detail, suggests that such a policy may have its own drawbacks. As we will show presently, giving attorneys more freedom in voir dire can be beneficial in terms of revealing otherwise unobservable juror biases and breaking attorney reliance on stereotypes based on observable characteristics like race and income, though trial outcomes will nonetheless be affected unless asymmetries in attorney skill are also addressed.

While we are therefore hesitant to make definitive policy recommendations at this stage, our present findings do make an important contribution to jury and broader legal reform debates by providing the most convincing evidence to date that the sources of jury bias run deeper than race. Our results suggest that these sources are nuanced and multi-faceted, encompassing several juror characteristics and attorney beliefs about them. The issue may thus require an equally nuanced
and multi-faceted policy response. Exploring how our results generalize to regions and types of cases not covered by our dataset is a crucial next step in eventually formulating such a policy response.
CHAPTER FOUR
Attorney Empowerment in Voir Dire and the Racial Composition of Juries

4.1 Introduction

Voir dire, more commonly known as jury selection, is the process by which a pool of potential jurors is questioned and pared down to the seated jury that will eventually hear a trial. The legal rationale for conducting voir dire, rather than simply drawing juries randomly from the population, is to ensure that juries can act impartially and competently in rendering their verdicts. However, there is a widespread notion that voir dire practices employed in many trials grant attorneys too much freedom to shape the jury, leading instead to the seating of biased juries and to the systematic exclusion of certain groups from jury service. Thus, jury selection is often perceived as harming defendants by leading to unfair trial outcomes, harming potential jurors by facilitating their discriminatory removal on the basis of race or other characteristics, and generally detracting from the trustworthiness of the legal system.

Our goal in this chapter is to discover what attorneys can actually accomplish in voir dire. When attorneys are granted more freedoms, will the jurors who make it through the jury selection process be biased against defendants? Will they be less likely to belong to certain identifiable groups?

---

91 This chapter is the product of joint work with Jee-Yeon K. Lehmann, of the University of Houston, Department of Economics.

92 King (1993) discusses these categories of harm from the perspective of the evolution of related Supreme Court decisions. Much of the popular and scholarly discussion on these matters focuses on exclusion from juries based on race, and even more specifically, on blacks excluded from juries in favor of whites. We take a broad perspective on the potential to exclude jurors on the basis of any observable characteristics. However, we also recognize the prevalence and importance of the focus on race, and adopt this focus ourselves in most of our empirical analysis.
We seek to answer these questions first with a simple model of attorney behavior. Attorneys on a given side are assumed to have the objective of retaining jurors who are favorably predisposed to their side. Such predispositions are assumed to be correlated with an observable juror characteristic, for example, race. When attorneys are unempowered – that is, when they have limited freedoms in voir dire – they must pursue their objective by striking potential jurors according to this characteristic. We model empowerment as having two potential effects: it can improve the ability of attorneys to strike whichever potential jurors they choose to; and it can allow attorneys to acquire more accurate information about the predispositions of potential jurors than can be inferred from the observable characteristic alone. We allow the magnitude of both of these effects to depend on the skill of the attorneys on each side.

The model predicts that, when attorneys are empowered, more skilled attorneys will have greater success in retaining jurors favorably inclined to their side. Therefore, empowerment can indeed lead to a seated jury that is more prejudiced against the defendant than a randomly-drawn jury would be – but only when the prosecuting attorney holds a skill advantage over the defense attorney. Empowerment can likewise lead to a seated jury that is prejudiced in favor of the defendant when the defense attorney holds a skill advantage.

However, the model does not make a clear prediction concerning the effect of empowerment on the distribution of seated jurors by the observable characteristic. If empowerment only increases striking ability and prosecutors have an \textit{a priori} preference for white jurors, they will leverage empowerment to strike non-white jurors at a higher rate. But defense attorneys will likewise leverage empowerment to strike white jurors at a higher rate. If the defense attorney holds a skill advantage, the overall effect of empowerment will be a greater proportion of
non-whites on the seated jury than in the pool of potential jurors. However, when the information revelation effect of empowerment is also present, the defense may be able to identify some white jurors who are preferred to some non-white jurors (and vice versa for the prosecutor). In this case, the interaction of empowerment and defense attorney skill need not lead to a higher proportion of non-whites on seated juries, and can, in fact, be associated with white jurors being retained at a very high rate.

Next, we test our theoretical predictions empirically, using a rich dataset on all non-capital felony trials in four large and diverse counties over a two-year period. The dataset includes information on the freedoms that were granted to attorneys during jury selection, on attorney skill, on observable characteristics of jurors, on the self-reported pre-deliberation leanings of jurors, and on numerous other characteristics of the trial and the parties involved.

The results of our empirical analysis are surprising. Attorney empowerment in jury selection, other things equal, is associated with a concentration of jurors who report having greater favoritism towards the defense prior to deliberations. This effect is almost entirely accounted for by trials in which the defense attorney held a skill advantage. We interpret this result as confirmation of the basic validity of our model.

However, neither skill nor empowerment nor their interaction is associated, in our data, with any substantial impact on the average composition of seated juries by observable characteristics, including, most notably, race. In the context of our model, this result implies that, for at least some of the trials in our dataset, attorneys leveraged empowerment in voir dire to learn more about potential jurors, rather than just to strike more effectively.
4.1.1 Related Literature and Our Contributions

Our work connects two branches of related literature to which economists have contributed. The first comprises theoretical studies of attorney behavior in jury selection and potential impacts of this behavior on verdicts. The second comprises empirical studies of the impacts of race and other juror characteristics on trial outcomes.

There is a common and natural framework underlying much of the previous theoretical work. The foundation is an assumption that potential jurors are drawn from a unidimensional distribution of predispositions. Other things equal, the defense attorney will want to strike the potential jurors with the strongest inclinations towards the prosecution, and the prosecuting attorney will likewise want to strike the potential jurors with the strongest inclinations towards the defense.

Brams and Davis (1978) explore this framework from a game theoretic perspective, and characterize attorney optimal responses when facing various constraints related to jury size, sequence of decisions, and number of strikes allowed. Neilson and Winter (2000) make use of the same underlying framework to gauge the impact of allowing attorneys a greater number of strikes on two types of injustice: false convictions and false acquittals. A common finding of this literature is that the probability of a conviction will often be too high – relative either to a case in which juries are chosen randomly or to a case in which the

---

93 Roth, Kadane and Degroot (1977), Degroot and Kadane (1980) and Kadane, Stone and Wallstrom (1999) provide technical extensions on this theme.

94 Kadane and Kairys (1979) design an algorithm to find the optimal number of strikes allowed under a slightly different definition of injustice. Ford (2010) examines the impact of varying the number of strikes allowed on conviction rates in a similar framework, without reference to an explicit definition of injustice. Feddersen and Pesendorfer (1998) consider the same types of injustice as Neilson and Winter (2000) in the context of strategic voting, but jurors in that context are ex ante identical, so there is no scope to examine aspects of jury selection.
likelihood of injustice is minimized – unless the defense attorney is allowed a
greater number of strikes than the prosecuting attorney.

There have been several attempts to establish empirical evidence that
certain observable juror characteristics are associated with different criminal trial
verdicts. Anwar et al. (2012b) find that juries formed from all-white jury pools
convict black defendants at a substantially and statistically significantly higher
rate than they convict white defendants, while Anwar et al. (2012a) find that
older jurors are more likely to convict. Lee (2010) uses state variation in the
timing of jury reforms, and finds that greater potential racial heterogeneity on
juries is associated with lower conviction rates of minority defendants. Our results
in Chapter 3 show that juries with higher average income and religiousness are
less likely to acquit in general, and that juries with a greater proportion of blacks
are less likely to convict when both defendant and victim are also black. On the
other hand, Hannaford-Agor et al. (2002) find little relation between jury racial
composition and the probability of a hung jury.95

This empirical literature establishes that attorneys who wish to maximize
the probability of a case being decided for their side should want to remove jurors
on the basis of race and perhaps other observable characteristics, at least in the
absence of other useful information. However, it has not satisfactorily explored
how attorneys actually behave and what they accomplish in voir dire.96 The

95Shayo and Zussman (2011), Alesina and La Ferrara (2011), Abrams et al. (2011), and
Iyengar (2011) provide evidence that the race of judges, victims, and defendants can affect trial
outcomes in other settings. There are many other studies on race and criminal trial outcomes
that rely on mock trials and case studies of a handful of trials at a time. A fairly recent and
comprehensive review of such studies is given by Sommers and Ellsworth (2003), and a more
in-depth and critical review of the earliest of these studies is given by Pfeifer (1990). Bowers et
al. (2001) include 340 trials in their analysis, but are unable to adequately control for
heterogeneity across these trials. They find that the death penalty is three times as likely for a
case with a black defendant whose victim is white and for which the jury consists of five or more
white male jurors than for similar cases with more mixed juries.

96One potential exception is Diamond et al. (2009), who find, in their sample of small claims
theoretical literature, on the other hand, has not incorporated empirical insights on juror race comprehensively, and can be usefully extended in this and other ways. We pursue such extensions theoretically, and we empirically confirm our model’s predictions concerning what attorneys can accomplish.

Our first contribution relative to previous theoretical work is that we make more realistic assumptions about the information available to attorneys. Brams and Davis (1978) assume that attorneys can observe juror predispositions, while Nielson and Winter (2000) assume that attorneys cannot observe predispositions but know that an observable juror characteristic is predictive of a stronger predisposition towards the prosecution. We assume instead that jurors are described by a two-dimensional set of characteristics: one, such as race, which is always observable; and a second, namely a predisposition to one side or the other, which can potentially be imperfectly observed if attorneys are empowered. When attorneys cannot observe predispositions at all, they have no choice but to rely on the common-knowledge correlation between the observable characteristic and unobservable leanings. But when empowerment allows attorneys to learn about juror inclinations, they can use this partial information in addition to the information already afforded by the observable characteristic. Our model thus includes the Nielson and Winter framework as a special case, while, in the general specification, avoiding the full-information assumption of Brams and Davis.

Our second contribution relative to previous theoretical work is that we broaden our focus beyond peremptory challenges. Peremptory challenges allow attorneys to strike any potential juror without stating a reason, with very few exceptions. In contrast, successful challenges for cause must be accompanied by an

civil suits, that attorneys on both sides appeared to strike potential jurors largely on the basis of race, but that these strikes mostly offset one another. The authors do not discuss the impact of jury composition on the outcomes of the trials.
acceptable argument that the targeted potential jurors have demonstrated an inability to be impartial. Previous studies only treat attorney decisions concerning which potential jurors to strike with a limited number of available peremptory challenges, assuming that all potential jurors for whom an inability to be impartial could be demonstrated have been removed via challenges for cause in an unmodeled earlier stage. Our modeling assumptions reflect a more realistic situation in which attorneys attempt to use both strikes for cause and peremptory strikes to remove jurors with unfavorable predispositions to their side. The ability of attorneys in our model to remove and object to the removal of potential jurors – and the effect of empowerment on this ability – is handled in a reduced-form manner that can incorporate strikes made via either type of challenge.

Our third contribution is the inclusion of heterogeneity in attorney skill, which follows naturally from these first two. There is little potential role for attorney skill when juror predispositions are fully observable and only peremptory challenges are considered. However, the acquisition and interpretation of information, and persuasiveness in semantic arguments over what constitutes grounds for dismissal for cause, can reasonably be expected to depend on attorney facility. We hence allow for the possibility that more skilled attorneys can reap greater benefits from empowerment than less skilled attorneys in our model, and test for this possibility empirically.97

Our final noteworthy contribution is the consideration of juror predispositions directly, rather than trial verdicts. To be sure, the verdict is what all parties involved in the trial care about most. However, verdicts are a product of more than the predispositions of individual jurors. Of specific concern for our

97Abrams and Yoon (2007) and Shinall (2010) find that attorney skill is an important determinant of trial outcomes, but do not discuss this in the context of jury selection.
work is that verdicts are determined by jury deliberations, which can entail complex group dynamics and decision-making. Previous theoretical contributions skirt this issue by calculating the jury’s overall probability of conviction as the simple product of each juror’s probability of conviction, while noting the unsatisfactory nature of the implicit independence assumption. Instead, we take individual juror predispositions as the outcome variable of interest to attorneys. We consider this to be a desirable approach based on our postulation that attorneys are much better at assessing individual juror leanings than they are at predicting how the seated jurors will interact with one another in arriving at their final decision. Indeed, Gobert et al. (2009) suggest that techniques available to attorneys for assessing individual jurors are much more advanced than those for predicting deliberation dynamics, and that selecting “juries, not jurors” is still very much an emerging concept.

This chapter is the most complete attempt to date to measure what attorneys can actually accomplish, given varying degrees of empowerment in jury selection, in terms of multiple characteristics of the seated jurors they manage to retain. We find that skilled and empowered attorneys can successfully retain jurors favorably inclined to their side on seated juries, but that they do not necessarily accomplish this simply by striking according to juror race. In fact, the attorneys in our sample end up altering the racial composition of the seated jury relative to the jury pool very little: equivalent to no more than one juror of a given race in either direction on a twelve-person jury. The implication is that empowerment in voir dire can allow attorneys to uncover valuable information about jurors and therefore to avoid using racial stereotypes as a crutch.

We believe that our results could lead to more constructive policy advice than recommendations that have been made to restrict attorney freedoms in jury
selection. We will leave specific discussion of policy implications for the conclusion, along with an attempt to address the important question of how our results might generalize to regions that we have not been able to include in our empirical analysis.

The chapter will proceed as follows. The Section 4.2 introduces the model and presents its main results. The Section 4.3 describes the dataset that we employ. The Section 4.4 presents and discusses the empirical findings and a number of robustness checks. Section 4.5 concludes by discussing remaining challenges to be addressed and offering policy suggestions.

4.2 A Model of Attorney Empowerment and Jury Selection

4.2.1 Voir Dire in Practice and in the Model

A very common organizational framework for voir dire employed in some variation in most American criminal trials is known as the strike-and-replace method. Under this framework, the “panel” or “venire” of potential jurors – which we can take as being approximately randomly drawn from the county population – is first randomly ordered. Sometimes the entire panel is initially examined at this stage, and strikes for cause exercised. In either case, the first several panelists from the random order are provisionally seated and potentially examined in more depth. Once the attorneys are satisfied that none of the provisionally seated jurors can be removed for cause, they exercise peremptory challenges in some pre-determined alternating sequence, usually with the prosecutor going first. Those provisionally seated jurors who are thus excused are then replaced by the panelists next in line. This cycle then repeats itself.

Our understanding of the voir dire process is based on Gobert et al. (2009) and Starr and McCormick (2001), as well as informal interviews with a number of acquaintances in the legal profession. A somewhat more detailed description was provided in the previous chapter.
sometimes with the restriction that seated jurors who have already been “passed” on by both attorneys cannot be further examined or struck. The process ends when there are twelve jurors in the box who both attorneys have passed on, or when all peremptory challenges have been exhausted by both sides. In either case, the final seated jury is composed of the twelve jurors then seated, with the remaining venirepersons who have not already been struck also sent home.

The examination of potential jurors can be carried out primarily by the judge, by the judge with suggestions or a greater degree of participation from the attorneys, or primarily by the attorneys. Attorneys may also have access to the results of questionnaires filled in by potential jurors: either rudimentary surveys used by the court or, sometimes, “Supplementary Juror Questionnaires” (SJQs) that the attorneys have designed. The judge has full discretion over the decisions of how much attorney participation to allow and whether an SJQ is allowed. Individual judges tend to have strong views on how these decisions should be made, and to hold adamantly to them.

The primary concerns that are examined in any voir dire are the existence of personal relationships between panelists and any other parties involved in the trial, and the capacity to understand and contemplate the salient legal issues and evidence in a dispassionate and impartial manner. But many other subjects can also be explored, especially if attorneys have a high degree of participation. As voir dire questioning proceeds, attorneys challenge for cause by arguing that the responses of a given venireperson have revealed an inability to be impartial. The attorney on the opposite side is permitted to object to this argument and raise counter-arguments, and the judge must ultimately rule in favor of or against the strike. If the judge rules in favor of the strike, the potential juror is dismissed. Attorneys can strike an unlimited number of potential jurors for cause, as long as
the judge can be sufficiently convinced that disqualification has been
demonstrated.

Attorneys also have a limited number of peremptory strikes at their
disposal. These allow the removal of potential jurors without the requirement that
bias or any other deficiency be demonstrated. States and some counties provide
guidelines on the number of peremptory challenges that should be made available,
but judges are under no obligation to follow these guidelines, and exceptions are
often made for various reasons. Nonetheless, the total number of peremptory
strikes available is generally small relative to the size of the panel. There is also a
chance that a peremptory strike can be disallowed: if an attorney suspects that
the motivation for a strike is discriminatory in nature, a “Batson objection” can
be raised. If the opposing attorney cannot respond by stating a
non-discriminatory justification for the strike, it will be disallowed.

We model a simplified strike-and-replace voir dire procedure. The panel is
assumed to be infinite, with a composition across all venireperson characteristics
fixed and mirroring that of the population.99 The panel is arranged according to a
random ordering that attorneys do not observe, and the single panelist at the top
of this order is provisionally seated. If either attorney chooses to strike this
panelist and is successful, the panelist is dismissed, and the next panelist from the
random ordering is provisionally seated. This process continues until an
attempted strike is unsuccessful or neither attorney chooses to attempt a strike.
At that time, the provisionally-seated panelist becomes the sole member of the

---
99Courthouses can, in principle, bring in new panels for a given trial if the initially-drawn
panel is exhausted by strikes, though this is rarely necessary. However, there have been concerns
raised that panel composition does not accurately reflect population characteristics due to
insufficient methods to construct and maintain master juror lists. Neither of these issues is
important for our analysis.
one-person seated jury that will hear the trial.\textsuperscript{100}

There is a probability that any strike attempt will be successful, and we allow this probability to differ by attorney and to depend on skill and empowerment. This is mostly meant to represent strikes for cause. With a certain probability, the panelist under examination will exhibit some traits that one of the attorneys can represent as demonstrating impartiality, at least to the point of rebutting any objections raised by the opposing counsel and convincing the judge. Empowered and skilled attorneys can better uncover such traits, and better argue over the semantics of how they reveal impartiality. However, this framework can be slightly modified to incorporate peremptory strikes too. First of all, peremptory strikes can have a less-than-perfect success rate due to Batson objections. More generally, though, we could allow attorneys to have a much higher (or perfect) success rate for a limited number of strikes. This is an extension that we consider worthwhile and that we are currently pursuing. In the meantime, we propose that our present approach is capable of providing most of the insights that the more general approach will.

We will proceed for the rest of this section within this simplified framework. After formalizing some notation and concepts, the roles of attorney empowerment and skill will be clarified. Our main goal is to evaluate the probability that the seated juror will possess various attributes, and characterize how this probability depends on attorney empowerment and skill.

\textsuperscript{100}The one-person jury model can be interpreted as a single round of decisions in a series of rounds in which attorneys decide how to fill a 12-person jury one juror at a time, which is essentially the approach taken by Brams and Davis (1978). The relaxation of this assumption, though cumbersome, is therefore straightforward.
4.2.2 Preliminaries

Suppose that an $\alpha \in (0, 1)$ proportion of the panel is non-white and a $1 - \alpha$ proportion is white. $^{101}$ Each potential juror is described by two characteristics: his/her

1. race, $i = \{\text{non-white (n), white (w)}\}$ and

2. standard required to convict, $s = \{\text{high (h), low (l)}\}$.  

Race is fully observable to both attorneys, but standards required to convict are not. We interpret someone with a high standard required to convict – that is, someone requiring very strong evidence against the defendant to vote guilty – as someone with a predisposition or inclination towards the defense. We therefore refer to $h$-types as those with a predisposition towards the defense, and $l$-types as those with a predisposition towards the prosecution. The defense attorney will always prefer seating an $h$-type to seating an $l$-type on the jury, and vice versa for the prosecution. $p_i \in (0, 1)$ is the proportion of jurors of race $i$ who are predisposed to the defense (i.e. $h$-types), and the remaining $(1 - p_i)$ is the fraction who are predisposed towards the prosecution (i.e. $l$-types). These proportions are common knowledge to both attorneys.

We assume that the proportion of $h$-types is greater among non-whites than whites:

**Assumption 1 (Race and Predispositions)**

$$p_n > p_w. \quad (4.1)$$

$^{101}$ We will focus on race for simplicity of exposition, but note that all results hold for any observable characteristic.
Hence, without any additional information about the panelists, the defense will always prefer a non-white panelist over a white panelist, and the prosecution will always prefer to seat a white venireperson rather than a non-white.

Given the proportion of non-whites/whites and the fraction of their predisposition types in the potential juror population, we define

\[ p = \alpha p_n + (1 - \alpha) p_w \]  

(4.2)
as the proportion of panelists with a predisposition towards the defense.

### 4.2.3 Roles of Attorney Empowerment and Skill

In voir dire, the judge may grant greater power to the attorneys, for example by allowing them to participate to a greater degree in the examination of potential jurors or to use a Supplementary Juror Questionnaire. We assume that such greater power potentially allows attorneys to 1) find grounds to successfully strike any panelist with a higher probability than otherwise; and/or 2) to acquire additional information about the predispositions of panelists and thus to better identify potential jurors who are more favorably-inclined to their side. We will first discuss the expected racial composition and predispositions of the seated jury when attorneys are not empowered. We will then analyze the impact of attorney empowerment on these aspects of the seated jury when only the first mechanism is operative; when both mechanisms are operative; and when only the second mechanism is operative.

In our analysis, we allow the magnitude of both effects to depend on the relative skill levels of the attorneys. For simplicity of exposition, we assume that the defense attorney is more highly skilled than the prosecutor throughout our theoretical discussion. This is motivated by the fact that it is common in criminal
trials for the defense attorney to be more experienced than the prosecutor. However, the model is symmetric from the perspective of each side, and our general conclusions hold whether the defense or the prosecutor is more skilled.

4.2.4 No Attorney Empowerment

Let $\beta$ be the baseline probability that a strike by either side is successful. We assume that greater rhetorical skill is ineffective without some information beyond that revealed by a rudimentary judge-conducted examination of panelists. Thus, without empowerment, attorneys have no means to improve this success rate, regardless of any skill advantage.

With probability $\alpha$, the first potential juror to be provisionally seated will be non-white. The prosecution will want to strike this panelist, since no information is available besides the panelist’s observable race and the a priori knowledge that non-whites are more likely to be predisposed to the defense. With probability $(1 - \beta)$, the strike will be unsuccessful, and the non-white panelist will become the seated juror. With probability $\beta$, the strike will be successful, and the next panelist in the random order will be provisionally seated. There is then a probability of $(1 - \alpha)$ that this panelist will be white. The same reasoning can be repeated in this case, except that it will be the defense attorney who attempts to strike the provisionally-seated juror. It is clear from the symmetry of the reasoning and the equality of striking success rates that attorneys will be powerless to alter the probability that the seated juror will be non-white beyond the probability that any randomly-drawn panelist is non-white, namely $\alpha$. It follows directly that the expected probability that the seated juror will be predisposed towards the defense is just $p$. 

Proposition 1 When attorneys are unempowered, regardless of attorney skill,

(i) the probability that the seated juror will be non-white is the same, and

(ii) the probability that the seated juror will be predisposed towards the defense is
also the same

as if the jury were randomly drawn from the panel.

4.2.5 Empowerment Increases Striking Ability Only

Assume now that attorney empowerment allows attorneys to find a broader basis for striking a panelist, thereby increasing the probability that an attempted strike will be successful. Let $\beta_p$ be the prosecutor’s striking success rate with attorney empowerment, and likewise, let $\beta_d$ be the defense attorney’s striking success rate when attorneys are empowered, where both of these are greater than the baseline striking success rate without empowerment. We further assume that the effect of empowerment on striking success is greater for the more highly skilled attorney, because attorneys with a greater rhetorical facility will be better able to support their claims of juror impartiality in the face of objections from their opponents. This assumption also allows for greater attorney skill to improve the chance of blocking an opponent’s attempted strikes through more convincing counter-arguments. Since, as was discussed above, we are treating the defense attorney as the more highly skilled attorney for the sake of exposition, this implies that $\beta_d > \beta_p$. Combining these assumptions, we have:

Assumption 2 (Empowerment and Striking Ability)

$$0 < \beta < \beta_p < \beta_d < 1$$

The voir dire proceeds as follows. Each attorney decides whether to strike the panelist at the top of the random order or not. If the panelist is not struck or
an attempted strike fails, then the panelist will be seated. If the panelist is struck successfully, he/she is replaced by the next individual from the top of the randomly-ordered venire, and the attorneys decide to attempt a strike or not (and argue over the basis for the strike). The process continues until a juror is successfully seated.

Let $V$ be the probability that the panelist who is eventually seated is non-white. Additionally, let $V_n$ be the probability that the seated juror will be non-white if a non-white panelist is at the head of the queue. $V_w$ is, then, the probability that the seated juror will be non-white if a white panelist is at the head of the queue. Our assumption that the panel is of infinite size ensures that the identity of the juror at the head of the queue does not change the composition of the remaining panel.

With no additional information about the predispositions of the potential jurors aside from race, the defense (prosecuting) attorney will always try to strike when a white (non-white) panelist is at the head of the queue. Let’s first consider the case in which a non-white panelist is at the head of the queue. With probability $(1 - \beta_p)$, the panelist will survive the prosecutor’s strike request and end up as the seated juror. With probability $\beta_p$, he/she will be successfully struck and replaced by a random panelist at the head of the queue. Therefore, $V_n$ can be defined as

$$V_n = (1 - \beta_p) + \beta_p V.$$

Similarly, when a white is at the head of the queue, he/she will survive the defense attorney’s strike with probability $(1 - \beta_d)$, or with probability $\beta_d$, will be struck and be replaced by a random panelist at the head of the queue. $V_w$ is then
\[
V_w = \beta_d V. 
\]  
(4.5)

The overall expected probability that the panelist who is eventually seated is non-white is a weighted average of \(V_n\) and \(V_w\), with weights given by the probability that a non-white will be at the head of the queue, \(\alpha\), versus a white, \((1 - \alpha)\):

\[
V = \alpha V_n + (1 - \alpha)V_w. 
\]  
(4.6)

This probability is then defined completely by the system of three equations (4.4), (4.5), and (4.6). Solving this system yields

\[
V = \frac{\alpha(1 - \beta_p)}{(1 - \beta_d) + \alpha(\beta_d - \beta_p)}. 
\]  
(4.7)

It is trivial to show from (4.7) that \(V > \alpha\) as long as \(\beta_d > \beta_p\), which holds by Assumption 2.

It is also intuitive and can be easily shown that the probability that the seated juror is predisposed to the defense – which we call \(U\) – is higher than that in the panel, \(p\). This probability is defined as

\[
U = p_n V + p_w (1 - V), 
\]  
(4.8)

which, after substitution of (4.7) for \(V\), yields

\[
U = \frac{\alpha(1 - \beta_p)p_n + (1 - \alpha)(1 - \beta_d)p_w}{(1 - \beta_d) + \alpha(\beta_d - \beta_p)}. 
\]  
(4.9)

It can be shown that \(U > p\) as long as \(\beta_d > \beta_p\) and \(p_n > p_w\), which hold directly by Assumptions 1 and 2.
These results on the race and predisposition of the seated juror are intuitive. When empowerment increases striking ability only, the attorneys are simply able to more successfully strike on the basis of their \textit{a priori} racial preferences; and when the defense attorney has a skill advantage, the defense is relatively more successful at striking than the prosecution, and so the outcome will more frequently benefit the defense attorney on net.

**Proposition 2** When empowerment only increases attorney striking ability, and the defense is more skilled,

(i) the probability that the seated juror will be non-white is higher, and

(ii) the probability that the seated juror will be predisposed towards the defense is also higher

than if the jury were randomly drawn from the panel.

4.2.6 Empowerment Increases Striking Ability and Information

Now suppose that empowerment not only increases the striking ability of the attorneys, but also allows attorneys to acquire additional information about the panelist and to better identify panelists who are favorably-inclined to their side.

We assume that, when empowered, attorneys observe a dichotomous signal \( \theta \in \{ \text{High}(H), \text{Low}(L) \} \) correlated with the panelists’ predispositions. For simplicity, we assume that only the more skillful attorney of the two sides is able to observe \( \theta \).\(^{102}\) A panelist of type \( s \) exhibits \( H \) with probability \( \sigma_s \) and exhibits \( L \) with probability \( 1 - \sigma_s \). We assume that the \( \sigma \)'s are independent of race and that \( \sigma_h > \sigma_l \), so that any \( h \)-type panelist has a higher probability of exhibiting \( H \) than any \( l \)-type panelist. The probability that a panelist of race \( i \) exhibits \( H \) is then

\(^{102}\)One could also assume that both attorneys observe \( \theta \), but that the more skillful attorney can observe the signal more accurately. This alternative assumption, while more realistic, has not been employed here because it complicates the model without providing additional insight.
\[
\gamma_i = p_i \sigma_h + (1 - p_i) \sigma_l, \tag{4.10}
\]
and, under our parameters assumptions about \(p_i\) and \(\sigma_h\), we know that \(\gamma_n > \gamma_w\).

We can also express the conditional posterior probability that a juror of race \(i\) exhibiting \(H\) is actually an \(h\)-type as

\[
\delta_{iH} = \Pr[s = h | i, \theta = H] = \frac{p_i \sigma_h}{p_i \sigma_h + (1 - p_i) \sigma_l}, \tag{4.11}
\]

Similarly, the conditional posterior probability that a juror of race \(i\) and signal \(\theta = L\) is actually an \(h\)-type is

\[
\delta_{iL} = \Pr[s = h | i, \theta = L] = \frac{p_i (1 - \sigma_h)}{p_i (1 - \sigma_h) + (1 - p_i) (1 - \sigma_l)}. \tag{4.12}
\]

Because the less-skilled prosecutor does not see \(\theta\), his strategy in voir dire remains the same as before. That is, the prosecutor will try to strike non-white panelists and will attempt to block the striking of white panelists by the defense.\(^{103}\) However, the defense attorney, now having seen \(\theta\), updates her race-based \textit{a priori} belief about any panelist’s predisposition with information revealed by \(\theta\), according to (4.11) and (4.12).

Note that, given our assumptions that \(p_n > p_w\) and \(\sigma_h > \sigma_l\), a non-white panelist exhibiting \(H\) has the highest probability of being an \(h\)-type \((\delta_{nH})\), while a white panelist exhibiting \(L\) has the lowest probability of being an

\(^{103}\)In fact, we implicitly make a rather stronger set of assumptions about the skill-disadvantaged attorney in order to ensure that this behavior arises. In addition to not allowing the prosecutor to see the signal, we also assume that the prosecutor is unable to learn anything by observing the striking behavior of the defense attorney, or otherwise condition his own striking behavior on the anticipated or actual striking behavior of the defense attorney. The prosecutor thus behaves as though fully unaware of the potential existence of any signal. Allowing the less-skilled attorney to instead observe a less precise signal necessitates the relaxation of this package of assumptions, which can give rise to some anomalous cases.
$h$-type ($\delta_{wL}$). However, it remains unclear whether $\delta_{wH} > \delta_{nL}$. This inequality will hold if and only if the following condition is met.

**Condition 1 ($\delta_{wH} > \delta_{nL}$)**

$$\frac{\sigma_h 1 - \sigma_l}{\sigma_l 1 - \sigma_h} > \frac{p_n 1 - p_w}{p_w 1 - p_n}. \quad (4.13)$$

This condition states that the odds ratio of exhibiting the signal $H$ between $h$-types and $l$-types is greater than the odds ratio of being an $h$-type between non-whites and whites. Intuitively, $\delta_{wH} > \delta_{nL}$ – so that the defense attorney will prefer some white panelists to some non-white panelists – if and only if the signal $\theta$ is sufficiently more informative about juror predispositions than race alone.

**Defense Attorney’s Reservation Striking Threshold**

When empowerment enables the defense attorney to better identify panelists predisposed to the defense, the defense and prosecuting attorneys can now potentially “agree” on their preference to strike or to seat a panelist.\textsuperscript{104} We, therefore, slightly modify our assumptions on empowerment’s effect on attorneys’ striking abilities by assuming that 1) if neither attorney wants to strike the panelist at the head of the queue, that individual will be seated on the jury; and 2) if both attorneys want to strike a panelist, that panelist will be struck with probability 1. If the prosecution and the defense disagree, then the probabilities of successful strikes are defined by $\beta_p$ and $\beta_d$, respectively, as before.

Given the primitive parameters of the model and the informativeness of $\theta$, the defense attorney will take the static behavior of the prosecutor as given, and choose her own striking behavior to maximize the expected probability that the seated juror will be predisposed to the defense. The defense attorney’s strategy

---

\textsuperscript{104}For example, the prosecution will always want to seat a white panelist, as before. However, now the defense may want to seat a white panelist if her posterior belief that the panelist is an $h$-type is sufficiently high.
can be represented as the choice of a reservation threshold \( \delta_R \) such that she will attempt to strike a panelist if \( \delta_i \theta < \delta_R \) and argue against any attempted strikes otherwise. This can be described more formally as follows:

\[
\delta_R \in \arg \max_{\hat{\delta}} U(\hat{\delta})
\]  

(4.14)

where \( U(\hat{\delta}) \) is the probability that the seated juror will be predisposed towards the defense and is defined implicitly by the following system of five equations:

\[
U_{nH}(\hat{\delta}) = (1 - \beta_p)\delta_{nH} + \beta_p U(\hat{\delta})
\]  

(4.15)

\[
U_{wL}(\hat{\delta}) = (1 - \beta_d)\delta_{wL} + \beta_d U(\hat{\delta})
\]  

(4.16)

\[
U_{nL}(\hat{\delta}) = \begin{cases} 
(1 - \beta_p)\delta_{nL} + \beta_p U(\hat{\delta}) & \text{if } \delta_{nL} \geq \hat{\delta} \\
U(\hat{\delta}) & \text{otherwise}
\end{cases}
\]  

(4.17)

\[
U_{wH}(\hat{\delta}) = \begin{cases} 
(1 - \beta_d)\delta_{wH} + \beta_d U(\hat{\delta}) & \text{if } \delta_{wH} \geq \hat{\delta} \\
(1 - \beta_d)\delta_{wH} + \beta_d U(\hat{\delta}) & \text{otherwise}
\end{cases}
\]  

(4.18)

\[
U(\hat{\delta}) = \alpha \gamma_n U_{nH}(\hat{\delta}) + \alpha (1 - \gamma_n) U_{nL}(\hat{\delta}) + (1 - \alpha) \gamma_w U_{wH}(\hat{\delta}) + (1 - \alpha)(1 - \gamma_w) U_{wL}(\hat{\delta}).
\]  

(4.19)

The structure of equations (4.15) to (4.19) can be explained as follows.

Since, as was mentioned above, a non-white exhibiting \( H \) has the highest probability of being predisposed to the defense and a white exhibiting \( L \) has the lowest, it is always optimal for the defense attorney to try to argue against attempted strikes of a non-white \( H \) and to attempt to strike a white \( L \). The less-skilled prosecutor will always want to strike any non-white and argue against
attempted strikes of any white, as explained above, and the defense attorney anticipates this fully. These choices are reflected in equations (4.15) and (4.16) defining $U_{nH}$ – the expected probability that the seated juror will be predisposed to the defense when a non-white exhibiting $H$ is at the head of the queue – and $U_{wL}$ – the corresponding probability when a white exhibiting $L$ is at the top of the order – respectively.

The probabilities that the seated juror will be an $h$-type when a non-white exhibiting $L$ or a white exhibiting $H$ is at the head of the queue depend on the threshold $\delta$ relative to the $\delta_{i\theta}$’s. For example, if a white exhibiting signal $H$ is at the head of the queue, $\delta_{wH} \geq \bar{\delta}$ represents a case in which neither the defense nor the prosecution attempt to strike him. Then the expected probability that the seated juror will be an $h$-type when a white exhibiting $H$ is at the head of the queue in (4.18) is exactly equal to $\delta_{wH}$, since he will be seated with probability 1 given our assumptions about attorney striking success. However, if $\delta_{wH} < \bar{\delta}$, this represents a case in which the defense attempts to strike, and the panelist will hence only be seated with probability $(1 - \beta_d)$, or will be replaced by a random panelist with probability $\beta_d$.

Finally, the overall expected probability that a seated juror will be predisposed towards the defense, $U$, for a given threshold $\bar{\delta}$, is a weighted average of (4.15) through (4.18), with weights given by the population proportions of the four $i\theta$ groups. There will in general be a range for $\bar{\delta}$ over which $U$ is maximized; we do not need to characterize $\delta_R$ more precisely than noting that it belongs to this range. This probability is hence maximized when the defense attorney follows the strategy of attempting to strike any panelist with $\delta_{i\theta} < \delta_R$ and arguing against any attempted strikes otherwise.
Race and Predisposition of the Seated Juror

We will limit our discussion to two key cases regarding the striking behavior of the defense attorney and the strength of the signal.

**Case 1:** $\delta_{nH} > \delta_{nL} > \delta_{R} > \delta_{wH} > \delta_{wL}$

In this case, Condition 1 does not hold, so that the information revealed by the signal is not enough to overcome the defense attorney’s *a priori* preference for non-whites. Further, non-whites represent a large enough fraction of the population and are likely enough to be predisposed to the defense that the defense attorney chooses to set her reservation threshold $\delta_{R}$ above $\delta_{wH}$. The net result is that the defense attorney will attempt to strike any white panelist, regardless of $\theta$, and will object to attempts by the prosecutor to strike any non-white panelists, again regardless of $\theta$. Therefore, this case is effectively the same as when empowerment only increased the striking abilities of the attorneys. The probability that the seated juror will be non-white is higher than if a randomly selected juror were seated, as is the probability that the seated juror will be predisposed towards the defense.

**Case 2:** $\delta_{nH} > \delta_{wH} > \delta_{R} > \delta_{nL} > \delta_{wL}$

A more interesting case is when Condition 1 is met, so that the defense attorney is able to identify some white panelists with a higher probability, conditional on the signal, of being predisposed to the defense than some non-white panelists. Furthermore, assume that the fraction of the population exhibiting signal $H$ is moderate enough that it is optimal for the defense attorney to set her reservation striking threshold $\delta_{R}$ above $\delta_{nL}$. Thus, the defense attorney will argue against the prosecutor’s attempted strikes of non-white panelists exhibiting $H$, as usual; but will herself attempt to strike non-white panelists exhibiting $L$, joining
in with the prosecutor’s efforts in this regard. Furthermore, the defense attorney will not attempt to strike any white panelists exhibiting $H$, and since the prosecutor will never attempt to strike any white panelist, a white panelist exhibiting $H$ that is at the top of the random order of panelists will become the seated juror without objection from either side. Therefore, in this case, $V$, the expected probability that the panelist who eventually becomes the seated juror will be non-white, is defined by the following system of equations:

\[
\begin{align*}
V_{nH} &= (1 - \beta_p) + \beta_p V \\
V_{nL} &= V \\
V_{wH} &= 0 \\
V_{wL} &= \beta_d V \\
V &= \alpha \gamma_n V_{nH} + \alpha (1 - \gamma_n) V_{nL} + (1 - \alpha) \gamma_w V_{wH} + (1 - \alpha) (1 - \gamma_w) V_{wL}.
\end{align*}
\]

Solving for $V$, we have

\[
V = \frac{\alpha \gamma_n (1 - \beta_p)}{1 - \alpha [1 - \gamma_n (1 - \beta_p)] - (1 - \alpha) [(1 - \gamma_w) \beta_d]}.
\]

It can be shown that $V < \alpha$ always, as long as Condition 1 and our basic parameter assumptions hold.\(^{105}\) Therefore, the probability that the seated juror will be non-white is less in this case than if the jury were chosen randomly.

\(^{105}\)The demonstration of this claim is tedious, as the expression involves all of the parameters of the model through the definition of $\gamma_i$ in equation (4.10). It is more straightforward to show a sufficient but not necessary condition in terms of the striking success rates of each attorney. This sufficient condition is $\beta_d > p_n p_w \times (1 - \beta_p)$. Intuitively, this condition is saying that it must be sufficiently more likely that the defense attorney successfully strikes a white panelist exhibiting $L$ (hence obtaining another draw from the panel) than it is for the prosecutor to fail to strike a non-white panelist exhibiting $H$ (which would result in the voir dire terminating with a non-white seated juror). This condition is fairly stringent for many parameter values. However, as noted, the only necessary conditions are the baseline parameter assumptions we have made.
However, as in previous cases, \( U \), the probability that the seated juror will be predisposed towards the defense, is higher than in the panel overall. The most straightforward way to calculate \( U \) is to use the specific form of equations (4.15) through (4.19) for this case, which is

\[
\begin{align*}
U_{nH} &= (1 - \beta_p)\delta_{nH} + \beta_p U \\
U_{nL} &= U \\
U_{wH} &= \delta_{wH} \\
U_{wL} &= (1 - \beta_d)\delta_{wL} + \beta_d U \\
U &= \alpha \gamma_n U_{nH} + \alpha(1 - \gamma_n)U_{nL} + (1 - \alpha)\gamma_w U_{wH} + (1 - \alpha)(1 - \gamma_w)U_{wL}.
\end{align*}
\]

Solving for \( U \) yields

\[
U = \frac{\alpha \gamma_n (1 - \beta_p)\delta_{nH} + (1 - \alpha)\gamma_w \delta_{wH} + (1 - \alpha)(1 - \gamma_w)(1 - \beta_d)\delta_{wL}}{1 - \alpha(1 - \gamma_n(1 - \beta_p)) - (1 - \alpha)(1 - \gamma_w)\beta_d}
\]

where it can be shown that \( U > p \) under our parameter assumptions.\(^{106}\)

In Case 2, we have thus shown an example for which empowerment (which increases attorney striking abilities and allows them to more precisely identify juror predispositions in this case) leads to a lower probability that the seated juror will be non-white, compared to the proportion of non-whites in the population.

**Other Cases**

Cases 1 and 2 are the most extreme of the many possible cases. Others are nonetheless interesting. When the proportion of the panel exhibiting the High

---

\(^{106}\)Again, this claim is tedious to demonstrate algebraically. However, it is clear that it must hold, since this is the objective that the defense attorney is attempting to maximize. If \( U \) were less than \( p \), the defense attorney would change striking behavior – for example, to that of Case 1 above – and be able to obtain a different \( U \) greater than \( p \).
signal is small, the defense attorney may choose to refrain from striking all but white panelists exhibiting $L$, achieving a probability that the seated juror is non-white slightly greater than $\alpha$. On the other hand, when this proportion is large, the defense attorney may wish to attempt to strike all but non-white panelists exhibiting $H$; depending on the specific value of $\sigma_h$, the probability that the seated juror is non-white can be slightly less than or slightly greater than $\alpha$.

We will avoid a detailed exposition of these cases for the present, as Cases 1 and 2 are sufficient to support the following Proposition.

**Proposition 3** When empowerment increases attorney striking ability and allows attorneys to better identify panelists’ predispositions, and the defense attorney is more skilled,

(i) the probability that the seated juror will be non-white may be higher or lower, yet

(ii) the probability that the seated juror will be predisposed towards the defense is always higher

than if the juror were randomly drawn from the panel.

4.2.7 Empowerment Increases Information Only

For completeness, we can also consider the set of cases for which only the second effect of empowerment is present. Specifically, assume now that empowerment does not increase attorney striking ability, but does allow the more skilled attorney to more accurately identify juror inclinations. Mechanically, this just involves setting $\beta_p = \beta_d = \beta$ in equations (4.15) to (4.19) and repeating the analysis of the previous section.\textsuperscript{107} We proceed here with an intuitive discussion of the salient results, without revisiting the formal analysis.

\textsuperscript{107}We retain the assumption that the attorneys have a perfect striking success rate when they would both attempt to strike a panelist, and a perfect success rate at seating a panelist whom neither of them attempts to strike.
In all cases, the probability that the seated juror will be non-white is less than the population proportion non-white. Whenever the attorneys differ in their desire to strike a panelist, the panelist will be struck successfully with probability $\beta$ on net. The attorneys always disagree on white panelists exhibiting $L$ and on non-white panelists exhibiting $H$. However, the attorneys can potentially agree on the two intermediate groups, depending on the quality of the signal and the defense attorney’s reservation striking threshold. The result is either that white panelists exhibiting $H$ will be struck with probability zero (rather than $\beta$); non-white panelists exhibiting $L$ will be struck with probability one (rather than $\beta$); or both.\textsuperscript{108} The net effect of any of these, in combination with the fixed and common probability of a panelist from the extreme groups being struck, is to reduce the probability that a white panelist will be struck relative to that of a non-white panelist being struck. The probability that the seated juror will be predisposed to the defense must always be at least $p$, since the skilled defense attorney can always replicate the outcome that would arise if there were no empowerment by disregarding the information obtained through empowerment.\textsuperscript{109}

Thus, we have our final proposition.

\textsuperscript{108}The less skilled prosecutor does not observe the signal, as per our assumption stated earlier, but rather “agrees” with the defense attorney in these cases by virtue of his usual strategy of only attempting to strike non-whites.

\textsuperscript{109}For the case in which Condition 1 does not hold and the defense attorney’s reservation striking threshold is greater than $\delta_{wH}$ but less than $\delta_{nL}$, the results collapse to those when attorneys are unempowered. The defense attorney will attempt to strike any white panelist (and argue against attempted strikes by the prosecutor of any non-white panelist), regardless of the signals exhibited by the panelists. The probability that any given panelist will be successfully struck is $\beta$, and therefore, the probability that the seated juror will be non-white is $\alpha$ and the probability that the seated juror will be predisposed towards the defense is $p$. We consider this as somewhat of a degenerate case: empowerment only provides information, but so little information that it is as though attorneys are not empowered at all. We therefore group this case together with the baseline case in which attorneys are not empowered.
**Proposition 4** When empowerment only allows attorneys to better identify panelists’ predispositions, and the defense attorney is more skilled,

(i) the probability that the seated juror will be non-white is lower, and

(ii) the probability that the seated juror will be predisposed towards the defense is higher

than if the juror were randomly drawn from the panel.

4.2.8 Summary

Comparing Propositions 2 through 4 to Proposition 1, the model produces a clear and intuitive testable implication. When the defense attorney is relatively more skilled than the prosecutor, attorney empowerment in voir dire is associated with an increase in the probability that the seated juror will be predisposed to the defense. Predictions concerning the race of the seated juror are less clear. If the nature of attorney empowerment is only such that it improves the striking ability of attorneys, empowerment will be associated with an increase in the probability that the seated juror will be non-white. In direct contrast, if the nature of attorney empowerment is only such that it allows attorneys to learn more about juror predispositions, empowerment will be associated with a *decrease* in the probability that the seated juror will be non-white. And, therefore, if both mechanisms are operative, empowerment could be associated with either an increase or a decrease in the probability that the seated juror will be non-white. Since the precise nature of attorney empowerment and of any information uncovered by attorneys in voir dire will be unobserved to us in our empirical analysis, we will not be able to test these predictions concerning race directly. Instead, our approach will be to attempt to draw inferences about the nature of empowerment from our results. Before presenting these results, we will describe our dataset.
4.3 The NCSC/ICPSR Hung Juries Dataset Revisited

Our empirical analysis relies on the same dataset that was employed in the previous chapter. As discussed there, the NCSC study provides a comprehensive look at the trial and all the parties involved: the defendant(s) and the victim(s) (if any), the judge, the attorneys, and, most importantly for our analysis, the seated jurors. Unfortunately, we do not have information about the initial jury pool from which the seated jury was formed, except for the initial panel size.\(^{110}\)

Regrettably, although the NCSC dataset contains information on 351 cases and 3,497 jurors in total, not every individual in our data answered every survey question. In our main specifications, which use responses from the case, judge, and juror modules of the dataset, we have data on 1,846 jurors across 251 trials. The attorney module has the lowest response rate, and in some specifications with attorney controls we have as few as 152 trials represented.

All data were collected confidentially, and no information was made available to any of the trial participants during the trial or during any subsequent appeals.\(^{111}\)

4.3.1 Variables of Interest

Before presenting summary statistics, we will give a detailed discussion of the empirical analogues we propose for some of the important theoretical variables from the model.

\(^{110}\)We can, however, infer the average characteristics of the panel from those of the county as a whole, since a panel is essentially a random sample of the county population. We include county fixed effects in all of our empirical specifications, in part to account for the composition of the panel as much as possible.

\(^{111}\)These were conditions of the data collection agreements. We do not observe if any of the trials in the dataset have since been retried or heard by an appeals court.
Juror Predispositions

Each juror was asked “Before you began deliberating with your fellow jurors at the end of the trial (after all of the evidence and the judge's instructions had been presented), which side did you favor?”, which was answered on a scale of 1 (“Strongly favored the prosecution”) to 7 (“Strongly favored the defense”). We interpret higher responses on this seven-point scale as manifesting greater doubts on the part of a juror that sufficient incriminating evidence had been presented in relation to that juror’s personal (and otherwise unobservable) standard required to convict. Thus, jurors with higher responses to this question are assumed to have entered deliberations with a general mindset more inclined towards the defense.

We believe that this variable closely resembles the concept of juror predispositions as it was conceived and employed in the construction of the model, and in addition that it is a reasonable objective that attorneys would want to target in practice. The wording of the question can encompass subtle juror leanings deriving potentially from common life experiences with the defendant, the resonance of certain pieces of evidence and argument, and so on. These are things that attorneys could conceivably foresee in jury selection and shape arguments to further cultivate during the trial.

Jurors were also asked their general opinions of the defendant and any victims outside the context of the facts of the case, as well as a series of questions concerning the evolution of their opinions during and after deliberations. We feel that these questions are less accurate representations both of our notion of juror predispositions and of traits that attorneys realistically target. Jurors could very likely find the defendant sincere and the victim untrustworthy and less believable, yet still lean towards believing the defendant guilty. Furthermore, attorney manuals have only very recently begun to espouse the idea that attorneys should
“select juries, not jurors,” and are more likely to advise that deliberations are akin to a black box, so that that the surer route to success is to focus on individual jurors during jury selection, rather than attempt to predict deliberation dynamics. This traditional view goes back to Clarence Darrow, who is said to have remarked that the outcome of a trial has virtually been decided by the time the jury has been sworn in.

All juror questionnaires were distributed at the conclusion of the trial. There is therefore a potential concern that this question on the side favored prior to deliberation is an inaccurate representation of jurors’ actual states of mind at that time. For example, if the final verdict was the most salient memory for jurors at the conclusion of the trial, it is possible that they engaged subconsciously in *ex post* rationalization in answering questions regarding earlier stages. We have not been able to satisfactorily diagnose this possibility in the data. However, we find reassurance in the general care taken by the survey designers, and in the fact that this section of the survey was carefully and transparently constructed to elicit information on the sequential nature of the opinion formation of jurors over the entire trial and deliberations. For example, the survey includes several reminders to respondents to consider precise times that they changed their minds.

**Attorney Empowerment**

The main case data contains an indicator for whether the voir dire was conducted by the judge alone, the judge with questions suggested by attorneys, the judge and attorneys together, or primarily attorneys. We combine the first two and the latter two categories to form a dummy variable for whether attorneys had high participation or low participation.\footnote{For all 345 trials for which a response to this question is available, only six had voir dire conducted primarily by attorneys, and only 36 had questions submitted by attorneys. Anecdotal evidence suggests that questions submitted by attorneys are often ignored by judges.}
The main case data also contains a dummy variable for whether or not there was a Supplemental Juror Questionnaire used by the attorneys. However, there is no other information available concerning SJQs. For example, we do not observe if it was requested by one side or the other if present; we do not know if a request was made and denied if not present; we do not observe if both sides had access to the responses if present; and we do not observe any of the specific information that it contained if present.

Two important points should be mentioned concerning these two variables. First, it is implicitly assumed that each side is equally empowered by either instrument (high participation or a questionnaire) in terms of the additional information about jurors that the defense and prosecuting attorneys can obtain with them. This is in line with the assumptions of our model. There is no information available concerning which attorney was more active in questioning or how any SJQs were used, so there does not appear to be any feasible alternative to making this uniformity assumption. As in the model, however, attorney skill will be examined as a way of identifying heterogeneity across sides in the use that was made of these freedoms when they were available.

Second, it seems plausible to assume that the presence (or lack thereof) of these attorney freedoms is exogenous to other aspects of the trial. Individual judges have wide latitude in determining how the jury selection process is conducted, and most judges strictly adhere to a set of personal preferences, especially in regard to the degree of participation afforded to attorneys in questioning potential jurors. Judges can occasionally be persuaded to change policies if an attorney on either side is sufficiently aware of the judge’s usual

---

113 To reiterate, the survey instrument is in no way related to whatever SJQs may or may not have been used in any trials, and the attorneys did not have access to any of the information from the survey instrument at any stage of the trial.
policies and submits a motion for a more extensive process than is typically allowed. This seems to be very rare in the case of attorney participation, but perhaps not unheard of in the case of supplementary questionnaires.

Without this exogeneity, our estimation would suffer a severe shortcoming. If, for example, judges always allowed high attorney participation in trials for crimes that had received extensive negative media attention, high participation could be highly correlated with jurors leaning away from the defense (based in part on preconceived notions of the trial that the voir dire did not successfully weed out). But this correlation would be driven by the fact that both high participation and unfavorable opinions of the defendant are associated with the (unobserved in the dataset) media attention. If there were some degree of endogeneity of this sort (for example, if, for some trials, a questionnaire were allowed based on the motions of a defense attorney desiring as much information as possible on a venire of potential jurors suspected of holding extreme prejudices against the defendant), the estimated effects of attorney empowerment on juror predispositions would be biased in a negative direction.

For most of our specifications, we use a single dummy variable for whether or not attorneys were empowered in voir dire. We construct this variable to take a value of 1 either if attorneys had high participation or if there was an SJQ in use, and to take a value of 0 if attorneys had neither freedom. However, we will also explore the robustness of our results to alternative definitions of attorney empowerment.114

114Two small idiosyncrasies of the dataset should be mentioned briefly. All of the trials taking place in Bronx County in the dataset had high attorney participation in questioning. On the other hand, none of the trials in the dataset taking place in Maricopa had a questionnaire available. All of our specifications control for the county in which the trial took place, and we do not otherwise foresee these issues presenting problems within our empirical strategy, which is more focused on the interaction of attorney skill and empowerment.
The dataset also contains additional information about the voir dire process. We observe the number of peremptory challenges actually exercised by each side, but not how many were available nor how this was determined. We also observe the total number of strikes for cause, but not how many suggested strikes for cause were objected to or denied. There is a dummy variable indicating whether any Batson objections were raised during voir dire, but if there were, we do not observe how many there were, which side raised them, whether they resulted in the finding of any actual violation, or what remedies were imposed, if any. There is also a dummy variable for whether the jury was anonymous, but there is no indication of whether this anonymity was from the perspective of record-keeping only or if the jury was literally empaneled in a separate room and voir dire conducted remotely. We observe the length of time voir dire took in hours, but not whether there was a time limit in place. Finally, we observe the total size of the initial panel. We have concerns about the completeness of the information furnished by these variables and about their exogeneity, but we do make use of them in some specifications.

Attorney Skill

The judge and each juror were asked separate questions of the form “How skillful was the [prosecutor/defense attorney] during the trial?” which were answered on a scale of 1 (“Not at all skillful”) to 7 (“Very skillful”).

In the reductionist spirit of the model, we would like to have an indicator for which attorney was the relatively more skilled. It would seem uncontroversial to assume that the judge would provide the most objective assessment of attorney skill, so that an attractive option would be to construct a dummy variable for attorneys were also asked similar questions about their own and their opponent’s skill. We do not make use of these variables due to concerns about subjectivity and for the pragmatic reason, as mentioned earlier, that attorneys had the lowest response rates.
whether or not the judge gave one attorney a higher skill rating than the other. Unfortunately, the judge assigned the same skill rating to both attorneys in about half the cases.

We have therefore decided to construct a hybrid skill indicator for each attorney by calculating an average of the judge’s rating on the one hand, and an average juror rating on the other hand. We then calculate a relative skill index by subtracting the hybrid skill rating for the prosecutor from the hybrid skill rating for the defense attorney. Finally, we construct a dummy variable that we refer to as ‘Defense More Skilled’ that takes a value of 1 when this relative skill index is greater than zero.

We are aware of the arbitrary nature and potential shortcomings of this measure of relative attorney skill. As we will discuss below, we do check for correlations between our relative skill index and the few objective indicators of attorney experience that we observe (namely years in legal practice, previous criminal trial experience, and age). We have smaller samples for this purpose due to the lower attorney response rates, but we find some reassuring results. We will also discuss a number of robustness checks we have attempted.

A conceptual shortcoming of any of these skill measures is that they refer to skill in the trial arena as a whole, whereas, it could be argued, our model suggests that we should be focusing on attorney skill in voir dire exclusively. We respond to this, first of all, by noting that skill in various aspects of litigation are likely to be highly correlated, so that one lawyer displaying more skill than another during a trial as a whole should be a good proxy for that same lawyer having shown superior skill in any individual stage of the trial. However, the more important point is that we do not necessarily want to disregard the possibility that attorney skill in argument and presentation of evidence could affect individual juror
predispositions, nor that attorneys could anticipate this during jury selection. This is neither ruled out by a more general interpretation of our theoretical framework nor by our empirical measurement of predispositions discussed above.

**Other Variables**

An important control variable in our analysis is the objective strength of evidence. For each trial, the judge was asked how close the case was based on the evidence that had been presented. The question was required to be answered on a scale of 1 (“Evidence strongly favored prosecution”) to 7 (“Evidence strongly favored defense”). Two dummy variables have been created from this variable. The first represents evidence strongly favoring the defense, and corresponds to responses of 6 or 7 (versus responses of 1 through 5). The second represents evidence strongly favoring the prosecution, and corresponds to responses of 1 or 2 (versus responses of 3 through 7). The overall excluded category therefore corresponds to a close trial in terms of the evidence favoring neither side very strongly.\(^{116}\)

For juror race, we focus on the broad white/non-white dichotomous distinction that we used to illustrate the model. We recognize that the more specific black/white distinction typically receives the greatest focus in discussions of public policy, but the more salient difference across our dataset as a whole is the broader one. The non-white category is predictably heterogeneous across the four counties, but the constituent members in each case are, in general, assumed by commentators and legal practitioners to hold greater sympathies for defendants than their white counterparts.

Other juror-level characteristics we can observe include education, income, 

\(^{116}\)We could be more flexible by treating all seven categories distinctly. As this approach does not affect our results, we have chosen to retain the more parsimonious treatment.
age, gender, and religiousness. These are encoded categorically in the dataset, and we have converted them to dummy variables in some cases and continuous variables in other cases.\footnote{See Chapter 3 for additional details.}

As has been mentioned before, we include county-level fixed effects in all of our specifications. We can also control for some other trial-level variates. The dataset provides a fairly detailed breakdown of trials by type of charge, which we coarsen into broad headings of violent, property and drug crimes. We can also control for defendant race and for the race and gender of any victims if present.\footnote{We can also observe the defendant’s gender, but we do not control for this because less than 10\% of trials in the dataset had a female defendant. A handful of trials in our estimation samples have multiple defendants, but always of the same race.} Finally, we can control for whether or not the defendant had counsel provided at no cost, which we interpret as a rough indicator of the defendant’s financial means.

4.3.2 Descriptive Statistics

Table 4.1 presents selected summary statistics for our dataset. The sample means are calculated for the observations that will be used in our main empirical specifications.

On average, jurors in our sample were mildly predisposed towards the prosecution prior to deliberations (comparing the mean response of 3.5 to the midpoint of 4 on the seven-point scale, with 7 representing “Strongly favored the defense”). However, individual responses are dispersed widely around this sample mean, and there was wide variation within many trials as well.

Slightly more than half of the seated jurors in our sample are non-white. This ranges from 84\% in the Bronx to 20\% in Maricopa. These differences across counties in the racial composition of seated juries largely correspond with the different racial compositions of the populations of these counties. The non-white
Table 4.1: Summary Statistics

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juror Predisposition (1 (pros.) - 7 (def.))</td>
<td>3.517</td>
<td>2.064</td>
<td>1846</td>
</tr>
<tr>
<td>Juror Non-White</td>
<td>0.520</td>
<td>0.500</td>
<td>1846</td>
</tr>
<tr>
<td>Bronx</td>
<td>0.842</td>
<td>0.365</td>
<td>310</td>
</tr>
<tr>
<td>LA</td>
<td>0.631</td>
<td>0.483</td>
<td>472</td>
</tr>
<tr>
<td>DC</td>
<td>0.507</td>
<td>0.500</td>
<td>609</td>
</tr>
<tr>
<td>Maricopa</td>
<td>0.202</td>
<td>0.402</td>
<td>455</td>
</tr>
<tr>
<td>High Participation</td>
<td>0.622</td>
<td>0.486</td>
<td>251</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>0.203</td>
<td>0.403</td>
<td>251</td>
</tr>
<tr>
<td>Either</td>
<td>0.685</td>
<td>0.465</td>
<td>251</td>
</tr>
<tr>
<td>Both</td>
<td>0.139</td>
<td>0.347</td>
<td>251</td>
</tr>
<tr>
<td>Defense Skill (judge; 1-7)</td>
<td>5.104</td>
<td>1.438</td>
<td>251</td>
</tr>
<tr>
<td>Defense Skill (juror average; 1-7)</td>
<td>4.418</td>
<td>1.134</td>
<td>251</td>
</tr>
<tr>
<td>Pros. Skill (judge; 1-7)</td>
<td>4.873</td>
<td>1.541</td>
<td>251</td>
</tr>
<tr>
<td>Pros Skill (juror average; 1-7)</td>
<td>5.203</td>
<td>1.111</td>
<td>251</td>
</tr>
<tr>
<td>Defense More Skilled (0-1)</td>
<td>0.450</td>
<td>0.499</td>
<td>251</td>
</tr>
</tbody>
</table>

Source: Authors’ tabulations from the NCSC/ICPSR Hung Juries Dataset (Hannaford-Agor et al., 2003).

category is primarily composed of blacks in D.C. and Hispanics in Maricopa, with both groups accounting for a large share in the Bronx and Los Angeles.

For nearly two thirds of the trials in our sample, attorneys enjoyed high participation in voir dire. This includes all of the trials in the Bronx, compared to 27% of trials in Los Angeles and about 60% of trials in D.C. and Maricopa. Supplemental Juror Questionnaires were available much less frequently (20% of trials), and were never available in Maricopa. Just over two thirds of the trials in our sample granted either privilege to attorneys, which is our main indicator of attorney empowerment.

Judges tended to give defense attorneys slightly higher skill ratings than prosecutors on average, while the average of juror ratings for the prosecutor for a given trial exceeded that for the defense attorney on average across the trials in our sample. Our main indicator of relative attorney skill suggests that the defense
attorney had a relative skill advantage over the prosecutor in 45% of trials.\textsuperscript{119}

4.4 Empirical Results

As has been discussed in detail, our theoretical model yields two predictions. The first is straightforward: when attorneys are empowered and the defense attorney is relatively more skilled, seated jurors will be more likely to be predisposed towards the defense. We will first present and discuss results that confirm this prediction for the trials in our dataset. We interpret this as confirmation of the basic validity of our model.

The second prediction is less amenable to empirical verification. The model suggests that skilled and empowered defense attorneys can end up manipulating the racial composition of juries in various ways in their pursuit of jurors with inclinations towards their side. The precise direction and magnitude of the effect on racial composition depends upon the mechanisms of empowerment that are operative and on the information that the attorneys for a given trial are able to learn about potential jurors, which we cannot observe. Therefore, if we observe skill and empowerment affecting the racial composition of juries in a certain direction, we cannot view this as confirmation of this prediction. Our approach will be to first test for a racial composition effect, and then to infer from this what we can about the benefits that attorneys are able to obtain from empowerment, under the assumption that our theoretical model is correct. We will also attempt

\textsuperscript{119}This is perhaps unexpectedly low, as there is a general perception of resource-constrained District Attorneys’ Offices being at a disadvantage relative to large private law firms representing defendants. Indeed, the defense attorneys in our dataset tend to be older and more experienced than their counterparts. However, over half of the trials in our sample involved a public defender (ranging from a third of trials in the Bronx to nearly three quarters in Los Angeles), and there is also a general perception of Public Defenders’ Offices competing ineffectively against private law firms for talented defense attorneys. In addition, anecdotal evidence suggests that exceptional law students are attracted to District Attorneys’ Offices out of law school because of the opportunity to gain trial experience early in their careers.
to support our inferences by examining evidence on the number of strikes exercised during voir dire.

4.4.1 Can Attorneys Retain Jurors Based on Predispositions?

Our chosen indicator of juror predispositions, as discussed in the previous section, is measured on an ordinal seven-point scale. The primary empirical strategy we pursue is to use this variable in its original form within an ordered logistic framework. The main explanatory variables of interest are our measure of attorney empowerment, our measure of relative attorney skill, and most importantly, their interaction. We also control for as many trial-level and additional juror-level characteristics as possible. Each observation is an individual juror, and standard errors are adjusted to account for intra-trial correlation. Table 4.2 presents our main results.

The first column of Table 4.2 purposely omits the interaction between attorney empowerment and relative skill. The results suggest that attorney empowerment and the presence of a more skilled defense attorney are separately associated with large and statistically significant increases in the probability that a seated juror will be predisposed towards the defense. Taking this at face value for the moment, this is an unexpected result: legal scholars and commentators typically assume that greater attorney involvement in voir dire is detrimental to the defense.

The second column of Table 4.2 reports what the theoretical model suggests is the “correct” specification, and the results are perhaps closer to

---

The table presents the ordered logistic estimated coefficients. These are marginal effects on the cumulative log-odds ratio of a juror’s response being higher on the seven-point scale than a given cut-off category. Table 4.3 below displays ordinary least squares regression results for the same specifications as Table 4.2, which provide a more straightforward interpretation of marginal effects.
Table 4.2: Prediction #1 (Ordered Logistic Regressions)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empowered</td>
<td>0.4262***</td>
<td>0.2248</td>
<td>0.1240</td>
</tr>
<tr>
<td></td>
<td>(0.1560)</td>
<td>(0.1995)</td>
<td>(0.1723)</td>
</tr>
<tr>
<td>EmpoweredXDefenseMoreSkilled</td>
<td></td>
<td>0.4654*</td>
<td>0.7072***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.2783)</td>
<td>(0.1685)</td>
</tr>
<tr>
<td>DefenseMoreSkilled</td>
<td>0.5545***</td>
<td>0.2443</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1333)</td>
<td>(0.2190)</td>
<td></td>
</tr>
<tr>
<td>EvidenceStrongerforDefense</td>
<td>0.3623</td>
<td>0.3740</td>
<td>0.3865</td>
</tr>
<tr>
<td></td>
<td>(0.2957)</td>
<td>(0.2986)</td>
<td>(0.3030)</td>
</tr>
<tr>
<td>EvidenceStrongerforProsecution</td>
<td>-0.6554***</td>
<td>-0.6554***</td>
<td>-0.6714***</td>
</tr>
<tr>
<td></td>
<td>(0.1425)</td>
<td>(0.1416)</td>
<td>(0.1417)</td>
</tr>
<tr>
<td>JurorIncome</td>
<td>0.0563</td>
<td>0.0644</td>
<td>0.0636</td>
</tr>
<tr>
<td></td>
<td>(0.0980)</td>
<td>(0.0983)</td>
<td>(0.0986)</td>
</tr>
<tr>
<td>JurorEducation</td>
<td>0.3384</td>
<td>0.3168</td>
<td>0.3212</td>
</tr>
<tr>
<td></td>
<td>(0.3118)</td>
<td>(0.3194)</td>
<td>(0.3208)</td>
</tr>
<tr>
<td>JurorRace(Non-White/White)</td>
<td>0.3228***</td>
<td>0.3335***</td>
<td>0.3310***</td>
</tr>
<tr>
<td></td>
<td>(0.0992)</td>
<td>(0.0999)</td>
<td>(0.0999)</td>
</tr>
<tr>
<td>JurorGender(Female/Male)</td>
<td>0.0402</td>
<td>0.0363</td>
<td>0.0351</td>
</tr>
<tr>
<td></td>
<td>(0.0745)</td>
<td>(0.0750)</td>
<td>(0.0751)</td>
</tr>
<tr>
<td>JurorReligiousness</td>
<td>0.0292</td>
<td>0.0115</td>
<td>0.0039</td>
</tr>
<tr>
<td></td>
<td>(0.0860)</td>
<td>(0.0865)</td>
<td>(0.0856)</td>
</tr>
<tr>
<td>JurorAge</td>
<td>-0.0017</td>
<td>-0.0019</td>
<td>-0.0018</td>
</tr>
<tr>
<td></td>
<td>(0.0033)</td>
<td>(0.0034)</td>
<td>(0.0034)</td>
</tr>
<tr>
<td>Wald $\chi^2$</td>
<td>139.55</td>
<td>142.59</td>
<td>140.33</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-3376.2</td>
<td>-3372.9</td>
<td>-3374.3</td>
</tr>
<tr>
<td># jurors</td>
<td>1846</td>
<td>1846</td>
<td>1846</td>
</tr>
<tr>
<td># trials</td>
<td>251</td>
<td>251</td>
<td>251</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors, clustered at the trial level, are in parentheses. All trials with non-missing data for at least one juror are included. Each equation includes controls for county, crime type, and defendant and victim characteristics, all of which have been suppressed in the table, along with the estimated ordered logistic cut points. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, based on two-tailed $z$-test.

The results in the second column provide qualified confirmation of the theoretical model. They suggest that when attorneys are empowered and the defense attorney is relatively more skilled, there is a significant and substantially higher probability that seated jurors will be favorably inclined towards the defense. The combination of skill and empowerment is crucial to deliver this effect
in the model: empowerment can provide both attorneys with a greater ability to
strike undesirable potential jurors, but it is the combination of this with the
greater skill of the defense attorney that sharpens the defense attorney's striking
ability and allows it to be directed more profitably.

The model also predicts, when taken very literally, that the defense
attorney's greater skill should have no effect on the probability that a seated juror
has favoritism for the defense when attorneys are not empowered. This is the case
because the model was constructed such that skill is only useful to attorneys in
utilizing the benefits that come with empowerment. The results in the second
column of Table 4.2 broadly confirm this aspect of the model in substance: the
estimated coefficient on 'Defense More Skilled' is positive but statistically
insignificant. On the other hand, it would also seem reasonable that the skill of
defense attorneys could be partially manifested in influencing juror inclinations
through arguments during trial, which the imprecisely-estimated positive
coefficient could be capturing.

The reason that the results in the second column are only a qualified
confirmation of the model lies in the construction of our measure of relative
attorney skill. The excluded category of this dummy variable represents trials in
which the prosecutor was strictly more skilled than the defense attorney. The
theoretical model is reversible according to which attorney is assumed to have the
greater skill, so the opposite prediction should hold when it is the prosecutor who
is more skilled: the probability that a seated juror is predisposed towards the
defense should be lower when attorneys are empowered and the prosecutor is
more skilled. We would therefore expect the estimated coefficient on the main
empowerment effect to be negative. Instead, the results in the second column
show that this effect is positive, though not statistically different from zero. This
suggests that prosecutors in our dataset are less able to leverage empowerment and a skill advantage in jury selection than their colleagues across the aisle, though potential explanations for this remain elusive.

The third column of Table 4.2 takes the theoretical model literally and imposes the zero effect of attorney skill when attorneys are not empowered. The effect of empowerment when the defense attorney is the more skilled becomes larger and more statistically significant, while the effect when the prosecutor is the more skilled remains positive but is closer to zero.

There are some important effects associated with the control variables that should be briefly examined. These effects are nearly identical across the three specifications. First, non-white jurors have a significantly and substantially higher likelihood of being predisposed to the defense than white jurors. This corresponds with a primitive assumption of our theoretical model and with much anecdotal and empirical evidence.

Second, when the evidence is strong for the prosecution’s case, seated jurors are much less likely to report having favored the defense. This is a sensible and encouraging result. On the other hand, when the evidence is strong for the defense’s case, seated jurors are only somewhat more likely to report a predisposition towards the defense, and this effect is not statistically significant. This result is sensible in direction, and though somewhat weak, should be considered in the context of the purposeful intention of the judicial system to avoid false convictions, which could lead many jurors to view weak or offsetting evidence (the excluded category) as being in favor of the defendant.

Perhaps the most interesting aspect of the estimated effects of strength of evidence is the context they provide for interpreting the magnitude of the effect of the interaction between empowerment and skill. The specification in the third
column suggests that the successes of skilled and empowered defense attorneys in jury selection can be substantial enough to offset the disadvantage from being faced with evidence supporting the prosecution. This is a potentially unsettling finding from the perspective of debates on the role of jury selection in the service of justice.

Controls for county, type of crime, and defendant and victim characteristics have been suppressed in the table, and for the most part are not associated with important effects. One interesting result is that, conditional on there being a victim of the alleged crime, if the victim was female, jurors were significantly less likely to report a predisposition towards the defense.

Table 4.3 presents ordinary least squares regression results corresponding to the specifications in Table 4.2. The sets of results do not differ substantively, but the latter give a better sense of magnitudes. In column (2), the estimated coefficient on the interaction term is 0.52, implying that, when the defense attorney is more skilled, attorney empowerment leads to an increase by about half a point on the seven-point favoritism scale for the average seated juror.

Alternative Measures of Attorney Empowerment

The results presented thus far make use of our main indicator of attorney empowerment, which, as discussed previously, is a dummy variable for whether either high attorney participation or a supplementary juror questionnaire were allowed. Table 4.4 repeats column (1) of Table 4.2 with several alternatives for capturing attorney empowerment. Column (4) of Table 4.4 reproduces column (1)

\footnote{The estimated ordered logistic cut-points for the specifications in Table 4.2 do not indicate a great deal of non-linearity in the underlying latent favoritism variable, lending some justification to treating the raw seven-point variable as a continuous, cardinal measure.}

\footnote{Results are also similar for random effects specifications, either GLS with the dependent variable in its original form, or probit, with the dependent variable transformed to a dummy variable for favoritism ratings of 5 or higher versus 4 or lower.}
Table 4.3: Prediction #1 (OLS Regressions)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empowered</td>
<td>0.5138***</td>
<td>0.2886</td>
<td>0.1689</td>
</tr>
<tr>
<td></td>
<td>(0.1687)</td>
<td>(0.2123)</td>
<td>(0.1831)</td>
</tr>
<tr>
<td>EmpoweredXDefenseMoreSkilled</td>
<td>0.5161*</td>
<td>0.7999***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3010)</td>
<td>(0.1829)</td>
<td></td>
</tr>
<tr>
<td>DefenseMoreSkilled</td>
<td>0.6360***</td>
<td>0.2864</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1460)</td>
<td>(0.2379)</td>
<td></td>
</tr>
<tr>
<td>EvidenceStrongerforDefense</td>
<td>0.4526</td>
<td>0.4668</td>
<td>0.4819</td>
</tr>
<tr>
<td></td>
<td>(0.3080)</td>
<td>(0.3071)</td>
<td>(0.3106)</td>
</tr>
<tr>
<td>EvidenceStrongerforProsecution</td>
<td>-0.7008***</td>
<td>-0.7017***</td>
<td>-0.7205***</td>
</tr>
<tr>
<td></td>
<td>(0.1506)</td>
<td>(0.1495)</td>
<td>(0.1489)</td>
</tr>
<tr>
<td>JurorIncome</td>
<td>0.0334</td>
<td>0.0415</td>
<td>0.0418</td>
</tr>
<tr>
<td></td>
<td>(0.1045)</td>
<td>(0.1047)</td>
<td>(0.1051)</td>
</tr>
<tr>
<td>JurorEducation</td>
<td>0.3519</td>
<td>0.3349</td>
<td>0.3382</td>
</tr>
<tr>
<td></td>
<td>(0.3024)</td>
<td>(0.3050)</td>
<td>(0.3058)</td>
</tr>
<tr>
<td>JurorRace(Non-White/White)</td>
<td>0.3775***</td>
<td>0.3803***</td>
<td>0.3801***</td>
</tr>
<tr>
<td></td>
<td>(0.1094)</td>
<td>(0.1099)</td>
<td>(0.1101)</td>
</tr>
<tr>
<td>JurorGender(Female/Male)</td>
<td>0.0691</td>
<td>0.0675</td>
<td>0.0659</td>
</tr>
<tr>
<td></td>
<td>(0.0823)</td>
<td>(0.0824)</td>
<td>(0.0826)</td>
</tr>
<tr>
<td>JurorReligiousness</td>
<td>0.0246</td>
<td>0.0099</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>(0.0943)</td>
<td>(0.0941)</td>
<td>(0.0932)</td>
</tr>
<tr>
<td>JurorAge</td>
<td>-0.0008</td>
<td>-0.0010</td>
<td>-0.0010</td>
</tr>
<tr>
<td></td>
<td>(0.0035)</td>
<td>(0.0035)</td>
<td>(0.0036)</td>
</tr>
<tr>
<td>(F)</td>
<td>7.95</td>
<td>7.67</td>
<td>7.80</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td># jurors</td>
<td>1846</td>
<td>1846</td>
<td>1846</td>
</tr>
<tr>
<td># trials</td>
<td>251</td>
<td>251</td>
<td>251</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors, clustered at the trial level, are in parentheses. All trials with non-missing data for at least one juror are included. Each equation includes a constant term and controls for county, crime type, and defendant and victim characteristics, all of which have been suppressed in the table. \(*p < 0.10, **p < 0.05, ***p < 0.01, based on two-tailed z-test.\)

The first two columns consider the two main instruments potentially available to attorneys separately. Comparing with column (4), it appears that the effect of a supplementary questionnaire is biased by the omission of the indicator for high participation. The third column presents a specification fully capturing the effects of the two instruments in any combination. The large and positive effect from high participation is again present, while the effect of having a questionnaire alone is positive but insignificant. The total net effect from having
Table 4.4: Alternative Measures of Attorney Empowerment (Ordered Logistic Regressions)

<table>
<thead>
<tr>
<th>Dep. Var. = Juror Predisposition (1 (pros.) to 7 (def.))</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HighParticipation</td>
<td>0.4186***</td>
<td>0.5175***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1611)</td>
<td>(0.1751)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questionnaire</td>
<td>-0.0221</td>
<td>0.2587</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1642)</td>
<td>(0.3005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>either</td>
<td>0.4262***</td>
<td>0.4517***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1560)</td>
<td>(0.1672)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>both</td>
<td>-0.5055</td>
<td>(0.3579)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PeremptoryStrikes(Def.)</td>
<td>-0.0544*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0287)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PeremptoryStrikes(Pros.)</td>
<td>-0.0021</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VenireSize</td>
<td>-0.0008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LengthVoirDire(hrs)</td>
<td>-0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald $\chi^2$</td>
<td>126.05</td>
<td>118.83</td>
<td>142.23</td>
<td>139.55</td>
<td>141.36</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-3376.9</td>
<td>-3385.7</td>
<td>-3374.0</td>
<td>-3376.2</td>
<td>-3246.5</td>
</tr>
<tr>
<td># of jurors</td>
<td>1846</td>
<td>1846</td>
<td>1846</td>
<td>1846</td>
<td>1775</td>
</tr>
<tr>
<td># of trials</td>
<td>251</td>
<td>251</td>
<td>251</td>
<td>251</td>
<td>239</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors, clustered at the trial level, are in parentheses. All trials with non-missing data for at least one juror are included. Each equation includes controls for attorney skill, strength of evidence, juror characteristics, county, crime type, and defendant and victim characteristics, all of which have been suppressed in the table, along with the estimated ordered logistic cut points. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, based on two-tailed z-test.

both instruments together is positive but not statistically different from zero.\(^{123}\)

The specification in column (4) – representing our preferred measure of empowerment – is equivalent to that in column (3) under two parameter restrictions: the high participation effect is equal to the questionnaire effect; and the effect from having both instruments is equal to the negative of the effect of having either individually. The null hypothesis that this composite parameter restriction holds cannot be rejected against the two-sided alternative.\(^{124}\)

The advantage of our chosen measure of empowerment is that it can

\(^{123}\) The $p$-value of the two-sided test is 0.20.

\(^{124}\) The Wald statistic with two degrees of freedom is 1.99, and the associated $p$-value of the test is 0.37.
parsimoniously yet comprehensively capture the potential that attorneys had to acquire information on jurors for any given trial. Table 4.4 at least provides no evidence of a statistical basis for preferring an alternative measure. On the other hand, the same broadly holds true for measuring empowerment with the indicator for high participation alone, though it is a less comprehensive measure.\footnote{The specification in column (3) collapses to that in column (1) under the restriction that the coefficients for ‘Questionnaire’ and ‘both’ are jointly equal to zero. The test of this null hypothesis against a two-sided alternative yields a Wald statistic with two degrees of freedom of 2.37 and a $p$-value of 0.31.} We have repeated selected specifications replacing our chosen measure of empowerment with high participation, and have found no substantive difference in results.

Finally, column (5) of Table 4.4 includes other variables related to attorney empowerment along with our baseline measure. The availability of a higher number of peremptory strikes unequivocally increases attorneys’ ability to strike additional potential jurors. However, we do not observe the potential number of peremptory strikes available, but only the number actually exercised by each side. When attorneys are able to strike the jurors they would like to for cause, they may have no need for the total number of peremptory strikes allowed them. Therefore, a lower number of peremptory strikes actually exercised could indicate a low number available and hence a low level of empowerment on the one hand, or a high level of empowerment and unexercised peremptory challenges on the other hand. In addition, judges have a high degree of discretion over the number of peremptory strikes allowed, and often increase the number based on the severity of the charge and other aspects of the case. The inclusion of these additional variables adds little explanatory power to the regression, and provides little insight in general.

**Evaluating our Measure of Attorney Skill**

As discussed in the previous section, our measure of attorney skill is based
on an index of relative skill constructed from survey questions asked of judges and jurors. We first argue that this index captures relevant aspects of actual attorney skill. Table 4.5 shows the results of linear regressions of the skill index on the few attorney characteristics for which data are available.\textsuperscript{126}

Previous legal experience of the defense attorney is associated with a greater relative skill rating. This effect is statistically significant but small. Previous legal experience of the prosecuting attorney has an offsetting negative effect, as would be expected, but this is small and not statistically significant. In addition, specific criminal experience for either attorney is not associated with an important effect on relative skill.

On the other hand, there is an overwhelming positive and significant effect on the skill index when the defense attorney is a woman. We interpret this as potentially embodying a selection effect: female defense attorneys entrusted with lead counsel duties are likely to be more highly skilled and talented than their male counterparts if they suffer statistical discrimination throughout their career.\textsuperscript{127} Thus, our chosen measure of relative attorney skill appears to be a valid indicator of at least some objective aspects of attorney skill.

We have attempted to gauge the robustness of our main results on juror predispositions to alternative treatments of attorney skill, with limited success. The primary concern with our measure of relative attorney skill, notwithstanding the results of Table 4.5, is that it is partially based on juror responses. Some jurors could have a tendency to report a predisposition towards the attorney they found to be skilled and vice versa, which would introduce simultaneity between

\textsuperscript{126}Each observation in the regressions in Table 4.5 is a trial, because all variables vary only at the trial level. As mentioned previously, attorney response rates were low compared to other parties, so the number of trials represented in this table is lower than in others.

\textsuperscript{127}See Lehmann (2011) for a discussion of this effect for blacks in large law firms.
Table 4.5: Determinants of Attorney Skill (OLS Regressions)

<table>
<thead>
<tr>
<th>Dep. Var. = Index of Attorney Skill (def. rating minus pros. rating)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>YearsPracticeDefense</td>
<td>0.0422***</td>
<td>0.0405**</td>
<td>0.0380**</td>
</tr>
<tr>
<td></td>
<td>(0.0156)</td>
<td>(0.0185)</td>
<td>(0.0175)</td>
</tr>
<tr>
<td>YearsPracticeProsec</td>
<td>-0.0117</td>
<td>-0.0143</td>
<td>-0.0112</td>
</tr>
<tr>
<td></td>
<td>(0.0261)</td>
<td>(0.0280)</td>
<td>(0.0239)</td>
</tr>
<tr>
<td>PrevCrimTrialDef</td>
<td>-0.0006</td>
<td>-0.0005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0010)</td>
<td>(0.0011)</td>
<td></td>
</tr>
<tr>
<td>PrevCrimTrialProsec</td>
<td>0.0007</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0027)</td>
<td>(0.0026)</td>
<td></td>
</tr>
<tr>
<td>AgeDefense</td>
<td>-0.0136</td>
<td>-0.0097</td>
<td>-0.0103</td>
</tr>
<tr>
<td></td>
<td>(0.0114)</td>
<td>(0.0111)</td>
<td>(0.0109)</td>
</tr>
<tr>
<td>AgeProsec</td>
<td>0.0057</td>
<td>0.0066</td>
<td>0.0089</td>
</tr>
<tr>
<td></td>
<td>(0.0185)</td>
<td>(0.0190)</td>
<td>(0.0187)</td>
</tr>
<tr>
<td>GenderDef(Female/Male)</td>
<td>0.4709**</td>
<td>0.5241**</td>
<td>0.5295**</td>
</tr>
<tr>
<td></td>
<td>(0.2006)</td>
<td>(0.2510)</td>
<td>(0.2473)</td>
</tr>
<tr>
<td>GenderProsec(Female/Male)</td>
<td>-0.1504</td>
<td>-0.0943</td>
<td>-0.1196</td>
</tr>
<tr>
<td></td>
<td>(0.1971)</td>
<td>(0.2145)</td>
<td>(0.2020)</td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>( F )</td>
<td>1.80</td>
<td>1.71</td>
<td>1.83</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.06</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td># of trials</td>
<td>174</td>
<td>152</td>
<td>152</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are in parentheses. Each equation includes a constant term, which has been suppressed in the table. In addition, the second and third columns include controls for county, crime type, and selected defendant and victim characteristics, which have also been suppressed in the table. See the text for a description of the dependent variable.

\*\( p < 0.10, \) **\( p < 0.05, \) ***\( p < 0.01, \) based on two-tailed z-test.

The same effects appearing in Table 4.5 also appear in specifications focusing on the separate ratings for the defense or prosecution by judges and jurors. In addition, some specifications examining skill ratings reported by jurors indicate that higher juror religiousness is associated with a greater likelihood of giving the prosecutor a high skill rating. Other juror characteristics – most notably race – are not significantly correlated with the skill ratings given to either attorney. Since juror predispositions are correlated with race but not religiousness in our data, this suggests that the skill ratings given by jurors are not necessarily biased by their predispositions.

Even though we are most focused on the effect of the interaction between skill and empowerment on juror leanings rather than the direct skill effect itself, this potential simultaneity problem is one that we would like to address. The issue that we have encountered in attempting to do so is that the interaction effect becomes very imprecisely estimated when the estimating sample changes.

---

128The table includes our dependent and explanatory variables. Even though we are most focused on the effect of the interaction between skill and empowerment on juror leanings rather than the direct skill effect itself, this potential simultaneity problem is one that we would like to address. The issue that we have encountered in attempting to do so is that the interaction effect becomes very imprecisely estimated when the estimating sample changes.
The results in Table 4.5 suggest a natural instrumental variables approach to diagnosing and correcting for any simultaneity between juror predispositions and our measure of attorney skill. Attorney experience and gender are correlated with the skill ratings of attorneys, but should be exogenous to the way that jurors responded to their surveys, making these variables good candidates to be used as instruments for our measure of skill in regressions explaining the predispositions of seated jurors. Two problems are encountered when attempting to implement such an IV strategy: the sample size is smaller because the number of trials for which attorney characteristics are available is lower than for other variables; and these attorney characteristics are not strong instruments for the attorney skill measure. Both of these problems lead to imprecisely-estimated coefficients, to such a degree that statistical comparisons with our main results on predispositions are not meaningful.

We have also attempted to replace our main skill measure with one based only on the skill ratings given by the judge, dropping all trials for which the attorneys received the same skill rating. This results in the loss of about half of the sample, and again, the coefficients are imprecisely estimated as a consequence. Modified Hausman tests against our main specifications indicate no statistical difference, but this is driven by the particularly wide confidence interval for the estimated interaction effect between empowerment and skill in the alternative specifications.

A less extreme variation on this is to base the indicator for a more highly skilled defense attorney on a skill index that varies by juror, with the jury component of the skill index calculated using only the responses of the other jurors from the same trial. This causes a small reduction in the sample size because some trials have just a single juror response. The results are not
statistically or qualitatively different from our main results on juror predispositions, but are less precisely estimated.

Finally, we have employed measures of absolute attorney skill in some specifications in order to broaden our focus beyond simple skill asymmetries. The results suggest that empowerment increases the likelihood that a seated juror will be predisposed to the defense to a greater degree when the defense attorney has a higher skill rating from the judge. The effects of the prosecutor’s skill rating are very imprecisely estimated, and more generally, these specifications have lower explanatory power than our main specifications. This is perhaps an indication that the attorney skill ratings, measured on a subjective seven-point scale, are more applicable for making relative comparisons within a trial.

Summary

We have discussed evidence that the more highly skilled attorney can indeed leverage empowerment in voir dire to retain jurors based on their predispositions. This result matches the main prediction of our theoretical model.

Our results have been shown to be moderately robust to different specifications and alternative measures of attorney empowerment. For attorney skill, it may perhaps be more accurate to say that we are simply not able to convincingly demonstrate robustness or non-robustness given our data.

We now proceed to examine further predictions of our model, retaining our preferred measures of attorney empowerment and skill.

4.4.2 Do Attorneys Retain Jurors Based on Race?

Table 4.6 examines the effect of attorney empowerment and skill on the probability that a seated juror is non-white. Our treatment of the two is the same as in Table 4.2: column (1) includes the main effects only; column (2) introduces
Table 4.6: Prediction #2 (Logistic Regressions)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empowered</td>
<td>0.1067</td>
<td>0.1227</td>
<td>0.1420</td>
</tr>
<tr>
<td></td>
<td>(0.1166)</td>
<td>(0.1596)</td>
<td>(0.1334)</td>
</tr>
<tr>
<td>EmpoweredXDefenseMoreSkilled</td>
<td>-0.0374</td>
<td>-0.0843</td>
<td>(0.2102)</td>
</tr>
<tr>
<td>DefenseMoreSkilled</td>
<td>-0.0715</td>
<td>-0.0468</td>
<td>(0.1033)</td>
</tr>
<tr>
<td></td>
<td>(0.1672)</td>
<td>(0.1297)</td>
<td></td>
</tr>
<tr>
<td>LA</td>
<td>-1.1702***</td>
<td>-1.1715***</td>
<td>-1.1713***</td>
</tr>
<tr>
<td></td>
<td>(0.1839)</td>
<td>(0.1843)</td>
<td>(0.1843)</td>
</tr>
<tr>
<td>Maricopa</td>
<td>-3.1249***</td>
<td>-3.1284***</td>
<td>-3.1311***</td>
</tr>
<tr>
<td></td>
<td>(0.1996)</td>
<td>(0.2030)</td>
<td>(0.2026)</td>
</tr>
<tr>
<td>DC</td>
<td>-1.4966***</td>
<td>-1.4974***</td>
<td>-1.4948***</td>
</tr>
<tr>
<td></td>
<td>(0.1689)</td>
<td>(0.1694)</td>
<td>(0.1696)</td>
</tr>
<tr>
<td>Wald χ²</td>
<td>339.13</td>
<td>345.83</td>
<td>340.73</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-1409.8</td>
<td>-1409.8</td>
<td>-1409.9</td>
</tr>
<tr>
<td># of jurors</td>
<td>2410</td>
<td>2410</td>
<td>2410</td>
</tr>
<tr>
<td># of trials</td>
<td>282</td>
<td>282</td>
<td>282</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors, clustered at the trial level, are in parentheses. All trials with non-missing data for at least one juror are included. Each equation includes a constant term and controls for crime type and defendant and victim characteristics, all of which have been suppressed in the table. *p < 0.10, **p < 0.05, ***p < 0.01, based on two-tailed z-test.

their interaction in addition; and column (3) retains the interaction effect but restricts the main skill effect to be zero. The dependent variable is a dummy variable, with whites as the excluded category. All specifications include several trial-level controls. Each observation is an individual juror, and standard errors are adjusted to account for intra-trial correlation. We have chosen the logistic functional form for the results we will focus on.

The main message of Table 4.6 is that there is no effect of attorney empowerment in voir dire on the racial composition of seated juries, regardless of attorney skill. The estimated logit coefficients for the main effects and the

---

129 Juror characteristics have been omitted from these regressions to avoid simultaneity with the dependent variable. This accounts for the larger sample size compared to Table 4.2

130 This functional form assumption does not appear to be important. Results are similar for linear specifications and for probit specifications with or without random effects. Further, multinomial logistic results do not indicate important differences for Blacks and Hispanics, the two primary non-white racial groups in the data.
interaction term are small in magnitude and statistically insignificant. Instead, the county fixed effects account for much of the explanatory power of the regressions. That is, whatever the attorneys accomplish in jury selection regarding the leanings of the jurors they retain, the racial composition of the seated jury is driven almost entirely by the demographics of the jury pool and other county-level effects.\footnote{These conclusions remain true when considering the alternative measures of attorney empowerment explored in the previous section.}

The table presents coefficients rather than marginal effects. The estimated marginal effect of attorney empowerment on the probability of a seated juror being non-white from the specification in column (2) – averaged over the estimating sample, with the defense as the more highly skilled attorney and other variables unchanged – is 1.7 percentage points, with a 95\% confidence interval of -4.3 percentage points to 7.7 percentage points. On a twelve-person jury, this amounts to the exclusion or inclusion of less than one additional non-white juror. In contrast, moving a trial from L.A. County (with a 50\% non-white population) to Bronx County (with a 90\% non-white population) would increase the estimated probability of a seated juror being non-white by 21.3 percentage points on average, or 2.5 non-white jurors on a twelve-person jury.\footnote{The change in racial composition due to empowerment, though small in a practical sense, may nonetheless be important for trial outcomes. Our findings in Chapter 3, however, do not suggest that this is the case: as reported there, we find that changes in the proportion of blacks on seated juries have a smaller effect on and explain less of the variation in convictions and acquittals per trial than measures of evidentiary strength and complexity.}

Our theoretical model does not predict a zero effect on racial composition of seated juries. Rather, it predicts an effect – potentially large – in either direction depending on how attorneys leverage empowerment. If empowerment is effective simply by allowing attorneys to find convincing grounds to have jurors dismissed for cause, a more highly skilled defense attorney will strike more white panelists than otherwise. On the other hand, if empowerment provides attorneys
with strong information regarding juror leanings (in addition to increasing striking ability or not), a more highly skilled defense attorney could choose to retain only those jurors appearing to be predisposed to the defense, regardless of race, and thus end up with a higher proportion of whites on the seated jury than would otherwise be the case. In the context of these predictions, we interpret our empirical results as indicating a mix of these various empowerment effects across the trials in our sample. That is, in at least some of the trials, attorneys were able to leverage empowerment to learn valuable information about jurors that would otherwise remain unobserved.

4.4.3 Supporting Evidence from the Number of Panelists Struck

Corollary predictions concerning the number of strikes exercised during voir dire can also be extracted from our model. As with predictions about race, these predictions about number of strikes are dependent upon the specific functions of empowerment. For example, when empowerment only increases attorney striking ability, Assumption 2 implies that more panelists will be struck in expectation, as both attorneys will have a greater striking success rate than otherwise. If, instead, empowerment only provides information on juror predispositions and the defense attorney chooses to strike white panelists exhibiting the low signal only, there will be fewer panelists struck in expectation, as no white panelists exhibiting the high signal will be struck.

Table 4.7 examines the effect of attorney empowerment and skill on the number of strikes exercised during voir dire. The same specifications as shown in Table 4.6 for race are employed, with the inclusion of additional controls for other aspects of the voir dire process. The dependent variable is the total number of strikes exercised, including all strikes for cause and the number of peremptory
Table 4.7: Number of Strikes (OLS Regressions)

<table>
<thead>
<tr>
<th></th>
<th>Dep. Var. = Total Number of Strikes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Empowered</td>
<td>0.8289</td>
</tr>
<tr>
<td></td>
<td>(1.5546)</td>
</tr>
<tr>
<td>EmpoweredXDefenseMoreSkilled</td>
<td>-3.5136</td>
</tr>
<tr>
<td></td>
<td>(3.3140)</td>
</tr>
<tr>
<td>DefenseMoreSkilled</td>
<td>-1.7173</td>
</tr>
<tr>
<td></td>
<td>(1.6996)</td>
</tr>
<tr>
<td>LA</td>
<td>-12.2622***</td>
</tr>
<tr>
<td></td>
<td>(3.8777)</td>
</tr>
<tr>
<td>Maricopa</td>
<td>-10.0843**</td>
</tr>
<tr>
<td></td>
<td>(4.1083)</td>
</tr>
<tr>
<td>DC</td>
<td>1.1096</td>
</tr>
<tr>
<td></td>
<td>(2.9243)</td>
</tr>
<tr>
<td>$F$</td>
<td>15.29</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.45</td>
</tr>
<tr>
<td># of trials</td>
<td>276</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are in parentheses. All trials with non-missing data for at least one juror are included. Each equation includes a constant term and controls for crime type, defendant and victim characteristics, and length of voir dire, all of which have been suppressed in the table. *p < 0.10, **p < 0.05, ***p < 0.01, based on two-tailed z-test.

strikes exercised by each side.\textsuperscript{133} Each observation is a trial, and the estimation procedure is ordinary least squares.

As with race, we find little significant effect of empowerment, skill, or their interaction on the number of strikes exercised during voir dire. The results in column (2) suggest that defense attorneys strike slightly fewer panelists when they have a skill advantage compared to prosecutors. However, neither the main empowerment effect (representing cases in which the prosecutor was relatively more skilled) nor the interaction with the indicator for the defense attorney being

\textsuperscript{133} Strikes for cause outnumber total peremptory strikes for many trials, but for some cases, the number of strikes for cause is reported as zero. We suspect that these are cases for which the entire venire was examined for cause out of the presence of court clerks, before the sequential exercise of peremptory challenges was begun (and also that the panel size reported for these trials refers to the size of the remaining venire after strikes for cause). We caution that the data on strikes for cause are therefore likely to be incomplete. However, results are similar when each of these three categories of strikes is examined individually, so we err on the side of using the more comprehensive dependent variable.
more skilled is statistically different from zero. Both coefficients are estimated imprecisely, so that fairly substantial effects in either direction cannot be ruled out. Most of the explanatory power of the regressions comes from the county fixed effects: judge tendencies and state guidelines generally lead to fewer potential peremptory challenges being available to attorneys in Los Angeles and Maricopa compared to the other counties.

In principle, we could attempt to better characterize how the attorneys in our dataset leveraged empowerment in voir dire by examining the model’s joint predictions concerning race and number of strikes, and comparing them with our empirical results. For example, if we had found that empowerment was associated both with a lower proportion of non-whites on seated juries and with a lower number of strikes in trials for which the defense attorney had a skill advantage, we could infer that attorneys were often in the situation like the second example mentioned above: i.e. with the information aspect of empowerment dominating, and with the high signal rare enough in the population that the defense attorney attempts to retain all but white panelists exhibiting the low signal.

Unfortunately, our present results do not shed further light on the role played by empowerment in the trials in our dataset. The small and insignificant effects on both racial composition and number of strikes could indicate varying prominence of the striking ability and information functions of empowerment across trials, along with variation in the quality of the signals of juror predispositions that skilled and empowered attorneys were able to obtain. In this sense, the results for number of strikes do provide some corroborating evidence in support of the results for racial composition. A strong effect on number of strikes

---

134Indeed, the coefficients are jointly insignificant ($p=0.51$) and not statistically different from one another ($p=0.25$).
in either direction combined with no effect on racial composition, on average for
the sample, could not be convincingly explained within our model.

4.4.4 Alternative Interpretations

We interpret our results as providing reasonable confirmation of our model,
and of the general hypothesis that attorney empowerment in voir dire can increase
the ability of attorneys both to have potential jurors removed and to learn more
about them. However, there are alternative explanations for why we might see the
interaction of empowerment and skill leading to an effect on juror favoritism but
not on juror race.

One such explanation is that attorneys may use voir dire less for striking
and learning about jurors, and more for testing and laying the groundwork for
arguments to be raised during trial. Skilled attorneys afforded this opportunity
could potentially be indifferent about the race and any other characteristics of the
jurors hearing the case, so long as those jurors could be sufficiently groomed and
prepared beginning in voir dire. If attorneys often made use of empowerment in
voir dire to this effect, one might expect to see fewer strikes for trials in which
attorneys were empowered. However, as discussed above, there are also cases for
which our model predicts that the combination of skill and empowerment will be
associated with fewer strikes than otherwise. It is therefore not clear that these
alternative hypotheses about the role of empowerment could be convincingly
tested against one another based on number of strikes alone. However, it can be
noted that column (1) of Table 4.7 shows a small and positive (though highly
insignificant) net effect of empowerment on number of strikes in our data, which
perhaps indicates that attorneys in our dataset did not always use empowerment
in voir dire simply to prepare potential jurors.
Another concern is that our measure of attorney empowerment could be capturing unobserved judge effects. For example, judges who routinely grant in-depth attorney participation in voir dire may hold subtle sympathies for defendants and defense attorneys in general, which could also be reflected in other relevant aspects of the trial, such as the instructions given to the jury, the evidence admitted, and the attorney motions allowed. Unfortunately, the data include neither judge identifiers nor even basic demographic information. We plan to address this question in the future by leveraging cross-county variation in the data to test selected comparative statics of our model. If the empirical results of this exercise correspond more with the political leanings of judges in certain counties than with the model’s predictions, such judge effects may deserve more serious consideration.

Finally, it should be mentioned that few criminal charges end up going to trial and terminating with a jury verdict.\textsuperscript{135} There are manifold factors at play determining which cases see the courtroom, from the presence of mandatory minimum sentences to chance interactions between attorneys attempting to work out a plea bargain. It is unclear how we might meaningfully assess which of many potential selection effects might be affecting our results, and to what degree. We leave a more detailed investigation of this matter for future work.

\section*{4.5 Conclusion}

This chapter is the most complete attempt to date to measure what attorneys can actually accomplish, given varying degrees of empowerment in jury selection, in terms of multiple characteristics of the seated jurors they manage to

\textsuperscript{135}The figure in 2004 was about 4\% of nearly 84,000 felony defendants, according to the Bureau of Justice Statistics.
retain. We have constructed a model of attorney behavior in which empowerment can operate through multiple channels, and in which the benefits of empowerment can depend on attorney skill. We have demonstrated that, in this model, skilled and empowered attorneys can successfully stack the jury with favorably-inclined jurors, but that this is equally true for the defense as for the prosecution, and further, that this is not necessarily accomplished by manipulating the racial composition of juries in predictable ways. Finally, we have verified the validity of our model empirically, and found that its main prediction holds in our sample of felony trials in four large and diverse counties.

Our results are surprising in relation to previous work and to received wisdom, but less so in light of the more realistically sophisticated attorney behavior that we allow for. The attorneys in our sample end up altering the racial composition of the seated jury relative to the jury pool very little when empowered: equivalent to no more than one juror of a given race in either direction on a twelve-person jury. The implication in the context of our model is that, for at least some of the trials in our dataset, empowerment in voir dire allowed attorneys to uncover valuable information about jurors and to avoid using racial stereotypes as a crutch.

An important question is how our results might generalize to other regions not represented in our dataset. History is replete with examples of black defendants being convicted by all-white juries, and with first-hand accounts of prosecutors using peremptory challenges to remove each of the few black panelists summoned for potential jury duty only to be perfunctorily sent back home. Such examples do not invalidate our model or its predictions so much as they underline the necessity of exploring its comparative statics more carefully. In some counties, blacks may have very strong predispositions towards the defense on average, while
making up a very small fraction of the population and juror roll. Other counties may be very racially diverse, but have very little variation in average predispositions across racial groups. In either of these examples, the benefit from empowerment for even the most skilled defense attorney is likely to be minimal, even if empowerment yields very detailed information about the jurors. This would especially be the case if minorities are underrepresented in jury pools and typical court procedures routinely make a large enough number of peremptory challenges available that the prosecution can easily remove any minority panelists. This situation may aptly describe many small jurisdictions, particularly in the south. It cannot be handled well within our current modeling framework. We therefore suspect that the strength of our empirical results – and perhaps the broader validity of our model – will not extend outside of large and racially diverse counties such as the ones represented in our dataset, and are leery of making any claims to the contrary without substantial further work.

Notwithstanding these concerns about generalizability, however, we do think that our present work can make a small contribution to policy discussions. Many commentators have raised concerns about jury selection practices, both from the perspective of potential unfairness to the defendant and potential discrimination against minority panelists. These commentators are virtually unanimous in their calls for the abolition of peremptory strikes, which are currently employed by the American justice system to a much greater degree than anywhere else in the world.\textsuperscript{136} We submit that peremptory challenges are something of a red herring in these discussions. Specifically, our results suggest that inequality in attorney skill may be of equal importance – where this can be interpreted broadly as inequality in the resources and support available to each side for a given trial.

\textsuperscript{136}See, for example, Hoffman (2006).
It would seem that collecting greater amounts of juror information should only aid in accomplishing the intended aim of voir dire – literally, to induce panelists to “speak the truth” about their biases and ability to be impartial. However, if one side lacks the skill or resources to interpret such information, this opens the door for opposing counsel to abuse this advantage by, far from seeking impartial jurors, seeking the most favorably-inclined jurors that can be slipped past the other side. Gobert et al. (2009) have suggested that large courthouses should employ analysts that both sides for a given trial can have equal access to in gaining assistance with jury selection. This seems like a worthwhile proposal to explore further, in that it could preserve the benefits of collecting a large amount of information from panelists, while leveling the playing field in how attorneys are able to use this information. On the other hand, if this only aided attorneys in arguing in greater detail over semantic points rather than in actually learning about the panelists, the time and resource costs of the court would perhaps need to be given greater consideration. Again, however, our results suggest that attorneys in our dataset utilized empowerment to acquire useful information rather than just to bolster striking success in at least some cases.

We have focused on juror race for the majority of the chapter, but it should be reiterated that the model only relies on the existence of any observable characteristic that is correlated with juror leanings. What constitutes observability to attorneys is not altogether clear, but virtually all courthouses collect rudimentary information on panelists’ occupations and employment. Therefore, most basic socio-demographic information should be at least partially observable. In Chapter 3, we found an especially robust relation between average jury income and acquittals, which suggests income as an obvious additional characteristic to incorporate in our analysis. More immediately, we plan to continue generalizing
our model, and to explore its comparative statics theoretically and empirically.

In a sense, the present chapter can be thought of as one step towards theoretically linking the empirical strategies of Anwar et al. (2012a, b) and our work in Chapter 3. The former study documents a correlation between the composition of *jury pools* and verdicts; and the latter documents a correlation between the composition of *seated juries* and verdicts. The present chapter is about how jury selection transforms jury pools into seated juries. Can attorneys cause the distributions by race and predispositions of people on seated juries to differ from those of people in the corresponding starting jury pools? The second step would be to go from the distributions of race and predispositions on seated juries to the verdicts that those seated juries deliver, which would entail an examination of group dynamics in jury deliberations. As an extreme example, the work of Sunstein (2000) suggests that a jury with only a few members exhibiting weak pre-deliberation leanings towards conviction could lead to “group polarization” and a unanimous decision to convict under various deliberation conditions. There are indications that group decision-making effects were at play in the results discussed in Chapter 3, and their potential to arise is also discussed by Anwar et al. (2012b) in interpreting their results. Bowers et al. (2001) provide a valuable catalogue of types of racial interaction that occurred during deliberations in their sample of capital trials. We also plan to address this question of deliberation dynamics more carefully in future work as part of our overall project on jury selection.

The interest amongst economists in jury selection and legal findings, though for the most part very recent, should perhaps not be surprising. The legal system presents a high-stakes environment in which a number of individuals are making constrained choices, sometimes with strategic considerations. Such an
environment provides a wide array of interesting questions to economists and decision scientists. Unfortunately – and unexpectedly given the volume of legal activity passing through courtrooms on a daily basis\textsuperscript{137} – data that would enable such questions to be satisfactorily addressed are for the most part lacking. The few econometric studies discussed in the introduction have all made use of customized self-collected datasets of one sort or another. The dataset that we have employed is unique and of great value in the richness of the information it contains, but the number of trials it covers leaves something to be desired. The more general point is that growth in this part of the literature is currently severely constrained by data availability, not by any dearth of interesting questions remaining to be answered.

\textsuperscript{137}Mize et al. (2007) estimate that state courts alone conduct nearly 150,000 jury trials per year, despite a steady increase in the frequency of various non-trial dispositions – itself an interesting phenomenon.
Appendix

A.1 Bandwidth

The trade-off involved with increasing the bandwidth is as follows: on the positive side, the precision of the estimate is improved; on the negative side, a bias is imparted on the estimate of the effect at the threshold by including observations further away from the threshold. As discussed by Lee and Lemieux (2010), when the relationship between the forcing variable and the outcome variable is approximately linear on both sides of the threshold, the bias concern becomes less prominent (and, therefore, the optimal bandwidth exercise less useful).

Lee and Lemieux (2010) suggest a plug-in rule-of-thumb bandwidth that we implement in order to derive the optimal bandwidths used in the text and figures. Imbens and Kalyanaraman (2012) provide a completely data-driven approach to selecting an optimal bandwidth, which we propose to implement in the future. In either case, we wish to adopt a uniform bandwidth for every month, dependent variable, and estimator (ITT or treatment effect) for a given experiment. To do so, we apply the two-stage rule-of-thumb procedure with a quartic form common to each side of the threshold repeatedly for various treatment months of a given experiment and the ITT specification with total usage and total expenditure as dependent variables. From the set of optimal bandwidth estimates thus produced, we informally choose one in the lower range to apply uniformly in the estimation of all ITTs and treatment effects for that experiment.

Figure A.1 shows the ITT on the total bill in July 2008 for the 4000 kWh experiment, with 95% confidence bounds, for bandwidths ranging from 200 kWh to 3200 kWh. The graph shows a rapid tightening of the confidence interval and
Figure A.1: Sensitivity to Bandwidth at the 4000 kWh Threshold

Note: Each point represents the intent to treat effect on the total bill in July 2008 corresponding to a given bandwidth.

relative stability in the absolute magnitude of the point estimate up to a bandwidth of about 1000 kWh; followed by a decrease in the absolute magnitude of the point estimate and moderate tightening of the confidence interval thereafter. Correspondingly, as shown in Figure 1.3, beyond a value of the forcing variable of about 1000 kWh to the right of the threshold, the relation with expenditure becomes quite non-linear, indicating, along with Figure A.1, that bias is becoming a more prominent concern than precision.

A.2 Bootstrapped Standard Errors

We use nonparametric bootstrap methods to perform statistical inference at several points in our analysis. The treatment effects for total usage and total bills are estimated in levels but reported as percent changes, the standard errors for which are calculated via bootstrap. Inference on imputed treatment effects for
on-peak and off-peak usage, and the substitution elasticities which they imply, are also retrieved via bootstrap. These statistics are calculated from estimates derived from two different datasets, making analytical solutions for standard errors difficult (if not impossible) to retrieve. Bootstrap methods allow us to circumvent these challenges, and in this section we describe the sampling method that we have used.\textsuperscript{138}

Let us first discuss the estimation of standard errors for the 2SLS treatment effects specifications, which rely solely on the billing dataset. Let \( w_i \) denote the full time series of data for household \( i \), \( w_i = (X_i, E_i, TOU_i, C_i, \tilde{X}_i) \) (corresponding to the notation in equation 1.9, where \( Y \) referred generically to either total usage (\( X \)), total expenditure (\( E \)), or the treatment indicator (\( TOU \))). We draw a bootstrap sample of size \( N \) by sampling \( w_1, \ldots, w_N \) with replacement at the household level from the subsample of the billing dataset corresponding to the optimal bandwidth restriction for a given experiment. Denoting the bootstrap sample by \( w^*_1, \ldots, w^*_N \), we calculate an estimate, \( \hat{\theta}^* \), of the vector of parameters of interest, \( \theta \), and repeat this for \( B \) separate bootstrap samples.\textsuperscript{139} Given the \( B \) bootstrap estimates, \( \hat{\theta}^*_1, \ldots, \hat{\theta}^*_B \), we calculate the bootstrap estimate of the variance-covariance matrix according to

\[
\hat{V}_{\text{Boot}}(\hat{\theta}) = \frac{1}{B - 1} \sum_{b=1}^{B} \left( \hat{\theta}^*_b - \bar{\hat{\theta}}^* \right) \left( \hat{\theta}^*_b - \bar{\hat{\theta}}^* \right)^t \tag{A.1}
\]

where \( \bar{\hat{\theta}}^* = B^{-1} \sum_{b=1}^{B} \hat{\theta}^*_b \).

For analyses that draw upon both the billing data and the load profile data, an additional layer of simulation is required. The billing data bootstrap sample is as described above, and we follow an identical protocol for the load

\textsuperscript{138}In both notation and procedure, what follows draws upon Cameron and Trivedi (2009).
\textsuperscript{139}For all of the estimates in the text, we draw \( B = 1000 \) bootstrap samples.
profile bootstrap sample. Note that we draw with replacement at the household level for the load profile samples even when the analysis is based on pooled observations, allowing us to account for within-household variation in an appropriate (conservative) way. Let $v_{it}$ correspond to an observation for load-profile household $i$ in month $t$, such that $v_{it} = (\bar{x}_{it}, X_{it}, month)$. We use the load profile data to examine the peak-to-off-peak usage ratio ($\bar{x}$) by calendar month, as discussed in more detail in the following section.

Having retrieved $B$ bootstrap samples by once again sampling with replacement, we randomly create pairs of billing and load profile bootstrap samples such that each of the samples is used once (as is appropriate for independent draws).\textsuperscript{140} Together, these allow us to calculate $B$ bootstrap estimates, $\hat{\gamma}^*_1, \ldots, \hat{\gamma}^*_B$, of the true parameter vector $\gamma$. Having now retrieved $B$ bootstrap estimates $\hat{\theta}^*_1, \ldots, \hat{\theta}^*_B$ and $\hat{\gamma}^*_1, \ldots, \hat{\gamma}^*_B$, we are able to combine individual parameters as necessary to calculate $B$ draws from the distribution of peak and off-peak usage changes and the corresponding substitution elasticities. We then calculate the second moment of the distribution of these random variables non-parametrically as in equation A.1.

A.3 Estimation of the Peak-to-Off-Peak Usage Ratio

We cannot estimate treatment effects for on-peak and off-peak usage directly because we do not observe these variables in our billing dataset. However, we can impute on-peak and off-peak usage levels for TOU household-months, and can take advantage of our load profile data to infer some information about the peak/off-peak usage patterns of non-TOU households.

We first note that we can use the structure of customers’ electric bills to

\textsuperscript{140}See Efron and Tibshirani (1993, 88-90).
impute a household’s on-peak and off-peak usage for months that it is on TOU.

When household $i$ is on TOU, its total billed amount $E$ in month $t$ is

$$E_{itT} = p_{itT}^on x_{itT}^on + p_{itT}^off x_{itT}^off + g_{itT} \tag{A.2}$$

where $T$ indicates the TOU pricing regime and $x^on$ and $x^{off}$ represent the household’s on-peak and off-peak usage respectively. That is, bills depend on a fixed fee $f$, and on on-peak and off-peak per-kWh charges of $p^on$ and $p^{off}$ respectively. Combining this with the fact that on-peak and off-peak usage must sum to the household’s observed total usage, $X$, i.e.

$$X_{its} = x_{its}^on + x_{its}^{off} \tag{A.3}$$

(for either pricing regime $s \in \{T, N\}$), gives two equations in two unknowns. This allows us to solve for on-peak and off-peak usage as functions only of variables that we observe:

$$x_{itT}^on = \frac{E_{itT} - g_{itT} - p_{itT}^{off} X_{itT}}{p_{itT}^{on} - p_{itT}^{off}} \quad \text{and} \quad x_{itT}^{off} = \frac{p_{itT}^{on} X_{itT} - g_{itT} - E_{itT}}{p_{itT}^{on} - p_{itT}^{off}}. \tag{A.4}$$

Note that this imputation is, unfortunately, impossible for non-TOU household-months, as the non-TOU rate is the same for on-peak and off-peak usage, and the non-TOU analogues to the expressions in (A.4) are hence undefined. Instead, we can estimate average peak/off-peak usage patterns for non-TOU households employing our load profile dataset.

Specifically, we first define the peak-to-off-peak usage ratio for non-TOU

---

141 We index usage and bills by $T$ because, even though a household is on only one pricing regime in a given month in reality, we wish to contemplate how a household’s behavior would change in a given month, varying the pricing regime only.
household-months as

$$\hat{x}_{itN} \equiv \frac{x_{itN}^{on}}{x_{itN}^{off}},$$  \hspace{1cm} (A.5)$$

where \(N\) indicates the non-TOU pricing regime. We then estimate this ratio for representative non-TOU households by making use of our load profile dataset. Recall that this dataset is a random sample of approximately 1,300 households, present for between 2 and 48 months, with hourly usage data. These hourly usage data can easily be aggregated to the calendar-month level by time of day, from which monthly levels of \(\hat{x}\) can be calculated for any household-month.

We mostly restrict ourselves to examining the peak-to-off-peak usage ratio for a given calendar month pooled across years and non-TOU households. One concern that this simple approach mitigates is the sparseness of data in the load profile dataset, especially in the total usage ranges we are most interested in. This sparseness is illustrated in Figure A.2, which shows that most observations in the load profile data for all months of July between 2006 and 2011 inclusive correspond to total monthly usage below 1500 kWh. This has the disadvantage of treating a calendar month in one year symmetrically with other years, which may not be warranted if average temperatures over a season vary from year to year; but it at least improves the precision of summary statistics we could calculate from these data and inferences we could draw from them. It should also be noted that the load profile households are, in general, not present in the sample for a long enough period to facilitate controlling for intra-household variation across months or years.
Figure A.2: Peak-to-Off-Peak Usage Ratio for Non-TOU Households

Note: The graph displays all non-TOU observations from the load profile dataset for the calendar month of July, pooled across the years 2006 through 2011 inclusive.

A.4 Billing Cycles

There were 17 distinct billing cycles for residential customers over the period covered by our dataset. Each billing cycle corresponds to a given day of the month (which can change by a couple of days in either direction depending on month and year, due to weekends and holidays) on which the meter is read and the billing period for customers on that billing cycle closes. For customers on billing cycle 1, the total usage and total bill data for “July 2008”, for example, correspond to usage that mostly happened in the calendar month of June; on the other hand, for customers on billing cycle 17, total usage and total bill data for “July 2008” correspond to usage that mostly happened in the calendar month of July. There is thus heterogeneity in our billing data in what “July 2008” (and every other month) refers to. This is relevant because we only have rate
information on a calendar-month basis. So the total bill in “July 2008” depends on a weighted average of the rates that were in place in the calendar month of June and those that were in place in the calendar month of July, with the appropriate weight depending on which billing cycle a household is on. We describe here how we retrieve billing cycle weights by household-month, and how we apply the weights thus retrieved to align variables observed on a calendar-month basis with variables observed on a billing-month basis.

We reconstruct the total billed amount for all non-TOU household months based on the observed rates, total usage, and the unknown weight; then solve for the weight that exactly aligns the reconstructed total billed amount with the observed total billed amount for each individual household-month. (We cannot do the same for TOU household-months because we do not observe the on-peak/off-peak breakdown of total usage. We can also not perform the calculation for months in which there was no rate change from the previous month.) A few households chronically had weights outside the sensible 0-1 range in the months for which weights could be calculated, and have been dropped completely from all analysis; a few remaining households occasionally had a month with a nonsensical weight, in which case it was just the single household-month observation that was dropped.

Finally, we calculate average billing cycle weights by billing cycle-month-year group over all household-months we could calculate the initial weights for; fill in the missing month-years (i.e. months across which there were no rate adjustments) with annual averages; then apply the appropriate billing cycle-month-year average to every corresponding non-TOU and TOU household-month. (We observe which billing cycle each household was on in September 2010, and know that households are supposed to always stay on the
same billing cycle.)

We need to account for billing cycles in the imputation of on-peak and off-peak usage for the TOU household at the threshold. We align billing-month estimates with calendar-month rates by taking a weighted average of the latter across the relevant months. The weight we use in the calculation must be the billing cycle weight for a TOU household at the threshold. This is furnished by once more applying 2SLS estimation to equations (1.11)-(1.12), this time with average billing cycle weights as the outcome variable of interest.

We estimate total expenditure based on bundled generation-inclusive rates in a similar fashion. We first impute on-peak and off-peak usage levels for TOU household-months using the method described in the previous section. We then align billing-month usage levels with calendar-month bundled rates by individual household-month using average billing cycle weights. Finally, we use the weighted rates and observed usage to estimate what the total generation-inclusive billed amount would have been had the household had the utility as supplier in addition to distributor. Expenditure levels at the threshold based on bundled rates are then estimated via the usual application of 2SLS to equations (1.11)-(1.12).
Bibliography


Curriculum Vitae

JEREMY BLAIR SMITH
171 York Street
Napanee, Ontario K7R 2Y8
Canada
Cell: (617) 314-1526
Email: jersmith@bu.edu
Web site: http://www.jeremyblairsmith.ca
Year of Birth: 1979

EDUCATION
Ph.D., Economics, Boston University, Boston MA, May 2013 (expected)
M.A., Economics, Queens University, Kingston ON, Canada, 2003
Bac. Social Sciences, Honours Economics (summa cum laude), University of Ottawa, Ottawa ON, Canada, 2002

FIELDS OF INTEREST
Primary: Environmental Economics, Law & Economics
Secondary: Labor Economics, Industrial Organization

TEACHING EXPERIENCE
Lecturer, Economic Analysis of Legal Issues, Department of Economics, Boston University, Fall 2012 and Spring 2013
Lecturer, Intermediate Macroeconomics, Department of Economics, Boston University, Spring 2011 and Spring 2012
Lecturer, Environmental Economics, Department of Economics, Boston University, Spring 2007-Fall 2011 (nine semesters)
Lecturer, Intermediate Microeconomics, Department of Economics, Boston University, Summer 2007
Teaching Fellow, Intermediate Macroeconomics (Professor Laurence J. Kotlikoff), Department of Economics, Boston University, Fall 2006
Teaching Assistant, Intermediate Microeconomics (Professor Art Stewart), Department of Economics, Queens University, Fall 2002 and Spring 2003

WORK EXPERIENCE
Economist, Centre for the Study of Living Standards (Executive Director: Andrew Sharpe), Ottawa ON, Canada, 2003-2005
Research Assistant, Centre for the Study of Living Standards, 2001-2002