

Dr. Sherod Thaxton,
Disentangling Disparity

Article

DISENTANGLING DISPARITY: EXPLORING RACIALLY DISPARATE EFFECT AND TREATMENT IN CAPITAL CHARGING

Sherod Thaxton*

One hundred and thirty years ago, in Yick Wo v. Hopkins, the U.S. Supreme Court ruled that racially discriminatory enforcement of facially-neutral laws violated defendants' equal protection rights. Since then, a voluminous body of research has documented persistent and unjustified racial disparities in charging and sentencing. Yet not a single claimant has prevailed in a race-based discriminatory prosecution action in federal court since Yick Wo. This seeming conflict—widespread evidence of racial discrimination coupled with claimants' inability to satisfy the Courts' evidentiary thresholds to prevail on the discriminatory prosecution claim—can be attributed to deep disagreements among the Supreme Court Justices over a uniform and workable evidentiary standard for social scientific evidence of discrimination. Although the Court has increasingly signaled its willingness to rely on statistical evidence to demonstrate racial discrimination, the majority of Justices have simultaneously found such evidence lacking in particular cases and failed to specify what types of evidence would be sufficient. Recently, members of the Court most skeptical of statistical evidence of discrimination have emphasized that claimants must show racial differences in outcomes are connected to racial differences in process, and not merely that there was an opportunity for discriminatory decision-making.

This article contributes to the understanding of discriminatory prosecutorial charging behavior by carefully disentangling the racial disparity into two separate components: the part that is explained by racial differences in case characteristics predictive of the charging decision (disparate effect) and the part explained by the racial differences in prosecutors' behavioral response to those characteristics (disparate treatment). By way of illustration, I apply the analytical approach to data on capital charging decisions in Georgia. I discover that between 60%-80% of the race-of-victim gap in capital charging behavior in Georgia is attributable to disparate treatment. I further show how prosecutors' differential treatment of specific case characteristics based on the victim's

* Assistant Professor of Law, UCLA School of Law. Former Staff Attorney, Capital Habeas Unit, Office of the Federal Defender for the Eastern District of California. I am grateful to Tendayi Achiume, Devon Carbado, Beth Colgan, Kristen Eichensehr, Lisa Griffin, Cheryl Harris, Jill Horwitz, Jack Katz, Jasleen Kohli, Maximo Langer, Gerald Lopez, Jyoti Nanda, Benjamin Nyblade, Jason Oh, J.J. Prescott, Richard Re, Joanna Schwartz, Rebecca Stone, Alex Wang, Noah Zatz, and participants at the Culp Colloquium at Duke Law School for their suggestions on drafts of this Article. Tori Clark and the staff of the American Journal of Criminal Law provided outstanding editorial assistance. Naturally, all remaining errors are my own.

race contributes to the overall racial disparity, thereby providing a more granular analysis of discriminatory decision-making than previously available. I conclude by discussing the legal implications of my findings in light of the Court's governing equal protection and anti-discrimination jurisprudence.

“As long as only Negroes are concerned and no whites are disturbed, great leniency will be shown in most cases. . . . The sentences for even major crimes are ordinarily reduced when the victim is a Negro.”

—Justice William J. Brennan, *McCleskey v. Kemp* (1987), quoting Gunnar Myrdal, *An American Dilemma: The Negro Problem and Modern Democracy* (1944)

Introduction.....	97
I. Conceptualizing Discrimination.....	104
II. Litigating Discrimination.....	109
A. INFERRING INTENTIONAL DISCRIMINATION FROM STATISTICAL EVIDENCE	109
B. EVALUATING STATISTICAL MODELS OF DISCRIMINATION.....	116
III. Empirical Scholarship.....	120
A. QUANTIFYING RACIAL DISCRIMINATION IN CAPITAL CHARGING.....	121
B. PERSISTENT PITFALLS OF MODELS OF RACIAL DISCRIMINATION IN CAPITAL CHARGING.....	125
1. Homogenous Effects.....	125
2. The “Black Box” of Disparities.....	128
IV. Analytical Approach.....	130
A. DISAGGREGATING THE SOURCES OF RACIAL DISPARITY.....	130
B. GEORGIA CAPITAL CHARGING DATA.....	134
V. Results.....	139
VI. Discussion and Implications.....	144
A. MUTUAL EXCLUSIVITY.....	146
B. COMMON SUPPORT.....	147
C. PREDICTIVE ACCURACY.....	147
D. POTENTIAL OMITTED VARIABLES.....	149
Conclusion.....	153

INTRODUCTION

Empirically oriented legal scholars and social scientists have developed a voluminous literature documenting racial disparities in sentencing at both the state and federal levels.¹ With very few exceptions, these studies demonstrate the persistence of racial disparities across time, place, and offense type, even after accounting for a wide range of nonracial factors purported to influence sentencing.² Of course, judges' and jurors' sentencing decisions come at the tail end of the adjudicative process, and earlier discretionary choices by legal actors—primarily prosecutors—also influence final outcomes.³ As a consequence, there has been increased emphasis on, and scrutiny of, prosecutorial decision-making because prosecutors are generally less constrained by the law—and their choices are less visible to the public—than judges and juries.⁴ The adjudicative process begins with the charging decision, and not only does research suggest that racial disparities are strongest at this stage, but also that racial disparities are not rectified during sentencing.⁵ Furthermore, studies that focus exclusively on

¹ For reviews of the extant literature, see, e.g., Todd Sorensen et al., *Race and Gender Differences Under Federal Sentencing Guidelines*, 102 AM. ECON. REV. 256 (2012) (federal non-capital sentencing); Barbara O'Brien et al., *Untangling the Role of Race in Capital Charging and Sentencing in North Carolina, 1990-2009*, 94 N.C. L. REV. 1997 (2016) (state-level capital sentencing); Kevin McNally, *Race and the Federal Death Penalty: A Nonexistent Problem Gets Worse*, 53 DEPAUL L. REV. 1615 (2004) (federal capital sentencing); Darrell Steffensmeier & Stephen Demuth, *Ethnicity and Judges' Sentencing Decisions: Hispanic-Black-White Comparisons*, 39 CRIMINOLOGY 145 (2001) (state-level non-capital sentencing).

² African American and Latino/Hispanic defendants receive more severe sentences than their Caucasian counterparts for the same criminal conduct and with similar criminal backgrounds. Sorensen et al., *supra* note 1. Defendants, irrespective of race/ethnicity, charged with committing crimes against Caucasians also receive harsher punishments than defendants charged with committing crimes against non-Caucasians. O'Brien et al., *supra* note 1.

³ I do not mean to suggest that unexplained racial disparities first emerge in the adjudicative process. In fact, there is a substantial research literature documenting racial discrimination in the investigative process. See, e.g., Civ. Rts. Div., U.S. Dep't of Just., *Pattern and Practice Police Reform Work: 1994-Present* (Gov't Printing Office 2017) (discovering widespread patterns and practices of racially biased policing); Andrew Gelman et al., *An Analysis of the New York City Police Department's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias*, 102 J. AM. STAT. ASS'N 813, 821–22 (2007) (discovering that minority group members were disproportionately stopped by police, relative to their levels of crime participation, but less likely to be arrested, suggesting that standards were more relaxed for stopping minority group members).

⁴ See, e.g., Stephanos Bibas, *Prosecutorial Regulation versus Prosecutorial Accountability*, 157 U. PA. L. REV. 959 (2009) (describing the immense, and often unreviewable, power of prosecutors); Maximo Langer, *Rethinking Plea Bargaining: The Practice and Reform of Prosecutorial Adjudication in American Criminal Procedure*, 33 AM. J. CRIM. L. 223 (2006) (distinguishing between coercive and non-coercive plea bargaining and arguing that defendants "have a moral right that prosecutors do not make plea proposals in weak cases, that prosecutors' plea proposals do not include unfair trial sentences, and that prosecutors do not overcharge.").

⁵ M. Marit Rehavj & Sonja B. Starr, *Racial Disparity in Federal Criminal Sentences*, 122 J. POL. ECON. 1320, 1343 (2014) (attributing a significant portion of racial disparities in sentencing to racial disparities in charging that are not attenuated through the adjudicative process). A recent study of policing behavior in homicide cases also suggests that racial disparities in policing are not ameliorated in the post-investigative stages. Nick Petersen, *Examining the Sources of Racial Bias in Potentially Capital Cases: A Case Study of Police and Prosecutorial Discretion*, 7 RACE & JUST. 1, 13–17 (2016) (reporting evidence of the interrelationship between racial bias in policing and prosecution for potentially capital cases).

sentencing and ignore earlier discretionary choices, which are vulnerable to racially discriminatory practices, tend to mask racial disparities.⁶

Any abuse of discretion by prosecutors creates serious cause for concern, but unjustified racial disparities in charging decisions in the capital punishment context is especially alarming. “One of the enduring arguments in Supreme Court death penalty jurisprudence is that the death penalty is ‘qualitatively different’ from all other punishments in ways that require extraordinary procedural protection against error.”⁷ And the omnipresent influence of impermissible racial considerations on the administration of capital punishment has figured prominently in the Court’s decisions. In fact, the case credited with “launching one hundred years of federalism”⁸ involved an African American defendant who, *inter alia*, challenged the legality of his death sentence (from state court) under the Due Process and Equal Protection Clauses of the U.S. Constitution because of overt racism at the pretrial, trial, and appellate stages.⁹ The vast majority of statistically sophisticated studies examining capital charging have discovered that race still exerts an impact: all else equal, African American defendants are more likely to be charged with the death penalty than Caucasian defendants *and* defendants of any race charged with killing Caucasian victims are significantly more likely to face a capital charge than defendants charged with killing non-Caucasian victims.¹⁰

⁶ See, e.g., U.S. Gov’t Accountability Office, *Death Penalty Sentencing: Research Indicates Patterns of Racial Disparities*, GGD-90-57 (Gov’t Printing Office 1990) 4 (remarking that “discretion exercised early in the process may have the effect of concealing [masking] race effects if analysis is limited only to the later stages”).

In what some scholars have labeled the “black premium,” African American arrestees—when compared to similarly situated Caucasian arrestees—are much more likely to: be charged with a felony; face a mandatory minimum sentence; denied pre-trial release; denied release on their own recognizance; serve their pre-trial detention in prisons (rather than jails); and be charged higher bond amounts when granted bail. Rehavi & Starr, *supra* note 5, at 1323–36, 1350; Cassia C. Spohn et al., *The Impact of Ethnicity and Gender of Defendants on the Decision to Reject or Dismiss Felony Charges*, 25 CRIMINOLOGY 175 (1987); John Wooldredge, *Distinguishing Race Effects on Pre-Trial Release and Sentencing Decisions*, 29 JUST. Q. 41, 53–64 (2012).

⁷ Jeffrey Abramson, *Death-is-Different Jurisprudence and the Role of the Capital Jury*, 2 OHIO ST. J. OF CRIM. L. 117, 117 (2004). See, e.g., *Furman v. Georgia*, 408 U.S. 238, 286–89 (1972) (Brennan, J., concurring) (“Death is a unique punishment...[it] is in a class by itself.”); *Ring v. Arizona*, 536 U.S. 584, 614 (2002) (Breyer, J., concurring) (noting that the Constitution “requires States to apply special procedural safeguards when they seek the death penalty”).

⁸ MARK CURRIDEN & LEROY PHILLIPS JR., CONTEMPT OF COURT: THE TURN-OF-THE-CENTURY LYNCHING THAT LAUNCHED A HUNDRED YEARS OF FEDERALISM (1999); accord Rachel F. Moran, *Race, Representation, and Remembering*, 49 UCLA L. REV. 1513, 1514 (2002).

⁹ *United States v. Shipp*, 203 U.S. 563, 572–73 (1906) (announcing the Supreme Court’s authority to review state criminal court decisions).

¹⁰ See Part III. Prosecutors routinely use a capital charge as a bargaining chip in order to compel defendants to waive their trial rights and agree to a sentence of life imprisonment, even when the prosecutor does not believe the defendant’s crime merits the death sentence or the evidence against the defendant is weak. Sherod Thaxton, *Leveraging Death*, 103 J. CRIM. L. & CRIMINOLOGY 475 (2013) (summarizing the empirical research on prosecutors’ use of the death penalty as leverage in plea negotiations); see also James S. Liebman, *The Overproduction of Death*, 100 COLUM. L. REV. 2030, 2097–98 (2000). The use of capital punishment in this fashion is associated with higher reversals of convictions, more wrongful convictions, and greater economic waste. Thaxton, *supra*; Andrew Gelman et al., *A Broken System: The Persistent Patterns of Reversals of Death Sentences in the United States*, 1 J. EMPIRICAL LEGAL STUD. 209 (2004); James S. Liebman, *Opting for Real Death Penalty Reform*, 63 OHIO ST. L. 315 (2002) [hereinafter Liebman, *Opting for Real Death Penalty Reform*].

These two troublesome processes—racially discriminatory charging *and* legal errors resulting from the overly aggressive use of capital charging—potentially contribute to the situation where non-

Despite near consensus in the scholarly literature about the persistence of racial disparities in the criminal justice system, judges, attorneys, legislators, and the general public continue to debate whether these racial disparities in criminal justice outcomes are primarily a function of differential criminal culpability (disparate effect)¹¹ or discriminatory legal decision-making (disparate treatment). Defendants have raised claims of racially discriminatory capital charging practices, referred to as selective prosecution, in federal court, as violative of their rights to equal protection under the law as guaranteed by the Fifth and Fourteenth Amendments, and have relied on statistical evidence to support their assertions.¹² The Supreme Court has uniformly rejected these claims,¹³ underscoring a very troubling fact about the Court's selective prosecution jurisprudence: the Court has not ruled in favor of a defendant raising a selective prosecution claim based on racial discrimination in over 130 years.¹⁴ In 1886, in *Yick Wo v. Hopkins*,¹⁵ the Court ruled for the first time that racially biased enforcement of a facially neutral law violated the Equal Protection Clause of the Fourteenth Amendment. *Yick Wo* was also the last time a race-based selective

Caucasian defendants who are sentenced to death for killing Caucasian victims are significantly more likely to have their cases reversed during appellate proceedings (direct and collateral review) for serious legal error compared to other defendant/victim combinations, even after taking into account a wide range of factors relevant to defendant culpability. Alberto Alesina & Eliana La Ferrara, *A Test of Racial Bias in Capital Sentencing*, 104 AM. ECON. REV. 3397, 3397 (2014).

Results from an analysis of capital charging decisions at the federal level were equally disquieting: over a twenty-year period, 80% of capital charged defendants who were either acquitted of the capital offense or found innocent were African American or Latino/Hispanic. McNally, *supra* note 1.

¹¹ "Disparate effect" is a legal term of art, and as such, it has multiple meanings. For the purposes of this article, I define disparate effect as the distributional consequences of a policy or practice; therefore, it is simply an empirical claim. As I explain below, my definition is analogous to an "endowment effect" in the economics literature, *see infra*, note 209. My definition of disparate effect is distinct from two other common understandings of the term found in constitutional and anti-discrimination law litigation and scholarship. The first meaning pertains to the adverse effect of a policy or practice that falls disproportionately on a racial group when the policy lacks substantial justification and there is an alternative to the policy or practice that would be comparably effective without creating the racial disparity (also labelled *disparate impact* or *adverse impact*). The second meaning relates to the disproportionate application of a legal sanction to members of a protected class (e.g., racial group) compared to similarly situated individuals (also referred to as *discriminatory effect/impact*). *Wayte v. United States*, 470 U.S. 598 (1985).

¹² *See* Part II.A; *Weinberger v. Wiesenfeld*, 420 U.S. 636, 638 (1975) ("This Court's approach to Fifth Amendment equal protection claims has always been precisely the same as to equal protection claims under the Fourteenth Amendment.").

¹³ *E.g.*, *McCleskey v. Kemp*, 481 U.S. 279 (1987); *United States v. Bass*, 536 U.S. 862 (2002).

¹⁴ Kristin E. Kruse, *Proving Discriminatory Intent in Selective Prosecution Challenges-An Alternative Approach to United States v. Armstrong*, 58 SMU L. REV. 1523, 1535 (2005).

¹⁵ *Yick Wo v. Hopkins*, 118 U.S. 356 (1886). The City of San Francisco enacted an ordinance requiring all laundries in wooden buildings to hold a permit issued by the city's Board of Supervisors. The Board refused to issue permits to owners of Chinese descent. As a result, *Yick Wo* and *Wo Lee*, both of Chinese descent, continued to operate laundries in wooden buildings without a permit. *Yick Wo* and *Lee* were initially fined for violating the ordinance, but ultimately imprisoned after refusing to pay the fine. They appealed their convictions on the grounds that the ordinance was enforced in a racially discriminatory manner in violation of the Equal Protection Clause. The Court ruled that, despite the impartial wording of the ordinance, its biased enforcement was unconstitutional. *Yick Wo* is, perhaps, more notable for the Court's ruling that the Equal Protection Clause applied to non-citizens than its implications for discriminatory prosecution claims. *See, e.g.*, HIROSHI MOTOMURA, *AMERICANS IN WAITING: THE LOST STORY OF IMMIGRATION AND CITIZENSHIP IN THE UNITED STATES* 63 (2007).

prosecution claim was successfully argued before the Court.¹⁶ This fact is especially mystifying given the weight of social scientific evidence of unjustified systemic racial disparities not only in prosecutorial charging decisions, but in virtually all aspects of criminal justice legal decision-making that has emerged in the aftermath of *Yick Wo*.¹⁷

The ineffectiveness of this body of research in racial discrimination litigation in the criminal context can be primarily attributed to the Court's anti-discrimination jurisprudence which has simultaneously failed to specify the type of statistical evidence necessary to support an inference of discrimination *and* increasingly suggested that, in order to be successful with statistical evidence, claimants must show how systemic racial bias leads to racial disparities, and not merely that there was an opportunity for discriminatory decision-making.¹⁸ In other words, the Court has emphasized that claimants must do a better job of demonstrating the manner in which racial discrimination influences decision-making, while at the same time neglecting to provide guidance to judges and litigants as to what kinds of circumstantial evidence would be demonstrative of a constitutional violation.¹⁹ The purpose of this article is to not only answer the Court's

¹⁶ Kruse, *supra* note 14, at 1535.

¹⁷ See Part III. Less than twenty years after the Court issued its ruling in *Yick Wo*, the first rigorous social scientific investigation of racial disparities in criminal sentencing was spearheaded by Atlanta University's *Atlanta Sociological Laboratory* under the direction of sociologist William Edward Burghardt ("W.E.B.") Du Bois. 9 SOME NOTES ON NEGRO CRIME, PARTICULARLY IN GEORGIA (William Edward Burghardt Du Bois ed., 1904). Du Bois and colleagues discovered inequities in both the length of sentences and assignment to the convict-lease system between African Americans and Caucasians convicted of criminal conduct.

Scholars began building upon Du Bois and colleagues' work beginning in the 1920s, and while the scope and methodological rigor of these studies varied considerably, a pattern pertaining to the defendant's race, the victim's race, and the interaction between them immediately emerged: (1) African American defendants received longer sentences than Caucasian defendants for similar offenses; (2) defendants, irrespective of their race, charged with crimes against Caucasian victims received more severe charges and harsher sentences; and (3) discrimination against African American defendants was most pronounced when accused of committing crimes against Caucasians—most notably for capital offenses. See, e.g., Thorsten Sellin, *The Negro Criminal: A Statistical Note*, 140 ANNALS OF THE AM. ACAD. OF POL. & SOC. SCI. 52 (1928); Guy B. Johnson, *The Negro and Crime*, 217 ANNALS OF THE AM. ACAD. OF POL. & SOC. SCI. 93 (1941); Harold Garfinkel, *Research Note on Inter- and Intra-Racial Homicides*, 27 SOC. FORCES 369 (1949). Nearly a century later, these particular racial dynamics have remained remarkably durable. See Part III.

¹⁸ *McCleskey v. Kemp*, 481 U.S. at 312 (statistical evidence must identify the source of the disparity, rather than simply indicate that a discrepancy appears correlated with race); see generally *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 357 (2011) (merely proving that the discretionary system has produced a racial or sexual disparity is insufficient for a plaintiff to prevail).

¹⁹ Chief Justice William Rehnquist, for example, explained "it should not [be] an insuperable task to prove that persons of other races [are] being treated differently [by prosecutors]," but believed the defendants' evidence showing that every single prosecution for that same crime over the past five years involved African Americans was insufficient to warrant the federal trial court's motion granting the defendants access to the prosecution's files. *United States v. Armstrong*, 517 U.S. 456, 470 (1996). The defendants sought specific information from the prosecution that would allow them to meet the Court's burden of racially differential treatment: (1) a list of all cases from the last three years in which the Government charged both cocaine and firearms offenses, (2) the identity of the race of the defendants in those cases, (3) the levels of law enforcement were involved in the investigations of those cases, and (4) explanations of the criteria for deciding to prosecute those defendants federally rather than allow the state to handle those cases. Absent the requested information, it seems highly implausible that the defendants could make the requisite showing for the underlying selective prosecution claim. Yet the evidence that the U.S. Attorney's office in question had only pursued federal charges against African Americans over a five year time span was remarkably similar to the evidence the Court found persuasive in cases involving racial discrimination in the selection of the jury venire. *Castaneda v. Partida*, 430 U.S. 482 (1977) (ruling

clarion call—that is, carefully demonstrating how racially differential treatment produces a racially disparate outcome—but also bring attention to some peculiarities of the current doctrine that have made it unduly burdensome on claimants to prevail in selective prosecution actions for well over a century.

The primary contribution of this article pertains to the *operationalization* of systemic discriminatory treatment,²⁰ subjecting prosecutorial decision-making in the capital charging process to a more granular analysis that is directly responsive to several of the Supreme Court's prior concerns about the use of statistical analyses of capital charging-and-sentencing behavior to provide evidence of racially disparate treatment. The statistical models described in this Article provide a template for the investigation of discriminatory charging dynamics in capital *and* non-capital cases. Concretely, my analytical approach carefully separates an observed racial disparity in capital charging into two components. The first component pertains to differences in the distribution of aggravating and mitigating evidence across Caucasian-victim and African American-victim cases and is analogous to a disparate effect (as defined in this Article).²¹ The second component captures the differences in the returns on that aggravating and mitigating evidence; in other words, differences in prosecutors' behavioral responses to that evidence. This latter component is a measure of discriminatory treatment. Under this analytical framework, the discriminatory treatment component does not purport to directly capture racial animus on the part of the decision-maker, although such effects may be highly probative of such animus and support an inference that it exists.²² This article is the first to apply the analytical approach to capital charging decisions, and decision-making in the capital punishment process, more generally.

Prior research has yet to sufficiently disentangle the sources of racial differences in capital charging at a descriptive level, even though such differences have been explored by legal scholars and social scientists for more than 70 years.²³ This shortcoming may partly stem from the fact that

that evidence of the gross underrepresentation of Mexican Americans on the grand jury that convicted the defendant was unconstitutional).

²⁰ Operationalization is “the transformation of an abstract, theoretical concept into something concrete, observable, and measurable in an empirical research project.” OXFORD DICTIONARY OF SOCIOLOGY 464 (John Scott & Gordon Marshall eds., 3d ed. 2005). This entails the development of specific research procedures that will result in empirical observations representing the previous defined concepts. See Part IV.

²¹ See *supra* note 11.

²² Daniel R. Taber et al., *Oaxaca-Blinder Decomposition of Disparities in Adolescent Obesity: Deconstructing Both Race and Gender Differences*, 24 *OBSIDITY* 719, 725 (2016) (explaining that the analytical framework I employ “explores potential [causal] mechanisms in more detail than conventional analysis”).

The magnitude of the effect strengthens an inference of a causal link between the race/ethnicity and the charging decision, irrespective of racial animus, after accounting for other plausible explanations. *Washington v. Davis*, 426 U.S. 229, 254 (1976) (Stevens, J., concurring) (explaining that the distinction between disparate effect and disparate treatment may be immaterial depending on the size of the racial disparity). See also Part II.B.

²³ See Part III.

scholars have been primarily concerned with measuring overall differences in criminal justice outcomes between racial groups that remain after taking into account a wide range of legally relevant variables.²⁴ Under this approach, racial discrimination is said to exist because no other valid explanation accounts for the observed differences.²⁵ These studies have been helpful in highlighting the fact that purely legal justifications fail to explain why, in the aggregate, members of certain groups are routinely subject to harsher punishments than others, net of the actual social harm they cause. Yet judges, lawyers, legislators, and scholars still lack an understanding of how prosecutors differentially assess legally relevant (and legally suspect) factors across different racial groups—that is, how prosecutors’ evaluations of seemingly objective criteria may *shift* based upon race. Put differently, this existing scholarship has failed to inform the legal community about the potential ways in which race modifies the impact of legally relevant (and legally suspect) factors on legal behavior, independent of the distribution of these characteristics across the various racial groups.²⁶ My analytical approach provides a more nuanced understanding of racially disparate treatment, which is especially necessary in light of the Supreme Court’s sparse case law that has failed to articulate a uniform and workable evidentiary standard for statistical evidence of discrimination.²⁷

Part I explores the differing conceptualizations of race-based discrimination present in the U.S. Supreme Court’s constitutional and statutory anti-discrimination jurisprudence, describing and evaluating the rationales for these differing conceptions—both theoretical and practical—as well as the critiques of those rationales.²⁸ The Court has blurred the line between these seemingly opposing notions of discrimination, raising important questions about the appropriateness of various types of evidence in particular contexts, the standards governing its applications, and the Court’s competency in assessing such evidence. Part II describes both the

²⁴ Todd E. Elder et al., *Unexplained Gaps and Oaxaca–Blinder Decompositions*, 17 LABOUR ECON. 284, 285 (2010).

²⁵ DAVID C. BALDUS ET AL., *EQUAL JUSTICE AND THE DEATH PENALTY: A LEGAL AND EMPIRICAL ANALYSIS* (1990).

²⁶ As noted earlier, merely relying on statistics of systemic disparities without explaining the story that the statistical representation is telling appears unlikely to be sufficient for a successful claim to courts, or even a compelling argument to legislators. Michael Selmi, *Theorizing Systemic Disparate Treatment Law: After Wal-Mart v. Dukes*, 32 BERKELEY J. EMP. & LAB. L. 477, 477 (2011) (discussing the Supreme Court case, *Wal-Mart v. Dukes*, and the heightened evidentiary standard required by the Court compared to prior systemic disparate treatment cases).

²⁷ *McCleskey v. Kemp*, 481 U.S. 279, 349 (1987) (Blackmun, J., dissenting) (“In analyzing an equal protection claim, a court must first determine the nature of the claim and the responsibilities of the state actors involved to determine what showing is required for the establishment of a prima facie case.”); accord Kruse, *supra* note 14 (discussing case law pertaining to discriminatory prosecution claims).

²⁸ Analyzing the constitutional and statutory frameworks, collectively, is warranted because the Court has explained that its “cases discussing constitutional principles provide helpful guidance in [the] statutory context” when disparate treatment is alleged. *Ricci v. DeStefano*, 557 U.S. 557, 582 (2009) (discussing petitioners’ statutory claim under both the disparate-treatment prohibition of Title VII and the Equal Protection Clause of the Fourteenth Amendment). See also Civ. Rts. Div., U.S. Dep’t of Just., *Title VI Legal Manual* (2017) 3 (noting that the elements of a statutory claim of discriminatory intent “derive from and are similar to the analysis of cases decided under the Fourteenth Amendment’s Equal Protection Clause”); Cheryl I. Harris, *Limiting Equality: The Divergence and Convergence of Title VII and Equal Protection*, 2014 U. CHI. LEGAL F. 95, 104 (2014) (describing the convergence of the statutory and constitutional interpretations of disparate treatment).

Court's embrace and skepticism of statistical evidence of intentional discrimination over the past forty years. The Court's sharply divided opinions and uneven approach to statistical evidence has failed to provide workable standards for lower courts to apply. Part III discusses and assesses the empirical social scientific literature on capital charging dynamics over the past quarter-century. This literature has almost unequivocally identified racial disparities in charging decisions based on the victim's race and the combination of the defendant's and victim's race, but the analytical frameworks utilized in these studies have impaired the ability of analysts to ask and answer questions that now appear to be of central interest to courts and legislators—namely, how are racial differences in outcomes *connected* to racial differences in process? I explain, both mathematically and in plain English, how prior studies have measured racial discrimination, their specific findings, and why their methodologies prevent addressing more fundamental questions that often lie at the heart of courts' inquiries. Part IV presents a set of statistical tools—again, both mathematically and in plain English—capable of disentangling disparate effect from disparate treatment in capital charging.²⁹ After discussing the statistical model, I describe an originally compiled dataset of capital charging decisions from Georgia over an eight-year period to which I apply the aforementioned analytical approach. Part V explains the results of the statistical analyses. My findings make it unequivocally clear that race still very much matters for capital charging decisions. I find that 60%-80% of the race-of-victim gap in capital charging behavior in Georgia is attributable to disparate treatment. In addition to the overall disparate treatment effect, I demonstrate how much the racially differential treatment of *specific case characteristics* contributes to the race-of-victim gap in capital charging. This aspect of my analysis demonstrates how unjustified racial differences in *process* directly contribute to racial differences in *outcomes*, and is thereby directly responsive to several Supreme Court Justices' heightened evidentiary standard for statistical evidence of discrimination. Part VI discusses the legal implications of my findings for discriminatory prosecution claims and examines the durability of the results in the presence of potential uncertainty about the underlying statistical model and measurement of key variables.

²⁹ I adapt methodological insights from sociology and labor economics to explain which factors account for the race-of-victim gap in death-noticing among death-eligible defendants. These models have been used for several decades to examine race- and gender-based discrimination in hiring, wages, and promotion. Evelyn M. Kitagawa, *Components of a Difference Between Two Rates*, 50 J. AM. STAT. ASS'N 1168 (1955) (sociology); Otis Dudley Duncan, *Inheritance of Poverty or Inheritance of Race*, in ON UNDERSTANDING POVERTY: PERSP. FROM THE SOC. SCI. 85 (Daniel P. Moynihan ed., 1969) (sociology); Alan S. Blinder, *Wage Discrimination: Reduced Form and Structural Estimates*, 8 J. HUM. RESOURCES 436 (1973) (economics); Ronald Oaxaca, *Male-Female Wage Differentials in Urban Labor Markets*, 14 INT'L ECON. REV. 696 (1973) (economics). More recently, this approach has been applied to racial disparities in criminal justice outcomes such as prison/drug treatment commitments and sentence length. John MacDonald et al., *Decomposing Racial Disparities in Prison and Drug Treatment Commitments for Criminal Offenders in California*, 43 J. LEGAL STUD. 155 (2014); Todd Andrew Sorensen et al., *Do You Receive a Lighter Prison Sentence Because You Are a Woman or a White? An Economic Analysis of the Federal Criminal Sentencing Guidelines*, 14 B.E. J. OF ECON. ANALYSIS & POL'Y 1 (2013).

I. CONCEPTUALIZING DISCRIMINATION

There is disagreement, among *both* jurists and social scientists, over the centrality of intentional bias in explanations of racial discrimination,³⁰ although the debate appears to be most contentious in the legal context rather than in the scientific one.³¹ According to a sizable number of judges and legal analysts, discrimination results from actions intentionally designed to favor or disfavor another individual (or collection of individuals) because of race.³² Well known examples in the adjudicative context are the *de jure*³³ or *de facto*³⁴ exclusions of otherwise eligible African Americans and Latinos/Hispanics from serving on juries. This view of discrimination, commonly referred to as discriminatory intent/motive, focuses on the racial animus residing in the decision-maker(s).³⁵ Discriminatory intent is understood to imply more than a mere awareness of the distributive consequences that correlate with race/ethnicity—it requires that race/ethnicity is a motivating factor.³⁶ A detailed legal analysis of the discriminatory intent doctrine in the context of prosecutorial decision-making is provided elsewhere,³⁷ but it suffices to say that the primary

³⁰ Clearly, anti-discrimination law, both constitutional and statutory, extends to more classifications than race. Although there is considerable overlap between the constitutional and statutory classifications, the definitions are not completely congruous. This article focuses, primarily, on race-based discrimination, but many (if not all) of the arguments would be applicable to other classifications.

³¹ See, e.g., Orley Ashenfelter & Ronald Oaxaca, *The Economics of Discrimination: Economists Enter the Courtroom*, 77 AM. ECON. REV. 321, 322 (1987) (noting that discriminatory motives are of central importance to many jurists, but motivations are irrelevant to determining the existence of discrimination for most economists); Barbara F. Reskin, *Including Mechanisms in Our Models of Ascriptive Inequality*, 68 AM. SOC. REV. 1, 4 (2003) (describing several fundamental limitations of motive-based explanations by sociologists for racial and gender inequality).

³² See *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1988) (emphasizing that intentional consideration of race, whether for malicious or benign motives, is subject to the most careful judicial scrutiny); George Rutherglen, *Disparate Impact, Discrimination, and the Essentially Contested Concept of Equality*, 74 FORDHAM L. REV. 2313, 2313 (2005) (remarking that dispute over whether purposeful discrimination is necessary to establish a claim of racial discrimination turns on what individuals believe anti-discrimination law is meant to achieve).

³³ *Strauder v. West Virginia*, 100 U.S. 303 (1880); *Neal v. Delaware*, 103 U.S. 370 (1881).

³⁴ *Norris v. Alabama*, 294 U.S. 587 (1935) (de facto exclusion of African Americans from the venire panel); *Batson v. Kentucky*, 476 U.S. 79 (1986) (racially discriminatory use of peremptory challenges to remove African American jurors); *Edmonson v. Leesville Concrete Co.*, 500 U.S. 614 (1991) (racially discriminatory use of peremptory strikes in civil cases); *Hernandez v. New York*, 500 U.S. 352 (1991) (racially discriminatory use of peremptory challenges to remove Hispanic/Latino jurors); *Castaneda v. Partida*, 430 U.S. 482 (1977) (de facto exclusion of Hispanics/Latinos from the venire panel). This logic was extended to defense counsel's use of racially motivated peremptory challenges in *Georgia v. McCollum*, 505 U.S. 42 (1992).

³⁵ *Washington v. Davis*, 426 U.S. 229 (1976). Courts and scholars have used the terms discriminatory, disparate, and adverse interchangeably. They have also used the terms intent and purpose interchangeably. See *supra* note 11.

³⁶ See *Pers. Adm'r of Mass. v. Feeney*, 442 U.S. 256, 279 (1979) (explaining that the Equal Protection Clause is violated only when laws are passed because of, not merely in spite of, their adverse effects upon an identifiable group); *accord* *Wayte v. United States*, 470 U.S. 598, 610 (1985).

Justice Clarence Thomas, for example, has recently called for a repudiation of the view that Congress intended to authorize claims of racial discrimination not based on intentional racial animus when it enacted Title VII of the Civil Rights Acts of 1964. *Tex. Dep't of Hous. & Cmty. Affairs v. Inclusive Communities Project, Inc.*, 135 S. Ct. 2411 (2015) (Thomas, J., dissenting).

³⁷ See Part II.

justification articulated for this doctrine is its utility as a limiting principle.³⁸ According to this perspective, because much government action harbors some risk of discrimination, it may be unmanageable to compensate for all such risks.³⁹ In adopting the discriminatory intent conceptualization of discrimination in certain contexts, the U.S. Supreme Court announced several reasons that “the invidious quality of the law claimed to be racially discriminatory must ultimately be traced to racially discriminatory purpose.”⁴⁰ The concerns most significant to the Court appeared to be institutional—namely, separation of powers and federalism. The Court explained that, because many facially neutral policies impact vulnerable racial groups, the evidence of *scienter* is required in order to avoid improperly expanding the scope of the judiciary’s power at the expense of Congress and state legislatures.⁴¹ The *scienter* requirement invokes both process (intent) and outcome (effect) in the determination of whether constitutionally or statutorily impermissible discrimination had occurred,⁴² and therefore makes the decision-maker’s purpose to discriminate the fulcrum of the inquiry.⁴³ In Justice Antonin Scalia’s view, for example, the magnitude of an unjustified racial disparity is irrelevant to its (un)constitutionality when the cause of the racial disparity is “the unconscious operation of irrational sympathies and antipathies[.]”⁴⁴

Yet a growing number of jurists, legal scholars, and social scientists have underscored the limited relevance of racial animus in explaining discrimination, and instead have espoused the social-scientific view: the presence of unexplained/unjustified differences in outcomes in aggregate data as evidence of discrimination.⁴⁵ While acknowledging that legal and

³⁸ *Davis*, 426 U.S. at 239 (“The central purpose of the Equal Protection Clause of the Fourteenth Amendment is the prevention of official conduct discriminating on the basis of race. [...] [O]ur cases have not embraced the proposition that a law or other official act, without regard to whether it reflects a racially discriminatory purpose, is unconstitutional solely because it has a racially disproportionate impact.”).

³⁹ Edward K. Cheng, *Constitutional Risks to Equal Protection in the Criminal Justice System* (Note), 114 HARV. L. REV. 2098, 2102 (2001) (recognizing traditional equal protection doctrine focuses on particularized harms, but neglects subtler systemic risks).

⁴⁰ *Davis*, 426 U.S. at 240.

⁴¹ *Id.* at 248.

⁴² Some statutory-based causes of action do not require proof of disparate treatment, only disparate impact. The statutory standard requiring proof of discriminatory treatment is significantly more demanding. *Id.* at 239; Rutherglen, *supra* note 32, at 2313–23.

⁴³ *Davis*, 426 U.S. at 240.

⁴⁴ Memorandum from Antonin Scalia, To the Conference Re: No. 84-6811, *McCleskey v. Kemp* [Thurgood Marshall Papers] (Jan. 6, 1987); *but see* *McCleskey v. Kemp*, 481 U.S. 279, 292 (1987) (Powell, J.) (noting that the magnitude of the racial disparity is an important component of proof of a constitutional violation under the Equal Protection Clause), *accord id.* at 352–53 (Blackmun, J., dissenting); *Washington v. Davis*, 426 U.S. at 254 (Stevens, J., concurring) (same).

⁴⁵ *United States v. Tuitt*, 68 F. Supp. 2d 4, 10 (D. Mass. 1999) (noting that holding defendants to actual knowledge of a discriminatory choice on the part of a prosecutor would make the equal protection standard for discovery and the underlying selective prosecution claim impossible to satisfy); Ashenfelter & Oaxaca, *supra* note 31, at 322 (most economists believe that evidence of discriminatory motives is irrelevant to determining the existence of discrimination). *See also* *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338 (2011) (Ginsburg, J., dissenting) (explaining that gender discrimination was the only plausible explanation for gender disparities in pay and promotion after the statistical models took into account a long list factors relevant to the discretionary process).

scientific definitions of proof are not the same, these scholars highlight that “the inevitable progress of scientific research raises important questions about the role of scientific advancements in the evolution of legal standards and doctrines.”⁴⁶ Legal scholar Noah Zatz, for example, has argued that it is inappropriate to focus on individualized, nonstatistical evidence of discrimination because “causal processes are typically too complex and the evidentiary uncertainties too great to show persuasively why any one person’s . . . race played a significant role somewhere along the way.”⁴⁷ Another commentator characterized the Court’s equal protection jurisprudence as endorsing a “cramped view of constitutional harm [that] forces courts to examine only individual cases, which cannot reveal or redress patterns of racial discrimination [because] considered in isolation, nearly all decisions can be rationalized using permissible explanations. . . . It is only when these decisions are considered in the aggregate that patterns may emerge that indicate the presence of impermissible discrimination.”⁴⁸ Social scientists have offered similar critiques. Sociologist Barbara Reskin, emphasized that “theories about actors’ motives guide the search for the explanation [of race and gender disparities] . . . [however,] the product of this approach is not explanation, but never-ending and unprofitable debate over the role of unobserved motives.”⁴⁹ In other words, the focus should shift from uncovering evidence of discriminatory purpose to carefully assessing whether alternative explanations (i.e., rival hypotheses) explain the observed racial disparity.⁵⁰ This framework does not presuppose scienter, but retains the requirement of a causal connection between race and racial disparity.⁵¹ This causal attribution tells us that something is to be expected; however, it is silent as to *why* something occurred.⁵² The best that can be expected is a careful process of elimination in which circumstantial evidence is used to remove the usual non-discriminatory reasons for the observed disparity, leaving the inference that the real reason was

⁴⁶ ANGELO N. ANCHETA, *SCIENTIFIC EVIDENCE AND EQUAL PROTECTION OF THE LAW* 14 (2006). See, e.g., *Roper v. Simmons*, 543 U.S. 551 (2005) (ruling the death penalty for juveniles unconstitutional, in part, because of growing scientific evidence of juvenile’s cognitive limitations vis-a-vis adults); *Graham v. Florida*, 560 U.S. 48 (2010) (abolishing life without the possibility of parole for juveniles for similar reasons as in *Roper v. Simmons*).

⁴⁷ Noah Zatz, *Disparate Impact and the Unity of Equality Law*, 96 B.U. L. REV. (forthcoming, 2017); accord Devah Pager & Bruce Western, *Identifying Discrimination at Work: The Use of Field Experiments*, 68 J. SOC. ISSUES 221, 230 (2012) (“[H]iring decisions are influenced by a complex range of factors, conscious racial attitudes being only one.”).

⁴⁸ Cheng, *supra* note 39, at 2103-04; accord Selmi, *supra* note 26 (aggregated statistics might reveal patterns that would not be evident by focusing on individual cases).

⁴⁹ Reskin, *supra* note 31, at 15.

⁵⁰ David C. Baldus & James W.L. Cole, *Quantitative Proof of Intentional Discrimination*, 1 EVALUATION Q. 51, 56-77 (1977); See also Zatz, *supra* note 47 (“Inferring disparate treatment from the observed disparity requires eliminating [] alternative explanations.”).

⁵¹ Sheila R. Foster, *Causation in Antidiscrimination Law: Beyond Intent Versus Impact*, 41 HOUS. L. REV. 1469, 1470 (2004) (arguing that a causal connection between race and the outcome of interest has always animated antidiscrimination law); RICHARD A. BERK, *REGRESSION ANALYSIS: A CONSTRUCTIVE CRITIQUE* 211 (2003) (noting that a causal link between race and legal decision-making does not require racial animus on the part of the decision-maker).

⁵² ERNEST NAGEL, *THE STRUCTURE OF SCIENCE: PROBLEMS IN THE LOGIC OF SCIENTIFIC EXPLANATION* 26-27 (1979).

discrimination.⁵³ In the words of linguist Benjamin Whorf, “the WHY of understanding may remain for a long time mysterious but the HOW . . . of understanding . . . is discoverable.”⁵⁴

For the lawyers and scholars subscribing to the social-scientific view of discrimination, the misalignment between the function equal protection/anti-discrimination law purportedly serves⁵⁵ and the extant standards of proof for these causes of action becomes especially apparent when examining the evidence from social scientists’ field experiments on employment discrimination. Despite strong evidence of differential treatment, employers remain adamant that race does not affect their decision to hire and maintain that they simply select the best available candidate.⁵⁶ Yet when these same employers are “asked to step back from their own hiring process to think about race differences more generally, [they are] surprisingly willing to express strong opinions about the characteristics and attributes they perceive among different groups of workers.”⁵⁷ The majority of employers, when “considering Black men independent of their own workplace, characterize this group according to three common tropes: as lazy or having a poor work ethic; threatening or criminal; or possessing an inappropriate style of demeanor.”⁵⁸ Even in situations when “employers seem genuinely interested in evaluating the qualifications of a given candidate [their] evaluations themselves appear to be influenced by race [because they] perceive real-skill or experience differences among applicants despite the fact that the [applicants] resumes were designed to convey identical qualifications.”⁵⁹ More flexible, inclusive standards are used to evaluate Caucasian applicants than in the case of minority applicants, and this “suggests that even the evaluation of ‘objective’ information can be affected by underlying racial considerations.”⁶⁰ This “shifting standards” phenomenon is “less consistent with a model of traditional prejudice than with a more contingent and subtle conceptualization of racial attitudes”⁶¹ that is attributable to “a high level of

⁵³ Int’l Bhd. of Teamsters v. United States, 431 U.S. 324, 358 fn.44 (1977); NAT’L RESEARCH COUNCIL, MEASURING RACIAL DISCRIMINATION 141–42 (Rebecca M. Blank et al. eds., 2004).

⁵⁴ Benjamin Lee Whorf, *Languages and Logic*, in LANGUAGE, THOUGHT, & REALITY: SELECTED WRITINGS 233, 239 (John B. Carroll ed., 1964) (capitalization in original).

Political scientists Donald Green, Shang Ha, and John Bullock underscored that “even when causal relationships are firmly established, demonstrating the mediating pathways is far more difficult—practically and conceptually—than is usually supposed. . . . [T]he impatience often express[ed] with studies that fail to explain why an effect obtains [is unwarranted]. . . . Just as it took more than a century to discover why limes cure scurvy, it may take decades to figure out the mechanisms that account for the causal relationships observed in social science.” Donald P. Green et al., *Enough Already about “Black Box” Experiments: Studying Mediation Is More Difficult than Most Scholars Suppose*, 628 ANNALS OF THE AM. ACAD. OF POL. & SOC. SCI. 200, 202 (2010).

⁵⁵ Rutherglen, *supra* note 32, at 2313.

⁵⁶ Pager & Western, *supra* note 47, at 229.

⁵⁷ *Id.*

⁵⁸ *Id.*

⁵⁹ *Id.* at 230.

⁶⁰ *Id.*

⁶¹ *Id.* at 231.

generalized anxiety or discomfort with Blacks than can shape decision-making.”⁶²

Notwithstanding this growing evidence of the evolving character of racial bias in modern society,⁶³ the racial animus conception of legally actionable racial discrimination remains the dominant view in constitutional law, as well as much of statutory anti-discrimination law;⁶⁴ however, the social-scientific view of discrimination has made considerable headway in courts, proceeding through an “accretion of decisions that have placed more and more reliance on [statistical] methods in the determination of whether there is evidence of discrimination.”⁶⁵ Nearly forty years ago, in *International Brotherhood of Teamsters v. United States*, the Court famously remarked, “[O]ur cases make it unmistakably clear that statistical analyses have served and will continue to serve an important role in cases in which the existence of discrimination is a disputed issue.”⁶⁶ As a corollary, the Courts’ increasing willingness to consider statistical evidence to infer intentional discrimination has blurred the lines between these seemingly opposing schools of thought—racial animus versus causation—raising important questions about the appropriateness of social scientific evidence in particular contexts, the standards governing its applications, and the Court’s competency in assessing such evidence.⁶⁷ The next section provides a brief discussion of the Court’s seemingly tenuous embrace of statistical evidence of discrimination cases and the challenges litigants have encountered when presenting such evidence to the Court, particularly in the criminal context.

⁶² *Id.* I argue that the shifting standards phenomena is also present in the capital charging context, and I employ an analytical approach capable of quantifying the degree of this race-based differential assessment. See *infra* Parts III.B and IV.A. See also, Mona Lynch & Craig Haney, *Discrimination and Instructional Comprehension: Guided Discretion, Racial Bias, and the Death Penalty*, 24 LAW & HUM. BEHAV. 337, 351-53 (2000) (reporting that Caucasian jurors were significantly more likely to undervalue, disregard, and even improperly use mitigation evidence in cases involving African American defendants as opposed to Caucasian defendants when imposing a death sentence); Joseph Rand, *The Demeanor Gap: Race, Lie Detection, and the Jury* 33 CONN. L. REV. 1, 5 (2000) (arguing that jurors are more likely to be distrusting of witnesses of another race).

⁶³ See generally EDUARDO BONILLA-SILVA, *RACISM WITHOUT RACISTS: COLOR-BLIND RACISM AND THE PERSISTENCE OF RACIAL INEQUALITY IN THE UNITED STATES* 8 (2003) (arguing that overt resentment or hostility towards racial minorities is largely irrelevant to racially discriminatory behavior in the modern era).

⁶⁴ *Tex. Dep’t of Hous. & Cmty. Affairs v. Inclusive Communities Project, Inc.*, 135 S. Ct. 2411 (2015) (recognizing important limits on causes of action resulting from alleged non-intentional discrimination in order to guard against abuse).

⁶⁵ *Ashenfelter & Oaxaca*, *supra* note 31, at 322; Douglas Laycock, *Statistical Proof and Theories of Discrimination*, 49 LAW & CONTEMP. PROBS. 97, 99 (1986) (“When properly used, multiple regression can measure the impact of all factors suspected to contribute to differences in employment history, and can show how much of the difference is due to each cause.”).

⁶⁶ *Int’l Bhd. of Teamsters v. United States*, 431 U.S. 324, 339 (1977) (quoting *Mayor of Philadelphia v. Educational Equality League*, 415 U.S. 605, 620 [1974]) (internal quotation marks omitted); see, e.g., Kruse, *supra* note 14, at 1540 (“The use of statistics in proving discrimination has a long history, and dates back to the 1970s in employment discrimination cases.”).

⁶⁷ See *supra* note 84.

II. LITIGATING DISCRIMINATION

A. *Inferring Intentional Discrimination from Statistical Evidence*

Writing for the majority in *Washington v. Davis*, Justice White explained that “invidious discriminatory purpose may often be inferred from the totality of relevant facts . . . [and] discriminatory impact . . . may for all practical purposes demonstrate unconstitutionality because in various circumstances the discrimination is very difficult to explain on nonracial grounds.”⁶⁸ Early cases that considered statistical evidence of discrimination, however, “neither offered sophisticated statistical analyses or a deep discussion of the theory for why statistics can prove intent.”⁶⁹ Evidence of discriminatory treatment consisted of differences in raw percentages and whether the magnitude was substantial.⁷⁰ For example, in *Castaneda v. Partida*,⁷¹ a criminal defendant alleged systematic exclusion of Latinos from the venire panel and provided statistics of their serious underrepresentation over an extended period of time.⁷² The Court deemed that the raw statistics, coupled with a selection process susceptible to abuse, were sufficient to support a *prima facie* case of intentional discrimination that violated the Equal Protection Clause. Over the next decade, the Court became receptive to more complex and sophisticated statistical methods to establish evidence of discrimination—namely the popular statistical technique of multiple regression modeling.⁷³ Multiple regression is capable of simultaneously measuring the impact of all factors suspected to contribute to group differences in an outcome.⁷⁴ Although the statistical models presented to the Court failed to include all plausible variables that could plausibly account for the observed racial and gender disparities,⁷⁵ the Court repeatedly reasoned that such models are still probative and capable of proving a plaintiff’s case.⁷⁶ Because the governing standard of proof for discrimination claims is *preponderance of evidence*, the Court has explained that a statistical model could permit a court to “fairly [] conclude that it [was] more likely than not that impermissible discrimination [existed] [and] the plaintiff [was] entitled to prevail.”⁷⁷

⁶⁸ *Washington v. Davis*, 426 U.S. 229, 242 (1976).

⁶⁹ *Selmi*, *supra* note 26, at 487.

⁷⁰ *Id.*

⁷¹ *Castaneda v. Partida*, 430 U.S. 482, 482 (1977).

⁷² The applicable jurisdiction was 79% Latino, yet the venire panels during the time in which the grand jury that indicted Mr. Partida were only 45% Latino (and only 39% Latino over an eleven-year period).

⁷³ *Bazemore v. Friday*, 478 U.S. 385, 400 (1986).

⁷⁴ Barbara A. Norris, *Multiple Regression Analysis in Title VII Cases: A Structural Approach to Attacks of “Missing Factors” and “Pre-Act Discrimination,”* 49 *LAW & CONTEMP. PROBS.* 63, 66 (1986) (“Because multiple regression statistics have the technical capacity to identify discriminatory influences from among the combined effects of a set of factors acting simultaneously, they have powerful and useful potential in [discrimination] litigation.”).

⁷⁵ *See* Part V.

⁷⁶ *Bazemore* 478 U.S. at 400.

⁷⁷ *Id.*

Recently, in *Wal-Mart Stores, Inc. v. Dukes*,⁷⁸ the Court reaffirmed the important role of statistical evidence in proving discrimination.⁷⁹ Although the majority and the dissenting opinions in the case differed as to whether the statistical evidence of gender disparities presented by the plaintiffs established a *prima facie* case of a pattern or practice of discrimination, neither side challenged the general utility of using statistics in claims alleging systemic intentional discrimination. The dissenting opinion, written by Justice Ruth Bader Ginsburg and joined by three other Justices, provided a detailed discussion of the multiple regression statistical models presented by the plaintiffs and how these models revealed gender disparities in pay and promotion after taking into account factors such as job performance, tenure, and store location.⁸⁰ As one scholar has noted, “the methods now presented to the courts look remarkably similar to the kinds of studies that once appeared in [economics] journals.”⁸¹

The primary appeal of multiple regression modeling is its ability to provide answers to “what if” questions, such as “what is the likelihood that a defendant’s case would have been noticed for the death penalty if the victim was Caucasian rather than African American?” Multiple regression-based evidence of the statistical pattern of discrimination is also an efficient means of aggregating individual decisions, even if each individual discriminatory decision could be identified, because doing so would be time consuming and cost prohibitive.⁸² The more accurately the statistical model is able to approximate reality by including the key determinants of the outcome of interest, the stronger the *prima facie* case of intentional discrimination.⁸³ This does not imply, however, that courts have been equally receptive to statistical evidence of discrimination across all contexts, even when the data under investigation are very detailed and the statistical models account for a wide range of non-discriminatory rival hypotheses. For example, some members of the Supreme Court have demanded “exceptionally clear proof” to infer racial discrimination from statistical evidence that would, essentially, allow for the identification of an abuse of discretion in a specific instance.⁸⁴ Under this standard, which other members of the Court have described as an unwarranted departure from its basic equal protection

⁷⁸ 564 U.S. 338 (2011).

⁷⁹ *Id.* at 356–57 (acknowledging that statistical evidence can be probative of discriminatory treatment when the correct comparison groups are identified).

⁸⁰ *Id.* at 372 (Ginsburg, J., joined by Breyer, Sotomayor, & Kagan, JJ., dissenting) (agreeing with the trial court that the statistical results were sufficient to raise an inference of gender discrimination.).

An equally comprehensive of evaluation of statistical evidence of systemic disparate effect and disparate treatment in the capital punishment charging-and-sentencing process was conducted nearly twenty-five years earlier in *McCleskey v. Kemp*, 481 U.S. 279 (1987).

⁸¹ Ashenfelter & Oaxaca, *supra* note 31, at 323.

⁸² Selmi, *supra* note 26, at 508.

⁸³ See NAT’L RESEARCH COUNCIL, *supra* note 53, at 142.

⁸⁴ *Cf. McCleskey*, 481 U.S. at 292 (Powell, J.) (an equal protection claim requires a defendant to prove the decision-makers in his [sic] case acted with discriminatory purpose), *with id.* at 352 (Brennan, J., dissenting) (defendants can satisfy the discriminatory intent prong with a multiple regression analysis of aggregate data that takes into account a large number of relevant factors).

framework,⁸⁵ statistical evidence would likely be of dubious value unless the statistical disparity was so stark that intentional discrimination was the only remaining reasonable inference.⁸⁶

To the extent that courts deem statistical models an appropriate, and even sometimes a necessary, component of a claimant's allegation of discrimination in particular contexts, there remains the key question of when particular models will be considered probative by courts. To the dismay of scores of litigants, the answer to this central question has remained especially elusive because the Supreme Court has neglected to provide workable standards. On occasions when the Court has addressed the issue, its routinely sharply-divided opinions have been incapable of announcing a coherent approach by which statistical evidence shall be assessed.⁸⁷ Statistical evidence of discrimination is now commonplace for cases populating the federal dockets, so the need for a transparent framework is unavoidable. Sociologist Arthur Stinchcombe famously remarked that it is "prefer[able] to be wrong than misunderstood [because] being misunderstood shows sloppy theoretical work."⁸⁸ The lack of clarity boils down to a refusal to discuss certain critical questions that must be answered in order for the doctrine and the theory that underlies it to be intelligible.⁸⁹ This jurisprudential sloppiness has left plaintiffs, attorneys, and judges with insufficient guidance to effectively litigate and resolve discrimination claims.⁹⁰ Admittedly, anti-discrimination law is not unique in terms of the existence and persistence of ambiguous legal standards. But the lack of clarity is especially troublesome in the anti-discrimination context because, as a practical matter, only circumstantial evidence is available to prove discriminatory intent and this circumstantial evidence is inherently difficult to verify.⁹¹ This is especially true when the plaintiff is also a criminal defendant alleging an equal protection violation based on the prosecution's discriminatory charging practices. These claimants lack access to internal documents necessary to elucidate the foundation upon which the charging

⁸⁵ *McCleskey*, 481 U.S. at 348 (Brennan, J., dissenting) ("the [majority] relies on the very fact that this is a case involving capital punishment to apply a lesser standard of scrutiny under the Equal Protection Clause.").

⁸⁶ *Baldus & Cole*, *supra* note 50, at 56 (explaining that statistical evidence provides indirect proof of intentional discrimination).

⁸⁷ *Kruse*, *supra* note 14, at 1548.

⁸⁸ ARTHUR L. STINCHCOMBE, *CONSTRUCTING SOCIAL THEORIES* 6 (1968); *accord* 4 FRANCIS BACON, *THE WORKS OF FRANCIS BACON*, 210 (1875) ("Truth more readily emerges from error than from confusion.").

⁸⁹ Bruce H. Mayhew, *Structuralism versus Individualism, Part II: Ideological and Other Obfuscations*, 59 SOC. FORCES 627, 629 (1981).

⁹⁰ *Kruse*, *supra* note 14, at 1548. *See also* *United States v. Thorpe*, 471 F.3d 652, 658 (6th Cir. 2006) (noting the Supreme Court's failure to provide a standard for the discriminatory intent prong in discriminatory prosecution cases); *United States v. Tuitt*, 68 F. Supp. 2d 4, 10 (D. Mass. 1999) (same).

⁹¹ *Kruse*, *supra* note 14, at 1526-28 ("So long as the prosecutor has probable cause to prosecute, based on the elements of the offense set forth in the statute, the decision to prosecute is solely within the discretion of the prosecutor [and] [p]rosecutors are capable of finding a violation in almost anyone."); *Ashenfelter & Oaxaca*, *supra* note 31, at 322 ("It is not hard to see that the appearance of disparate treatment is easy for an employer to eliminate without making any change in behavior at all. Differential hiring or pay scales may be supported by simply asserting that all hiring and pay is determined by merit, and merit is determined by employee supervisors.").

decision rested and permit the claimant, when applicable, to argue that such non-racial assertions are pretextual. And unlike many employment discrimination cases, a criminal defendant often lacks any contact with prosecutors and their staff that could lead to the discovery of non-statistical corroborative evidence, such as racially offensive remarks and other behaviors indicative of racial animus. Consequently, statistics are often the only avenue through which a claimant can prove clandestine and covert discrimination.⁹²

Complicating matters is the fact that some members of the Court have emphasized the necessity of non-statistical corroborative evidence in all but the most extreme circumstances,⁹³ but have declined to offer any general guidance about the character of acceptable evidence (e.g., scope and intensity). And to date, only in the rare case of palpable racism have defendants been deemed by lower federal courts to have established “some evidence” of discriminatory intent in criminal charging.⁹⁴ In *United States v. Jones*, the Sixth Circuit granted an African American plaintiff’s motion for discovery in a selective prosecution case after the plaintiff presented evidence that his arresting police officers wore T-shirts emblazoned with inappropriate images of the defendant and his wife, as well as mailed the defendant racially insulting postcards after the arrest.⁹⁵ The plaintiff also presented evidence showing that no other similarly situated defendants had been referred for federal prosecution in the preceding five years. In another case, *United States v. Gordon*, the Eleventh Circuit ruled in favor of an African American plaintiff’s discovery request from the prosecution after the plaintiff provided evidence that similarly situated Caucasian defendants had not been prosecuted *and* the prosecutor told an intern that the

⁹² *Washington v. Davis*, 426 U.S. 229, 253–54 (1976) (Stevens, J., concurring) (“It is unrealistic [] to require the victim of alleged discrimination to uncover the actual subjective intent of the decision-maker.”); *Gross v. FBL Fin. Services, Inc.*, 557 U.S. 167, 190 (2009) (Breyer, J., dissenting) (explaining that causal attribution is particularly difficult when assessing non-physical causal forces, such as motives).

⁹³ In his concurring opinion in *Washington v. Davis*, Justice Stevens remarked that evidence of discriminatory effect, alone, would be sufficient for an equal protection challenge if the disproportionate impact was drastic. *Davis*, 426 U.S. at 254. Justice Stevens cited two cases supporting his assertion: *Yick Wo v. Hopkins*, 118 U.S. 356 (1886) and *Gomillion v. Lightfoot*, 364 U.S. 339 (1960).

Whereas *Yick Wo* was decided on equal protection grounds, *Gomillion* was decided under the Fifteenth Amendment. The Alabama legislature re-drew the electoral district boundaries from a region with a square shape to a twenty-eight sided figure in order to exclude all African Americans from the city limits of Tuskegee. The plaintiffs presented their claim on both Fourteenth (equal protection) and Fifteenth Amendment grounds. Justice Charles Whittaker, concurring in judgement, argued that the case should have been decided on equal protection grounds. Both the equal protection and the Fifteenth Amendment arguments address infringements of rights based on racial classifications, but the Fifteenth Amendment is specific to voting. Justice Whittaker believed the Fifteenth Amendment claim was inapplicable because the re-districting scheme did not deprive African Americans the right to vote, rather it was “an unlawful segregation of races of citizens, in violation of the Equal Protection Clause of the Fourteenth Amendment.” *Gomillion*, 364 U.S. at 349.

In the statutory context, the Court also emphasized the importance of non-statistical evidence to substantiate the statistical evidence. *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338 (2011) (discussing the relevance of anecdotal evidence of gender bias and sexism, in addition to statistical results, in a gender discrimination claim concerning pay and promotion).

⁹⁴ James Babikian, *Cleaving the Gordian Knot: Implicit Bias, Selective Prosecution, and Charging Guidelines* [Note], 42 AM. J. CRIM. L. 139, 152 (2015).

⁹⁵ *United States v. Jones*, 159 F.3d 969, 975–77 (6th Cir. 1998).

investigation of the plaintiffs had been “brought on by the arrogance on the part of blacks in [the jurisdiction].”⁹⁶ In both of these cases, however, the courts did not rule on the adequacy of the underlying constitutional violation; they only decided that the plaintiffs presented enough evidence of discriminatory intent to prove a “colorable entitlement” to a claim of a constitutional violation that warranted the discovery of government documents relating to the decision whether to file charges against similarly situated defendants.⁹⁷ The Eleventh Circuit explained that the prosecutors racially discriminatory statement “standing alone would not be enough, but assumes significance in light of other evidence suggesting a [racially biased] pattern of Government activity in [] cases that were prosecuted.”⁹⁸

Extensive statistical evidence of racially disparate capital charging decisions was presented to the Court in the landmark case, *McCleskey v. Kemp*,⁹⁹ yet the implications of *McCleskey* for the use of statistical models of discrimination is quite ambiguous. The defendant, Warren McCleskey, an African American, presented evidence from a comprehensive examination of capital charging and sentencing practices in Georgia. The study revealed, *inter alia*, the odds a prosecutor sought the death penalty against a defendant accused of killing a Caucasian victim was 3.1 times greater than a defendant accused of killing an African American victim, all else equal.¹⁰⁰ The Court, by the slimmest of margins, five-to-four, rejected Mr. McCleskey’s claim. The Court accepted the statistical evidence as valid, but held that the evidence was incapable of showing that race was a motivating factor in Mr. McCleskey’s specific case.¹⁰¹ The majority’s refusal to accept the statistical evidence as sufficient to warrant relief was in stark contrast to its earlier decisions involving the adequacy of statistical evidence in jury selection and employment discrimination.¹⁰² Dissenting in *McCleskey*, Justice Blackmun noted that the statistical evidence presented in the case would have satisfied

⁹⁶ *United States v. Gordon*, 817 F.2d 1538, 1540 (11th Cir. 1987).

⁹⁷ *E.g., id.* (“The evidence submitted indicates that [the plaintiff] has sufficiently established the essential elements of the selective prosecution test to prove a ‘colorable entitlement’ to the defense [and the defendant] is entitled to discovery of the relevant government documents.”).

⁹⁸ *Id.* In other words, even in light of apparent “smoking gun” evidence of racial motive, courts require further inquiry into the “concurrence of elements” of a discriminatory treatment claim—in essence, the plaintiff is required to prove that the elements of the violation happened at the same time of the cause of the harm. See JOSHUA DRESSLER, *UNDERSTANDING CRIMINAL LAW* § 15 (7th ed. 2015) (describing the concurrence of elements requirement for establishing criminal liability).

⁹⁹ 481 U.S. 279 (1987).

¹⁰⁰ Mr. McCleskey’s statistical evidence also revealed that the odds of receiving a death sentence at trial were 4.3 higher in Caucasian victim cases compared to non-Caucasian victim cases, all else equal. *Id.*

¹⁰¹ *Id.* at 287. See also Part V.

¹⁰² *McCleskey v. Kemp*, 481 U.S. 279, 347–48 (Blackmun, J., dissenting) (noting that the majority improperly watered down the equal protection doctrine in the capital context). Cf. Marc Price Wolf, *Proving Race Discrimination in Criminal Cases Using Statistical Evidence* (Note), 4 HASTINGS RACE & POVERTY L. J. 395, 396 (2006) (noting that the Supreme Court has uncritically adopted the conclusions of social science studies in death penalty cases not dealing with race, such as the juvenile death penalty and intellectual disability).

even the most “crippling” burden of proof for an equal protection violation that the Court erected and, wisely, rescinded a year before *McCleskey*.¹⁰³

The implications of the Court’s holding in *McCleskey* for selective prosecution actions remain unclear because the majority opinion focused on Mr. McCleskey’s racially discriminatory death sentencing claim, and the inherent problems in identifying racially biased actors when there are multiple decision points with different actors. The Court’s ruling did not specifically address the selective prosecution claim, although Mr. McCleskey’s statistical evidence appeared most probative for this question given the Court’s prior rulings on prosecutorial misconduct—a point emphasized by Justice Blackmun in his dissent.¹⁰⁴ Even assuming that the majority was correct in rejecting the claim pertaining to racially discriminatory capital sentencing, Mr. McCleskey would have still been entitled to relief had he prevailed on the selective prosecution claim. This distinction is far from trivial because there was considerable disagreement between the majority and the dissent as to whether the statistical evidence of racially disparate charging and racially disparate sentencing should be evaluated according to different standards.¹⁰⁵ Interestingly, post-*McCleskey*, two of the justices who rejected Mr. McCleskey’s claims appeared to offer recantations. Justice Lewis Powell authored the majority opinion in *McCleskey*, but would subsequently remark that he had an extremely limited understanding of statistical analysis and regretted his decision in that case after he retired from the Court.¹⁰⁶ Nearly fifteen years after the ruling, Justice Sandra Day O’Connor, who also joined the majority in *McCleskey*, openly stated that she had serious concerns as to whether the death penalty was being administered fairly.¹⁰⁷

The year following Justice O’Connor’s public statement expressing doubts about the even-handed administration of the death penalty, the Court decided *United States v. Bass*, a case in which it was presented nationwide statistics of federal prosecutors’ death penalty charging decisions that suggested Africans Americans were being targeted for capital prosecutions in the Eastern District of Michigan.¹⁰⁸ Specifically, the evidence revealed that none of the 17 defendants charged with the death penalty in the Eastern District of Michigan were Caucasian (14 were African American and 3 were

¹⁰³ *McCleskey*, 481 U.S. at 364 (arguing that Mr. McCleskey’s evidence would have satisfied the exceedingly difficult standard for proving racially discriminatory jury selection developed in *Swain v. Alabama* and overruled in *Batson v. Kentucky*).

¹⁰⁴ *Id.* at 350 (Blackmun, J., dissenting) (noting that Mr. McCleskey’s evidence of racial bias in capital charging decisions was especially strong, but the majority purposefully ignored this claim and focused on other decision points).

¹⁰⁵ *Cf. id.* at 293 (Powell, J.) (recognizing that the Court has accepted the use of statistics as proof of intent to discriminate in limited contexts, such as the venire pool and employment discrimination and has permitted a finding of a constitutional violation even when the statistical pattern of discrimination is extreme), *with id.* at 350 (Blackmun, J., dissenting) (noting the majority’s mischaracterization of the defendant’s selective prosecution claim to “distinguish [the] claim from the venire and employment cases, which have long accepted statistical evidence and has provided an easily applicable framework for review”).

¹⁰⁶ JOHN C. JEFFRIES JR., JUSTICE LEWIS F. POWELL, JR.: A BIOGRAPHY (1994).

¹⁰⁷ Brian Bakst, *O’Connor Questions Death Penalty*, ASSOCIATED PRESS, Jul. 2, 2001.

¹⁰⁸ *United States v. Bass*, 536 U.S. 862 (2002).

Latino/Hispanic), but Caucasians comprised nearly 20% of capital prosecutions nationally and African Americans only comprised 48%. The Court rejected Mr. Bass' use of raw statistics, reasoning that he was unable to show discriminatory effect—that is, he failed to provide evidence that charges were not brought against similarly situated Caucasian defendants. The perplexing aspect of the Court's *per curiam* opinion, besides its brevity,¹⁰⁹ was that Mr. Bass was merely requesting information from the prosecution about its charging practices so he could identify similarly situated defendants. The lower federal appellate court had granted Mr. Bass' request to obtain discovery from the Department of Justice, reasoning that the evidence Mr. Bass presented to the court satisfied the less-stringent standard for access to the prosecution's files in order to more fully develop an equal protection violation claim (i.e., "some evidence...of discriminatory effect and discriminatory intent").¹¹⁰ The federal appellate court acknowledged that the statistical evidence, standing on its own, would be insufficient to support a *prima facie* case of selective prosecution to merit dismissal of the capital charge. The Supreme Court's seemingly cursory opinion omitted any careful articulation of what a credible showing entailed that could provide guidance to lower courts, while at the same time providing tacit confirmation that statistical evidence could be sufficient to justify, at minimum, the defendant's discovery request.¹¹¹

But even in the absence of court-defined standards for proving intentional discrimination with statistical evidence, careful attention must be given to the validity of the statistical results because courts have rejected statistical evidence of intentional discrimination based on either the perceived inaccuracy of the statistical models or on the merits of the legal claim.¹¹² The probative value of any statistical technique is highly dependent on the plausibility of the assumptions underlying the model.¹¹³ When rejecting statistical evidence of discrimination, courts commonly highlight two potential shortcomings of the underlying statistical models: (1) the low predictive power of statistical models in determining the outcome *and* (2) the potential correlation between unobserved factors and the race of the victim (or defendant).¹¹⁴ To be sure, a careful inquiry into the adequacy of the statistical model is indispensable, but as I explain below, these criticisms are often misplaced and courts may be overly cautious when evaluating

¹⁰⁹ The opinion was 532 words and included a single footnote. *Id.* at 862–64.

¹¹⁰ *United States v. Bass*, 266 F.3d 532, 536 (6th Cir. 2001). *Cf.* *United States v. Armstrong*, 517 U.S. 456, 481 (1996) (Stevens, J., dissenting) ("I thought it was agreed that defendants do not need to prepare sophisticated statistical studies in order to receive mere discovery in [selective prosecution] cases like this one.").

¹¹¹ *Bass*, 536 U.S. at 862 (*per curiam*) (assuming, but not deciding, that national statistics of capital charging decisions could be sufficient for a discovery request).

¹¹² *Cf.* *McCleskey v. Kemp*, 481 U.S. 279, 279–80 (1987) (accepting the validity of Mr. McCleskey's statistical results, but rejecting the underlying constitutional claim) *with* *McCleskey v. Zant*, 580 F. Supp. 338 (N.D. Ga. 1984) (rejecting the validity of Mr. McCleskey's statistical results).

¹¹³ *BERK*, *supra* note 51 (discussing the assumptions underlying regression-based statistical models and the consequences of violating those assumptions).

¹¹⁴ *Ashenfelter & Oaxaca*, *supra* note 31, at 323; *Baldus & Cole*, *supra* note 50, at 76.

statistical evidence, leading them to inappropriately reject extremely probative statistical evidence of intentional discrimination. In Part VI, I specifically address the possible influence of the aforementioned shortcomings of statistical models of discrimination on the results presented in this Article.

B. Evaluating Statistical Models of Discrimination

The first critique, the low-predictive power of statistical models, is related to the under-determinacy of the statistical model. By construction, the statistical model assumes that there remain some unmeasured factors that influence the outcome of interest after taking into account the effects of the included factors in the model, as well some degree of randomness in the data. This is commonly referred to as “residual error.” A concern associated with large residual error is the inability of the statistical model to adequately capture the underlying process that generated the outcomes. If a statistical model fails to fit the data particularly well (based on measures that emphasize predictability), critics contend that the model inadequately approximates the discretionary process, and thus is of limited utility.¹¹⁵ Courts have held that “the explanatory power of a model is a factor that may legitimately be considered [when] deciding whether the model may be relied upon,”¹¹⁶ but they have generally avoided “establishing a particular predictive capacity as a *sine qua non* for a model to pass muster.”¹¹⁷ In *McCleskey*, the trial court reasoned that “the validity of the model depends upon a showing that it predicts the variations in the dependent variable to some substantial degree,”¹¹⁸ although what qualifies as “substantial” remains elusive. There is no consensus in the scientific community as to what qualifies as the minimally acceptable predictive capacity of a statistical model.¹¹⁹ The probative value of a statistical model will largely depend on the comprehensiveness of the relevant explanatory variables included in the model. A model with low predictive power may still establish a *prima facie* case of discrimination when the model incorporates “information central to understanding the causal relationships at issue.”¹²⁰

The second objection, which is related to the first critique, centers on the possible influence of omitted variables on any statistical measure of racial discrimination. This is commonly known as the *unconfoundedness assumption*. When there are omitted factors in determining the outcome of interest (e.g., wages or charging decisions), there are no guarantees that these factors are uncorrelated with the race of the plaintiff, defendant, or

¹¹⁵ See, e.g., Stephen P. Klein et al., *Race and the Decision to Seek the Death Penalty in Federal Cases* (RAND 2006) (discussing differing interpretations of models with low predictive power).

¹¹⁶ Griffin v. Bd. of Regents of Regency Univ., 795 F.2d 1281, 1292 (7th Cir. 1986).

¹¹⁷ *Id.* at 1291–92.

¹¹⁸ *McCleskey v. Zant*, 580 F. Supp. 338, 351 (N.D. Ga. 1984).

¹¹⁹ JEFFREY M. WOOLDRIDGE, *INTRODUCTORY ECONOMETRICS: A MODERN APPROACH* 43–43 (2d ed. 2003).

¹²⁰ *Valentino v. U. S. Postal Serv.*, 674 F.2d 56, 71 (D.C. Cir. 1982).

victim.¹²¹ When an omitted factor is related to both the status characteristic of interest (e.g., race) and the outcome of interest, the statistical association between the status characteristic and the outcome may simply be an artifact of its association with the omitted factor.¹²² And even in the event that the inclusion of the omitted variable in the statistical model does not render the relationship between the status characteristic and the outcome null and void, the omitted variable's inclusion in the model may substantially attenuate the effect of the status characteristic on the outcome.¹²³

Social scientists have readily acknowledged these potential shortcomings, but emphasize that the methodological rigor of any particular study, which primarily pertains to how well the model approximates the underlying discretionary process that generated the alleged racial disparity, must be judged on a case-by-case basis, and “[social scientists] will be more or less convinced by the findings of a particular non-experimental study according to how well it is done[.]”¹²⁴ The more convinced the trier-of-fact is that members of the defendant's racial group and the individuals who are not in the defendant's racial group are similarly situated, the stronger the claim of intentional discrimination.¹²⁵ With respect to specific criticisms articulated, *supra*, social scientists have offered several responses. The common rejoinder to the first critique—i.e., the under-determinacy of the statistical models—is the recognition that all models, by definition, are “wrong” because they are simplifications of a more complex process.¹²⁶ The goal of scientific explanation is to supply a useful approximation of reality that is illuminating and useful.¹²⁷ As psychologist Stephen Klein and colleagues have noted, “[f]ew systems as complex as the criminal justice system lend themselves to high-accuracy statistical modeling.”¹²⁸ Statistician George E. P. Box famously remarked, “just as the ability to devise simple but evocative models is the signature of the great scientist, so over-elaboration [] is often the mark of mediocrity.”¹²⁹ Sociologist Guillermina Jasso has explained, “the goal is to develop a [model] that is at once simple and fruitful, that is, with a minimum number of postulates and a maximum number of predictions.”¹³⁰ The important question is not whether the model predicts the data perfectly—the answer to that question is clearly (and unequivocally) “no.” Rather, the key inquiry is whether, based on the existing research literature and the litigants' plausible arguments, the model includes the essential features hypothesized to govern the discretionary

¹²¹ Ashenfelter & Oaxaca, *supra* note 31, at 323.

¹²² BERK, *supra* note 51, at 81–101.

¹²³ *Id.* at 13.

¹²⁴ Ashenfelter & Oaxaca, *supra* note 31, at 324.

¹²⁵ Baldus & Cole, *supra* note 50, at 63.

¹²⁶ George E.P. Box, *Science and Statistics*, 71 J. AM. STAT. ASS'N 791, 792 (1976).

¹²⁷ *Id.*

¹²⁸ Klein et al., *supra* note 115, at 40.

¹²⁹ Box, *supra* note 126, at 792; Donald Black, *The Epistemology of Pure Sociology*, 20 L. & SOC. INQUIRY 829, 838 (1995) (remarking that “science loves simplicity and despises generality”).

¹³⁰ Guillermina Jasso, *Principles of Theoretical Analysis*, 6 SOC. THEORY 1, 1 (1988).

process to permit reasonable inferences based on the model.¹³¹ Indeed, the view that the under-determinacy of statistical models does not automatically preclude employing the model in anti-discrimination cases has been expressly recognized by the Court on numerous occasions.¹³²

The second objection to statistical evidence of discrimination—omitted variable bias—has been levied so frequently by critics that it may be deemed the “lowest hanging fruit” of methodological scrutiny of models of legal behavior because as previously explained, by definition, theoretically relevant variables are omitted from statistical models.¹³³ Nearly all social scientists acknowledge, at the outset, that omitted variable bias is possible,¹³⁴ but they also emphasize that the actual critique implies its own underlying theory of the interrelationships between the observed and unobserved factors of interest. Researchers need not control for every conceivable variable possibly influencing the outcome of interest.¹³⁵ The excluded variables must satisfy four conditions: (1) correlation with the key explanatory variable of interest (e.g., race or gender); (2) causal effect on the outcome variable (e.g., plea-bargaining decision); (3) not proxied by any other variable or combination of variables already included in the model; and (4) not caused by the explanatory variable of interest (e.g., race or gender).¹³⁶ If any one of these four conditions is absent, then controlling for the omitted variable is unnecessary when examining the causal impact of the key variable of interest.¹³⁷ And even when omitted variables satisfy these conditions, the impact of the excluded of the variable(s) on the statistical measure of discrimination is far from obvious. For example, research has repeatedly revealed that, in the death penalty context, evidence of racial disparities can be *stronger or weaker* when the statistical models expressly

¹³¹ Baldus & Cole, *supra* note 50, at 76 (assessing a model requires, inter alia, an examination of whether the data fit the model adequately).

¹³² See generally Castaneda v. Partida, 430 U.S. 482 (1977); Bazemore v. Friday, 478 U.S. 385 (1986) (a statistical model may help prove discrimination even though it does not incorporate every conceivable relevant variable).

¹³³ Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 78 (2002).

¹³⁴ Ashenfelter & Oaxaca, *supra* note 31, at 322.

¹³⁵ Epstein & King, *supra* note 133, at 78.

¹³⁶ *Id.*; BERK, *supra* note 51, at 81. The fourth condition is often overlooked by critics of statistical models of discrimination. If the omitted variable is, itself, influenced by the status characteristic (e.g., race or gender), then the controlling for the omitted variable actually removes some of the true or total effect of the status characteristic because part of the effect of the status characteristic operates through its influence on the omitted variable. GARY KING ET AL., *DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH* 78 (1994) (“By holding constant something that is itself affected by the causal variable[s] of interest, one removes precisely the effect one is attempting to study.”).

There are methods available to identify the total effect of a status characteristic on an outcome variable by uncovering both its direct and indirect effects, but these methods rely on additional model assumptions that can be difficult to justify in many situations. Green et al., *supra* note 54. It should be obvious that many of the variables included in models of discrimination are susceptible to this critique. Common practice in the research literature is to acknowledge this fact yet treat these intermediate variables as being exogenously determined (i.e., not determined by race). See D. James Greiner & Donald B. Rubin, *Causal Effects of Perceived Immutable Characteristics*, 93 REV. OF ECON. & STAT. 775 (2011).

¹³⁷ Epstein & King, *supra* note 133, at 78.

take into account more variables that could account for the observed relationship between race and the discretionary decision.¹³⁸

A recent evaluation of the use of statistical models in employment discrimination cases by economists Joni Hersch and Blair Druhan Bullock note that criticism of multiple regression models is overblown and violation of the underlying assumptions of these models typically have very little influence on the overall results, yet courts routinely decide in favor of the defendants in cases when these common criticisms are raised.¹³⁹ Accordingly, the authors underscore the severe consequences of courts giving undue weight to these critiques when they lack merit.¹⁴⁰ It is often the case that all a defendant can show is that, after multiple-regression analyses have accounted for plausible non-racial explanations, disparities still remains.¹⁴¹ But “parties [attempting to refute an allegation of intentional discrimination] must do more than speculate about possible flaws to invalidate statistical evidence. The key question is whether the omission of potentially explanatory factors creates sufficient doubt in a study’s accuracy to warrant the denial of all relief.”¹⁴²

In the landmark case *Washington v. Davis*, Justice John Paul Stevens noted that “[N]ormally the actor is presumed to have intended the natural consequences of his deeds. [...] The line between discriminatory purpose and discriminatory impact is not nearly as bright, and perhaps not quite as critical, as [one] might assume. [...] [A] constitutional issue does not arise every time some disproportionate impact is shown [but] when the disproportion is [] dramatic [] it really does not matter whether the standard is phrased in terms of purpose or effect.”¹⁴³ As a result, advocates of statistical evidence of discrimination opine that the task of the courts is to determine whether the possibility of prejudice influencing legal decision-making is so high as to render that particular process constitutionally unacceptable.¹⁴⁴ Indeed, the Court’s early equal protection cases emphasized that systematic discrimination in the enforcement of laws violates the equal protection clause when coupled with the absence of rules to adequately

¹³⁸ David C. Baldus & George Woodworth, *Race Discrimination in the Administration of the Death Penalty: An Overview of the Empirical Evidence with Special Emphasis on the Post-1990 Research*, 39 CRIM. L. BULL. 194 (2003) (noting that studies examining the influence of race on capital punishment decision-making tended to find stronger effects when they included a wider range of explanatory variables); see also Kevin Lang & Michael Manove, *Education and Labor Market Discrimination*, 101 AM. ECON. REV. 1467, 1492 (2011) (explaining that failing to control for educational attainment, which is highly correlated with race/ethnicity, led to an underestimation the impact of discrimination in wages by 66%).

¹³⁹ Joni Hersch & Blair Druhan Bullock, *The Use and Misuse of Econometric Evidence in Employment Discrimination Cases*, 71 WASH. & LEE L. REV. 2365 (2014).

¹⁴⁰ *Id.*

¹⁴¹ Cheng, *supra* note 39, at 2098.

¹⁴² *Id.* at 2103.

¹⁴³ *Washington v. Davis*, 426 U.S. 229, 254 (1976) (concurring opinion) (internal citations omitted).

¹⁴⁴ *McCleskey v. Kemp*, 481 U.S. 279, 364 (1987) (Blackmun, J., dissenting) (“The issue is whether the constitutional guarantee of equal protection limits the discretion in the [criminal justice] system.”); Memorandum from Scalia, *supra* note 44 (“The task in *McCleskey* was to determine whether the possibility for racial prejudice influencing legal decision-making had become so high that Georgia’s system for inflicting capital punishment was constitutionally unacceptable.”).

guide or control the exercise of discretion.¹⁴⁵ And, in fact, the legislative history for the Fourteenth Amendment reveals that the framers specifically intended for it to prohibit the unequal enforcement of the states' *criminal laws* based on racial distinctions.¹⁴⁶

Statistical models of discrimination, when used correctly, are able to provide important insights into potentially discriminatory decision-making. Over the last four decades, courts have repeatedly engaged with quantitative data and statistical models when determining whether a claimant's constitutionally- or statutorily-based violation was meritorious. Many of the earlier concerns pertaining to the use of statistical evidence have subsided in the face of significant advances in statistical methodology and the rules of statistical inference developed to specifically address those concerns.¹⁴⁷ The statistics literature is now replete with tools designed to assist scholars with carefully examining the sensitivity of statistical evidence of discrimination to violations of assumptions of the statistical models.¹⁴⁸ Some scholars have even argued that we may now be at a point where courts have become overly cautious of statistical evidence discrimination in light of widespread agreement in the scientific community over appropriate levels of methodological rigor, as well as the emergence and persistence of patterns of legally illegitimate racial disparities—especially in the criminal justice context. In the following section, I describe and evaluate the empirical literature on racial discrimination in capital charging decision-making over the past quarter-century. As noted, *supra*, social scientific inquiry into the influence of race on capital charging dates back much further than twenty-five years, but the recent scholarship is the most methodological sophisticated. The U.S. Government Accountability Office conducted a detailed review of pre-1990 empirical research on capital charging and sentencing,¹⁴⁹ and the results of those earlier studies are consistent with the more recent research I discuss and critique.

III. EMPIRICAL SCHOLARSHIP

The continuity in racial disparities in capital sentencing, in light of intense and sustained attention to its sources and consequences, shares stark similarity to racial disparities in the employment context.¹⁵⁰ Race may exert

¹⁴⁵ *Yick Wo v. Hopkins*, 118 U.S. 356, 372–74 (1886) (highlighting that nothing in the challenged municipal ordinance guided or controlled the discretionary authority); *accord Castaneda v. Partida*, 430 U.S. 482, 494, 500 (1977) (emphasizing that statistical evidence of disparate impact, coupled with a selection/enforcement scheme that is susceptible to abuse, is adequate for an equal protection challenge).

¹⁴⁶ *McCleskey*, 481 U.S. at 346 (Blackmun, J., dissenting) (“[T]he legislative history of the Fourteenth Amendment reminds us that discriminatory enforcement of States’ criminal laws was a matter of great concern for the drafters.”).

¹⁴⁷ See, e.g., Daniel E. Ho & Donald B. Rubin, *Credible Causal Inference for Empirical Legal Studies*, 7 ANN. REV. OF L. & SOC. SCI. 17 (2011).

¹⁴⁸ I utilize these tools to examine the robustness of my statistical results in Part VI.

¹⁴⁹ U.S. Gov’t Accountability Office, *supra* note 6.

¹⁵⁰ Labor economist Pedro Carneiro and colleagues have explained: “In spite of 40 years of civil rights and affirmative action policy, substantial gaps remain in the market wages of African American males and females compared to white males and females.” Pedro Carneiro et al., *Labor Market Discrimination and Racial Differences in Premarket Factors*, 48 J. L. & ECON. 1, 1 (2005).

an influence on decision-making in the capital charging-and-sentencing process at nearly every stage: prosecutor's decision to charge a defendant with capital murder; grand jury's decision to indict a defendant for capital murder; prosecutor's decision to seek the death penalty; prosecutor's use of peremptory challenges; prosecutor's willingness to offer a favorable plea deal; prosecutor's decision to advance the case to the penalty phase; jury's decision to impose a death sentence; and governor's (or pardon board's) decision to grant clemency.¹⁵¹ Several of these decision stages require collective decision-making (e.g., grand and petit juries), so it may be difficult to identify conscious or unconscious racial bias, or a particular pattern or practice, responsible for the racially disparate result;¹⁵² yet the vast majority of these decisions are controlled by the prosecutor, and often with very little oversight or constraints on these discretionary decisions.¹⁵³ As explained, *supra*, racial bias in earlier decision stages, such as charging, are unlikely to be rectified during sentencing.¹⁵⁴

A. *Quantifying Racial Discrimination in Capital Charging*

Over the past twenty-five years, there have been at least a dozen statistical analyses of capital charging decisions. This research literature pales in comparison to the number of studies examining the discretionary choices of actors in the capital punishment process *after* a death notice has been already filed—namely, capital sentencing.¹⁵⁵ Scholars have primarily focused on post-capital charge decision-making because they lack adequate information on the population of defendants at risk for a capital charge. Information on potentially capital cases can be extremely difficult to obtain because of many local law enforcement agencies' sub-optimal record-

¹⁵¹ Scott Phillips, *Continued Racial Disparities in the Capital of Capital Punishment: The Rosenthal Era*, 50 HOUS. L. REV. 131, 149 (2012) (describing the various stages at which racial bias can influence legal decision-making in the death penalty system); *Gregg v. Georgia*, 428 U.S. 153 (1976) (noting that improper considerations, such as race, may theoretically impact the decision-making at various points throughout the capital adjudication process, but deciding a constitutionally permissible death penalty statute need not address all of the stages); *McCleskey*, 481 U.S. 279 (same).

There is not much cause for concern about racial disparities in pre-trial release and bail determinations for individuals who may potentially face the death penalty because most jurisdictions either prohibit pretrial release for alleged murders or require that bail be set at the very top of the bail schedule. THOMAS H. COHEN & BRIAN A. REAVES, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, PRETRIAL RELEASE OF FELONY DEFENDANTS IN STATE COURTS 3, 6 (2007); BRIAN A. REAVES & JACOB PEREZ, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, PRETRIAL RELEASE OF FELONY DEFENDANTS, 1992, at 2 (1994).

¹⁵² See *McCleskey*, 481 U.S. 279 (discussing the difficulty in determining the source of discrimination when a collective is responsible for making the decision).

¹⁵³ Liebman, *Opting for Real Death Penalty Reform*, *supra* note 10; *McCleskey*, 481 U.S. at 350 (Blackmun, J., dissenting) (referring to the prosecutor as the "quintessential state actor in a criminal justice proceeding").

¹⁵⁴ See *supra* notes 5 & 6 and accompanying text.

¹⁵⁵ See, e.g., U.S. Gov't Accountability Office, *supra* note 6 (conducting a review of 28 empirical studies of the death penalty process and concluding that over 80% of the studies reported a race of victim effect for charging and/or sentencing).

keeping practices.¹⁵⁶ As noted, *supra*, the examination of charging dynamics is extremely important because it not only provides valuable insights about the “front-end” of the process, but it is also key to properly understanding down-stream legal error. Studies conducted in California, Colorado, Connecticut, Georgia, Kentucky, Maryland, Missouri, North Carolina, and South Carolina all reveal racial disparities in capital charging decisions, based on the race of the victim, the offender/victim racial combination, or both.

The vast majority of these studies report a measure of racial discrimination referred to as an *odds ratio* (*OR*).¹⁵⁷ The odds ratio for a race-of-victim effect represents the odds that a prosecutor will file a death penalty notice against a defendant accused of killing a Caucasian victim, compared to the odds the prosecutor will file the death penalty against a defendant accused of killing a non-Caucasian victim, holding constant other factors relevant to the charging decision. This statistic is often used as a measure of discriminatory effect.¹⁵⁸ The formula is $OR = [P_W \div (1 - P_W)] \div [P_B \div (1 - P_B)]$, where P_W is the probability a death penalty eligible defendant charged with killing a Caucasian victim is noticed for the death penalty and P_B is the probability that a death-eligible defendant charged with killing an African American victim is noticed for the death penalty. By way of example, assume that $P_W = .5$ and $P_B = .3$. The odds ratio is $[(.5 \div .5) \div (.3 \div .7)]$ or 2.33. In other words, the *odds* of receiving a death notice are 2.33 times larger (or 133% more likely) if the victim is Caucasian than if the victim is African American. Alternatively, one might inquire about the relative odds of *not* receiving a death notice based on the victim’s race. One simply reverses the numerator and denominator for both odds calculations and the odds ratio becomes $[(.5 \div .5) \div (.7 \div .3)]$ or .428. This translates to a death eligible defendant accused of killing of a Caucasian victim having odds roughly 57% lower of *not* receiving a death notice than a similarly situated defendant accused of killing an African American victim.¹⁵⁹

The odds ratios for Caucasian-victim cases compared to African American-victim cases—or African American-defendant/Caucasian-victim compared to African American-defendant/African American-victim cases—reported in the statistical studies of capital charging decisions, range from

¹⁵⁶ James A. Fox, *Missing Data Problems in the SHR: Imputing Offender and Relationship Characteristics*, 8 HOMICIDE STUD. 214 (2004) (describing problems with official homicide data kept by local law enforcement agencies across the country); accord Wendy C. Regoeczi & Marc Riedel, *The Application of Missing Data Estimation Models to the Problem of Unknown Victim/Offender Relationships in Homicide Cases*, 19 J. QUANTITATIVE CRIMINOLOGY 155 (2003).

¹⁵⁷ RONET BACHMAN & RAYMOND PATERNOSTER, *STATISTICAL METHODS FOR CRIMINOLOGY AND CRIMINAL JUSTICE* 574 (1997).

¹⁵⁸ As explained *supra*, note 11, the Article distinguishes “disparate effect” from “discriminatory effect.”

¹⁵⁹ Judges and attorneys often misinterpret odds ratios as risk ratios (*RR*), but the two are distinct. Both statistics describe the likelihood that an event will occur, but they measure this likelihood on difference scales—somewhat akin to measuring temperature in terms of Fahrenheit versus Celsius. Risk ratios capture relative differences in probabilities, not odds. For this reason, risk ratios are often referred to as *relative risks*. See John M. Conley & David W. Peterson, *Science of Gatekeeping: The Federal Judicial Center’s New Reference Manual on Scientific Evidence*, 74 N.C. L. REV. 1183, 1219 (1996).

1.26 (Kentucky)¹⁶⁰ to 5.66 (Durham County, North Carolina).¹⁶¹ In between these jurisdictions, researchers discovered odds ratios of 1.48 (Los Angeles County, California),¹⁶² 2.0 (Maryland),¹⁶³ 2.21 (federal government),¹⁶⁴ 2.3 (Connecticut),¹⁶⁵ 2.38 (Missouri),¹⁶⁶ 2.64 (North Carolina),¹⁶⁷ 2.78 (Los Angeles County, California),¹⁶⁸ 3.0 (South Carolina),¹⁶⁹ 3.1 (Georgia),¹⁷⁰ 4.2 (Colorado),¹⁷¹ and 5.0 (San Joaquin County, California) (mean = 3.03; std. dev. = 1.3).¹⁷² The studies significantly varied in terms of the number of cases comprising their sample ($N = 120$ to $N = 4,929$; mean = 1,574; std. dev. = 1,706), the years investigated (from as early as 1969 to as late as 2010), the number of years covered (from 5 years to 35 years; mean = 13.8; std. dev. 9.6), the jurisdictional scope (from a single county to the entire nation), and the breadth of relevant non-racial variables accounted for in the statistical models (ranging from less than five non-racial controls to over 200);¹⁷³ nonetheless, the studies report striking consistency as it pertains to the effect of the victim's race on capital charging.

It is important to emphasize that the odds ratios reported above represent the *adjusted* racial gap in capital charging. In other words, the odds ratio is a measure of the magnitude of the difference in the odds of a death penalty notice between racial groups, holding constant other factors included the model; therefore, the aforementioned studies explicitly take into account rival non-racial explanations for the observed racial gap. As noted, *supra*, the race-of-victim effect is a simple measure of the residual

¹⁶⁰ Thomas J. Keil & Gennaro F. Vito, *Race and the Death Penalty in Kentucky Murder Trials: 1976-1991*, 20 AM. J. OF CRIM. JUST. 17 (1996).

¹⁶¹ Isaac Unah, *Choosing Those Who Will Die: The Effect of Race, Gender, and Law in Prosecutorial Decision to Seek the Death Penalty in Durham County, North Carolina*, 15 MICH. J. OF RACE & L. 135 (2009).

¹⁶² Robert E. Weiss et al., *Death Penalty Charging in Los Angeles County: An Illustrative Data Analysis Using Skeptical Priors*, 28 SOC. METHODS & RES. 91 (1999).

¹⁶³ Raymond Paternoster et al., *Justice by Geography and Race: The Administration of the Death Penalty in Maryland, 1978-1999*, 4 U. OF MD. L. J. OF RACE, RELIGION, GENDER & CLASS 1 (2004).

¹⁶⁴ U.S. Dep't of Just., *The Federal Death Penalty System: Supplementary Data, Analysis and Revised Protocols for Capital Case Review*, 14 FED. SENT'G REP. 40 (2001); Memorandum from David C. Baldus, To the Honorable Russell D. Feingold: Re DOJ Report on the Federal Death Penalty System (Jun. 11, 2001).

¹⁶⁵ John J. Donohue, *An Empirical Evaluation of the Connecticut Death Penalty System Since 1973: Are There Unlawful Racial, Gender, and Geographic Disparities?*, 11 J. EMPIRICAL LEGAL STUD. 637 (2014).

¹⁶⁶ Jon R. Sorensen & Donald H. Wallace, *Prosecutorial Discretion in Seeking Death: An Analysis of Racial Disparity in the Pretrial Stages of Case Processing in a Midwestern County*, 16 JUST. Q. 559 (1999).

¹⁶⁷ O'Brien et al., *supra* note 1.

¹⁶⁸ Nick Petersen, *Cumulative Racial and Ethnic Inequalities in Potentially Capital Cases*, CRIM. JUST. REV. 9 (forthcoming, 2017).

¹⁶⁹ Michael J. Songer & Issac Unah, *The Effect of Race, Gender, and Location on Prosecutorial Decisions to Seek the Death Penalty in South Carolina*, 58 S.C. L. REV. 161 (2006).

¹⁷⁰ BALDUS ET AL., *supra* note 25.

¹⁷¹ Meg Beardsley et al., *Disquieting Discretion: Race, Geography and the Colorado Death Penalty in the First Decade of the Twenty-First Century*, 92 DENV. U. L. REV. 15 (2015); Justin Marceau et al., *Death Eligibility in Colorado: Many Are Called, Few Are Chosen*, 84 U. COLO. L. REV. 1069 (2013).

¹⁷² Catherine Y. Lee, *Hispanics and the Death Penalty: Discriminatory Charging Practices in San Joaquin County, California*, 35 J. CRIM. JUST. 17 (2007).

¹⁷³ The mean and standard deviation for the number of statistical controls is not reported because the studies differ in how controls are reported and counted, so precise comparisons are precluded.

gap in death charging behavior, and can be interpreted as a measure of discriminatory effect.¹⁷⁴ The standard approach when investigation capital charging dynamics is to specify a logistic regression model positing that the probability of receiving a death notice is a function of set of explanatory variables:

$$\log_e \left(\frac{P(N)}{1 - P(N)} \right) = \beta_0 + \beta_k X_k + \beta_2 CV, \quad (1)$$

where $P(N)$ is the probability that a death notice is filed, X_k is a vector of k explanatory variables, CV is a binary variable that indicates whether the victim is Caucasian, β_k (*beta*) is a vector of k regression coefficients, $\log_e \left(\frac{P(N)}{1 - P(N)} \right)$, is the natural logarithm of the odds that a death notice is filed, conditional on the explanatory variables, X and CV .¹⁷⁵ The vector of explanatory variables typically includes a wide range of aggravating and mitigating evidence relevant to the crime and the defendant's background. The inverse natural logarithm (the anti-logarithm) of the coefficient for CV , e^{β_2} , is the odds ratio reported in the aforementioned studies.

The intuitive appeal of this framework is that it provides a single measure of the unexplained racial gap based on systemic disparate treatment.¹⁷⁶ The model is formulated to take into account factors purported to drive the death-noticing calculus, so it performs the function of assessing rival hypotheses. Greater confidence in the inferences of discrimination drawn from these models is achieved when the analyst includes a wide of variables that could potential explain the race gap—that is, variables that are likely correlated with race and the likelihood of a receiving a death notice.¹⁷⁷ As mentioned, *supra*, a model need not take into account every conceivable variable, but the inclusion of key explanatory variables should be guided by both doctrine and the extant empirical literature.¹⁷⁸ Despite its intuitive appeal, the model specification suffers from two significant shortcomings that potentially undermine our ability to better understanding racial disparities: (1) the assumption of homogenous treatment effects and (2) the inability to unpack the observable behavioral dynamics responsible for the generating the racial gap. These two shortcomings are discussed below.

¹⁷⁴ See *supra* note 158

¹⁷⁵ J. SCOTT LONG, REGRESSION MODELS FOR CATEGORICAL AND LIMITED DEPENDENT VARIABLES 49 (1997).

¹⁷⁶ Elder et al., *supra* note 24.

¹⁷⁷ See Part II.B.

¹⁷⁸ See Part I. *McCleskey v. Kemp*, 481 U.S. 279, 328 (1987) (Brennan, J., dissenting) (“[A] multiple-regression analysis need not include every conceivable variable to establish a party’s case, as long as it includes those variables that account for the major factors that are likely to influence decisions.”); accord *Bazemore v. Friday*, 478 U.S. 385, 400 (1986) (“[I]t is clear that a regression analysis that includes less than ‘all measurable variables’ may serve to prove a plaintiff’s case. A plaintiff...need not prove discrimination with scientific certainty; rather, his or her burden is to prove discrimination by a preponderance of the evidence.”).

B. *Persistent Pitfalls of Models of Racial Discrimination in Capital Charging*

1. *Homogenous Effects*

The standard statistical model employed to examine discrimination in capital charging-and-sentencing rests on the questionable assumption that the case factors have the same influence (i.e., coefficients) across both groups of cases—i.e., homogenous effects.¹⁷⁹ The group indicator variable, *CV*, captures the differences in the average value of the outcome variable, $\log_e \left(\frac{P(N)}{1-P(N)} \right)$ (i.e., the log odds of a death notice), after holding the effects of other variables in the model constant, but it says nothing about the potential heterogeneous effects of the non-racial explanatory variables across the groups. The regression coefficients, β_k , in these analyses represent a weighted average of the effects across the groups, but will fail to capture the true effects for either group when those effects differ.¹⁸⁰ And as a consequence, the model that investigates the groups together (i.e., the pooled model) may misrepresent *both* the size of the racial disparity¹⁸¹ and the predictability of charging-behavior for each group based on relevant non-racial variables.¹⁸² Prior work on inconsistency and irrationality in capital charging behavior has revealed that relevant aggravation and mitigation evidence does a much better job of explaining prosecutorial decision-making when the victim is Caucasian than when the victim is African American.¹⁸³ Differences in predictability likely stem from the fact that the decision-making process for one group is more idiosyncratic than the other, and this can be interpreted as the level of rationality governing the process is dependent on the victim's race.¹⁸⁴ Properly analyzing racially-heterogeneous effects not only provides better measures of the effects of the explanatory variables across racial groups, but also improves the overall predictive power of the statistical model—an important concern of many courts evaluating statistical evidence.¹⁸⁵ The implausibility of the homogenous effects assumption is underscored by both qualitative¹⁸⁶ and

¹⁷⁹ Elder et al., *supra* note 24.

¹⁸⁰ *Id.*

¹⁸¹ For example, David Baldus and his colleagues discovered that the effect of the victim's race in the capital charging-and-sentencing process was not uniform across the spectrum of homicide cases. Racial disparities were strongest in the mid-range of cases—i.e., cases that were *neither* the least aggravated *nor* the most aggravated, but somewhere in the middle. BALDUS ET AL., *supra* note 25, at 145, 154.

¹⁸² DON HEDEKER & ROBERT D. GIBBONS, LONGITUDINAL DATA ANALYSIS 158, 195–96 (2006) (emphasizing the importance of looking beyond regression coefficients when comparing groups and determining whether groups differ in terms of the degree of unexplained variation).

¹⁸³ Sherod Thaxton, *Disciplining Death: Assessing and Ameliorating Arbitrariness in Capital Charging*, 49 ARIZ. ST. L.J. 137, 179–80 (2017) (reporting that models of capital charging behavior have different predictive power depending on the race of the victim). See also Part VI.C.

¹⁸⁴ Thaxton, *supra* note 183.

¹⁸⁵ See Part II.B.

¹⁸⁶ See *supra* notes 59–61 and accompanying text.

quantitative¹⁸⁷ studies of racial discrimination in the employment context that has discovered that impact of job qualifications on hiring and wages systematically varies across racial groups.

The homogenous effects assumption can be relaxed by permitting the group variable to condition (i.e., moderate) the impact of one or more explanatory variables, but the model becomes difficult to estimate and interpret when the group variable is believed to condition the effect of more than a couple of explanatory variables because the number of regression coefficients, β_k , becomes too large. For example, if the group variable, such as whether the case involved a Caucasian victim (*CV*), is believed to condition the impact of four explanatory variables, then four additional parameters must be estimated in the model that represent the interaction among these factors, as well as the five “main effects”:

$$\begin{aligned} \log_e \left(\frac{P(N)}{1 - P(N)} \right) = & \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \\ & + \beta_5 CV + \beta_6 (X_1 \times CV) + \beta_7 (X_2 \times CV) \\ & + \beta_8 (X_3 \times CV) + \beta_9 (X_4 \times CV), \end{aligned} \quad (2)$$

where $\beta_k X_k$ and $\beta_5 CV$ and the main effects, and $\beta_k (X_k \times CV)$ are the interaction effects that capture the differences in the impact of the nonracial explanatory variables across cases with Caucasian victims and non-Caucasian victims. So, for example, if we assume that X_1 is a variable representing whether the defendant had a monetary motive for the homicide (assume the case involves a single offender), then β_1 is the impact of a monetary motive on the odds of the prosecutor filing a notice of intent to seek the death penalty and $\beta_5 CV$ is the effect of the presence of a Caucasian victim in the case (assume the case involves a single victim), holding the other variables constant. $\beta_6 (X_1 \times CV)$ represents the *difference in the effect* of β_1 (monetary motive) on the log odds of a death penalty notice being filed when the case involves a Caucasian victim compared to when the case involves a non-Caucasian victim. In other words, β_6 does not have an independent effect on the log odds of the prosecutor filing a death penalty notice, rather it must be combined with the β_1 to determine the effect of β_1 on Caucasian-victim cases: $(\beta_1 + \beta_6 CV)$. If the victim in the case is non-Caucasian, $CV = 0$, then the effect of monetary motive is $(\beta_1 + \beta_6 CV) = (\beta_1 + [\beta_6 \times 0]) = \beta_1$.

The central problem with estimating the aforementioned model is that the new parameters included in the model to capture effect heterogeneity rely on the inclusion of variables that are the product terms of the nonracial explanatory variables and the group variable: $\beta_k (X_k \times CV)$. These newly included variables are highly correlated with their constituent variables, thus

¹⁸⁷ Roland G. Fryer et al., *Racial Disparities in Job Finding and Offered Wages*, 56 J. L. & ECON. 633, 635 (2013) (reporting that one-third of the black-white wage gap is attributable to differential treatment); accord Lang & Manove, *supra* note 138, at 1490. See also Lynch & Haney, *supra* note 62 at 351 (utilizing a mock juror design and discovering that Caucasian jurors differentially assess mitigation evidence according to the race of the victim when deciding to impose the death penalty).

making it extremely difficult, and sometimes impossible, to estimate the separate effects of each variable.¹⁸⁸ Rather than estimating a statistical model with interactive effects between the victim's race and the nonracial variables, a more sensible approach is to estimate separate models for each race-of-victim group. This approach minimizes the collinearity problem and is feasible when each group is sufficiently large to accommodate a wide range of relevant cases characteristics.

An additional advantage of estimating the models separately for African American- and Caucasian-victim cases has to do with the interpretation of the race-of-victim effect when the model assumes that the race-of-victim influences the impact of multiple case characteristics in the model. In the previous example, $\beta_5 CV$ is interpreted as the effect of the case having a Caucasian victim on the odds of a capital charge when $X_1 = 0$; that is, when the defendant did *not* have a monetary motive: $(\beta_5 + [\beta_6 \times 0]) = \beta_5$. When β_5 is also interacted with the three other variables in the model (X_2, X_3, X_4), the race-of-victim effect is interpreted as the effect of having a Caucasian victim on the probability of a death penalty notice when *all* of those other variables are "zero": $X_1 = X_2 = X_3 = X_4 = 0$. In some situations, the interpretation will be straightforward because a "zero" value on the variable has substantive meaning (e.g., a binary variable representing the presence or absence of a case characteristic); however, in other situations, a "zero" value for a variable will lack any substantive meaning.¹⁸⁹ In either case, the race-of-victim is, itself, a conditional effect rather than the average effect of the variable across the range of other variables in the model.

Another important, yet often overlooked, shortcoming with the aforementioned statistical model is that it does not take into account the fact that race-of-victim differences in case characteristics (data) and differences in the effects of those characteristics (parameters) may occur simultaneously—that is, these components can influence the racial gap in capital charging *jointly* rather than independently. In other words, the influence of race-of-victim differences in case characteristics and treatment of those characteristics on capital charging is greater than their simple summation. The analytical approach that I advocate in this Article explicitly takes this possibility into account.¹⁹⁰

A key advantage of the heterogeneous effects approach is that one can observe how prosecutors differentially treat the various non-racial case characteristics based on the victim's race. But the approach, standing alone, cannot reveal how much of the racial disparity in capital charging is attributable to differential treatment and how much is attributable to the fact that the groups, on average, differ in terms of those relevant non-racial case

¹⁸⁸ JOHN FOX, APPLIED REGRESSION ANALYSIS, LINEAR MODELS, AND RELATED METHODS 22, 425 (1997).

¹⁸⁹ A popular approach is to subtract the average value of each variable from itself, so the "zero" represents the average value of the variable. LEONA S. AIKEN & STEPHEN G. WEST, MULTIPLE REGRESSION: TESTING AND INTERPRETING INTERACTIONS (1991).

¹⁹⁰ See Part IV.A.

characteristics. It may still be the case that, even if Caucasian-victim and African American-victim cases were treated similarly by prosecutors, a substantial racial disparity would remain because, on average, the nonracial characteristics of those cases—relating to aggravation and mitigation—substantially differ. This highlights the other limitation of existing research on capital charging dynamics which I describe in more detail below.

2. *The “Black Box” of Disparities*

The second significant shortcoming of prior scholarship examining death penalty charging-and-sentencing dynamics has been its inability to empirically unpack the observed racial gaps in a manner most useful for litigation and legal reform. The Court’s recent decisions suggest that a more nuanced and targeted analysis of system-wide disparities may be required in order for claimants to prevail.¹⁹¹ As noted, *supra*, scholars are now suggesting that the Court’s current systemic discrimination framework requires plaintiffs to provide deeper meaning to observed patterns.¹⁹² In other words, claimants must both explain how a system is vulnerable to discriminatory practices *and* how the discrimination has influenced actual decision-making. Generally speaking, motive-based theories of racial discrimination cannot be empirically tested; however, as sociologist Barbara Reskins has argued, “explanation requires including the specific processes that link groups’ ascribed characteristics to variable outcomes [and] redirecting our attention from motives to [these specific processes] is essential for understanding inequality and—equally important—for contributing meaningfully to social policies that will promote social equality.”¹⁹³

The bulk of the discussion of empirical research finding racial disparities in charging-and-sentencing has focused on issues of model misspecification rather than attempting to carefully link racial status to outcomes.¹⁹⁴ While it is beyond dispute that potential bias from the omission of important explanatory variables that may be correlated with the victim’s race must be carefully considered,¹⁹⁵ “it is difficult to imagine that a few covariates exist that if included as predictors would lead to clear and justified distinctions between defendants who are charged with a capital crime and defendants who are not; likewise for death sentences.”¹⁹⁶ My prior discussion of the extant research literature underscores the fact that the race-of-victim effect on capital charging is largely consistent across study

¹⁹¹ *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 357 (2011) (explaining that the “plaintiff must begin by identifying the specific employment practice that is challenged”); *see also McCleskey v. Kemp*, 481 U.S. 279 (1987) (rejecting evidence of systemic racial discrimination in the capital charging-and-sentencing process, in part, because of the inability of the plaintiff to specify the source of the disparity).

¹⁹² *See* Part I; Selmi, *supra* note 26.

¹⁹³ Reskin, *supra* note 31, at 1.

¹⁹⁴ *See* Part II.B.

¹⁹⁵ *See* Part I; Richard A. Berk, *Randomized Experiments as the Bronze Standard*, 1 J. EXPERIMENTAL CRIMINOLOGY 417, 428 (2005).

¹⁹⁶ Richard A. Berk et al., *Statistical Difficulties in Determining the Role of Race in Capital Cases: A Reanalysis of Data from the State of Maryland*, 21 J. QUANTITATIVE CRIMINOLOGY 365, 387 (2005).

designs, so it does not appear that quantitative estimates of racial discrimination are unduly sensitive to the inclusion or exclusion of relevant factors when a modest number of non-racial variables are taken into account.¹⁹⁷ A similar concern over an apparent preoccupation with potential omitted variables when investigating persistent gender differences in earnings led sociologist Thomas Daymont and economist Paul Andrisani to offer the following admonition:

[S]uch attempts [to incorporate potentially relevant omitted variables] have not produced any substantial reductions in the size of the unexplained earnings gap. Differences in college majors, training, individual personality traits and tastes failed to account for the gender gap. Thus, after many empirical attempts spanning more than a decade, researchers are still unable to account for more than about half of the male-female difference in earnings through differences in productivity-related variables. For some, this constitutes compelling evidence that labor market discrimination is the primary factor producing earnings inequality. Others remain unconvinced, however, believing that some important productivity-related factors have either been omitted or measured imprecisely.¹⁹⁸

Rather than remaining embroiled in this “explanatory stalemate,”¹⁹⁹ more attention should be devoted to quantifying the relative contributions of disparate effect and disparate treatment in explaining racial disparities in capital charging.²⁰⁰ This is possible through identifying the specific processes that link differences in race to differences in capital charging decisions. Racial disparities in capital charging can be attributed to the shifting standards phenomenon in which “the evaluation of ‘objective’ information can be affected by underlying racial considerations.”²⁰¹ This explanation can take two different, yet complementary, forms. The first is a general assessment of disparate effect and disparate treatment that examines relevant case characteristics in the aggregate. That is, an inquiry into the differential influence of the group-specific composition of case attributes and behavioral responses to those attributes evaluated as a whole for, respectively, the disparate effect and disparate treatment components. The

¹⁹⁷ See notes 160-172 and related text.

¹⁹⁸ Thomas N. Daymont & Paul J. Andrisani, *Job Preferences, College Major, and the Gender Gap in Earnings*, 19 J. HUM. RESOURCES 408, 409-10 (1984).

¹⁹⁹ Reskin, *supra* note 31, at 1 (noting that social scientists’ pre-occupation with motive-based explanations of race and sex disparities have contributed to an “explanatory stalemate”).

²⁰⁰ See *supra* note 11.

²⁰¹ See *supra* note 62 and accompanying text. See generally Monica Biernat & Melvin Manis, *Shifting Standards and Stereotype-Based Judgments*, 66 J. PERSONALITY & SOC. PSYCHOL. 5, 5 (1994) (explaining that shifting standards occur when evaluators make judgements based on subjective criteria that maximize differentiation between groups based on race and gender). Some scholars have argued that racial balance on juries is necessary to limit bias against African American criminal defendants because “jurors of one race, even those well-intended and free of racial animus, will be unable to dependably judge the demeanor of a witness of a different race because they are unable to accurately decipher the cues that the witness uses to communicate sincerity.” See, e.g., Rand, *supra* note 62, at 5.

second, and potentially more illuminating, approach is a detailed assessment of the unique contribution of each case attribute in terms its disparate effect and disparate treatment. This analytical framework, which I describe in the following section, is known as “regression decomposition” because it partitions an observed racial disparity into discriminatory and non-discriminatory components. This approach permits closer examination of the “black box” of racial disparities, providing improved insight into *how* racial discrimination influences death charging behavior because the “decomposition [framework] explores potential mechanisms in more detail than a conventional analysis.”²⁰²

IV. ANALYTICAL APPROACH

A. *Disaggregating the Sources of Racial Disparity*

Multivariate decomposition techniques, also called “regression standardization,” have been used for well over a half-century in social scientific research to quantify the contributions to group differences. These techniques were initially introduced by sociologists in the 1950s but popularized by economists in the early 1970s.²⁰³ The most popular iteration of the approach is attributed to economists Ronald Oaxaca and Alan Blinder, and as a result many social scientists refer to the technique as the Oaxaca-Blinder Decomposition. The approach parcels out the components of a group difference in a statistic, such as a mean or proportion, into compositional differences between groups (i.e., differences in the characteristics of the groups) and differences in the returns on the characteristics (i.e., differences in behavioral responses by the decision-makers). The group differences in returns on those characteristics can be further disaggregated into a component that accounts for the fact that differences in characteristics and differences in returns on those characteristics exist simultaneously between the groups. Stated differently, this third component indicates how much of the gap can be accounted for by the fact that the returns to one group (e.g., Caucasians) tends to be greater for those characteristics for which compositional differences are the strongest.²⁰⁴ Decomposition techniques have been most commonly applied to research on wage differentials for the purpose of understanding the relative importance of group differences in levels of certain characteristics (e.g., education, tenure, prior work experience) and group differences in the returns on those characteristics. More recently, the decomposition approach

²⁰² Taber et al., *supra* note 22, at 725.

²⁰³ See *supra* note 29 and accompanying text.

²⁰⁴ Hailman H. Winsborough & Peter Dickinson, *Components of Negro-White Income Differences*, PROC. OF THE AM. STAT. ASSOCIATION, SOC. STAT. SEC. 6 (Edwin G. Goldfield ed., Washington, D.C. 1971). An alternative decomposition results from the concept that the coefficients from the pooled model represents the nondiscriminatory coefficient vector, and this vector should be used to determine the contribution of differences in the predictor variables. This results in a two-fold decomposition where the first component and second components are differentials relative to the overall baseline. Ben Jann, *The Blinder-Oaxaca Decomposition for Linear Regression Models*, 8 STATA J. 453, 455 (2008).

has been applied to racial differences in sentencing,²⁰⁵ the use of post-acute rehabilitation care,²⁰⁶ alcohol treatment completion,²⁰⁷ and drug treatment commitments.²⁰⁸ To the best of my knowledge, the approach has yet to be applied to capital charging or sentencing.

In the context of capital charging, the first component of the decomposition, commonly referred to as an “endowment effect,” can be interpreted as the proportional change in the likelihood of receiving a death notice that would result if the average African American-victim case had the same characteristics as the average Caucasian-victim case, but there was no change in the manner in which African American-victim cases were treated by prosecutors.²⁰⁹ The second component, called the “coefficient effect,” is interpreted as the proportional change in the likelihood of receiving a death notice that would occur if African American-victim cases were treated similarly as Caucasian-victim cases, but there was no change in the average characteristics of African American-victim cases. The (optional) third component examines the simultaneous change in both the endowment effect and the coefficient effect, and describes how much of the racial disparity can be accounted for by the fact that racially differential treatment by prosecutors tends to be stronger in situations where racial differences in observable case characteristics are most pronounced.²¹⁰ The three components can examine the variables in the aggregate (i.e., summing over all case characteristics) or individually.²¹¹ A more formal treatment of the decomposition technique is presented below.

Assume there are two groups of death penalty-eligible cases, W and B , representing cases with Caucasian victims and African American victims, respectively; an outcome variable, N , that takes on the value of “1” if a death notice is filed against the defendant in the case, and “0” if otherwise; and a set of explanatory variables for the death penalty charging decision, X , that indexes aggravating and mitigation factors relevant the defendant’s degree of culpability. The gap, G , in the average outcome, \bar{N} , between W and B is: $G = \bar{N}_W - \bar{N}_B = P(N_W) - P(N_B)$, where $P(\cdot)$ is the probability that cases in each group receive a notice for the death penalty. G can also be expressed as the difference in the regression predictions of the group specific means:

$$G = P(N_W) - P(N_B) = \mathcal{F}(\overline{X_W \times \beta_W}) - \mathcal{F}(\overline{X_B \times \beta_B}), \quad (3)$$

²⁰⁵ Sorensen et al., *supra* note 29.

²⁰⁶ George M. Holmes et al., *Decomposing Racial and Ethnic Disparities in the Use of Postacute Rehabilitation Care*, 47 HEALTH SERVICES RES. 1158 (2012).

²⁰⁷ Jerry Owen Jacobson et al., *A Multilevel Decomposition Approach to Estimate the Role of Program Location and Neighborhood Disadvantage in Racial Disparities in Alcohol Treatment Completion*, 64 SOC. SCI. & MED. 462 (2007).

²⁰⁸ MacDonald et al., *supra* note 29.

²⁰⁹ For the purposes of this Article, I define disparate effect as the endowment effect. *See supra* note 11.

²¹⁰ Winsborough & Dickinson, *supra* note 204.

²¹¹ *See, infra*, notes 218-220 and accompanying text.

where the subscripts index the groups, β is a vector of k regression coefficients corresponding to $k - 1$ explanatory variables, X , and F is the logistic function, $F(\cdot) = \frac{e^{(\cdot)}}{1+e^{(\cdot)}}$, that relates the effects of β to changes in the probability of observing a particular outcome.²¹² To identify the contribution of group differences in predictors to the overall outcome difference, the terms can be rearranged as follows:

$$G = \underbrace{\mathcal{F}(X_W \times \beta_W) - \mathcal{F}(X_B \times \beta_W)}_E + \underbrace{\mathcal{F}(X_W \times \beta_W) - \mathcal{F}(X_W \times \beta_B)}_C, \tag{4.1}$$

where E and C represent the endowment and coefficient effects, respectively. The terms in Equation 4.1 can be rearranged to underscore the distinctions that each component captures:

$$G = \underbrace{\mathcal{F}(X_W - X_B)\beta_W}_E + \underbrace{\mathcal{F}\{X_W \times (\beta_W - \beta_B)\}}_C. \tag{4.2}$$

Equations 4.1 and 4.2 are alternative ways of representing the two-fold composition because they do not consider the portion of the gap that is attributable to the simultaneous influence of E and C .²¹³ The three-fold composition is:

$$G = E + C + \underbrace{\mathcal{F}\{(X_W - X_B) \times (\beta_W - \beta_B)\}}_{CE}, \tag{5}$$

where CE is the interaction between the group differences in endowment and coefficient effects. Both C and CE may be attributed to discrimination,²¹⁴ but it is also important to recognize that the terms capture the potential effects of group differences in unobserved variables.²¹⁵ The decompositions in Equations 4.1, 4.2, and 5 are formulated from the viewpoint of group W , meaning that the group differences in predictors are weighted by the coefficients of group W to determine the endowment effect, E . Similarly, for the coefficient effect, C , the differences in the coefficients are weighted by group W 's predictor levels. The differential could also be expressed from the viewpoint of group B .

An alternative decomposition is possible that uses coefficients from the pooled model as the nondiscriminatory coefficients, and these coefficients

²¹² Alternatively, one could model the gap in the likelihood of a prosecutor filing a notice to seek the death penalty as the difference in the log odds: $G = N_W - N_B = \log_e \left(\frac{P(N_W)}{1-P(N_W)} \right) - \log_e \left(\frac{P(N_B)}{1-P(N_B)} \right)$, which would be consistent with the logistic regression formulation in Equation 1. I prefer the modeling the gap in terms of the differences in probabilities rather than log odds because this differential is both easier to understand and becomes necessary when comparing the effects of specific explanatory variables across groups. For an accessible discussion of the decomposition method for dichotomous variables, see Daniel A. Powers et al., *Mvdcmp: Multivariate Decomposition for Nonlinear Response Models*, 11 STATA J. 556, 564-69 (2011).

²¹³ Recall that E is calculated while holding C constant; similarly, C is calculated while holding E constant.

²¹⁴ See *infra* note 255 and accompanying text.

²¹⁵ See *infra* Part VI.D.

are used to determine the contribution of differences in the predictor variables.²¹⁶ Some scholars advocate using the pooled coefficients as the baseline because there usually is no *a priori* reason to select one group as the baseline over the other when measuring discrimination.²¹⁷ This yields in a two-fold decomposition where the first component and second components are differentials relative to the overall baseline, β_P :

$$G = \underbrace{\mathcal{F}(X_W \times \beta_P) - \mathcal{F}(X_B \times \beta_P)}_E + \underbrace{\mathcal{F}\{[X_W \times (\beta_W - \beta_P)] - [X_B \times (\beta_P - \beta_B)]\}}_C. \quad (6.1)$$

Similar to Equation 4.1, the terms in Equation 6.1 can be rearranged in order to more clearly emphasize what is being measured by each component on the right-hand side of the equation:

$$G = \underbrace{\mathcal{F}(X_W - X_B)\beta_P}_E + \underbrace{\mathcal{F}\{[X_W \times (\beta_W - \beta_P)] - [X_B \times (\beta_P - \beta_B)]\}}_C. \quad (6.2)$$

The above decompositions have been described at the aggregate level, but as mentioned, *supra*, understanding the unique contribution of each explanatory variable may also be of interest.²¹⁸ For example, one might want to know how much of the race-of-victim gap in death noticing behavior is due to differences in aggravating evidence or mitigation evidence. And even within those categories, one might be interested in uncovering how much of the gap is explained by the number of statutorily defined special circumstances present in the case and how much is due to the defendant's criminal history (i.e., endowment effects). Similarly, it might be useful to determine how much of the unexplained gap is related differences in prosecutors' behavioral response to those particular case factors (i.e., coefficient effects). The intuition underlying the detailed decomposition is that the total component is a summation of the individual contributions, although the specific approaches to detailed decompositions differ depending on whether the outcome variable in model is continuous or categorical.²¹⁹ This makes it possible to also examine the endowment and coefficient effects of predictors in batches, which may be especially appealing in the context of capital charging because one is able to examine the impact of a collection of thematically related variables (e.g., several

²¹⁶ Jann, *supra* note 204, at 455. The pooled decomposition includes race-of-victim in the model as an additional control variable to account for differences in group-specific intercepts. Failing to do so would cause the influence of endowments to be overstated and unexplained differences to be understated. Elder et al., *supra* note 24, at 285, 288.

²¹⁷ Jann, *supra* note 204, at 457.

²¹⁸ Daniel A. Powers & Myeong-Su Yun, *Multivariate Decomposition for Hazard Rate Models*, 39 SOC. METHODOLOGY 233, 238 (2009) (discussing potential uses for detailed decompositions).

²¹⁹ *Id.*

factors related to the defendant's family background or multiple factors related to the commission of the crime).²²⁰

In Part V, I apply the aforementioned multivariate decomposition approach to actual death penalty charging data in Georgia. These analytical techniques permit me to ascertain (a) the magnitude of victim-based racial disparities in Georgia, (b) the influence of the differential distribution of relevant aggravation and mitigation evidence across Caucasian-victim and African American-victim cases on the racial disparity, and (3) the influence of prosecutors' differential behavioral responses to the aggravating and mitigation evidence on the racial disparity. As explained, *supra*, social scientists interpret this differential behavioral response as racial discrimination. The decomposition technique, like any other regression-based approach, is unable to directly test whether this differential behavioral response is attributable to racial animus, and the Court has never adopted such a requirement;²²¹ although, the analytical tools I employ may make such an inference even more plausible than inferences based on prior statistical studies.²²² But before applying the model to the Georgia data, I describe the specific information contained in the data that is relevant to carefully scrutinizing prosecutorial charging decisions.

B. Georgia Capital Charging Data

I collected data on 1,238 potential death penalty cases in Georgia over an eight-year period (1993-2000) to examine the extent and sources of victim-based racial disparities impacting the capital charging process. Relevant case-level data on all potentially capital cases from which prosecutors could identify and select defendants for the death penalty were compiled from five separate sources: the Georgia Bureau of Investigation, the Georgia Department of Corrections, the Office of the Georgia Capital Defender, the Clerk's Office of the Georgia Supreme Court, and the Atlanta Journal-Constitution newspaper.²²³ Death-eligibility was determined by the presence of *at least* one of eleven crime elements listed in Georgia's death penalty statute for defendants 17 years of age or older.²²⁴ As required by

²²⁰ Jann, *supra* note 204 (explaining the aggregation of individual variables into subsets in order to capture the collective contributions of those variables to the endowment and coefficient components).

²²¹ *McCleskey v. Kemp*, 481 U.S. 279, 293-94 (1987) (acknowledging that statistics can be used to prove intentional discrimination if the evidence is compelling).

²²² Taber et al., *supra* note 22, at 725 (noting that the decomposition framework explores potential causal mechanisms in more detail than conventional analysis).

²²³ For a detailed description of these sources, see Thaxton, *supra* note 183; Thaxton, *supra* note 10.

²²⁴ GA. CODE ANN. § 17-10-30(a), b(1)-(10). The enumerated aggravating circumstances in the Georgia capital statute are: (a) the death penalty may be imposed for the offenses of aircraft hijacking or treason in any case; (b1) the offense of murder, rape, armed robbery or kidnapping was committed by a person with a prior record of conviction for a capital felony; (b2) the offense of murder, rape, armed robbery or kidnapping was committed while the offender was engaged in the commission of another capital felony or aggravated battery, or the offense of murder was committed while the offender was engaged in the commission of burglary or arson in the first degree; (b3) the offender, by his act of murder, armed robbery, or kidnapping, knowingly created a great risk of death to more than one person in a public place by means of a weapon or device which would normally be hazardous to the lives of more than one person; (b4) the offender committed the offense of murder for himself or another, for the purpose of receiving money or any other thing of monetary value; (b5) the murder of a judicial officer,

law, whenever a prosecutor intends to seek the death penalty against a defendant, the prosecutor must file a formal notice with the Clerk's Office of the Georgia Supreme Court, and the Clerk's Office keeps a record of all notices submitted—prosecutors filed notices in 400 cases.²²⁵ The data consist of the *entire* population, and not a mere sample, of homicide cases during the years under investigation. The major benefit of analyzing the entire population of homicide cases is that statistical inference based on sample statistics (e.g., *p*-values, significance tests, confidence intervals, etc.) does not apply in the convention sense, so the focus is on the direction and magnitude of the statistical parameters and quantities of interest derived from these parameters.²²⁶ I selected 1993 as a starting point because the Georgia legislature enacted its life without the possibility of parole (LWOP) statute in 1993, and the law was specifically designed as a sentencing alternative reserved only for capital murder trials.²²⁷ Because juries (and judges if the defendant opted for a bench trial) were only permitted to impose a sentence of LWOP if the prosecutor officially sought the death penalty against the defendant, the statute potentially had a substantial impact on prosecutors' calculi when deciding whether to seek the death penalty. I concentrate on cases after the statute was enacted so the governing statutory regime is consistent across all of the cases. The year 2000 was chosen as a cut-off point in order to allow sufficient time for all of the cases to advance from the charging phase through the initial sentencing phase. Of the 1,238 death-eligible murder cases, roughly 45% involved at least at least one Caucasian victim and 50% involved African American victims.²²⁸ Eighty-

former judicial officer, district attorney or solicitor-general, or former district attorney, solicitor, or solicitor-general was committed during or because of the exercise of his or her official duties; (b6) the offender caused or directed another to commit murder or committed murder as an agent or employee of another person; (b7) the offense of murder, rape, armed robbery, or kidnapping was outrageously or wantonly vile, horrible, or inhuman in that it involved torture, depravity of mind or an aggravated battery to the victim; (b8) the offense of murder was committed against any peace officer, corrections employee or firefighter while engaged in the performance of his official duties; (b9) the offense of murder was committed by a person in, or who has escaped from, the lawful custody of a peace officer or place of lawful confinement; and (b10) the murder was committed for the purpose of avoiding, interfering with or preventing a lawful arrest or custody in a place of lawful confinement, of himself or another.

In 2006, Georgia's capital statute was amended to include an additional aggravating circumstance: "the offense of murder, rape, or kidnapping was committed by a person previously convicted of rape, aggravated sodomy, aggravated child molestation, or aggravated sexual battery." 2006 Ga. Laws 571, § 22; GA. CODE ANN. § 17-10-30(b)(11).

²²⁵ For a description of Georgia's capital punishment process from initial appearance through execution, see Appendix B.

²²⁶ Steffensmeier & Demuth, *supra* note 1, at 160 (explaining that sample-based significance tests are inappropriate when analyzing the entire population of cases). Another source of uncertainty in the estimation of the model parameters is the specification of the model—e.g., the choice and measurement of variables. Modifying the features of the models will result in many plausible models and yields a distribution of estimates. Cristobal Young & Katherine Holsteen, *Model Uncertainty and Robustness: A Computational Framework for Multimodel Analysis*, 46 SOC. METHODS & RES. 3, 32 (2017). I address this form of uncertainty in VI.D. See also *infra* notes 236 and 307 and accompanying text.

²²⁷ 1993 Ga. Laws 569, § 4; Ga. Code Ann. § 17-10-30.1 (1993). The 1993 statute was modified in 2005 to allow LWOP as a sentencing option in non-death penalty cases. This statutory change occurred after the period under investigation in the current study.

²²⁸ Similar with prior studies, I code cases involving at least one Caucasian victim as a Caucasian-victim case. The results do not appreciably change when I coded a multi-victim case with victims of different races/ethnicities as a multi-racial case. See also Part VI.A.

three percent of the cases involved a single victim, and when I limited my analyses to these cases, 44% involved a Caucasian victim and 50% involved an African American victim.

As is standard in the extant literature on capital charging, I model the likelihood that a prosecutor files a death penalty notice against a defendant as a function of defendant characteristics, crime characteristics, and victim characteristics. The largest model includes 40 case-level variables indexing the heinousness of the crime and the culpability of the defendant.²²⁹ The first category, *crime-related factors*, includes statutorily defined aggravating factors, circumstances of the murder, type of murder weapon, motive for killing, type of evidence, strength of evidence,²³⁰ and jurisdiction where killing occurred. The second category, *defendant-related factors*, encompasses the number of defendants, defendant's sex, age, race/ethnicity, level of education, employment status, marital status, number of children, military service, history of drug use, psychiatric status, IQ score, troubled family history, prior felony conviction, county of residence, and trigger-person status.²³¹ The third and final category, *victim-related factors*, contains the number of victims, sex, age, race/ethnicity, and prior relationship with defendant. These factors can also be grouped in terms of

²²⁹ For a description of the variables, see Appendix A. The Georgia dataset includes much more information than the 40 variables included in the model specification. Moreover, the model actually includes more information than the 40 variables imply because I employ a conservative counting method in order to reduce the number of parameters that must be estimated in the model. For example, in terms of inculpatory/aggravation evidence, I have information on the presence or absence of the eleven statutorily defined special circumstances enumerated in Georgia's capital statute, but rather than count them separately, I combined them into a single variable that indexes the total number of statutory aggravating circumstances present in the case. A potential complication with this approach is that it implies that all of the aggravating factors have equal weight in the overall composite measure, and this may not accurately reflect how the factors influence capital charging. It is common practice in statistics to use a summation scale when the individual items have low variability or are highly correlated (or both) – this is the case with the individual items in the statutory aggravating circumstance scale. Table 1 reports that the average number of statutory aggravating circumstances in a case is 2.2 and the range is 1-7; however, only three of the eleven statutory aggravating circumstances were present in more than 10% of the cases. As a result, there is little variability in the factors to access the bulk of the cases and many of the cases were charged with identical aggravating circumstances. Thus, the structure of the data made estimating the individual effects for all of the factors in a single model infeasible. Furthermore, the research literature suggests that the number statutory aggravating circumstances is a better predictor of death penalty charging and sentencing behavior than the individual items, see *infra* note 232. I re-examined the Georgia with the individual items, rather than the composite scale, and obtained results that were nearly identical across the two models with respect to race-of-victim effect. See *infra* note 304. Equally important is that the model with the composite measure fit the data better than the model with the individual items when taking into account model complexity.

Similarly, with respect to exculpatory/mitigation evidence, for example, I have information on the presence or absence of five types of "troubled family background" factors. I combine these factors into a single variable, capturing the total number of problematic family features occurring in a defendant's background.

²³⁰ Consistent with prior research, we limit our analysis to cases that ultimately resulted in a conviction for murder as a proxy for the strength of evidence in the case. BALDUS ET AL., *supra* note 25, at 40-42, 477.

²³¹ I include legally impermissible/legally suspect factors—e.g., defendant and victim's race/ethnicity, sex, and age—in my models in order to stay consistent with prior studies of capital charging and make direct comparisons to those studies possible. See Part III.A. The lone exception involves defendant's race, where I examine two separate model specifications: one including defendant's race and the other excluding defendant's race. The models meaningfully differ because the defendant's race accounts for approximately 35% of the effect of group differences in case-level attributes (i.e., disparate effect). See Part V.

their inculpatory or mitigating character. Important inculpatory/aggravating evidence includes the total number of statutorily defined aggravating circumstances present in the case, defendant's contemporary convictions and prior criminal history, money- or sex-related motive, the number of victims, the relationship between the defendant and the victim(s), and the age of victim. Potentially mitigating evidence includes the defendant's age, marital status, educational background, and employment history, troubled family history, military service, history of drug and alcohol use/abuse, psychiatric status, IQ, and religious affiliation. Many of the specific variables included in the model have been identified in the literature as having the strongest associations with capital charging and sentencing decisions.²³² Several of the factors labeled as inculpatory might be deemed as mitigating in some situations. Similarly, some of the variables categorized as mitigating may be viewed as aggravating depending on the situation. This does not present a problem for the current analysis because the direction of the effect in any individual case is immaterial. The overall effect of each of these variables is estimated from the data and, therefore, reflects the manner in which prosecutors, on average, treat these factors for each race-of-victim group.²³³

Table 1 provides the summary statistics for the variables included in the model for the pooled data, and Table 2 presents the data disaggregated by the victim's race.²³⁴ Because nearly 95% of cases in the dataset involve

²³² Among the most important factors influencing death sentencing behavior in Georgia are the number statutory aggravating circumstances present in the case; the number of victims killed by the defendant; the commission of a contemporaneous felony; a prior felony conviction or record of violence personal crimes; the presence of multiple mitigating factors (e.g., history of alcohol/drug abuse); and a female victim. David C. Baldus & George Woodworth, *Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience*, 74 J. CRIM. L. & CRIMINOLOGY 661, 685–86 (1983).

It is nearly impossible to know what information is available to the prosecutor (or to defense counsel) at the time of the charging decision and, as I have argued elsewhere, “many of the factors impacting *capital sentencing* are unknown to prosecutors or defense attorneys at the time of *capital charging*, and specifics about aggravation and mitigation evidence come to light in preparation for trial.” See *infra* note 325 at 165. The potential complication arising from this fact is that the statistical model may inappropriately assume the prosecutor was aware of a particular piece of information at the time of charging decision, and this may impact the analysis. But in order for this issue to bias my results, it would need to be the case that Caucasian-victim and African American-victim cases differed in terms of what information was actually known at the time of the charging decision *and* the information had the effect of: (a) making Caucasian-victim cases appear more aggravated (or less mitigated) at the time of the charging decision or (b) making African American-victim cases seem less aggravated (or more mitigated) at the time of the charging decision than the model suggests (or both).

²³³ See *Penry v. Lynaugh*, 492 U.S. 302, 323 (1989) (acknowledging the ambiguous effect of aggravating and mitigating evidence); *Buchanan v. Angelone*, 522 U.S. 269, 275–79 (1998) (holding that the Constitution does not require jurors to be told how they should consider specific evidence offered as mitigation, and an instruction to the jury to consider all relevant evidence is sufficient).

²³⁴ Specific information on case-level variables is missing for a significant number of the cases. Only 29% of the cases have complete information on every variable included in the model, but approximately 75% of the cases are missing data on three or fewer variables. The degree of missing values across all of the variables ranged from 0% to 12.2%. In other words, no single variable had less than approximately 88% of the available information.

My statistical models require that all cases included in the analysis have complete information for every variable analyzed. Discarding cases with missing data will bias the results unless the data are missing completely at random (i.e., missing values cannot be predicted from available information in the dataset—an assumption that the data do not satisfy). Rather than discard nearly 70% of the cases in the data and bias the results because that data are not missing completely at random, I adopt the “fully

African American or Caucasian victims, Table 2 only includes information for these cases. The column on the far-right of Table 2 shows the differences in the averages of each case characteristic between African American-victim and Caucasian-victim cases. As noted, *supra*, the data consist of the entire population of death eligible homicide cases in Georgia from 1993-2000, and not a mere sample, so tests of statistical significance are inapplicable in this context—the differences in means/proportions are the population differences. It is clear from the column reporting differences in observed case-level characteristics across these two groups that, on average, the cases differ along several important dimensions, but the differences are not especially stark for the vast majority of variables. The notable exception is the observed difference in death notices filed (43.4% of Caucasian cases were noticed for the death penalty compared to 18.9% of African American victim cases—a difference of 24.5 percentage points). With respect to defendant's race/ethnicity, 51.6% of Caucasian-victim cases have a Caucasian defendant and 46.1% have an African American defendant, 96.7% of African American-victim cases have an African American defendant (and 2.8% have a Caucasian victim).

The basic structure of the analysis is the estimation of two separate logistic regression equations examining prosecutorial death charging behavior—one for Caucasian-victim cases and one for African American-victim cases:

$$P(N_R) = \mathcal{F}(\overline{X_R \times \beta_R}), \quad (7)$$

where $R = W, B$. Both model specifications include the defendant, crime, and victim factors described above. The right-hand side and left-hand side elements of the equations are defined as above (see Table 1 and Appendix A). The results from Equation 7 are imported into Equations 4.1, 5, and 6.1 to perform the necessary multivariate decompositions.²³⁵

conditional specification" (FCS) approach to multiple imputation to address the missing data concern. STEF VAN BUUREN, FLEXIBLE IMPUTATION OF MISSING DATA 108 (2012). The FCS algorithm makes educated guesses about the missing values based on the observed interrelationships between the variables in the data. The process is repeated M times to create M completed datasets. Each individual dataset contains slightly different values for each educated guess to account for uncertainty in the guesses. The datasets are analyzed separately and the M results are combined to provide the final estimates reported in the analyses. The efficiency of the parameter estimates is given by: $1 \div [1 + (F \div M)]$, where F is the fraction of missing data. *Id.* at 49. I used the FCS algorithm to create and analyze twenty complete datasets ($M=20$), which yields estimates that are approximately 97% as efficient as those based on an infinite number of imputations. I examined the robustness of the estimates by varying M from 10 to 30 and the results were virtually indistinguishable.

I also examined the models on data without missing cases. Disparate treatment accounted for 53% of the gap with race-of-offender included in the model and 89% of the gap with race-of-offender excluded from the model. But these results are only based on 360 cases instead of the full 1,238 cases available via multiple imputation. The magnitude of the differences between multiple imputation and non-multiple imputation estimates is standard in the social science literature when data are not missing completely at random. See generally Gary King et al., *Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation*, 95 AM. POL. SCI. REV. 49 (2001).

²³⁵ The pooled decomposition in Equation 6.1 requires the model to be estimated on an aggregation of African American-victim and Caucasian-victim cases to obtain β_p . Elder et al., *supra* note 24, at 285.

V. RESULTS

Table 3 reports the effects of the case characteristics on the capital charging decision for African American-victim and Caucasian-victim cases based on Equation 7.²³⁶ Marginal effects, rather than log-odds are reported for ease of interpretation.²³⁷ A marginal effect is the predicted change in the probability of a capital charge for an incremental change (if continuous) or a unit change (if discrete) in that variable, holding other variables constant.²³⁸ The models are estimated separately for each race-of-victim group, so they do not provide a general estimate for the racial disparity based on a pooled model that calculates the racial gap while holding constant the effects of the other case characteristics.²³⁹ The column on the far-right of Table 3 (“Difference in Effect”) captures prosecutors’ racially differential behavioral response to each case characteristics. In the words of Sorensen et al., the far-right column “pertain[s] to [prosecutorial] preferences for each group in a binary comparison.”²⁴⁰ For example, the marginal effect of an incremental increase in the number of statutory aggravating circumstances is nearly twice as large for Caucasian-victim cases compared to African American-victim cases ($0.178 \div 0.093 = 1.91$), all else being equal. And not only does the magnitude of the effects of these variables differ across cases, but the direction of the effects for some of the variables also differ. For example, the number of co-defendants, number of contemporary felonies committed by the defendant, the number of prior felonies, the defendant’s marital status, high school graduation status, military service status,

²³⁶ The parameters presented in Tables 3, 4, and 5 are based on the *entire* population of death-eligible homicide cases, and not a mere sample, so statistical inference based upon uncertainty from sampling distribution (e.g., *p*-values and confidence intervals) is inapplicable in the convention sense. In other words, there is no uncertainty arising from limitations of the *data*. Nevertheless, to measure the reliability of my estimates of *E* and *C* assuming they were based on a mere sample, I calculated a measure of uncertainty for the race-of-victim effect via bootstrapping. In brief, the bootstrapping algorithm randomly samples cases from the data (with replacement) and calculates the variability of *E* and *C* across the samples. ADRIAN COLIN CAMERON & PRAVIN K. TRIVEDI, *MICROECONOMETRICS: METHODS AND APPLICATIONS* 254 (2005). The standard errors for *E* and *C* were, respectively, 0.024 and 0.028. This suggests that the estimates of *E* and *C* are robust to random variations in the selection of cases from the population and those effects would be statistically significant even if my data were a sample rather than the entire population of cases.

Whereas bootstrapping addresses potential uncertainty arising from limitations with the *data*, another source uncertainty stems from potential limitations of the *model*. Every model rests on certain assumptions about which variables to include, how those variables are measures, and the form of the relationship between the explanatory variables and the outcome variable (e.g., linear versus curvilinear). See *supra* note 226 and accompanying text. I address this type of uncertainty in Part VI.D.

²³⁷ See *supra* note 159 and accompanying text. Furthermore, the comparison of logit coefficients across groups for binary regression models is inappropriate because differences in the magnitude of the coefficients may be an artifact of the differences in the degree of residual variation between the groups. This problem is avoided when using marginal effects. Paul Allison, *Comparing Logit and Probit Coefficients Across Groups* 28 *SOCIOLOGICAL METHODS AND RESEARCH* 186, 189 (1999).

²³⁸ BACHMAN & PATERNOSTER, *supra* note 157, at 574.

²³⁹ As noted *supra*, the pooled models are based on problematic assumptions about heterogeneous effects of case characteristics. See Part III.B.1.

A pooled model was estimated and the race-of-victim disparity is 16.9 percentage point difference in the probability of a death penalty notice. The odds-ratio for the race-of-victim effect is 3.3, which is remarkably close to the 3.1 odds-ratio reported by Baldus and colleagues in *McCleskey v. Kemp*.

²⁴⁰ Sorensen et al., *supra* note 29, at 11.

defendant's psychiatric status, and whether a firearm was used in the homicide *all* have opposite-sign effects across the Caucasian-victim and African American-victim cases.

Tables 2 and 3 reveal that African American-victim and Caucasian-victim cases differ with respect to both their observable characteristics relevant to aggravation and mitigation, as well as how prosecutors respond to these characteristics. I now use these differences to determine how much of the racial gap in death charging behavior is attributable to racially disparate effect ("endowment") and racially disparate treatment ("coefficient").²⁴¹ Table 4 presents the first set of decomposition results.²⁴² The top panel in Table 4 decomposes the racial gap into the endowment, *E*, and coefficient, *C*, effects. The probability of a defendant receiving a death penalty charge in a Caucasian victim case is 43.4%, whereas the probability for a defendant in an African American-victim case is 18.9%—a racial gap, *G*, of 24.5 percentage points. Of the 24.5 percentage point gap, 8.4 percentage points, or 34.4% of the total racial gap is due to disparate effect—that is, observable differences in case characteristics between the African American-victim and Caucasian-victim cases. The overwhelming majority of the total race gap, 65.6% (or 16.1 percentage points) is attributable to disparate treatment. Stated differently, in the absence of racially disparate treatment, 35% of African American-victim cases would receive a capital charge—much closer to the 43.4% of Caucasian-victim cases receiving a capital charge.²⁴³ This finding is consistent with research on racial discrimination in the employment context: differential treatment is responsible for 50%-70% of the black-white wage gap.²⁴⁴

The bottom panel of Table 4 provides results of the detailed decomposition, which indicates the contribution of each variable to the racial gap based on disparate effect and disparate treatment.²⁴⁵ Column "*E*" reveals the proportion of the predicted racial gap attributable to group differences in each variable. So, for example, if African American-victim cases had, on average, the same number of statutory aggravating circumstances as Caucasian-victim cases, then the racial gap would decrease by 4 percentage points. Perhaps a more intuitive way to understand the effect of group differences in statutory aggravating circumstances present in

²⁴¹ See *supra* note 11.

²⁴² The decompositions in Table 4 use the pooled coefficients as the baseline, so the "coefficient effect" compares differences prosecutors' responses to the case characteristics between African American- and Caucasian-victim cases relative to the non-discriminatory coefficients. David Neumark, *Employers' Discriminatory Behavior and the Estimation of Wage Discrimination*, 23 THE J. OF HUM. RESOURCES 279, 282 (1988) (advocating the use of the coefficients from a pooled regression over both groups as the baseline as opposed to selecting a particular group for the baseline)]. See *supra* notes 216 & 217 and accompanying text.

²⁴³ I also examined the robustness of the estimates to individual observations (i.e., outliers) in the data. For each case in the data, I calculated an influence statistic, $\Delta\beta$ (*delta-beta*), measuring the impact of each case on overall effects of the variables in the model. Any case with a value of $\Delta\beta$ over 1 in considered to have undue influence on the results. The largest $\Delta\beta$ value for any case in the data was 0.38 and the median value was 0.002. DAVID W. HOSMER & STANLEY LEMESHOW, APPLIED LOGISTIC REGRESSION (1989).

²⁴⁴ Fryer et al., *supra* note 187, at 637-39 (citing studies).

²⁴⁵ See *supra* note 236.

a case is to calculate the percentage of the overall racial gap due to group differences in that particularly variable. Column “Prop. Change (*E*)” provides the answer: 15.5% of the total racial gap is because of African American-victim cases and Caucasian-victim cases differ, on average, in their number of statutory aggravating circumstances ($0.038 \div 0.245 = 0.155$).

Columns “(*C*)” and “Prop. Change (*C*)” can be interpreted in a similar fashion. Returning attention to the statutory aggravating circumstance variable, the racial gap in capital charging would decrease by 2.5 percentage points if statutory aggravating circumstances has the same *effect* in African American-victim cases as they had in Caucasian-victim cases. Stated differently, if prosecutors treated statutory aggravating circumstances in African American-victim cases in the same fashion they treated statutory aggravating circumstances in Caucasian-victim cases, the racial gap would decrease by 10.2% ($0.025 \div 0.245 = 0.101$).

Some of the endowment and coefficient effects have a negative (“-”) sign, so the interpretation is opposite of the previous discussion. The variable indicating whether the defendant had a monetary motive for the homicide has a negative sign for the endowment effect in Column (*E*). This suggests that the racial gap in capital charging would *increase* by 0.6 percentage points (or 2.4%) if homicides in African American-victim cases were equally motivated by money as Caucasian-victim cases. With respect to Columns “(*C*)” and “Prop. Change (*C*),” the interpretation is similar. The racial gap in capital charging would increase by 1.1 percentage points (or 4.5%) if defendants’ history of drug use in African American-victim cases was treated by prosecutors the same way as in Caucasian-victim cases.

It is also worth noting that the total endowment effect is comprised of variables that are legally impermissible or, at minimum, legally suspect.²⁴⁶ As a result, the endowment effect does not solely capture non-discriminatory dynamics influencing capital charging decisions. For example, the defendant’s race/ethnicity accounts for 13.1% of the total capital charging racial gap (see Table 4, bottom panel). The magnitude of the effect of defendants’ race is larger than the strength of evidence in the case (8.9%), and second only to the number of statutory aggravating circumstances (15.5%). So, even assuming, *arguendo*, that that the defendant’ race is treated the same by prosecutors across African American- and Caucasian-victim cases, the race-of-defendant endowment effect is likely a measure of racial discrimination.²⁴⁷ The detailed decompositions displayed in the bottom panel of Table 4 shows that the race-of-defendant endowment effect is 0.032, thus nearly two-fifths of the alleged non-discriminatory component of the capital charging cap is attributable to a legally impermissible factor ($0.032 \div 0.084 = 0.381$), all else equal. When the model is estimated *without* race-of-defendant, the total endowment

²⁴⁶ See *supra* note 231.

²⁴⁷ The variable indicating the defendant’s race, while itself an impermissible factor, is also subject to disparate treatment. See Tables 4 and 5.

effect accounts for 21.2% of the racial charging gap ($0.052 \div 0.245 = 0.212$).²⁴⁸

Table 5 presents results from the three-fold decomposition.²⁴⁹ Recall, *supra*, that this analysis accounts for the fact that racial differences in endowment and coefficient effects may occur simultaneously.²⁵⁰ The three-fold decomposition isolates the source of the racial disparity that would otherwise be arbitrarily attributed to both the endowment and coefficient effects.²⁵¹ The additional component in the three-fold decomposition, *CE*, captures the difference between what is expected from the two individual differences—disparate effect, *E*, and disparate treatment, *C*—and the observed result. In other words, it measures the effect *beyond* a simple summation of the effects of *E* and *C*. Demographers Hailman Winsborough and Peter Dickinson explain that the third component “is the increment (or decrement) in effect due to modifying both aspects of the situation simultaneously [...] over the effect of changing each singly.”²⁵² The top panel of Table 5 displays the percentage of the total racial charging gap attributable *E*, *C*, and the interaction between the two, *CE*.²⁵³ Disparate effect accounts for 37.3% of the gap, disparate treatment comprises 61.4% of the gap, and the interaction effect constitutes 1.3% the gap (or 0.3 percentage points).²⁵⁴ The magnitude of *CE* will be determined by the size of differences in its component parts, *E* and *C*. Recall from Table 2 that the differences in the case characteristics across the race-of-victim groups are mostly trivial, even though Table 3 reveals that the race-of-victim differences the effects of those case characteristics can be quite stark for many variables. *CE* is a multiplicative term, $CE = (X_W - X_B) \times (\beta_W - \beta_B)$, so the small value of *CE* can be attributed to the fact that racial differences in *E* are minor across most variables.

It must be reemphasized that the three components reduce to a two-component solution in either of two ways: some place *CE* in the disparate effect part, while others place *CE* in the disparate treatment part.²⁵⁵ Analysts differ on the proper interpretation of the *CE* effect because it has both

²⁴⁸ See *supra* note 231.

²⁴⁹ See *supra* note 236.

²⁵⁰ See Part IV.A.

²⁵¹ Daymont & Andrisani, *supra* note 198, at 420–21 (describing the three-fold decomposition).

²⁵² Winsborough & Dickinson, *supra* note 204, at 7.

²⁵³ The estimates for *E* and *C* reported in Table 4 used the pooled (i.e., non-discriminatory) coefficients for as the baseline, see *supra* note 242. The pooled coefficients cannot be used for the three-fold composition because the calculation of *CE* precludes the inclusion of the non-discriminatory baseline (β_p) in the same model. See Part IV.A. The three-fold decomposition in Table 5 is expressed from the point of view of African American-victim cases. Using Caucasian-victim cases as the baseline yields similar results: 31.1% of the gap is explained by *E*; 61.3% percent is explained by *C*, and 7.6% is explained by *CE*.

²⁵⁴ The bootstrapped standard errors for *E*, *C*, and *CE* were, respectively, 0.062, 0.034, 0.059. This suggests that, if the population were a mere sample, *E* would fail to achieve statistical significance when taking into account the portion of the *E* that is conditional on *C*. In other words, the coefficient effect (i.e., disparate treatment) is the sole phenomenon accounting for the racial disparity in capital charging, as evidenced by both the statistical significance of *C* and the statistical insignificance of *CE*. See *supra* note 243.

²⁵⁵ See Frank L. Jones & Jonathan Kelley, *Decomposing Differences between Groups: A Cautionary Note on Measuring Discrimination*, 12 SOC. METHODS & RES. 323, 329 (1984) (noting that there is no unambiguous way of allocating the interaction effect to endowment or coefficient effects).

discriminatory and non-discriminatory components. *CE* may be a consequence of differences in the case characteristics, and would disappear if Caucasian-victim and African American-victim cases had the same characteristics. But *CE* may also be a consequence of the differential treatment and would disappear if Caucasian- and African American-victim cases were treated similarly. So it is an interaction term in the sense of depending *jointly* on both differences. Researchers have explained that “the choice between [interpreting *CE* as an endowment or coefficient effect] depends on whether or not there is a clear argument for including the interaction as an aspect of discrimination.”²⁵⁶ The preference for a particular interpretation will turn on whether changes in endowment and coefficient effects are independent—that is, whether changes in one component is likely to affect the other.²⁵⁷ Stated differently, the key question is whether one believes that (a) differences in the treatment of case characteristics are likely to result in differences in the compositions of those characteristics between the race-of-victim groups? *or* (b) differences in composition of those case characteristics between race-of-victim groups are likely to result in differences in the prosecutors’ behavioral response to those characteristics by prosecutors?

It appears that logic would dictate that the most plausible interpretation of *CE* is that it is a component of discrimination: the interaction effect captures the percentage of the capital charging gap accounted for by the fact that prosecutorial treatment of African American (Caucasian) victim cases tends to be more punitive—or more lenient, depending on the sign—for those case characteristics for which the differences between African American-victim and Caucasian-victim cases tend to be most pronounced.²⁵⁸ The contrasting interpretation—that is, the racially differential treatment of case characteristics by prosecutors produces race-of-victim differences in the distribution of objective aggravation and mitigation evidence—seems highly implausible. Due to the relatively small *CE* effect, the placement of *CE* does not meaningfully alter the results from Table 4—namely, at least three-fifths of the racial gap in capital charging is attributable to disparate treatment.²⁵⁹

The bottom panel of Table 5 provides the decompositions for the individual case-level factors. The *CE* effect is most pronounced for the number of statutory aggravating circumstances, defendant’s WRAT Score, defendant having a monetary motive, firearm homicide, and victim’s age. Returning to the effect of the number of statutory aggravating circumstances example, *CE* is the difference in prosecutorial racially differential responses

²⁵⁶ *Id.* at 333 (internal quotations marks omitted).

²⁵⁷ Kitagawa, *supra* note 29, at 1179.

²⁵⁸ Winsborough & Dickinson, *supra* note 204, at 7 (explaining that the *CE* component “indicates how much of the gap can be accounted for by the fact that the returns to one group [e.g., whites] tends to be greater for those characteristics for which members of that group have higher average values.”).

²⁵⁹ The results from the two-fold and three-fold decompositions reveal that the defendant’s race accounts for 38%-42% of the endowment effect, so it is highly likely that the disparate treatment effect is significantly understated in my models.

to the number of aggravators *multiplied* by the difference in the average number of aggravators across the groups. If *CE* is interpreted as evidence of disparate treatment, then the total disparate treatment effect of statutory aggravating circumstances is: $C + CE = 0.020 + 0.005 = 0.025$. This would account for 10.1% of racial gap in capital charging ($0.025 \div 0.245 = .101$), all else being equal.

VI. DISCUSSION AND IMPLICATIONS

The statistical models described in this Article provide a template for the investigation of discriminatory charging dynamics in capital and non-capital cases. My analysis of detailed information on dozens of legally relevant variables indexing the level of aggravation and mitigation present in potentially capital cases reveals,²⁶⁰ consistent with prior research, that defendants accused of murdering Caucasians have odds of being noticed for the death penalty that are 3.3 times greater than a similarly situated defendant accused of murdering African Americans (or an increase of 230% in the odds). The magnitude of this racial disparity is very close to the findings reported to the Court in *McCleskey* (3.1),²⁶¹ although the data analyzed for *McCleskey* were nearly twenty years older than the data examined in this Article. The magnitude of the race-of-victim effect is also very similar to the average effect discovered across all studies of capital charging over the last twenty-five years (3.03).²⁶² The race-of-victim effect translates to an increase in the predicted probability of being noticed for the death penalty of 16.9 percentage points if the victim is Caucasian rather than African American, all else equal.²⁶³ These two measures of the likelihood of a capital charge are examples of the traditional metrics used to identify disparate impact and infer disparate treatment. But as explained, *supra*, these two measures are based on implausible assumptions about homogeneous effects of case characteristics for death-eligible Caucasian-victim and African American-victim homicides.²⁶⁴ Furthermore, they do not provide important insights into how race-of-victim differences plausibly generate the racial disparity in capital charging outcomes.²⁶⁵

My study advances our understanding of racial dynamics in capital charging by disaggregating the race-of-victim gap into disparate effect and disparate treatment components, potentially telling a more powerful and intuitive story about the role of race in capital charging.²⁶⁶ The Article provides answers to a pair of fundamental questions with which courts must wrestle when assessing the merits of a selective prosecution claim. First, how much would the race-of-victim gap change if the two groups were

²⁶⁰ See Table 1 and Appendix A.

²⁶¹ See generally BALDUS ET AL., *supra* note 25 (describing the statistical results presented to the Court in *McCleskey*).

²⁶² See Part III.A.

²⁶³ See, *supra*, note 159 and accompanying text.

²⁶⁴ See Part III.B.1.

²⁶⁵ See Part III.B.2.

²⁶⁶ See *supra* note 11.

identical in terms of their level of culpability, but treated in the current racially-differential manner? And second, how much would the race-of-victim gap change if the two groups were treated in a similar fashion by prosecutors, but retained their current differences in culpability?

There is nearly a 25 percentage point racial gap in capital charging between Caucasian victim and African American-victim cases (43.4% versus 18.9%), and approximately 61% of this gap is attributed to disparate treatment. In other words, less than 39% of difference in charging behavior between Caucasian- and African American-victim cases is accounted for by differences in the case characteristics; the remainder of the difference is due to prosecutor's racially differential behavioral response to those characteristics. The magnitude of disparate treatment reported is likely to be a conservative estimate because the disparate effect measure includes the defendant's race, which compromises a sizable portion of the total disparate effect (approximately 33%). When race-of-defendant is excluded from the models, the differences in case characteristics between Caucasian- and African American-victim cases account for approximately 22% of the racial gap, thereby leaving approximately 80% of the racial gap attributable to disparate treatment.²⁶⁷ The magnitude of disparate treatment in capital charging is eerily similar to the magnitude of the disparate treatment effect reported in studies of the racial gap in wages.²⁶⁸

The detail decompositions, which focus on the disparate effect²⁶⁹ and disparate treatment components of the individual case factors, are also illuminating.²⁷⁰ As explained earlier, the descriptive statistics provided in Table 2 clearly reveal that the differences in case characteristics between Caucasian-victim and African American-victim cases are rather insubstantial. For example, the typical Caucasian-victim case has 2.37 statutorily defined aggravating circumstances present, compared to 2.11 for the typical African American-victim case. Differences in the number of contemporary felonies, criminal history, number of defendants, and number of victims are equally trivial. It is only by examining Table 3, which reports racial differences in prosecutors' behavioral responses to these characteristics, do we begin to understand how the racial status of the victim impacts charging behavior. Racially disparate treatment is evident for both aggravating evidence (e.g., number of statutory aggravators, criminal history, monetary motive, trigger-person status, use of firearm, strength of evidence), and mitigating evidence (defendant's employment status at the time of the crime, defendant's marital status, defendant's military service, and defendant's education).²⁷¹ Tables 4 and 5 describe how these differences in case attributes and prosecutorial behavior, in the aggregate and uniquely,

²⁶⁷ See *supra* note 231.

²⁶⁸ Fryer et al., *supra* note 187, at 637–39.

²⁶⁹ See *supra* note 11.

²⁷⁰ Taber et al., *supra* note 22, at 725 (noting that decompositions provide insights into causal mechanisms for disparities).

²⁷¹ See *supra* note 233 and accompanying text.

contribute the overall race-of-victim gap. The “Prop. Change (*E*)” and “Prop. Change (*C*)” columns in Table 4 provide clear evidence that, for most case characteristics, the influence of disparate treatment on the racial gap is larger than the influence of disparate effect. And even after accounting for the fact that differences in disparate effects and disparate treatments exist simultaneously between the race-of-victim groups (see Table 5), we notice that differences in case characteristics are only able to explain a small fraction of the racial gap.

Yet, as illuminating as the aforementioned analyses may be, the results will only be convincing to courts and other empirical legal scholars if the key assumptions underlying statistical models are defensible.²⁷² There are four key assumptions that I address below: (1) mutual exclusivity of race-of-victim groups;²⁷³ (2) overlapping distribution case characteristics across race-of-victim groups (i.e. common support);²⁷⁴ (3) adequate representation of the underlying discretionary process (i.e., model fit);²⁷⁵ and (4) the conditional mean for unobservable case characteristics, given observed characteristics is equal to zero (i.e., unconfoundedness/“no omitted variable bias”).²⁷⁶ As I explained in Part II.B, the latter two assumptions have received the most attention from the courts, so I devote the bulk of my discussion to them.

A. *Mutual Exclusivity*

The first assumption is that race-of-victim groups are mutually exclusive: that is, a case can only enter the model as having either a Caucasian victim or an African American victim, but not both. This is a potential problem for cases with multiple victims who are racially heterogeneous. Approximately 17% of the cases in the Georgia data involve multiple victims, and 7.6% of those multiple-victim cases involved victims of different races. Consistent with prior research, cases with *at least* one Caucasian victim were coded as having a Caucasian victim for the purposes of this study.²⁷⁷ Cases involving at least one African American victim and a non-Caucasian victim (i.e., Asian/Pacific Islander, Latino/Hispanic, or Native American) were coded as having an African American victim. The results were substantively identical when these racially heterogeneous-victim cases were removed from the analyses, which is to be expected given the extremely small number of cases that fell into that category.²⁷⁸

²⁷² See Part II.B.

²⁷³ Nicole Fortin et al., *Decomposition Methods in Economics*, in 4A HANDBOOK OF LAB. ECON. 1, 14 (Orley Ashenfelter & David Card eds., 2011).

²⁷⁴ *Id.* at 17.

²⁷⁵ Part II.B.

²⁷⁶ Part II.B; Fortin et al., *supra* note 273, at 21.

²⁷⁷ See, e.g., Raymond Paternoster & Robert Brame, *An Empirical Analysis of Maryland's Death Sentencing System with Respect to the Influence of Race and Legal Jurisdiction* (Univ. of Maryland, College Park 2003).

²⁷⁸ See *supra* note 228 and accompanying text.

B. Common Support

The second assumption is that the distribution of values of the case characteristics across the two groups overlap. In statistics parlance, the cases analyzed in the model must share the “region of common support.”²⁷⁹ This means that cases with the same values for the case characteristics have a non-zero probability for being in either group. This is a crucial assumption because it must be reasonable to use the observed outcomes from one group to construct counterfactuals for the other group. When one group has no comparables in the other group in the data, any attempted comparisons between the groups are based on extrapolating the data from where it is observed to where it is needed rather than what the data actually are.²⁸⁰ In other words, the statistical model assumes what the data “should be” based on parametric assumptions of the model, and as a consequence, the results are extremely dependent on the idiosyncratic features of the model. Political scientists Gary King and Langche Zeng refer to this phenomena as the “dangers of extreme counterfactuals.”²⁸¹ Only seven cases fell outside the region of common support (0.5%). The results were identical whether or not these cases were included in the analyses.

C. Predictive Accuracy

The third assumption is that the statistical model provides an “adequate” representation of the underlying discretionary process—that is, the model does an acceptable job of predicting outcomes. One must exercise caution when interpreting the adequacy of a statistical model, especially in the criminal justice context, because the discretionary choices may not lend themselves to highly accurate statistical modeling, irrespective of the comprehensiveness model.²⁸² So even when the predictive power is not particularly strong, it may be difficult to imagine that a few case characteristics, if they exist, would lead to clear distinctions between defendants who are noticed for the death penalty and defendants who are not.²⁸³ Idiosyncrasies associated with charging decisions may be evidence of an arbitrary process, and not misspecification of the statistical model because model fit statistics tend to be small or modest when the “true” model has a large residual variance (i.e., a lot of inherent unpredictability).²⁸⁴ The difficulty associated with traditional model fit measures is often magnified when analyzing micro-level (e.g., court cases), cross-sectional data (i.e., data taken at a single point in time, rather than data

²⁷⁹ Gary King & Langche Zeng, *The Dangers of Extreme Counterfactuals*, 14 POL. ANALYSIS 131, 146–51 (2006).

²⁸⁰ *Id.*

²⁸¹ *Id.*

²⁸² Berk et al., *supra* note 196.

²⁸³ *Id.*

²⁸⁴ Gary King, *How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science*, 30 AM. J. OF POL. SCI. 666, 675 (1986).

that track changes over time), and non-continuous outcome variables (e.g., yes/no capital charging decisions).²⁸⁵ Statistical models of cross-sectional micro-level data will typically have lower predictive power because of the greater overall variability in the phenomenon under investigation,²⁸⁶ and model fit statistics for non-continuous outcomes typically do not scale to unity, even when the model fits the data perfectly, so the predictive power will be lower than an equally predictive model for continuous data.²⁸⁷

With the aforementioned caveats in mind, I calculated several different model fit statistics. The first measure, Tjur's *D*, compares the predicted probability of observing an outcome when the outcome is actually observed to the predicted probability of observing an outcome when the outcome is not observed.²⁸⁸ The statistic has a range from 0% to 100%, and the larger the statistic, the more accurately the model predicts charging decisions. Tjur's *D* for the pooled (i.e., the model that includes both African American- and Caucasian-victim cases), African American-victim, and Caucasian-victim models are, respectively, 33.2%, 28.6%, and 36.9%. Another model fit statistic, R^2 , quantifies the percentage of variation in capital charging decisions explained by the model based on a transformation of the outcome variable rather than the natural binary metric of the outcome.²⁸⁹ This statistic is most analogous to the traditional R^2 for continuous outcomes.²⁹⁰ The R^2 for the pooled, African American victim, and Caucasian victim models are, respectively, 46.8%, 39.6%, and 50.3%. These R^2 statistics are very similar to the predictive power of the 230 variable model that was the centerpiece of the statistical evidence offered in *McCleskey* ($R^2 = 47\%$), and although the federal trial court criticized the model's predictive capacity,²⁹¹ both an *en banc* Court of Appeals²⁹² and the Supreme Court²⁹³ assumed the model was valid.

A third, and perhaps a more intuitive, measure of model fit is the percentage of capital charging decisions that were correctly classified. For the pooled, African American-victim, and Caucasian-victim models, the classification rates are, respectively, 79.4%, 84.5%, and 78.2%. The primary shortcoming of the classification measure is that it tends to overestimate model fit when the binary outcome is extremely skewed. Nearly 70% of the cases did not result in a death penalty notice, so there is significant skew

²⁸⁵ WOOLDRIDGE, *supra* note 119, at 43–44, 536.

²⁸⁶ *Id.* at 43–44.

²⁸⁷ *Id.* at 536.

²⁸⁸ Tue Tjur, *Coefficients of Determination in Logistic Regression Models—a New Proposal: The Coefficient of Discrimination*, 63 AM. STATISTICIAN 366, 369 (2009). Formally, $Tjur's D = Pr(y = 1|y = 1) - Pr(y = 1|y = 0)$. The first term on the right-hand side of the equation is defined as the sensitivity of the model (i.e., how well the model predicts the presence of a death penalty notice in a case when, in fact, the case has been noticed for the death penalty) and the second term is the false positive rate.

²⁸⁹ Richard D. McKelvey & William Zavoina, *A Statistical Model for the Analysis of Ordinal Level Dependent Variables*, 4 J. MATHEMATICAL SOC. 103, 111–12 (1975).

²⁹⁰ Frank A.G. Windmeijer, *Goodness-of-Fit Measures in Binary Choice Models*, 14 ECONOMETRIC REV. 1995 (2007).

²⁹¹ *McCleskey v. Zant*, 580 F. Supp. 338, 361 (N.D. Ga. 1984).

²⁹² *McCleskey v. Kemp*, 753 F.2d 877, 895 (11th Cir. 1985).

²⁹³ *McCleskey v. Kemp*, 481 U.S. 279, 291 n.7 (1987).

present. This also explains why the classification rate for the African American-victim model (84.5%) is higher than the pooled and Caucasian-victim models, but the African American-victim model explains the least amount of variance (39.6%). Only 18% of African American-victim cases resulted in a death penalty notice, so there was much less variability in the outcome variable, whereas 44% of Caucasian-victim cases received a death notice. For this reason, the Tjur's D and R^2 statistics are generally preferable to the simple classification measure.²⁹⁴

As I noted in Part IV.B, the statistical models analyzed in this study included nearly all of the case characteristics deemed to be primary determinants of capital charging decisions: statutorily defined death eligibility factors, concurrent criminal charges, defendant's prior criminal history, and the relationship between the defendant and the victim. Recent litigation over Connecticut's capital punishment system included statistical models with nearly an identical set of variables, and such models were deemed probative by the state supreme court.²⁹⁵ It is unlikely, then, that model fit could be substantially improved by including some heretofore elusive legally relevant variable.²⁹⁶ Moreover, the fundamental task of the statistical models is to include all theoretically relevant variables to the charging decision. Once that task has been accomplished, the Court's equal protection jurisprudence requires the prosecutor to demonstrate the decision was based upon reason rather than caprice or emotion.²⁹⁷

D. Potential Omitted Variables

Clearly a statistical model's predictive power and the inclusion of theoretically relevant variables are closely connected, although low explanatory power does not necessarily imply that important variables have been omitted.²⁹⁸ There is no way to directly test the unconfoundedness assumption analyzing non-experimental data.²⁹⁹ Other approaches to address potential omitted variable bias, such as the instrumental variable framework popularized by econometricians which isolates the effect of an explanatory variable from possible omitted variables, are generally inappropriate for

²⁹⁴ Two additional measures of model fit, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were used to assess whether the inclusion of information about the victim's race substantially improved the fit of model. These statistics do not evaluate any particular model in an absolute sense, rather they permit an assessment of competing models. The smaller the AIC and BIC statistics, the better the model fits the data. The AIC and BIC for the race-inclusive models were both lower than the race-exclusive models (race-inclusive: AIC = 1058, BIC = 1243; race-exclusive: AIC = 1124, BIC = 1304). WILLIAM H. GREENE, *ECONOMETRIC ANALYSIS* 306 (4th ed. 2000).

²⁹⁵ Donohue, *supra* note 165, at 646.

²⁹⁶ See *supra* note 196 and accompanying text.

²⁹⁷ *Wayte v. United States*, 470 U.S. 598, 608 (1985) ("It is appropriate to judge selective prosecution claims according to ordinary equal protection standards."); *accord McCleskey*, 481 U.S. at 292; *United States v. Armstrong*, 517 U.S. 456, 465 (1996).

²⁹⁸ See Part II.B.

²⁹⁹ Pager & Western, *supra* note 47, at 222 ("[I]n the contemporary United States where acts of discrimination are likely to be subtle and covert, it is extremely difficult to measure discrimination directly.").

examining the effects of immutable characteristics, such as race/ethnicity and gender.³⁰⁰ A recent report from the National Research Council on measuring racial discrimination underscored this fact: “[t]he most common approach for dealing with omitted variable bias is to use an instrumental variables estimator . . . [but] [t]his strategy is not likely to be available in observational studies in the case of race . . . the best we are likely to be able to do with observational studies of racial discrimination is to specify the model as completely as possible.”³⁰¹ There are, however, other approaches that permit an examination of the sensitivity of the results to alterations to the statistical model. One cannot state with certainty whether the omitted variable bias exists; nevertheless, these approaches underscore the robustness of our results. I adopt two general approaches to assess the robustness of my findings: (a) model uncertainty test and (b) causal bounds test.

Model Uncertainty Test. The first approach examines the stability of the magnitude of the race-of-victim effect across various combinations of the explanatory variables in the model, as well as different measurements of those explanatory variables.³⁰² The rationale for this typical type of test is that there are many plausible statistical models, but researchers typically only report a small number of preferred causal estimates and neglect to inform the audience about the sensitivity of the results stemming from changes in the model specification. If the results reported can be nullified by small, sensible changes in the model specification, then one should be cautious about the existence of a “true” causal relationship. Sociologists Cristobal Young and Katherine Holsteen have explained that “[r]elaxing model assumptions makes the results more empirical, less model dependent, and focuses attention on the model ingredients that are critical to the results.”³⁰³ I examine the robustness of the race-of-victim effect by estimating thousands of statistical models across combinations of explanatory variables (and different measurements of some of those variables) and then calculating both a weighted and unweighted average causal estimate for race-of-victim.³⁰⁴ Roughly speaking, the weights are based upon the predictive capacity, i.e., model fit, with estimates from superior fitting models given greater weights.³⁰⁵ The weighted estimates are

³⁰⁰ Greiner & Rubin, *supra* note 136.

³⁰¹ NAT'L RESEARCH COUNCIL, *supra* note 53, at 141–42; *see also* Pager & Western, *supra* note 47, at 222–23 (advocating the use of audit studies to measure discrimination to limit the likelihood of omitted variable bias).

³⁰² Steven Deller et al., *Model Uncertainty in Ecological Criminology: An Application of Bayesian Model Averaging with Rural Crime Data*, 4 INT'L J. OF CRIMINOLOGY & SOC. THEORY 683, 684 (2011) (explaining the sources of model uncertainty).

³⁰³ Young & Holsteen, *supra* note 226, at 32.

³⁰⁴ An example of measurement uncertainty that I examine is the level of statutorily defined aggravation present in each case. There are ten aggravating circumstances enumerated in Georgia's capital statute, so measurement of the level of aggravation might include a summation scale indexing the presence of the various aggravating factors in a case. Alternatively, the individual aggravating circumstances could be included in the model. The former approach assigns equal weight to each aggravating factor, whereas the latter approach assigns an empirically-derived weight for each aggravating circumstance and the sum of those individual effects captures of “total effect” of the level of aggravation in the case.

³⁰⁵ *See generally* Young & Holsteen, *supra* note 226, at 30.

helpful in calculating a single measure that averages over the entire modeling distribution.³⁰⁶ The unweighted estimates, on the other hand, provide insight into the distribution of estimates that can be obtained from the data.³⁰⁷

The *weighted* average causal effect of the victim's race is a 18.4 percentage point increase the probability of a capital charge (for Caucasian-victim cases).³⁰⁸ This estimate is larger than the effect of "C" reported in Table 4 (16.1), which can be attributed to the weighting algorithm that privileges simpler models over more complex models, all else equal. According to this estimate, 74.1% of the race-of-victim gap results from disparate treatment. Of greater interest, however, are the features of the distribution of the unweighted causal estimates. The 95% confidence interval of the race-of-victim effect reported in Table 4 is [11.7, 20.4]. In other words, racially disparate treatment accounts for as low as 47.8% or as high as 83.5% of race-of-victim gap in capital charging. From this modeling distribution of estimates, I calculate the *robustness ratio* (*RR*), which is the race-of-victim effect from Table 4 ($\beta_{CV} = 0.161$) divided by the modeling standard error (*s. e.* = 0.022). The *RR* statistic is analogous to the *t*-statistic and examines the probability that the race-of-victim effect is "zero" in across the various model specifications. The critical value for the *t*-statistic is 1.98. The *RR* for the race-of-victim effect is 7.3, providing strong evidence that the effect of victim's race on capital charging decisions is not simply an artifact of my model specification ($RR = 7.3$; $p < 0.001$).

Causal Bounds Test. The second approach I used to assess the sensitivity of the causal effect of race-of-victim is a "bounds test."³⁰⁹ If there are unmeasured variables that simultaneously affect whether a case has a Caucasian victim and a prosecutor files a capital charge, even after holding case characteristics constant, then the causal estimate of the victim's race may be an artifact of these unmeasured factors. This is sometimes called "positive selection" and leads to upward bias in the estimated race-of-victim effect. A causal bounds test quantifies how strong this omitted variables bias

³⁰⁶ *Id.*; accord Gary King et al., *Making the Most of Statistical Analyses: Improving Interpretation and Presentation*, 44 AM. J. OF POL. SCI. 341, 350–51 (2000) (explaining that model averaging is the "best choice" when the researcher is interested in a single estimate of an explanatory variable because it removes modeling uncertainty by averaging over the modeling distribution of the estimate).

³⁰⁷ There is a "conceptual analogy between the *sampling* distribution and the *modeling* distribution. While the sampling distribution shows whether a point estimate is statistically significant (i.e., different from zero) [in the overall population], the modeling distribution shows whether it is different from those of other plausible models." Young & Holsteen, *supra* note 226, at 30 (emphasis in original).

The Georgia data comprise the entire population of death-eligible homicides, and not a mere sample, so the uncertainty in the estimate of the race-of-victim effect arises solely from model uncertainty. See *supra* notes 226 & 236 and accompanying text.

³⁰⁸ The estimate is based on models that potentially include defendant's race as a control variable. As I explained earlier, see *supra* note 231, there is good reason to exclude race-of-defendant from these models separate disparate impact from disparate treatment. When defendant's race is excluded from the model, the race-of-victim effect increases to a 21 percentage point increase in the likelihood of a capital charge.

³⁰⁹ PAUL R. ROSENBAUM, *OBSERVATIONAL STUDIES* 105 (2d ed. 2002).

must be in order to undermine the estimate of the causal effect.³¹⁰ If the results are very sensitive to the effect potential unmeasured factors, then the unconfoundedness assumption might be unwarranted. The bounds test differs from the model uncertainty tests discussed, *supra*, in that the focus is on the magnitude of the effect of unobserved characteristics rather than the sensitivity of the causal estimates to different combinations of observed factors.

The capital charging decision is a dichotomous variable (yes/no), so I use a variation of the bounds test tailored for this type of outcome. The procedure works as follows: I first match pairs of cases across the different race-of-victim groups that otherwise have the same observed case characteristics.³¹¹ I then manipulate the odds that the matched cases have the same probability of being selected into either race-of-victim group. Under the assumption of unconfoundedness, the cases have even odds of being in either group, so the odds ratio, γ (*gamma*), equals one. By changing $\gamma = 2$, defendants with similar observed case characteristics could differ in their odds of having a Caucasian victim as opposed to an African American victim by a factor of 2. Stated differently, γ is a measure of the degree of departure from a study that is free of hidden bias.³¹²

The bounds test produces two test statistics, Q_{MH}^+ and Q_{MH}^- , for each value of γ , that are used to test the null hypothesis that the model has, respectively, overestimated and underestimated the causal effect. For the purposes of this study, I only focus on Q_{MH}^+ because my interest is in the increased probability of a capital charge when the victim is Caucasian. The bounds test reveals that the race-of-victim estimate is insensitive to hidden bias even when that bias would increase the odds of differential selection up to a factor of 2.6 ($Q_{MH}^+ = 1.71; p < 0.05$). To provide some context, only two case characteristics increase the odds of capital charge by a factor greater than 2: the number of statutory aggravating circumstances (2.4) and victim's race (3.3). And when case characteristics are used to "predict" the race of the victim in the case, no case characteristic increases the odds of the victim being Caucasian by a factor greater than 2. These findings suggest that it is unlikely that there are unmeasured factors that have remained unidentified in the research literature for more than 40 years which satisfy the conditions that (1) the effect is not proxied by one or more legally relevant variables routinely included in statistical models of capital charging

³¹⁰ STEPHEN L. MORGAN & CHRISTOPHER WINSHIP, COUNTERFACTUALS AND CAUSAL INFERENCE: METHODS AND PRINCIPLES FOR SOCIAL RESEARCH 172–79 (2007).

³¹¹ Matching is based on each case's conditional probability (i.e., propensity score) of having either a Caucasian or African American victim, given the other case characteristics. The assumption of the algorithm is that the cases with the same propensity score have the same distribution of observable and (hopefully) unobservable characteristics, independent of the victim's race. In other words, for a given propensity score, the likelihood of the case having a Caucasian or African American victim should be, on average, observationally identical. A measure of racial disparity can be calculated from the average of the differences across all matched pairs. Paul R. Rosenbaum & Donald B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 BIOMETRIKA 41, 48 (1983). For the Georgia data, the matching algorithm reveals a 24.2 percentage point racial disparity in capital charging.

³¹² The bounds test relaxes the assumption that the matched cases are similar along both observed and unobserved characteristics by altering the degree of dissimilarity between the matched cases based on unobserved characteristics.

and (2) would increase the odds of a case having a Caucasian victim and being noticed for the death penalty by a factor significantly larger than the effect sizes of nearly all legally relevant case characteristics commonly included in models of capital charging.

As I emphasized earlier, the model uncertainty and bounds tests do not unequivocally preclude the potential of omitted variable bias, but they do attempt to quantify the degree of sensitivity of the race-of-victim effect to alterations in the underlying assumptions of the model. These robustness checks suggest that the data are not unduly delicate to the key assumptions of the statistical models. The results should be sufficient to give rise to an inference of discrimination that would require prosecutors to offer more than “general assertions that [they] did not discriminate or that they properly performed their official duties, [and require them to] demonstrate that the challenged effect [is] due to permissible racially neutral selection criteria.”³¹³ Granted, the statistical models do not include every conceivable variable relevant to a capital charging decision, but that standard has not been applied to statistical evidence of purposeful discrimination in jury selection and Title VII cases.³¹⁴ The relevant inquiry is whether the models “include those variables that account for the major factors that are likely to influence decisions.”³¹⁵ The statistical models I analyze in this study account for similar information as other models deemed probative of racial discrimination in the capital charging-and-sentencing process by state supreme courts.³¹⁶ And even if deemed insufficient to establish a *prima facie* case of an equal protection violation, the results, at minimum, should permit a defendant to “make a credible showing” of the existence of discriminatory effect and discriminatory treatment³¹⁷ to warrant an inspection of the prosecution’s files.³¹⁸

CONCLUSION

In his historic dissenting opinion in *Glossip v. Gross*, Justice Breyer remarked that the “arbitrary imposition of punishment is the antithesis of the rule of law. [...] How then can we reconcile the death penalty with the demands of a Constitution that first and foremost insists upon a rule of

³¹³ *McCleskey v. Kemp*, 481 U.S. 279, 352 (1987) (Blackmun, J., dissenting).

³¹⁴ *Id.* at 327–28.

³¹⁵ *Id.*

³¹⁶ See, e.g., Donohue, *supra* note 165; Paternoster & Brame, *supra* note 277; David C. Baldus, *Death Penalty Proportionality Review Project: Final Report to the New Jersey Supreme Court* (N.J. Judiciary 1991).

³¹⁷ *United States v. Bass*, 536 U.S. 862, 863 (2002); *United States v. Armstrong*, 517 U.S. 456, 470 (1996).

³¹⁸ *United States v. Bass*, 266 F.3d 532, 540 (6th Cir. 2001); *United States v. Jones*, 159 F.3d 969, 978 (6th Cir. 1998); *United States v. Gordon*, 817 F.2d 1538, 1540 (11th Cir. 1987); *United States v. Tuitt*, 68 F. Supp. 2d 4, 6 (D. Mass. 1999).

law?”³¹⁹ He described the vast social scientific literature over the past 40 years documenting the unconstitutional administration of the death penalty, including “numerous studies [that] have concluded that individuals accused of murdering white victims, as opposed to black or other minority victims, are more likely to receive the death penalty.”³²⁰ According to the Justice, the “circumstances and the evidence of the death penalty’s application have radically changed”³²¹ since the Court upheld the constitutionality of the death penalty forty years earlier in *Gregg v. Georgia*.³²² He “believe[s] that it is now time to reopen the question” of the constitutionality of the administration of the death penalty and invited “full briefing that would allow [the Court] to scrutinize [the empirical scholarship on the administration of the death penalty] with more care.”³²³

This Article accepted Justice Breyer’s invitation and set forth a framework that more carefully parses race-of-victim differences in capital charging than prior studies into the part explained by actual differences in the defendant’s level of culpability and the part explained by prosecutors’ racially discriminatory treatment of these cases. The model is directly responsive to the Court’s critique of much of the existing statistical evidence of racial discrimination—its inability to explicitly connect racial differences in *process* to racial differences in *outcomes*. The approach I adopt quantifies the extent of prosecutorial “shifting standards” in capital charging according to the victim’s race and establishes the foundation for an articulation of a more powerful and appropriately nuanced story about the role of race on prosecutorial decision-making.

Accompanying my methodological contribution is an important substantive one: race still matters *a lot* in capital charging decisions in Georgia. And there is good reason to believe that similar results would be obtained in other jurisdictions based on the similarity of empirical findings across studies,³²⁴ as well as the fact that many states modeled their own death penalty statutes after Georgia’s (which, itself, was based on the American Law Institute’s Model Penal Code).³²⁵ I discover that 60%-80% of the race-of-victim gap in capital charging results from disparate treatment.³²⁶ More importantly, I show that many of the case characteristics relevant to defendant culpability (i.e., aggravation and mitigation evidence) have radically different effects on the likelihood that the prosecutor seeks the death penalty *depending on the victim’s race in the case*. In other words, I

³¹⁹ *Glossip v. Gross*, 135 S. Ct. 2726, 2759, 2764 (2015). See Lincoln Caplan, *Richard Glossip and the End of the Death Penalty*, THE NEW YORKER, Sept. 30, 2015 (noting that Justice Breyer’s “widely commented-on dissent” will be a point of reference for the ultimate abolition of the death penalty).

³²⁰ *Glossip*, 135 S. Ct. at 2760.

³²¹ *Id.* at 2755.

³²² *Gregg v. Georgia*, 428 U.S. 153 (1976).

³²³ *Glossip*, 135 S. Ct. at 2755, 2759.

³²⁴ See, e.g., *McCleskey v. Kemp*, 481 U.S. 279, 328 (1987) (Brennan, J., dissenting) (the “evaluation of [statistical] evidence cannot rest solely on the numbers themselves. We must also ask whether the conclusion suggested by those numbers is consonant with our understanding of history and human experience”).

³²⁵ Sherod Thaxton, *Un-Gregg-Ulated: Capital Charging and the Missing Mandate of Gregg v. Georgia*, 11 DUKE J. OF CONST. L. & PUBLIC POL’Y 145, 145–46 (2016).

³²⁶ See *supra* notes 231 & 267 and accompanying text.

demonstrate how race influences the process of prosecutorial decision-making that leads to racially disparate outcomes.

Of course, any statistical approach to measuring discrimination will only be as reliable as the assumptions of the underlying statistical model are reasonable. So in addition to presenting a novel framework for examining racial discrimination in capital charging, I also describe and implement various diagnostic tools to examine the sensitivity of my results. These tools, like the statistical model to which they are applied, are also responsive to the Court's general concerns about the reliability of statistical evidence. The diagnostic tools, along with underlying framework, constitute a template for the investigation of discriminatory dynamics in the capital context, and therefore are critically important to how judges, lawyers, legislators, and legal scholars think about the constitutional constraints on prosecutorial decision-making and the courts' role in ensuring the rule of law remains operative. As Justice Brennan eloquently explained in *McCleskey*, the "diminished willingness to render [capital punishment] when blacks are victims, reflects a devaluation of the lives of black persons. [...] Race is a consideration whose influence is expressly constitutionally proscribed...and evidence that race may play even a modest role in levying [capital punishment] should be enough to characterize that [punishment] as [unconstitutional]."³²⁷

³²⁷ *McCleskey*, 481 U.S. at 326, 340-41 (Brennan, J., dissenting).

TABLE 1: SUMMARY STATISTICS

Variables	Mean/ Proportion	Std. Dev.	Min	Max
DP Notice Filed	0.301	--	0	1
Total Statutory Aggravators	2.224	1.091	1	7
Year of Offense	--	--	1993	2000
# of Defendants	1.793	1.109	1	7
Defendant White	0.248	--	0	1
Defendant Black	0.728	--	0	1
Defendant Latino	0.018	--	0	1
Defendant Asian/Pacific Islander	0.005	--	0	1
Defendant Male	0.946	--	0	1
Defendant Age	27.150	9.935	17	69
Defendant # of Violent Crimes	2.100	1.413	1	16
Defendant # of Contemp. Felonies	1.724	1.602	0	9
Defendant # of Prior Felonies	0.514	1.332	0	10
Defendant has Children	0.583	--	0	1
Defendant Employed	0.562	--	0	1
Defendant Married	0.179	--	0	1
Defendant High School Grad	0.262	--	0	1
Defendant Military Service	0.084	--	0	1
Defendant History of Drug Use	0.506	--	0	1
Defendant Psychiatric Status	1.219	0.508	1	4
Defendant IQ (Culture Fair)	100.110	14.833	50	151
Defendant WRAT	8.089	3.494	1	13
Defendant Family History	1.298	1.224	0	5
Monetary Motive	0.577	--	0	1
Sex-Crime Motive	0.053	--	0	1
Defendant is "Trigger Person"	0.853	--	0	1
Firearm Homicide	0.644	--	0	1
Strength of Evidence	0.729	0.778	0	3
Defendant Born in Georgia	0.639	--	0	1
# of Victims	1.185	0.504	1	6
Victim White	0.448	--	0	1
Victim Black	0.497	--	0	1
Victim Latino	0.034	--	0	1
Victim Asian/Pacific Islander	0.021	--	0	1
Victim Female	0.368	--	0	1
Victim Age	36.720	18.200	0	97
Victim Stranger	0.350	--	0	1
County	--	--	1	159
Judicial Circuit	--	--	1	46
Total Cases	1,238			

Note: Mean (average) values and standard deviations are reported for ordinal and continuous variables; proportions are reported for binary variables.

TABLE 2: SUMMARY STATISTICS (DISAGGREGATED BY RACE-OF-VICTIM)

Variables	Mean/ Proportion (White Victim)	Mean/ Proportion (Black Victim)	Mean/ Proportion (Difference)
DP Notice Filed	0.434	0.189	0.245
Total Statutory Aggravators	2.373	2.113	0.260
Year of Offense	--	--	--
# of Defendants	1.775	1.753	0.022
Defendant White	0.516	0.028	0.488
Defendant Black	0.461	0.967	0.506
Defendant Latino	0.020	0.003	0.017
Defendant Asian/Pacific Islander	0.003	0.002	0.001
Defendant Male	0.940	0.959	0.019
Defendant Age	27.554	27.494	0.060
Defendant Prior Violent Crimes	2.132	2.085	0.047
Defendant # of Contemp. Felonies	1.874	1.670	0.204
Defendant # of Prior Felonies	0.479	0.541	0.062
Defendant has Children	0.554	0.632	0.078
Defendant Employed	0.564	0.573	0.009
Defendant Married	0.174	0.188	0.014
Defendant High School Grad	0.256	0.295	0.040
Defendant Military Service	0.091	0.099	0.008
Defendant History of Drug Use	0.546	0.471	0.075
Defendant Psychiatric Status	1.271	1.152	0.119
Defendant IQ (Culture Fair)	102.409	98.956	3.453
Defendant WRAT	8.428	7.708	0.723
Defendant Family History	1.292	1.340	0.041
Monetary Motive	0.677	0.472	0.205
Sex-Crime Motive	0.046	0.068	0.021
Defendant is "Trigger Person"	0.828	0.874	0.045
Firearm Homicide	0.618	0.664	0.046
Strength of Evidence	0.863	0.615	0.248
Defendant Born in Georgia	0.618	0.664	0.046
# of Victims	1.233	1.214	0.019
Victim White	--	--	--
Victim Black	--	--	--
Victim Latino	--	--	--
Victim Asian/Pacific Islander	--	--	--
Victim Female	0.390	0.363	0.027
Victim Age	42.550	31.337	11.214
Victim Stranger	0.450	0.247	0.203

County	--	--	--
Judicial Circuit	--	--	--
Total Cases	554	613	

Note: Mean (average) values are reported for ordinal and continuous variables; proportions are reported for binary variables. The number in the final column is the difference in those means/proportions across the two groups.

TABLE 3: FACTS-OF-CASE EFFECTS (DISAGGREGATED BY RACE-OF-VICTIM)

Variables	Model $P(N_W)$ (White Victim)	Model $P(N_B)$ (Black Victim)	Difference in Effect
Total Statutory Aggravators	0.178	0.093	0.084
Year of Offense	--	--	--
# of Defendants	0.041	-0.013	0.054
Defendant White	0.117	0.044	0.074
Defendant Male	0.031	-0.007	0.038
Defendant Age	-0.002	0.001	0.003
Defendant # of Violent Crimes	0.036	0.028	0.009
Defendant # of Contemp. Felonies	0.004	0.005	0.001
Defendant # of Prior Felonies	0.007	-0.032	0.039
Defendant has Children	0.038	0.034	0.004
Defendant Employed	0.014	0.090	0.075
Defendant Married	0.051	-0.017	0.067
Defendant High School Grad	0.094	-0.037	0.131
Defendant Military Service	-0.095	-0.010	0.085
Defendant History of Drug Use	0.052	0.055	0.003
Defendant Psychiatric Status	-0.003	0.029	0.032
Defendant IQ (Culture Fair)	-0.003	-0.002	0.001
Defendant WRAT	0.001	0.016	0.015
Defendant Family History	0.027	0.006	0.022
Monetary Motive	-0.027	-0.067	0.040
Sex-Crime Motive	0.064	0.046	0.018
Defendant is "Trigger Person"	0.017	0.057	0.039
Firearm Homicide	0.087	-0.014	0.101
Strength of Evidence	0.105	0.070	0.035
Defendant Born in Georgia	-0.029	-0.044	0.015
# of Victims	0.042	0.062	0.020
Victim Female	0.065	0.044	0.021
Victim Age	-0.002	0.000	0.001
Victim Stranger	0.009	0.008	0.001

Note: The numbers in the first two columns are the effects of the corresponding case characteristics on the probability that a death penalty notice was filed in, respectively, Caucasian-victim $P(N_W)$ and African American-victim $P(N_B)$ cases. The number in the final column is the difference in those effects across the two groups.

TABLE 4: TWO-FOLD DECOMPOSITION OF THE RACE-OF-VICTIM GAP IN CAPITAL CHARGING

	Overall	% of Total Gap	
White-Victim $P(N_W)$	0.434		
Black-Victim $P(N_B)$	0.189		
Gap (G)	0.245		
Endowment (E)	0.084	34.4%	
Coefficient (C)	0.161	65.6%	

Variables	(E)	Prop. Change (E)	(C)	Prop. Change (C)
Total Statutory Aggravators	0.038	0.155	0.025	0.102
Year of Offense	--	--	--	--
# of Defendants	0.000	0.000	-0.001	-0.004
Defendant White	0.032	0.131	0.027	0.110
Defendant Male	0.000	0.000	0.040	0.163
Defendant Age	0.000	0.000	-0.001	-0.004
Defendant # of Violent Crimes	0.001	0.004	-0.001	-0.004
Def. # of Contemp. Felonies	0.001	0.004	-0.001	-0.004
Defendant # of Prior Felonies	0.001	0.004	0.001	0.004
Defendant has Children	-0.003	-0.012	-0.005	-0.020
Defendant Employed	0.000	0.000	-0.045	-0.184
Defendant Married	0.000	0.000	0.014	0.057
Defendant High School Grad	0.000	0.000	0.042	0.171
Defendant Military Service	0.000	0.000	-0.007	-0.029
Defendant History of Drug Use	0.003	0.012	-0.011	-0.045
Defendant Psychiatric Status	0.001	0.004	-0.001	-0.004
Defendant IQ (Culture Fair)	-0.009	-0.037	-0.000	0.000
Defendant WRAT	0.006	0.024	-0.000	0.000
Defendant Family History	-0.001	-0.004	0.000	0.000
Monetary Motive	-0.006	-0.024	0.031	0.127
Sex-Crime Motive	-0.001	-0.004	0.000	0.000
Defendant is "Trigger Person"	-0.001	-0.004	-0.054	-0.220
Firearm Homicide	-0.001	-0.004	0.067	0.273
Strength of Evidence	0.022	0.090	0.001	0.004
Defendant Born in Georgia	0.001	0.004	0.019	0.078
# of Victims	0.001	0.004	-0.054	-0.220
Victim Female	0.002	0.008	-0.001	-0.004
Victim Age	-0.008	-0.033	-0.002	-0.008
Victim Stranger	0.002	0.008	0.004	0.016

Note: Top Panel: $P(N_W)$ and $P(N_B)$ are, respectively, the probability a death penalty notice is filed in a Caucasian-victim and African American-victim case. Gap (G) is the

difference in the probability of a death notice between the two groups of cases. “Endowment (E)” is the predicted change in (G) that would occur if the two groups of cases had identical case characteristics, in the aggregate. “Coefficient (C)” is the predicted change in (G) if the two groups of cases were treated identically by prosecutors. Bottom Panel: “Column (E)” is the predicted change in (G) if the two groups of cases were identical on that specific case characteristic; “Column Prop. Change (E)” is the proportional change in (G). “Column (C)” is the predicted change in (G) if the two groups of cases were treated identically on that specific case characteristic; “Column Prop. Change (C)” is the proportional change in (G).

TABLE 5: THREE-FOLD DECOMPOSITION OF THE RACE-OF-VICTIM GAP IN CAPITAL CHARGING

	Overall	% of Total Gap	
White-Victim $P(N_W)$	0.434		
Black-Victim $P(N_B)$	0.189		
Gap (G)	0.245		
Endowment (E)	0.091	37.3%	
Coefficient (C)	0.151	61.4%	
Interaction (CE)	0.003	1.3%	

Variables	(E)	(C)	(CE)
Total Statutory Aggravators	0.027	0.020	0.005
Year of Offense	--	--	--
# of Defendants	-0.000	-0.001	-0.000
Defendant White	0.038	0.001	-0.000
Defendant Male	0.000	0.028	-0.001
Defendant Age	0.000	-0.000	-0.001
Defendant # of Violent Crimes	0.001	-0.000	0.000
Def. # of Contemp. Felonies	0.001	0.000	-0.001
Defendant # of Prior Felonies	0.003	0.002	-0.003
Defendant # of Children	-0.004	-0.014	0.001
Defendant Employed	-0.001	-0.049	0.002
Defendant Married	0.000	0.014	-0.001
Defendant High School Grad	0.002	0.038	0.000
Defendant Military Service	0.000	-0.007	-0.001
Defendant History of Drug Use	0.004	-0.012	-0.001
Defendant Psychiatric Status	0.004	0.002	0.002
Defendant IQ (Culture Fair)	-0.006	-0.001	0.002
Defendant WRAT	0.013	0.005	-0.010
Defendant Family History	0.000	0.001	-0.002
Monetary Motive	-0.013	0.021	0.001
Sex-Crime Motive	-0.001	0.001	-0.000
Defendant is "Trigger Person"	-0.002	-0.048	0.000
Firearm Homicide	0.001	0.064	0.002
Strength of Evidence	0.022	-0.001	0.001
Defendant Born in Georgia	0.003	0.023	-0.000
# of Victims	0.002	0.001	0.001
Victim Female	0.002	0.001	-0.000
Victim Age	-0.007	0.004	-0.011
Victim Stranger	0.005	-0.001	-0.005

Note: Top Panel: $P(N_W)$ and $P(N_B)$ are, respectively, the probability a death penalty notice is filed in a Caucasian-victim and African American-victim case. Gap (G) is the difference in the probability of a death notice between the two groups of cases. "Endowment (E)" is the predicted change in (G) that would occur if the two groups of cases had identical case characteristics, in the aggregate. "Coefficient (C)" is the predicted change in (G) if the two groups of cases were treated identically by

prosecutors. “Interaction (CE)” is the predicted change in (G) resulting from modifying (E) and (C) jointly rather than independently. Bottom Panel: “Column (E)” is the predicted change in (G) if the two groups of cases were identical on that specific case characteristic. “Column (C)” is the predicted change in (G) if the two groups of cases were treated identically on that specific case characteristic. “Column (CE)” is predicted change in (G) resulting from the simultaneous effect of (E) and (C) for that specific case characteristic.

APPENDIX A: DESCRIPTION OF VARIABLES

CRIME RELATED FACTORS
Statutorily defined aggravating factors; circumstances of murder (commission of felony, domestic altercation, other altercation, gang related, drug-related, sex-crime related); type of murder weapon (firearm, knife, automobile, poison, rope, etc.); motive for killing (jealousy, money, revenge, argument, etc.); confession evidence; weapon evidence; video evidence; date; location (home, business, street, bar, etc.); murder conviction.
DEFENDANT RELATED FACTORS ³²⁸
Number of defendants; race/ethnicity (African American, Asian/Pacific Islander, Caucasian, Hispanic, Other); sex; age; level of education (some high school, high school grad/GED, some college, college grad); employment status; marital status; number of children; religious affiliation (Catholic, Hindu, Jehovah Witness, Jewish, Mormon, Muslim, None, Protestant, Other); military service; history of drug use; psychiatric status (no impairment, minimal, serious, severe); IQ (Culture Fair Test); Wide Range Achievement Test (WRAT) (reading, math, spelling); troubled family history (alcoholism, criminality, drug abuse, absentee father, absentee mother, emotional/psychological abuse, physical abuse); prior felony conviction; prior murder conviction; trigger-person.
VICTIM RELATED FACTORS
Number of victims; race/ethnicity (African American, Asian/Pacific Islander, Caucasian, Hispanic, Other); sex; age; relationship with defendant (stranger, intimate partner, family, friend).

³²⁸ The Georgia Diagnostic and Classification Prison conducts diagnostic processing for the state's correctional system. Inmates undergo a battery of tests and diagnostic questionnaires, including the Culture Fair IQ test, Wide Range Achievement Test (WRAT) (reading, math, and spelling), history of substance abuse (summary & detailed report); latest mental health treatment; psychiatric test (based on PULHES Factor), assessment of inmate's family background, alcoholism and/or drug abuse, and presence/absence of parents absent during childhood.

**APPENDIX B: PROGRESSION OF GEORGIA DEATH PENALTY CASE
(ABRIDGED)**

STAGE	DESCRIPTION
<i>First Appearance</i>	Accused presented before a magistrate judge within 48 (warrant) or 72 (without warrant) hours. Unif. Super. Ct. R. 26.1 (2007).
<i>Grand Jury Indictment</i>	Grand jury returns an indictment charging a capital offense within 180 days. Ga. Code Ann. § 17-7-50.
<i>Appointment of Counsel</i>	Pursuant to the Georgia Indigent Defense Act of 2003 (GIDA), if the accused is eligible, she must be appointed two attorneys before she is called upon to plea to the charges, which generally occurs at the arraignment. Unif. App. R. II(A)(1).
<i>Pretrial Conference</i>	Pretrial conference must be held as soon as possible after indictment and before arraignment, and the conference must be recorded and transcribed. Prosecuting attorney must announce intention to seek the death penalty and then file a notice of intent with the clerk of the superior court. The superior court must then transmit the notice to the clerk of the Supreme Court of Georgia. Unif. App. R. IIC(1) (2007).
<i>Arraignment</i>	During the arraignment, the court must read the indictment and ask the defendant to plead to the capital felony and any lesser-included offenses charged. The defendant is allowed to plead guilty, not guilty, or mentally incompetent to stand trial; <i>nolo contendere</i> pleas are disallowed. Ga. Code Ann. § 17-7-95.
<i>Capital Voir Dire</i>	The court must empanel forty-two prospective jurors from which the state and defense must select a total of twelve jurors and one or more alternative jurors, if deemed necessary by the judge. Ga. Code Ann. §§ 15-12-160, 168.
<i>Capital Trial</i>	Capital cases are conducted in two phases. If the defendant is convicted of capital murder at the conclusion of the guilt/innocence phase, the case proceeds to the penalty phase where both the prosecutor and defense counsel may present witnesses and evidence regarding the statutory aggravating circumstances, as well as non-statutory aggravating and mitigating circumstances. The jury <i>may</i> sentence the defendant to death if, and only if, they find one or more statutory aggravating circumstance beyond a reasonable doubt. Ga. Code Ann. § 17-10-31.
<i>Post-Sentencing and Direct Appellate Proceedings</i>	Following a sentence of death, the defendant may challenge her conviction or death sentence by: filing a motion for a new trial with the superior court or filing a direct appeal with the Georgia Supreme Court. The appeal to the Georgia Supreme Court is automatic and may not be

	waived by the defendant. Ga. Code Ann. § 17-10-35.
<i>State Post-Conviction Proceedings (Habeas Corpus)</i>	A death-sentenced inmate may petition for a writ of habeas corpus to challenge the denial of her rights under the Georgia Constitution. A petitioner may appeal the denial of her petition to the Georgia Supreme Court. Ga. Code Ann. § 9-14-1.
<i>Federal Post-Conviction Proceedings (Habeas Corpus)</i>	A death-sentenced inmate may petition for a writ of habeas corpus to challenge the denial of her rights under the U.S. Constitution. A petitioner may appeal the denial of her petition to the federal appellate court. 28 U.S.C. § 2254.
<i>Clemency</i>	A death-sentenced inmate may apply for a pardon or commutation of her sentence to the State Board of Pardon and Paroles. Following the review of the case, each Board member will individually vote on the case. A majority vote is required in order to grant a pardon or commute a death sentence. Ga. Const. Art. 4, § 2, ¶ II(a).
<i>Execution</i>	Following exhaustion of her appeals and a denial of clemency by the State Board of Pardon and Paroles, the trial court must schedule an execution date. An inmate may not be executed if she is found to be mentally incompetent. Ga. Code Ann. § 17-10-40, 17-10-61.

Dr. Sherod Thaxton,
Disciplining Death

DISCIPLINING DEATH: Assessing and Ameliorating Arbitrariness in Capital Charging

Sherod Thaxton*

ABSTRACT

Justice Stephen Breyer recently made international headlines when he emphasized that reforms to the capital punishment process have apparently failed to ameliorate the rampant arbitrariness, capriciousness, and bias that led the U.S. Supreme Court to temporarily invalidate the death penalty over forty years ago. According to the Justice, the primary cause of this failure has been the Court's backpedaling on the very substantive and procedural protections it initially articulated as necessary for the constitutional administration of the death penalty. The Court's capital punishment jurisprudence initially underscored the importance of social scientific evidence in assessing the fairness of capital punishment systems, but now the Court routinely minimizes, or outright ignores, social science evidence on the operation of the death penalty. This has led to the growing disjunction between the Court's rhetoric and the reality of capital punishment. Justice Breyer underscored the Court's responsibility in holding death penalty systems accountable and called for full briefing on the basic question of the social realities of the administration of capital punishment.

Meaningful death penalty reform, if possible, requires a more prominent role for social science in death penalty decision-making. In this Article, I develop a doctrinally anchored statistical model that carefully disentangles and evaluates questions of arbitrariness, bias, and disproportionality in capital charging. I begin by discussing the Court's inconsistent efforts to rationalize and regulate capital punishment systems. I then adopt a framework of statistical inference in an effort to provide greater definitional and analytical clarity. Finally, I describe a set of analytical tools uniquely

* Assistant Professor of Law, UCLA School of Law. I would like to thank Anthony Alfieri, Stuart Banner, Bennet Capers, Devon Carbado, Beth Colgan, Ingrid Eagly, Laura Gomez, Cheryl Harris, Daniel Ho, Jill Horwitz, Irene Joe, Russell Korobkin, Sung Hui Kim, Máximo Langer, Douglas Lichtman, Mathew McCubbins, Hiroshi Motomura, Benjamin Nyblade, Jason Oh, Richard Re, Richard Sander, Joanna Schwartz, Seanna Shiffrin, Alex Wang, Adam Winkler, Noah Zatz, and the participants at Duke Law School's Culp Colloquium, UCLA's Criminal Justice Faculty Workshop, and UCLA's Faculty Colloquium for their valuable feedback. Naturally, all remaining errors are my own.

suited for diagnosing capital charging errors that closely aligns with the Court’s conceptualization of unacceptable arbitrariness. I illustrate the usefulness of the model on data involving actual death penalty-eligible defendants from Georgia.

My analysis reveals that death penalty charging practices are highly inconsistent, irrational, and disproportionate, both within and across jurisdictions in Georgia. The Article concludes by explaining how the empirical model might be used to improve accuracy and consistency in capital charging systems through empirically informed front-end charging screening.

INTRODUCTION.....	140
I. CONFRONTING CAPITAL ERROR.....	146
A. <i>Furman v. Georgia</i> and its Progeny.....	147
B. Conceptualizing and Operationalizing Capital Error.....	154
1. Reliability.....	155
2. Validity.....	156
3. Proportionality.....	158
C. Summary.....	160
II. DIAGNOSING CAPITAL ERROR.....	160
A. Modeling Capital Charging.....	161
B. Applying the Model to Georgia.....	166
1. Georgia Capital Charging Data.....	166
2. Arbitrariness as Unreliability.....	173
3. Arbitrariness as Invalidity.....	178
4. Arbitrariness as Disproportionality.....	183
5. Summary.....	186
C. Model Predictions for Case-Specific Outcomes.....	186
D. Generalizability of the Findings and Exportability of the Model.....	190
1. Generalizability.....	190

49:0137]	<i>DISCIPLINING DEATH</i>	139
2. Exportability.....		190
E. Summary		192
III. POLICY IMPLICATIONS		193
A. Designing Front-End Proportionality Review		195
B. Summary		202
IV. CONCLUSION.....		204
METHODOLOGICAL APPENDIX		205
FIGURE 1: MAP OF GEORGIA		214
FIGURE 2: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (UNADJUSTED)		215
FIGURE 3: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (UNADJUSTED)		215
FIGURE 4: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (ADJUSTED).....		216
FIGURE 5: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (ADJUSTED).....		216
FIGURE 6: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (UNADJUSTED & ADJUSTED ESTIMATES)		217
FIGURE 7: JURISDICTIONAL VARIABILITY IN THE RACE OF DEFENDANT EFFECT (CAUCASIAN)		217
FIGURE 8: JURISDICTIONAL VARIABILITY IN THE RACE OF DEFENDANT EFFECT (CAUCASIAN)		218
FIGURE 9: JURISDICTIONAL VARIABILITY IN THE RACE OF VICTIM EFFECT (CAUCASIAN)		218
FIGURE 10: JURISDICTIONAL VARIABILITY IN THE RACE OF VICTIM EFFECT (CAUCASIAN)		219
TABLE 1: SUMMARY STATISTICS.....		220
TABLE 2: DEATH NOTICES BY JUDICIAL CIRCUIT (1993-2000).....		221

INTRODUCTION

On June 28, 2015, by a vote of five-to-four, the U.S. Supreme Court rejected Oklahoma death row inmates' challenge to the constitutionality of the state's lethal injection protocol.¹ Richard Glossip, along with twenty other capitally condemned inmates, argued that Oklahoma's method of execution created an unacceptable risk of severe pain, thereby violating the Eighth Amendment's prohibition against cruel and unusual punishment.² Writing for the majority, Justice Alito noted that because the death penalty is constitutionally permitted, some risk of pain is inherent in execution and the petitioners were unable to identify a reasonable alternative that would entail a significantly lower risk of pain.³

In his dissenting opinion, Justice Breyer (joined by Justice Ginsburg) emphasized that the Court should stop "try[ing] to patch up the death penalty's legal wounds one at a time" and consider a more fundamental question: whether the current death penalty system is violative of the U.S. Constitution?⁴ After juxtaposing the Court's capital punishment jurisprudence with the voluminous social science literature on capital charging-and-sentencing practices across the nation, Justice Breyer concluded the Court has developed a large body of procedural regulations to govern the administration of the death penalty while simultaneously doing little to ensure that the panoply of protections that exist on paper are provided, in reality, to capital defendants.⁵ Justices Breyer and Ginsburg are not alone in their assessment of the previous four decades of capital charging-and-sentencing practices. In fact, the persistent obstacles to the fair administration of capital punishment have caused several current and recently retired justices to openly question whether efforts to fix the system should be finally abandoned and the country should move towards complete abolition.⁶

1. Glossip v. Gross, 135 S. Ct. 2726, 2727 (2015).

2. *Id.* at 2731.

3. *Id.* at 2733, 2738–39.

4. *Id.* at 2755 (Breyer, Ginsburg, JJ., dissenting).

5. *Id.* at 2755–77; see also James S. Liebman, *Slow Dancing with Death: The Supreme Court and Capital Punishment, 1963-2006*, 107 COLUM. L. REV. 1, 6 (2007).

6. In recent years, Justices Anthony Kennedy, Sandra Day O'Connor, and John Paul Stevens have made statements either expressly condemning the practice of capital punishment, or raising serious concerns as to its fair administration. See James S. Liebman & Lawrence C. Marshall, *Less Is Better: Justice Stevens and the Narrowed Death Penalty*, 74 FORDHAM L. REV. 1607, 1607 (2006); Brent E. Newton, *Justice Kennedy, the Purposes of Capital Punishment, and the Future of Lackey Claims*, 62 BUFF. L. REV. 979, 990–98 (2014); *O'Connor Questions Death Penalty*, CBS NEWS (Jul. 3, 2001), <http://www.cbsnews.com/news/oconnor-questions-death-penalty>. Justices John Paul Stevens, Lewis Powell, and Harry Blackmun all voted to uphold the constitutionality of capital punishment in *Gregg v. Georgia*, but publically criticized the death

Scholars have also described the death penalty system as being in a state of “perpetual malfunction.”⁷ There is strong evidence that death sentences are being imposed just as arbitrarily, discriminatorily, and excessively as they were prior to the Court expressly ruling that the death penalty must be administered fairly and evenhandedly, or not at all.⁸ Current capital charging-and-sentencing practices have resulted in a system marred by inexcusably high rates of reversals and retrials of capital verdicts, as well as extremely lengthy delays in executions.⁹ These problems have significantly undermined the credibility of the death penalty “whose chief function appears to be making mistakes, then taking years in a sometimes vain effort to correct them” rather than deterring potential killers and punishing those murders most deserving of the ultimate sanction.¹⁰

Some scholars have argued that, ultimately, a nationwide prohibition against the death penalty may be the only reasonable response to the chronic problems that have plagued the practice.¹¹ But death penalty abolitionists are unlikely to “unplug the machine” of death anytime soon given its continued popularity among legislatures and the general public.¹² Unless the Supreme Court unexpectedly reverses direction and decides that the death penalty is unconstitutional per se, both death penalty abolitionists and retentionists must

penalty after retiring from the Court. See Andrew Cohen, *Why Don't Supreme Court Justices Ever Change Their Minds in Favor of the Death Penalty?*, ATLANTIC.COM (Dec. 10, 2013), <http://www.theatlantic.com/national/archive/2013/12/why-dont-supreme-court-justices-ever-change-their-minds-in-em-favor-em-of-the-death-penalty/282100/>.

7. James S. Liebman, *Opting for Real Death Penalty Reform*, 63 OHIO ST. L.J. 315, 321 (2002); see also Carol S. Steiker & Jordan M. Steiker, *Lessons for Law Reform from the American Experiment with Capital Punishment*, 87 S. CAL. L. REV. 733, 733 (2014).

8. See *infra* Part I.

9. Carol S. Steiker & Jordan M. Steiker, *Report to the ALI Concerning Capital Punishment*, in REPORT OF THE COUNCIL TO THE MEMBERSHIP OF THE AMERICAN LAW INSTITUTE ON THE MATTER OF THE DEATH PENALTY annex B, at 7 (2009) (withdrawing its endorsement of the death penalty framework it had developed and promoted for over four decades “in light of the current intractable institutional and structural obstacles to ensuring a minimally adequate system for administering capital punishment”).

10. Liebman, *supra* note 7, at 320; see also *infra* Part I.

11. See, e.g., Liebman, *supra* note 7, at 342.

12. Although public support for the death penalty is at its lowest point in four decades, the majority of states still authorize the death penalty and the majority of Americans continue to endorse the practice. See Frank Newport, *In U.S., Support for Death Penalty Falls to 39-Year Low*, GALLUP (Oct. 13, 2011), <http://www.gallup.com/poll/150089/Support-Death-Penalty-Falls-Year-Low.aspx>; Rebecca Stewart, *CNN Poll: Number Who Prefer Death Penalty on Decline*, CNN (Oct. 12, 2011), <http://politicalticker.blogs.cnn.com/2011/10/12/cnn-poll-number-who-prefer-death-penalty-on-decline>. But see Kenneth E. Shirley & Andrew Gelman, *Hierarchical Models for Estimating State and Demographic Trends in US Death Penalty Public Opinion*, 178 J. ROYAL STAT. SOC'Y 1, 1 (2015) (noting significant racial/ethnic, gender, and geographic differences in support for capital punishment).

continue to focus their attention on whether the death penalty is capable of being administered in a manner that comports with the legal standards announced by the Court.¹³ “Only if it is not, is abolition a constitutional imperative” under the Court’s current death penalty jurisprudence.¹⁴

The legitimacy of the death penalty system rests, primarily, on the ability of applicable policies and procedures to reduce foreseeable errors that undermine fairness.¹⁵ These constitutional errors come in the form of systematic inconsistent, irrational, discriminatory, and excessive charging-and-sentencing outcomes. Developing a system for the management of these errors requires methods that “open an inquiry into the effects of changes in criminal justice standards, policies, and practices on the incidence of justice errors.”¹⁶ In this Article, I develop a doctrinally anchored empirical model that disentangles and evaluates questions that are central to the constitutionally permissible administration of capital punishment: how arbitrary, biased, and disproportionate is capital charging? Properly identifying, quantifying, and ultimately discouraging the inappropriate use of the death penalty at the charging stage is likely the most effective and efficient way to reduce the overall prevalence of constitutional error in capital charging-and-sentencing systems.¹⁷ My model is exportable, thus with minor modifications, it can be used to directly diagnose the level of arbitrariness at which a statute operates in any jurisdiction.¹⁸ Additionally, the proposed model evaluates the potential excessiveness of any individual capital charging decision. In concrete terms, the model is capable of predicting the likelihood that a defendant would face the death penalty based on prior capital charging decisions in the jurisdiction. This statistic provides meaningful and

13. Justice Ginsburg was the only other member of the Court to join Justice Breyer’s dissent in *Glossip*. See *Glossip v. Gross*, 135 S. Ct. 2726, 2755 (2015). Justice Sotomayor also authored a dissenting opinion, joined by Justices Breyer, Ginsburg, and Kagan, but her opinion was not as far-reaching in its condemnation of the capital punishment; rather she expressed her view that Oklahoma’s execution protocol was unconstitutional. *Id.* at 2780.

14. Liebman, *supra* note 5, at 5; see also *Glossip*, 135 S. Ct. at 2776–77 (Breyer, J., dissenting) (stating his belief that it is highly likely that the administration of capital punishment violates the Eighth Amendment).

15. BRIAN FORST, *ERRORS OF JUSTICE: NATURE, SOURCES, AND REMEDIES* 7 (2004); HANNAH QUIRK ET AL., *REGULATION AND CRIMINAL JUSTICE: INNOVATIONS IN POLICY AND RESEARCH* 35 (2010).

16. FORST, *supra* note 15, at 31–32.

17. See *infra* Parts II & III.

18. The model I describe in this Article focuses on arbitrariness and disproportionality. In other work, I expressly address the matter of the racially discriminatory administration of the death penalty in the context of capital charging. See Sherod Thaxton, *Disentangling Disparity: Exploring Racially Disparate Effect and Treatment in Capital Charging* (2017) (unpublished manuscript) (on file with author).

verifiable information to decision-makers when evaluating the appropriateness of pursuing the ultimate sanction against a particular defendant. I demonstrate the usefulness of the model by analyzing eight years of capital charging decisions from Georgia.

The Article is organized into three parts. Part I discusses the U.S. Supreme Court's inconsistent efforts to rationalize and regulate capital punishment systems. The Court's vague and, often internally incoherent, jurisprudence has undermined efforts to identify constitutional errors and develop systems capable, at least in theory, of eliminating them. On the one hand, the Court has repeatedly emphasized that "death is different" and capital punishment systems require heightened consistency and accuracy, even within the narrow class of death-eligible defendants.¹⁹ On the other hand, the Court has failed to provide sufficiently precise or consistent workable definitions of reliability, validity, and proportionality—all central concepts articulated in its doctrines. Legally meaningful definitions are not only crucial to properly evaluate capital punishment systems' conformity with the heightened standards announced by the Court, but also indispensable when considering adequate modifications to existing procedures and remedies for undesirable outcomes. In an attempt to provide greater definitional and analytical clarity, I adopt a framework of statistical inference that is particularly well-suited for identifying and quantifying the types of constitutional error that animated the Court's modern capital punishment jurisprudence.²⁰

Building upon the analytical framework described in the previous section, Part II develops a doctrinally anchored empirical model capable of assessing the level of arbitrariness—i.e., unreliability, invalidity, and disproportionality—in capital charging that exists both within and across jurisdictions. The model has two key features that are essential to properly studying capital punishment practices. First, by focusing on a single decision point at the very outset of the capital punishment process—where prosecutorial discretion is nearly unfettered—one can obtain a much clearer picture of the dynamics driving variability in capital charging practices

19. See *Furman v. Georgia*, 408 U.S. 238, 257–306 (1972) (Brennan, J., concurring).

20. Other scholars have also advocated for the adoption of a framework for statistical inference for such a task. See, e.g., QUIRK ET AL., *supra* note 15 (suggesting that criminal justice evaluation should be informed by principles of statistical design and inference); Alberto Alesina & Eliana La Ferrara, *A Test of Racial Bias in Capital Sentencing*, 104 AM. ECON. REV. 3397, 3398 (2014) (adopting a statistical framework for errors of inference to determine racial bias in sentencing by examining appellate review of capital cases); David C. Baldus et al., *Reflections on the "Inevitability" of Racial Discrimination in Capital Sentencing and the "Impossibility" of Prevention, Detection, and Correction*, 51 WASH. & LEE L. REV. 359, 363–64 (1994) (advocating a statistical-inference based approach to determining racial discrimination in death penalty sentencing).

observed under a single capital statute.²¹ Second, consistent with the decentralized and county-centric nature of death penalty charging authority, my proposed model mobilizes analytical tools uniquely developed to investigate hierarchically structured (“clustered”) data—that is, individual cases nested in counties or similar sub-state units. The Court has repeatedly emphasized the importance of explicitly examining both intra- and inter-jurisdiction processes when assessing the constitutionality of criminal punishments.²² Identifying and quantifying variation in this manner—i.e., single decision point and contextual influences—more closely aligns with the Court’s conceptualization of unacceptable arbitrariness under its capital punishment jurisprudence. To the best of my knowledge, this Article is the first to apply these tools to capital charging decisions.

Beyond diagnosing an entire system for macro-level constitutional errors of arbitrariness, I describe how the model can be used when making case-by-case capital charging assessments to detect potential micro-level errors. These micro-level errors are excessive charging decisions in cases that are technically eligible for the death penalty. Such errors have been definitively linked to another type of legal error: prosecutors who pursue the death penalty in cases that are not highly aggravated are significantly more likely to have those very same cases overturned on appeal because of serious trial-level errors.²³

The model is designed to be exportable, so its usefulness transcends its application in any particular jurisdiction. I illustrate the usefulness of the model on data involving actual death penalty-eligible defendants from Georgia, but there is good reason to believe the patterns I identify are present in other capital jurisdictions. The data reveal several constitutionally problematic features of Georgia’s capital charging practices. First, capital charging practices are highly inconsistent within jurisdictions, i.e., “between-case” heterogeneity, even for factually similar cases.²⁴ Second, capital charging practices for alike cases are highly inconsistent across jurisdictions

21. See *United States v. Bass*, 536 U.S. 862, 863–64 (2002) (per curiam) (holding that nationwide statistics of racial disparities in capital charging decisions are insufficient to support an equal protection violation claim); *McCleskey v. Kemp*, 481 U.S. 279, 342–43 (1987) (Brennan, J., dissenting) (statistical evidence of racial disparities derived from multiple decision-makers is insufficient to mount a constitutional challenge under the Court’s current death penalty jurisprudence).

22. *Solem v. Helm*, 463 U.S. 277, 290–92 (1983); *Enmund v. Florida*, 458 U.S. 782, 794–96 (1982).

23. See, e.g., Andrew Gelman et al., *A Broken System: The Persistent Patterns of Reversals of Death Sentences in the United States*, 1 J. EMPIRICAL LEGAL STUD. 209, 260–61 (2004).

24. See *infra* Part II.B.2.

for “between-jurisdiction” heterogeneity.²⁵ Third, the jurisdiction where a case is prosecuted exerts a strong influence on whether a defendant is charged with the death penalty.²⁶ Locales characterized by large deviations in death noticing behavior, relative to the statewide baseline, for similarly situated defendants, may be interpreted as being unjustifiably idiosyncratic given existing constitutional constraints on the capital punishment process. These jurisdictional effects are indicators of “institutional performance” and permit the comparative ranking of the jurisdictions.²⁷

Fourth, case-level characteristics only explain a small percentage of the variation in capital charging decisions both within and across jurisdictions.²⁸ This lack of a strong correspondence between the legally legitimate case characteristics and charging outcomes is a marker of the irrationality of the charging process, which is an important component of arbitrary government action.²⁹ Fifth, the race of defendants and victims is strongly associated with the level of inconsistency and irrationality in capital charging practices.³⁰ Caucasian-defendant and Caucasian-victim cases are handled more consistently than in non-Caucasian-defendant and non-Caucasian-victim cases. Similarly, with respect to the rationality of charging decisions, case-level characteristics explain a larger percentage of variation in outcomes in Caucasian-defendant and Caucasian-victim cases than in non-Caucasian-defendant and non-Caucasian-victim cases. Lastly, the race of the defendant and the race of the victim appear to have a direct influence on capital charging decisions. Specifically, when looking at Georgia as whole, cases involving Caucasian defendants and victims are more likely to be noticed for the death penalty, all else equal; however, this effect of race is highly variable across jurisdictions. The magnitude of the difference of the effect of race, from the

25. See *infra* Part II.B.2.

26. See *infra* Part II.B.2.

27. SOPHIA RABE-HESKETH & ANDERS SKRONDAL, MULTILEVEL AND LONGITUDINAL MODELING USING STATA: CONTINUOUS RESPONSES 50 (3d ed. 2012) (describing contextual effects as “valued added” by the location, all else being equal); Craig Duncan et al., *Context, Composition and Heterogeneity: Using Multilevel Models in Health Research*, 46 SOC. SCI. & MED. 97, 111 (1998) (discussing the use of predictions of cluster-specific effects to rank institutions).

28. See *infra* Part II.B.2.

29. *Burlington Truck Lines, Inc. v. United States*, 371 U.S. 156, 168 (1962) (defining arbitrary government action as the lack of a rational connection between the facts that should govern a decision and the choice being made); *accord* *Motor Vehicle Mfrs. Ass’n v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29, 43 (1983).

30. See *Furman v. Georgia*, 408 U.S. 238, 365 (1972) (Marshall, J., concurring) (arguing that arbitrary capital sentencing standards are an “open invitation to discrimination”); *infra* Part II.B.4.

statewide average, is as high as a factor of four for race-of-defendant and a factor of three for the race-of-victim.³¹

Part III sketches some ideas about how my model can be used to improve accuracy and consistency in capital charging systems. The development of a fully-specified policy proposal is beyond the scope of this project,³² but this section does lay a foundation, with rigorous social scientific inquiry at its core,³³ upon which meaningful death penalty reform can be erected. Specifically, I identify what I believe are some essential features of a feasible and effective reform policy. Generally speaking, any defensible reform must provide substantial disincentives for poor prosecutorial charge screening and save both state and federal governments much of the expense of error correction occurring at the appellate review stages.

I. CONFRONTING CAPITAL ERROR

Capitally condemned inmates have challenged the legality of the death penalty, as applied, under three provisions of the U.S. Constitution: the Eighth Amendment's Cruel and Unusual Punishment Clause, the Fourteenth Amendment's Due Process Clause, and the Fourteenth Amendment's Equal Protection Clause.³⁴ Both cruel and unusual punishment and due process violation claims have focused on the alleged arbitrary administration of capital punishment, whereas equal protection challenges highlighted the racially discriminatory application of the death penalty.³⁵ The Cruel and Unusual Punishment Clause also formed the basis of challenges to the excessiveness of the death penalty as applied in particular cases (disproportionality). A routinely divided court accepted some of these claims and rejected others. The consequence of these cases has been the

31. See *infra* Part II.

32. See Sherod Thaxton, *Death, Dollars, and Deference: Rethinking Prosecutorial Accountability in Capital Charging* (2016) (describing a system of financial, administrative, and reputational disincentives for poor charge screening in potentially capital cases) (unpublished manuscript) (on file with author).

33. See CRAIG HANEY, *DEATH BY DESIGN: CAPITAL PUNISHMENT AS A SOCIAL PSYCHOLOGICAL SYSTEM* 214–15 (2005) (advocating for a more prominent role for social science in the Court's death penalty jurisprudence).

34. Due Process Clause and Equal Protection Clause arguments, for challenges to the application of the death penalty for cases originating in federal court, are governed by the Fifth Amendment's Due Process Clause. See *Bolling v. Sharpe*, 347 U.S. 497, 499 (1954) (explaining that the Fifth Amendment's Due Process Clause implicitly incorporates an equal protection guarantee).

35. *Glossip v. Gross*, 135 S. Ct. 2726, 2762 (2015) (Breyer, J., dissenting) (noting that irrelevant factors, such as race, that determine who receives the death penalty is indicative of arbitrariness).

development of an increasingly complex, and often contradictory, capital punishment jurisprudence that has been derived, primarily, from the Court's interpretation of the Cruel and Unusual Punishment Clause.

Part A discusses the key aspects of the Court's jurisprudence as it pertains to the constitutionally permissible administration of the death penalty and state legislatures' responses to these rulings. The Court emphasized distributive justice as a key component to the fundamental fairness of capital punishment systems,³⁶ but this commitment has waned over the years as social facts about the actual operation of capital punishment unequivocally describe a system plagued by unpredictability, irrationality, and multiple forms of discrimination. Procedural justice concerns now dominate the Court's jurisprudence, but the Court has failed to identify any evidence suggesting the procedures developed by legislatures to promote consistency and accuracy in capital charging and sentencing are capable of satisfying the requisite constitutional standards.³⁷

I argue that meaningful death penalty reform can only be possible if courts and legislatures reengage with the social realities capital punishment practices. Central to this reengagement is the articulation of clear and workable standards that permit careful assessments of capital charging-and-sentencing behavior. The Court has consistently refused to provide the necessary guidance, which has both allowed it to insulate itself from the difficult task of enforcing its own doctrines and exacerbated the gulf between the rhetoric and reality of the death penalty. In Part B, I turn to the framework of statistical inference for identifying and expressly measuring the types of errors that are of primary concern to the Court: errors of arbitrariness, bias, and disproportionality. The field of statistics has historically been concerned with these types of errors, and offers useful guidelines on their detection and measurement. These guidelines provide the necessary foundation for a systematic inquiry into the social facts of capital punishment.

A. *Furman v. Georgia and its Progeny*

In the landmark case, *Furman v. Georgia*,³⁸ the Court held, by a vote of five-to-four, that all existing capital punishment statutes were

36. Distributive justice is generally defined as the perceived fairness of the allocation of rewards and costs. Procedural justice, on the other hand, focuses on the fairness and transparency of the processes that resolves disputes and allocates resources. *Justice, Social*, OXFORD DICTIONARY OF SOCIOLOGY 379 (John Scott & Gordon Marshall eds., rev. 3d ed. 2005).

37. *Glossip*, 135 S. Ct. at 2755–56 (Breyer, J., dissenting) (arguing that the administration of capital punishment remains constitutionally infirm).

38. *Furman v. Georgia*, 408 U.S. 238 (1972).

unconstitutional, as applied, because they failed to articulate to decision-makers any rational basis by which to distinguish those limited number of defendants sentenced to death from the thousands of other similarly situated defendants who were not subject to the death penalty.³⁹ The Court was primarily troubled by three glaring problems with the existing practice of capital punishment: (1) the small number of death sentences handed out relative to potentially capital crimes (infrequency and arbitrariness); (2) the lack of statutory restrictions upon sentencing discretion of judges and jurors (standardlessness); and (3) sentencing disparities based on race/ethnicity and social class (bias/discrimination).⁴⁰ *Furman* lacked a true holding because all nine Justices wrote separate opinions;⁴¹ nonetheless, Justices Brennan,⁴² Douglas,⁴³ Marshall,⁴⁴ Stewart,⁴⁵ and White,⁴⁶ comprising the majority, all expressed serious concern over the irrational and inconsistent imposition of the death penalty. The justices were split over whether racism still infected the death penalty process. Although the justices all acknowledged that racism in the administration of the death penalty was evident in the past, they disagreed as to its continuing relevance. Chief Justice Burger and Justice Powell (both dissenting), as well as Justice Douglas (joining the majority), all hinted that an equal protection challenge might prevail if there was compelling evidence of racial bias. The immediate impact of the Court's

39. *Id.* at 239. The Court agreed to hear four cases out of a pool of nearly two hundred pending capital cases: two non-homicidal rape cases from Georgia (*Jackson*) and Texas (*Branch*) and two murder cases from California (*Aikens*) and Georgia (*Furman*). The writ in *Aikens* was dismissed after the California Supreme Court ruled that the state's capital punishment violated the state's Constitution. *People v. Anderson*, 493 P.2d 880, 899 (Cal. 1972) *superseded by constitutional amendment*, CAL. CONST. art. I, § 27.

40. DAVID GARLAND, *PECULIAR INSTITUTION: AMERICA'S DEATH PENALTY IN AN AGE OF ABOLITION* (2010).

41. *Id.* at 225–30. Justices Stewart and White's opinions provided the narrowest ground for agreement, and therefore were deemed controlling. Both Justices focused, primarily, on the arbitrary administration of the death penalty. *Furman*, 408 U.S. at 309–10, 313; *see also* Liebman & Marshall, *supra* note 6 at 1608–10.

42. *Furman*, 408 U.S. at 293 (Brennan, J., concurring) (describing the death penalty system as “little more than a lottery system”).

43. *Id.* at 249–52 (Douglas, J., concurring) (explaining that equality in the administration of the death penalty is the key consideration for the Court).

44. *Id.* at 362–63 (Marshall, J., concurring) (noting that convicted murderers are seldom sentenced to death).

45. *Id.* at 309 (Stewart, J., concurring) (noting that “death sentences are cruel and unusual in the same way that being struck by lightning is cruel and unusual”).

46. *Id.* at 311 (White, J., concurring) (stating that there was no principled way to distinguish defendants who received the death penalty from those who did not).

ruling was the commutation of all defendants who were under the sentence of death and a *de facto* moratorium on executions.⁴⁷

The Court's decision to strike down prevailing death penalty statutes in *Furman* was a mixture of procedure and substance.⁴⁸ A year prior to the *Furman* decision, in *McGautha v. California*,⁴⁹ the Court ruled that statutes guiding juror discretion in capital cases were not constitutionally required under the Due Process Clause.⁵⁰ Justice Harlan, authoring the majority opinion for the Court, believed that it was impossible to develop a legal formula capable of distinguishing the worst-of-the-worst cases from the vast majority of murder cases that were not capitally prosecuted, so the unstructured sentencing authority of juries neither violated capital defendants' due process nor equal protection rights.⁵¹ The *Furman* Court expressly stated that its decision did not overrule *McGautha*, because that case had only considered due process and (arguably) equal protection objections.⁵² *Furman* deemed that the broad and unbridled discretion afforded to capital juries violated the Eighth Amendment's prohibition against cruel and unusual punishment precisely because a permissible process could generate an impermissible result. In other words, the Eighth Amendment's focus was on actual punishments, and not merely the process by which the punishment was decided. The Court declined to offer guidance, however, as to what types of procedures, if any, would produce outcomes satisfying the Eighth Amendment or how those outcomes would be policed. States were left to devise their own statutes that would, ostensibly, pass constitutional muster.

47. *Id.* at 239–40.

48. Steiker & Steiker, *supra* note 9, at 762.

49. *McGautha v. California*, 402 U.S. 183 (1971).

50. *Id.* at 207–08.

51. *Id.* at 205 (Justice Harlan reasoning that it was both unwise and futile to attempt to determine, *a priori*, the factors that would warrant a death sentence).

52. *Furman*, 408 U.S. at 310 (Stewart, J., concurring). Undoubtedly, due process concerns were at the core of *McGautha* and *Crampton*; nevertheless, several Justices also emphasized that due process and equal protection considerations were closely linked: “A vague statute may be applied one way to one person and a different way to another. Aside from the fact that this in itself would constitute a denial of equal protection the reasons underlying different applications to different individuals may in themselves be constitutionally impermissible.” *McGautha*, 402 U.S. at 259 n.9 (Brennan, J., dissenting) (internal citation omitted). *But see Furman*, 408 U.S. at 400 (Burger, J., dissenting) (“Although the Court's decision in *McGautha* was technically confined to the dictates of the Due Process Clause of the Fourteenth Amendment, rather than the Eighth Amendment as made applicable to the States through the Due Process Clause of the Fourteenth Amendment, it would be disingenuous to suggest that today's ruling has done anything less than overrule *McGautha* in the guise of an Eighth Amendment adjudication.”).

Following *Furman*, many states immediately revamped their capital statutes and resumed sentencing defendants to death.⁵³ Uncertain as to what was now constitutionally permissible under the newly revised statutes, defendants sentenced to death challenged their sentences in state and federal courts. The Court agreed to hear a group of five cases that, roughly, represented the range of post-*Furman* capital statutes: three of the cases involved guided-discretion statutes (Florida, Georgia, and Texas), while two others completely eliminated sentencing discretion and required the death sentence for a very narrow class of defendants (Louisiana and North Carolina). The Court ultimately approved the modified guided-discretion death penalty statutes in *Gregg v. Georgia*,⁵⁴ *Jurek v. Texas*,⁵⁵ and *Proffitt v. Florida*,⁵⁶ and invalidated the mandatory death penalty statutes in *Woodson v. North Carolina*⁵⁷ and *Roberts v. Louisiana*.⁵⁸ In each of these rulings, the Court reiterated that not only must the death penalty be reserved for the worst-of-the-worst offenses, but even among that limited group of persons, the death penalty is only permissible for the most culpable defendants. The Court was convinced that the guided-discretion statutes enacted after *Furman* would result in greater consistency and rationality/accuracy in the administration of the death penalty.

The statutes crafted by legislatures in Florida, Georgia, and Texas imposed different requirements on juries and reviewing courts. Under Florida's scheme, eight aggravating circumstances and seven mitigating circumstances were established to guide jury discretion.⁵⁹ Juries were required to weigh aggravating and mitigation evidence and impose a death sentence if the latter

53. See generally DAVID C. BALDUS ET AL., EQUAL JUSTICE AND THE DEATH PENALTY: A LEGAL AND EMPIRICAL ANALYSIS (1990); HERBERT H. HAINES, AGAINST CAPITAL PUNISHMENT: THE ANTI-DEATH PENALTY MOVEMENT IN AMERICA, 1972-1994 (1996). "The new sentencing schemes [adopted by many states after *Furman*] were immediately put to use. Only 42 people were sentenced to death in 1973, but there were 149 death sentences in 1974, probably more than any year since 1942. . . . In 1975, 298 people were sentenced to death—far more than any previous year for which data exist." STUART BANNER, THE DEATH PENALTY: AN AMERICAN HISTORY 270 (2002).

54. *Gregg v. Georgia*, 428 U.S. 153, 154 (1976).

55. *Jurek v. Texas*, 428 U.S. 262, 276-77 (1976).

56. *Proffitt v. Florida*, 428 U.S. 242, 259-60 (1976).

57. *Woodson v. North Carolina*, 428 U.S. 280, 305 (1976).

58. *Roberts v. Louisiana*, 428 U.S. 325, 336 (1976). The *Gregg* court expressly recognized that its primary concern in *Furman* was the arbitrary and capricious manner in which defendants were being condemned to death. *Gregg*, 428 U.S. at 195 (noting that the Court's concern about arbitrariness in *Furman* could be adequately addressed by "carefully drafted statute[s] that ensure[] that the sentencing authority is given adequate information and guidance"). Several justices also emphasized the risk of the discriminatory imposition of the death penalty. *Id.* at 206.

59. *Proffitt*, 428 U.S. at 251.

did not sufficiently outweigh the former.⁶⁰ The jury issued an “advisory” sentence by majority vote, and the judge was authorized to override the jury’s sentencing recommendation.⁶¹ All death sentences were automatically reviewed by the Florida Supreme Court.⁶² In Georgia, ten aggravating circumstances were developed, but no specific mitigating circumstances were specified.⁶³ Once the jury found at least one aggravating circumstance, it was required to weigh all of the aggravating and mitigating evidence when deciding whether to impose a death or life sentence.⁶⁴ The jury was required to be unanimous, and its sentencing recommendation was binding on the judge.⁶⁵ Defendants sentenced to death received a non-waivable review by the Georgia Supreme Court.⁶⁶ Texas’s statute included five categories of homicides and defendants were subject to the death penalty only if the killing was unprovoked, deliberate, and the defendant was likely to commit violent acts in the future. The jury’s vote for death needed to be unanimous, and a death-sentenced defendant received an automatic (and non-waivable) review by the Texas Court of Criminal Appeals.⁶⁷

Despite their differences, the important commonalities of the statutes were (1) a list of factors that would, ostensibly, narrow the reach of the death penalty and identify the most culpable defendants to the sentencing authority,⁶⁸ and (2) mandatory appellate review of death sentences by the jurisdiction’s highest criminal court that would assess the appropriateness of every death sentence imposed. Whereas Georgia’s statute clearly outlined the

60. *Id.* at 246.

61. *Id.* at 248–49.

62. *Id.* at 250–51.

63. *Gregg*, 428 U.S. at 154.

64. *Id.* at 193–95.

65. *Id.* at 153–54.

66. *Id.* at 156.

67. *Jurek v. Texas*, 428 U.S. 262, 269 (1976). Texas’ statute limited capital homicides to intentional murders committed in the following situations: murder of a peace officer or fireman; murder committed in the course of kidnaping, burglary, robbery, forcible rape, or arson; murder committed for remuneration; murder committed while escaping or attempting to escape from a penal institution; and murder committed by a prison inmate when the victim is a prison employee. *See* TEX. PENAL CODE ANN. § 19.03 (West 1974).

68. By extension, the factors would also constrain the charging authority because certain elements of the crime must be proven to the sentencing authority (e.g., killing of a police officer) in order for the death penalty be an available sentencing option. The Court would repeatedly (re)emphasize that capital statutes must “genuinely narrow” the death-eligible class to encompass only defendants materially more depraved than the average murderer. *See Atkins v. Virginia*, 536 U.S. 304, 319 (2002) (“Since *Gregg* our jurisprudence has consistently confined the imposition of the death penalty to a narrow category of the most serious crimes.”); *Zant v. Stephens*, 462 U.S. 862, 876–77 (1983).

appellate court's task,⁶⁹ the capital statutes from Florida and Texas did not clearly specify what appellate review would entail.⁷⁰ Nonetheless, the Court noted that, in practice, the reviewing courts in those states were determining whether each defendant's death sentence was arbitrarily imposed, disproportionate, or the product of any impermissible consideration.⁷¹ The Court, once again, signaled that the consideration of both procedure (i.e., narrowing death-eligibility) and results (i.e., appellate review of capital sentences irrespective of the whether the process was followed) were indispensable components of a constitutionally permissible death penalty system.⁷² In other words, a "fair" death penalty system must satisfy procedural and distributive justice concerns.

Glaring omissions from both the revised statutes and the Court's analysis of them, however, were workable definitions of arbitrariness, bias, and disproportionality. The Court and legislatures employed intuitive understandings of these concepts, but they failed to translate these general principles into terms that frontline legal actors—e.g., prosecutors, juries, and appellate courts—could actually put into operation. How were errors of arbitrariness, bias, and disproportionality to be measured in the capital sentencing context? What baselines should be used? What threshold showings must be made before these various claims of constitutional error were cognizable by the Court? These key unresolved questions jeopardized the heightened reliability required under the Court's "death is different" approach to the Eighth Amendment.

Several of the Court's subsequent rulings underscored its coarse and inelegant analysis in *Gregg*. In each of these cases, the Court appeared to retreat from its initial positions in *Furman* and *Gregg*, yet the Court neither expressly overruled those initial cases, redefined the core principles and standards articulated in those cases, nor provided meaningful clarity to

69. The Georgia Supreme Court was required to decide three things: (1) whether the death sentence was imposed under the influence of passion, prejudice or any other arbitrary factor; (2) whether the evidence supports the aggravating circumstance(s) found by the jury; and (3) whether the death sentence is excessive or disproportionate, relative to the penalty imposed in similar cases, considering the crime and the defendant. *Gregg*, 428 U.S. at 153–54.

70. All death sentences were automatically reviewed by the Florida Supreme Court. The statute did not specify what the mandatory review would entail, but the Court underscored that, in practice, inter-case review was conducted by the Florida Supreme Court. *Proffitt v. Florida*, 428 U.S. 242, 250–51 (1976). Texas' revised statute did not require comparative proportionality review, although each death sentence was to be reviewed, at least, on its own merits to ensure that death sentences "will not be 'wantonly' or 'freakishly' imposed." *Jurek*, 428 U.S. at 276.

71. See *Jurek*, 428 U.S. at 276 ("By providing prompt judicial review of the jury's decision in a court with statewide jurisdiction, Texas has provided a means to promote the evenhanded, rational, and consistent imposition of death sentences under law.").

72. *Gregg*, 428 U.S. at 206–07.

legislatures and frontline actors in the capital charging-and-sentencing process to assist in devising systems capable of minimizing errors. Complicating matters further, after emphasizing the centrality of narrowly-crafted capital statutes to guide the sentencing authority and minimize arbitrary (and capricious) decision-making, the Court ruled that states were prohibited from limiting the type of mitigation (i.e., exculpatory) evidence defendants could present at trial.⁷³ A few years later, the Court held that the prosecutor could present, and the sentencing authority could consider, aggravating (i.e., inculpatory) evidence that was not specifically enumerated in the capital statute.⁷⁴ The Court waited seven years to clarify its Eighth Amendment proportionality analysis framework,⁷⁵ although it deemed proportionality determinations by state reviewing courts critical features of the post-*Furman* statutes.⁷⁶ It took the Court an additional four years to announce that statistical evidence of pronounced and persistent racial disparities in death penalty charging and sentencing patterns was inadequate to prove an unacceptable risk of arbitrariness or intentional racial discrimination in the administration of capital punishment in a state.⁷⁷ More than four decades after *Furman*, the doctrinal haziness remains and there is

73. *Lockett v. Ohio*, 438 U.S. 586, 608 (1978).

74. *Zant v. Stephens*, 462 U.S. 862, 884–89 (1983); *Wainwright v. Goode*, 464 U.S. 78, 86–87 (1983).

75. *Solem v. Helm*, 463 U.S. 277, 290–92 (1983) (noting that reviewing courts were required to do three things when determining whether a punishment was disproportionate: (1) compare the nature and gravity of the offense and the harshness of the penalty; (2) compare the sentences imposed on other criminals in the same jurisdiction [intra-jurisdictional]; and (3) compare the sentences imposed for commission of the same crime in other jurisdictions [inter-jurisdictional]).

The Court acknowledged the difficulty reviewing courts face in attempting to draw distinctions between similar crimes (and criminals), but it did not believe this was an insurmountable obstacle because reviewing courts were competent to judge the gravity of the offense and the defendant’s culpability on a relative scale. *Id.* at 294.

Gregg did announce an analytical framework for assessing whether a punishment was disproportionate, and therefore violative of the Eighth Amendment. *Gregg*, 428 U.S. at 188–95. But the Court’s focus was on the constitutionality of capital punishment per se, and not the potential excessiveness of any individual death sentence. *Id.* at 176.

76. The Court previously ruled that punishments were “excessive,” and therefore prohibited by the Eighth Amendment, if not graduated and proportioned to the offense. *See Weems v. United States*, 217 U.S. 349, 366–67 (1910). This definition failed to provide reviewing courts with any meaningful guidance until *Gregg*. *Gregg*, 428 U.S. at 171–73.

77. *McCleskey v. Kemp*, 481 U.S. 279, 314–19 (1987). A growing body of evidence documenting persistent gender bias in the administration of the death penalty has emerged since the Court’s ruling in *McCleskey*. *See, e.g.*, Steven F. Shatz & Naomi R. Shatz, *Chivalry Is Not Dead: Murder, Gender, and the Death Penalty*, 27 *BERKELEY J. GENDER L. & JUST.* 64, 64–65 (2012) (capital cases involving female murder defendants are treated more leniently and cases involving female victims are treated more harshly, all else being equal).

little indication that the Court will offer any elucidation.⁷⁸ The post-*Furman* Court appears content with looking where it believes the light is brightest and never journeying too far from its perceived areas of competence: statutory interpretation and procedural innovation. However, there may be a strong shadow where there is much light.⁷⁹ The Court's "continu[ed] [] treat[ment] [of] the social facts and empirical data that document systemic failures in death penalty imposition as somehow irrelevant to constitutional decision making seems increasingly indefensible."⁸⁰ If repairing the death penalty system is possible, it requires the Court's reengagement with the social facts of capital charging-and-sentencing practices. Indispensable to this undertaking are clear and consistent rules and standards governing the administration of capital punishment from both a process and outcome perspective. To date, the Court has refused to perform these essential functions.

B. *Conceptualizing and Operationalizing Capital Error*

The Court's consistent refusal to provide clear standards and guidance to lower courts, lawyers, and legislatures for assessing the forms of constitutional error described in *Furman* and *Gregg* has allowed it to avoid upholding the very principles and rules it established.⁸¹ By sidestepping precise definitions of constitutional error, and by extension, the evidentiary thresholds that parties must satisfy to make a colorable claim to the Court, the justices were insulating themselves from the "excruciatingly difficult responsibility for deciding who the State may and may not constitutionally kill."⁸² The promise of *Furman* (and arguably *Gregg*) cannot be fulfilled without the Court, intelligibly, describing how systems may satisfy or fail the constitutional standards it developed. In the absence of these standards from the Court, scholars must look elsewhere for guidance. Even if the Court is skeptical of, or unreceptive to, analysts resorting to standards developed outside of its jurisprudence to identify and quantify the types of constitutional error announced in *Furman* and its progeny, the use of articulable and defensible standards may gain an audience and traction outside of the courtroom. This, in turn, could potentially exert pressure on the Court to

78. See *infra* Part B, where I provide improved clarity of the Court's doctrine utilizing the framework of statistical inference.

79. JOHANN WOLFGANG VON GOETHE, GÖTZ VON BERLICHINGEN act I, at 15 (1773).

80. HANEY, *supra* note 33, at 216.

81. See, e.g., Liebman, *supra* note 5 (describing the Court's avoidance of policing the standards it developed for the regulation of capital punishment systems).

82. *Id.* at 5.

adopt these general standards, or more precisely define its own. The widely accepted framework of statistical inference can provide the depoliticized clarity that is lacking from the Court's current death penalty jurisprudence.⁸³ "Statistics is the art of making numerical conjectures about puzzling questions."⁸⁴ Statistical inference, among other things, provides a useful framework for identifying and quantifying errors produced by a system. The Court's concern over errors of arbitrariness directly implicates the reliability, validity, and proportionality of capital decision-making. These three components of arbitrariness are discussed *in seriatim* below.

1. Reliability

The reliability of a system stems from the absence of random error, and is related to the consistency or stability of outcomes across decisions and/or decision makers (e.g., prosecutors).⁸⁵ Thus, when a repeated process gives highly similar results, the process is said to be reliable.⁸⁶ While random errors are inevitable and some may even be socially optimal given the costs of reducing such error relative to the overall benefit from the reduction, a system purporting to provide heightened reliability—as required by the Court's modern death penalty jurisprudence—must minimize the inconsistency of the outcomes that system generates. Random errors can often be reduced through systems that structure the exercise of discretion of criminal justice actors.⁸⁷

It may be difficult, if not impossible, to specify a priori how much inconsistency is allowable before a system should be deemed unconstitutionally arbitrary; nevertheless, a highly inconsistent system can hardly be characterized as a "fair and evenhanded" one.⁸⁸ Relatedly, if the inconsistency of the behavior of legal actors is strongly associated with legally illegitimate factors, such as race/ethnicity and gender, then there should be heightened concern about the risk of arbitrariness. The guided-

83. Oliver Wendell Holmes famously wrote, "[f]or the rational study of the law the blackletter man may be the man of the present, but the man of the future is the man of statistics and the master of economics." Oliver Wendell Holmes, *The Path of Law*, 10 HARV. L. REV. 457, 469 (1897).

84. DAVID A. FREEDMAN ET AL., STATISTICS xiii (4th ed. 2007).

85. GARY KING ET AL., DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH 25 (1994).

86. FORST, *supra* note 15, at 33; W. PAUL VOGT, DICTIONARY OF STATISTICS & METHODOLOGY: A NONTECHNICAL GUIDE FOR THE SOCIAL SCIENCES 274 (3d ed. 2005).

87. FORST, *supra* note 15, at 181.

88. *Furman v. Georgia*, 408 U.S. 238, 309–10 (1972) (Stewart, J., concurring); *id.* at 312–13 (White, J., concurring) (the death penalty must be meted out in a consistent, fair, and evenhanded manner).

discretion statutes approved by the Court in *Gregg* were designed to rationalize the capital punishment process. Such rationalization required the significant reduction of the arbitrary application of the death penalty. Statutes that do not meaningfully minimize erratic charging-and-sentencing outcomes have failed to sufficiently constrain frontline legal actors—namely prosecutors, judges, and jurors—to satisfy *Furman*.

2. Validity

The validity of a system refers to its truthfulness or accuracy.⁸⁹ That is, the system actually produces what it is purported and designed to produce.⁹⁰ Validity stems from both the absence of systematic error and justifiability of the interpretations about the data.⁹¹ A system typically generates invalid results when it is incorrectly calibrated or actors implementing the system use different procedures and considerations.⁹² Assessments of culpability should be aligned with the standards set forth in the applicable capital statutes.⁹³ Validity, then, refers to the degree to which theory and evidence support the interpretation of legal decision-makers.⁹⁴ The validity of an assessment is also closely related to its fairness. In the educational testing literature, assessments are deemed fair when they assess what is “taught.”⁹⁵ In the capital charging context, the death penalty statute provides prosecutors the standards/metrics upon which culpability assessments must be made in order to comport with the constitutional requirements announced in *Furman* and its progeny. So, in concrete terms, if the factors specified in a capital statute fail to predict actual outcomes within an acceptable range, then there is compelling evidence that the system permits an impermissible degree of arbitrariness and is, therefore,

89. There are many sub-species of “validity.” W. PAUL VOGT, *supra* note 86. In the present context, validity refers to “internal validity.” FORST, *supra* note 15, at 33 (“‘Threat to internal validity’ is generally defined in such a way as to include any factor that jeopardizes the accuracy of the test of a theory.”). *See generally* THOMAS D. COOK & DONALD T. CAMPBELL, *QUASI-EXPERIMENTATION: DESIGN AND ANALYSIS ISSUES FOR FIELD SETTINGS* (1979) (defining and identifying threats to internal validity).

90. KING ET AL., *supra* note 85.

91. The lack of reliability qualifies as a threat to the validity of a system, whereas the absence of validity is not necessarily a threat to the reliability (i.e., consistency) of a system. FORST, *supra* note 15, at 33; *see also* W. Steve Lang & Judy R. Wilkerson, *Accuracy vs. Validity, Consistency vs. Reliability, and Fairness vs. Absence of Bias: A Call for Quality* (Feb. 2008) (unpublished manuscript) (“[V]alidity means that assessors are making justifiable interpretations about their data and good decisions.”).

92. FORST, *supra* note 15, at 33–34.

93. *See generally* FORST, *supra* note 15, at 55.

94. Lang & Wilkerson, *supra* note 91, at 7–8.

95. *Id.* at 13.

constitutionally dubious. Of course, the definition of an “acceptable range” will be subject to debate, but relative assessments can be made by gauging the predictability of similarly structured systems with aligned objectives. Low explanatory power of statutorily relevant factors provides evidence of invalid charging decisions.⁹⁶

Unmeasured or improper factors that influence the functioning of a system also undermine its accuracy.⁹⁷ When the relative explanatory power, and therefore (ir)rationality, of capital charging decisions is associated with legally illegitimate factors, such as race/ethnicity, gender, and socioeconomic status, then there is reason to believe that those decisions are unconstitutionally arbitrary because prosecutors, judges, and juries are prohibited from considering those factors. A valid assessment is free of racial/ethnic bias and stereotypes so, by definition, decisions based in whole or in part on these factors are legally invalid.⁹⁸ The same would hold true for cases differing only with respect to the gender of the defendant or victim.⁹⁹

Admittedly, the strength of the association between legally impermissible factors and charging-and-sentencing outcomes necessary to qualify as constitutional error is far from obvious. Similar to the aforementioned assessment of arbitrariness, it will be difficult to specify, *a priori*, the magnitude of the association necessary to qualify as constitutionally impermissible error. Relative assessments, again, may be required. A rational death penalty system, however, clearly identifies the factors that decision-makers must not consider when making culpability determinations with respect to charging and sentencing. Conscious discrimination on the part of decision-makers need not be demonstrated in order to substantiate a claim of a legally arbitrary system. Not only is such evidence nearly impossible to obtain because actors have very little incentive to admit this type of wrong-

96. See Richard A. Berk et al., *Statistical Difficulties in Determining the Role of Race in Capital Cases: A Reanalysis of Data from the State of Maryland*, 21 J. QUANTITATIVE CRIMINOLOGY 365, 386–87 (2005) (the low explanatory power of statistical models of capital charging-and-sentencing decisions is attributable to an arbitrary and irrational process); Lang & Wilkerson, *supra* note 91, at 5–10 (valid assessments must comport with predefined standards).

97. FORST, *supra* note 15, at 37.

98. Lang & Wilkerson, *supra* note 91, at 15 (racial and ethnic bias undermine the validity of an assessment). *But see* Russell D. Covey, *The Unbearable Lightness of Batson: Mixed Motives and Discrimination in Jury Selection*, 66 MD. L. REV. 279 (2006) (describing some courts’ endorsement of a “mixed-motives” analysis in jury selection that permits racially-motivated reasoning).

99. See, e.g., *Furman v. Georgia*, 408 U.S. 238, 363–65 (1972) (Marshall, J., concurring) (noting that unjustified gender disparities in the administration of capital punishment would be unconstitutional).

doing,¹⁰⁰ but evidence of unconscious/implicit bias in capital punishment decision-making is equally discriminatory and violates the Court's heightened reliability standard.¹⁰¹ While it is true that the U.S. Constitution has never been interpreted to require identical punishments for similar situated defendants,¹⁰² evidence of a robust pattern of the influence of impermissible factors on these charging and sentencing outcomes is highly probative of the level of bias permitted by the system.¹⁰³ A justice policy, *de jure* or *de facto*, that biases outcomes towards over- or under-sanctioning produces systematic error. These "less honest errors appear all too common" and typically result from a lack of professionalism coupled with flawed systems of accountability.¹⁰⁴

3. Proportionality

Concerns about proportionality have both macro-level and micro-level dimensions. At the macro-level, a punishment can be excessive and therefore prohibited by the Eighth Amendment, if not graduated and proportioned to the offense.¹⁰⁵ The inquiry does not focus on the individual defendant, rather the gravity of the alleged offense and the harshness of the penalty is at issue.¹⁰⁶ The Court has narrowed the reach of the death penalty over the past

100. STEVE WEINBURG, HARMFUL ERROR: INVESTIGATING AMERICA'S LOCAL PROSECUTORS (2003) (noting that there is little incentive for prosecutors to admit wrongdoing in murder cases because misconduct is treated with great leniency).

101. See Scott Howe, *The Futile Quest for Racial Neutrality in Capital Selection and the Eighth Amendment Argument for Abolition Based on Unconscious Racial Discrimination*, 45 WM. & MARY L. REV. 2083 (2004); Sheri L. Johnson, *Unconscious Racism and the Criminal Law*, 73 CORNELL L. REV. 1016 (1988).

102. See generally *McCleskey v. Kemp*, 481 U.S. 279, 297–98 (1987) (finding that a correlation between race and the imposition of the death penalty, even after accounting for a host of legally relevant variables, was not indicative of a discriminatory purpose); Ernest van den Haag, *The Ultimate Punishment: A Defense*, 99 HARV. L. REV. 1662, 1662–64 (1986) (arguing that some inequality in the application of the death penalty is unavoidable as a practical matter, but such inequalities are constitutionally permissible if not the product of irrational discrimination).

103. E.g., Jon R. Sorensen & Donald H. Wallace, *Prosecutorial Discretion in Seeking Death: An Analysis of Racial Disparity in the Pretrial Stages of Case Processing in a Midwestern County*, 16 JUST. Q. 559, 576 (1999) (drawing inferences of intentional discrimination from statistical data of capital charging decisions by a single prosecutor).

104. FORST, *supra* note 15, at 17.

105. *Weems v. United States*, 217 U.S. 349, 366–67 (1910).

106. The Court's analytical framework for assessing whether a punishment was disproportionate, per se, and therefore violative of the Eighth Amendment was announced in *Gregg*. The three-step inquiry involved: (1) assessing whether the punishment employs cruel methods or involved unnecessary and wanton infliction of pain, (2) considering society's evolving standards of decency as reflected by legislative judgments and jury verdicts, and (3) determining

four decades, but has maintained that the death penalty is not, *per se*, excessive for the crime of murder when committed by an adult who does not suffer from extreme intellectual disability.¹⁰⁷ At the micro-level, a proportionality analysis for an individual defendant requires a comparison to sentences imposed on other criminals in the same jurisdiction (intra-jurisdictional) and sentences imposed for commission of the same crime in other jurisdictions (inter-jurisdictional).¹⁰⁸ An individual criminal charge or punishment may be excessive because it is used very infrequently against similarly situated defendants (i.e., arbitrary). Similarly, if a charge or punishment is used very frequently against defendants who share a trait that is prohibited from legal consideration, but not against other similarly situated defendants who do not share that trait, that charge or punishment may be excessive.

Disproportionality is a matter of degree and no clear guidelines exist for determining what is excessive in any particularly situation. As explained, *supra*, the Court has expressly endorsed capital punishment statutes that required reviewing courts to make proportionality determinations for each defendant receiving a death sentence. Irrespective of the specific designs of those proportionality review systems, meaningful proportionality review requires the development of a data-driven metric that permits the analyst to clearly situate each defendant vis-à-vis other defendants when making these determinations.¹⁰⁹

whether the punishment offends “human dignity.” *Gregg v. Georgia*, 428 U.S. 153 (1976). The Court would later note that “[S]tate legislatures and sentencing juries do not wholly determine this controversy [over what constitutes excessive punishment under the Eighth Amendment], for the Constitution contemplates that in the end our own judgment will be brought to bear on the question.” *Coker v. Georgia*, 433 U.S. 584, 597 (1977); *see also supra* text accompanying note 75.

107. *Roper v. Simmons*, 543 U.S. 551, 551 (2005) (outlawing capital punishment for juvenile offenders, but upholding its constitutionality for adult murderers); *see also Atkins v. Virginia*, 536 U.S. 304, 321 (2002) (prohibiting the death penalty for mentally disabled defendants); *Eberheart v. Georgia*, 433 U.S. 917, 917 (1977) (prohibiting the death penalty for non-homicidal kidnapping); *Coker*, 433 U.S. at 600 (prohibiting the death penalty for non-homicidal rape).

108. *See supra* text accompanying note 75; *see also Solem v. Helm*, 463 U.S. 277, 291–92 (1983).

109. I will discuss proportionality review systems in greater detail in Part III. *See infra* Part III.

C. Summary

Reliability, validity, and proportionality are empirical questions, not legal ones.¹¹⁰ They all relate to the patterns of measurement error. By clearly defining how different patterns of error correspond to these distinct, but often related concepts, it is possible to assess the performance of a system. Attorneys in *Furman* advanced factual claims about the death penalty that were subject to social scientific inquiry and, as a consequence, *Furman* contained the most extensive discussion of social science research in any decision before or since.¹¹¹ In response to these facts, the *Furman* court announced that systems must be modified to substantially reduce these patterns—the procedures that states developed were to be a means to an end: fundamental fairness. The careful collection and analysis of detailed data on death penalty charging-and-sentencing practices still provides the best opportunity to directly explore the level of arbitrariness at which a statute operates.¹¹² It can also provide useful insights into the excessiveness of individual charging or sentencing decisions. Part II describes and implements an exportable framework for accomplishing these tasks.

II. DIAGNOSING CAPITAL ERROR

This Section presents a statistical model capable of assessing the level of arbitrariness present in a capital charging system and implements the model using charging data from Georgia. Part A provides a very general overview of the model and explains why it is particularly well-suited for analyzing capital charging data. A more technical discussion of the model is provided in the Methodological Appendix. Part B illustrates the usefulness of the model in identifying macro-level (i.e., systemic) errors through an analysis of eight years of death penalty charging decisions in Georgia. Part C demonstrates the usefulness of the model for assessing micro-level errors—

110. HANEY, *supra* note 33, at 3–23 (explaining that the style of reasoning in *Furman* suggested that the real facts and actual operation of the death penalty would be at the forefront of any future litigation and judicial decisions that pertained to its constitutionality); see Robert J. Smith, *The Geography of the Death Penalty and Its Ramifications*, 92 B.U. L. REV. 227, 255 (2012).

111. HANEY, *supra* note 33, at 10. *Furman* contained over sixty footnotes citing published social science research on the realities of the death penalty process. Justices writing for both the majority and the dissent grappled with social science evidence.

112. Disaggregating death penalty practices to the county-level (or some other sub-state level unit) permits the identification of a state-level baseline (i.e., statewide average) governing by the same capital statute. Kyle Graham, *Overcharging*, 11 OHIO ST. J. CRIM. L. 701, 710–11 (2014) (noting the difficulty of defining a widely accepted baseline by which to compare prosecutors' charging decisions).

i.e., the (in)appropriateness of seeking the death sentence against a defendant in a particular case. Part D discusses the generalizability of the empirical findings as well as the exportability of the statistical model.

A. Modeling Capital Charging

Prosecutors wield nearly unbridled discretion when making charging decisions.¹¹³ The capital statutes approved by the Court in *Gregg* did very little to regulate prosecutorial behavior, although the defendants in *Gregg* specifically identified the absence of appropriate checks on charging and clemency decisions as a fatal flaw of those capital statutes. Not only does inadequate charge screening increase the risk of arbitrariness, bias, and excessiveness, it can also have strong ripple effects throughout the remainder of the death penalty trial and appellate process: acquittal and reversal rates are greatest in cases that are not highly aggravated or involve defendants with strong mitigation evidence.¹¹⁴ The practical consequences of this cycle are millions spent on correcting trial level error and tremendous delay and uncertainty of executions.¹¹⁵ Modeling capital charging with a specific focus on arbitrariness not only provides invaluable insights into the “front-end” of the death penalty process, but it is also relevant to understanding downstream legal error.¹¹⁶

113. STEPHANOS BIBAS, *THE MACHINERY OF CRIMINAL JUSTICE* 41–48 (2012) (describing the unchecked powers of prosecutors); James S. Liebman, *The Overproduction of Death*, 100 COLUM. L. REV. 2030, 2101–29 (2000) (same).

114. Nationally, one-third of capital trials result in a death sentence, one-third of those *initial* death sentences are upheld on appeal, and one-tenth of inmates sentenced to death are executed. RICHARD C. DIETER, DEATH PENALTY INFORMATION CENTER, *SMART ON CRIME: RECONSIDERING THE DEATH PENALTY IN A TIME OF ECONOMIC CRISIS* 14 (2009) (reporting conviction and execution statistics); James S. Liebman et al., *Capital Attrition: Error Rates in Capital Cases, 1973-1995*, 78 TEX. L. REV. 1839, 1862–63 (2000) (reporting that sixty-eight percent of death sentences initially imposed at trial were overturned on appeal); *see also* Part I.

115. RICHARD C. DIETER, DEATH PENALTY INFORMATION CENTER, *MILLIONS MISSPENT: WHAT POLITICIANS DON'T SAY ABOUT THE HIGH COSTS OF THE DEATH PENALTY* 6–8 (1994) (discussing the great financial costs of error correction in capital cases); Arthur L. Alarcon & Paula M. Mitchell, *Executing the Will of the Voters?: A Roadmap to Mend or End the California Legislature's Multi-Billion-Dollar Death Penalty Debacle*, 44 LOY. L.A. L. REV. S41 (2011) (describing federal expenditures on capital cases arising in state court).

116. The vast majority of studies examining the capital punishment process focus on the discretionary choices of actors in the criminal justice system *after* a death penalty notice has been filed (and frequently *after* a death sentence has been imposed). This limited focus largely stems from the fact that most analysts lack adequate information on the population of defendants who might be subject to capital punishment, so researchers concentrate on those cases once they formally enter the system. Focusing on capital charging decisions is especially important because many prosecutors charge defendants capitally in order to induce a plea agreement with the

Consistent with prior research, I model capital charging decisions as a function of legally permissible and impermissible case-level factors.¹¹⁷ The empirical relationships between these factors and capital charging decisions is described through the use of multivariate regression. Regression analysis is the primary technique employed by quantitatively-oriented lawyers and social/behavioral scientists for the identification general patterns of arbitrariness and bias in capital decision-making.¹¹⁸ Indeed, regression is most powerful (and least controversial) when used as a descriptive technique to identify conditional averages and variances based on these observed patterns.¹¹⁹ When factors specified in a capital statute are highly predictive of actual outcomes, one can infer that the process which generated the outcome is rational and, therefore, non-arbitrary. Relatedly, when like cases are treated similarly, then a process is considered consistent.¹²⁰ Prosecutors are prohibited from basing their decisions to seek the death penalty on defendants' or victims' race/ethnicity,¹²¹ so if these factors exert an influence on capital charging, all else being equal, one can infer the process that generated the outcomes is legally arbitrary. It is unnecessary to assume that the racial/ethnic disparities are motivated by intentional bias. Compelling evidence of the effects unconscious/implicit bias in capital decision-making may still support a constitutional challenge under the Eighth Amendment.¹²²

defendant and avoid the time and expense associated with trial. *See, e.g.*, Susan Ehrhard, *Plea Bargaining and the Death Penalty: An Exploratory Study*, 29 JUST. SYS. J. 313 (2008); Susan Ehrhard-Dietzel, *The Use of Life and Death as Tools in Plea Bargaining*, 37 CRIM. JUST. REV. 89 (2012); Sherod Thaxton, *Leveraging Death*, 103 J. CRIM. L. & CRIMINOLOGY 475 (2013).

117. *See, e.g.*, John J. Donohue, *An Empirical Evaluation of the Connecticut Death Penalty System Since 1973: Are There Unlawful Racial, Gender, and Geographic Disparities?*, 11 J. EMPIRICAL LEGAL STUD. 637 (2014).

118. *E.g.*, BALDUS ET AL., *supra* note 53; Donohue, *supra* note 117, at 650.

119. RICHARD A. BERK, REGRESSION ANALYSIS: A CONSTRUCTIVE CRITIQUE 12–13 (2003).

120. A model of prosecutorial charging behavior should approximate the actual data-generating process, so it is necessary to include legally impermissible factors when it is hypothesized that such factors are predictive of outcomes. To do otherwise would imply that the model precludes these variables from having any predictive power, and such an assumption would result in an improperly specified model. This also holds true when modeling discretionary choices at other stages of the criminal justice process. *See, e.g.*, John Wooldredge, *Distinguishing Race Effects on Pre-Trial Release and Sentencing Decisions*, 29 JUST. Q. 41 (2012) (modeling pretrial detention behavior and including race/ethnicity as an explanatory variable).

121. *See* *McCleskey v. Kemp*, 481 U.S. 279 (1987); *Furman v. Georgia*, 408 U.S. 238 (1972).

122. *See supra* note 100 and accompanying text. Although it remains unclear what constitutes “compelling” evidence, the New Jersey Supreme Court stated that it would be receptive to statistical proof of significant racial disparities in the administration of capital punishment and believed the evidence presented in *McCleskey* was very persuasive. *State v. Loftin*, 724 A.2d 129, 151 (N.J. 1999); *cf. United States v. Bass*, 266 F.3d 532, 540 (6th Cir. 2001) (holding that national statistics of racially disparate death penalty charging patterns were sufficient to satisfy the

The empirical strategy adopted in this paper uses a variant of regression analysis to develop a model for analyzing the arbitrariness in prosecutors' charging decisions.¹²³ The model, commonly referred to as a "multilevel model" (MLM) (or "variance components model"), has been under-utilized in death penalty research, but has important advantages over other types of regression when analyzing differences in case outcomes across jurisdictions.¹²⁴ I will demonstrate certain of these advantages in the application segment of the analysis, *infra*, Part B.¹²⁵ This Article is the first

threshold showing of a selective prosecution claim to merit discovery of the prosecution's files pertaining to charging decisions), *rev'd*, 536 U.S. 862 (2002).

123. A detailed discussion of these models is beyond the scope of this project, but very accessible descriptions of these models are readily available. *See generally* ANDREW GELMAN & JENNIFER HILL, DATA ANALYSIS USING REGRESSION AND MULTILEVEL/HIERARCHICAL MODELS (2007); JOSÉ C. PINHEIRO & DOUGLAS M. BATES, MIXED-EFFECTS MODELS IN S AND S-PLUS (2000); STEPHEN W. RAUDENBUSH & ANTHONY S. BRYK, HIERARCHICAL LINEAR MODELS: APPLICATIONS AND DATA ANALYSIS METHODS (2d ed. 2002).

124. MLMs have a long history in agriculture and educational statistics, where a nested data structure is natural in these settings (e.g., animals on plots, students in classrooms, etc.), and have become more popular in the social sciences in recent years. GELMAN & HILL, *supra* note 123, at 276. Recently, MLMs have been used to study state and demographic trends in public opinion about the death penalty. Shirley & Gelman, *supra* note 12, at 17 (discovering that death penalty support levels in northern and southern states have moved in opposite directions over the past fifty years).

125. In brief, these models offer significant improvements with respect to data reduction, prediction, and causal inference. GELMAN & HILL, *supra* note 123, at 246; Bradford S. Jones & Marco R. Steenbergen, *Modeling Multilevel Data Structures*, 46 AM. J. POL. SCI. 218, 219 (2002).

application of MLMs to capital charging decisions.¹²⁶ A more technical discussion of the model is provided in the Methodological Appendix.¹²⁷

The MLMs I employ allow a more accurate examination and quantification of the sources of variability in capital charging practices. Uniformity in prosecutorial charging decisions is not required under the Court's death penalty jurisprudence, but substantial variation in charging behavior both within and across jurisdictions that is not attributable to legally legitimate case characteristics may be indicative of arbitrariness or bias (or both) in a system.¹²⁸ Empirical studies of capital charging-and-sentencing decisions routinely discover that the death penalty is a "minority practice" that exerts a huge "majority burden."¹²⁹ Capital charging and sentencing activity remains concentrated among a small subset of counties.¹³⁰ The

126. A limited number of studies have examined capital charging decisions. Earlier research was criticized for failing to properly account for legally relevant differences across the cases. See, e.g., Gary Kleck, *Racial Discrimination in Criminal Sentencing: A Critical Evaluation of the Evidence with Additional Evidence on the Death Penalty*, 46 AM. SOC. REV. 783 (1981); Gary Kleck, Book Review, 20 CONTEMP. SOC. 598 (1991) (reviewing BALDUS ET AL., *EQUAL JUSTICE AND THE DEATH PENALTY: A LEGAL AND EMPIRICAL ANALYSIS*). More recently, methodologically rigorous studies revealed that prosecutors were four times more likely to seek the death penalty in white-victim cases compared to black-victim cases in Colorado, three times more likely in Georgia, twice as likely in Maryland, and three times more likely in South Carolina. See BALDUS ET AL., *supra* note 53; Stephanie Hindson et al., *Race, Gender, Region and Death Sentencing in Colorado, 1980-1999*, 77 U. COLO. L. REV. 549 (2006); Raymond Paternoster & Robert Brame, *Reassessing Race Disparities in Maryland Capital Cases*, 46 CRIMINOLOGY 971 (2008); Michael J. Songer & Issac Unah, *The Effect of Race, Gender, and Location on Prosecutorial Decisions to Seek the Death Penalty in South Carolina*, 58 S.C. L. REV. 161 (2006); Isaac Unah, *Choosing Those Who Will Die: The Effect of Race, Gender, and Law in Prosecutorial Decision to Seek the Death Penalty in Durham County, North Carolina*, 15 MICH. J. RACE & L. 135 (2009). Although these studies have contributed, considerably, to our understanding of arbitrariness and bias in capital charging, the research designs fail to fully exploit the county-centric nature of capital decision-making in order to obtain a deeper understanding of the inconsistency in death charging dynamics.

127. See *infra* Methodological Appendix. See generally Andrew Bell & Kelvyn Jones, *Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data*, 3 POL. SCI. RES. & METHODS 133, 139 (2015) (listing the key shortcomings of non-MLMs when applied to data that have a nested structure).

128. See, e.g., *Bush v. Gore*, 531 U.S. 98, 109 (2000) (per curiam) (explaining that the absence of specific standards for the equal application of the law may be unconstitutional under certain circumstances when a government body is empowered to assure uniformity).

129. James S. Liebman & Peter Clarke, *Minority Practice, Majority's Burden: The Death Penalty Today*, 9 OHIO ST. J. CRIM. L. 255 (2011).

130. There are currently 3,143 counties in the United States, but only 15% of these counties account for all executions since the Court lifted its de facto moratorium on executions in 1976. RICHARD C. DIETER, DEATH PENALTY INFO. CTR., *THE 2% DEATH PENALTY: HOW A MINORITY OF COUNTIES PRODUCE MOST DEATH CASES AT ENORMOUS COSTS TO ALL* (2013); Smith, *supra* note 110. From 2004–2009, only 10% of counties returned a single death verdict and approximately 1% of counties consistently produced one death verdict per year over the six year period. Smith,

regional clustering of capital charging decisions is important because “it permits a tailored and rigorous analysis for gauging the continued constitutionality of capital punishment.”¹³¹ Developing legally cognizable claims about the arbitrariness (and biasedness) of the death penalty can be buttressed by disaggregating statewide data of capital charging practices to sub-state units, and thereby permitting the careful examination of within- and between-county variability in the use of capital punishment for (legally) similarly situated defendants.¹³² County-disaggregated analyses may be relevant at both trial and appellate levels when raising challenges focusing on intra-county, intra-state, and inter-state arbitrariness and bias,¹³³ as well as

supra note 110; *see also* DIETER, *supra* note 130 (noting that only 2% of counties account for the vast majority of death sentences since capital punishment was reinstated in 1976); Richard Willing & Gary Fields, *Geography of the Death Penalty*, USA TODAY, Dec. 20, 1999, at A1 (fifteen counties account for nearly a third of all prisoners sentenced to death, but only one-ninth of the population of the states with capital punishment).

From 1995–2000, 42% of federal death penalty prosecutions submitted for authorization came from just 5 of the 94 federal districts. U.S. DEP’T OF JUST., THE FEDERAL DEATH PENALTY SYSTEM: A STATISTICAL SURVEY, 1988-2000, at T-14–17 tbl.5a (2000).

131. Smith, *supra* note 110, at 229.

132. *Id.* at 247; Carol S. Steiker & Jordan M. Steiker, *A Tale of Two Nations: Implementation of the Death Penalty in Executing Versus Symbolic States in the United States*, 84 TEX. L. REV. 1869 (2005).

Focusing to a single decision-maker (or smaller group of decision-makers at a particular stage of the capital charging-and-sentencing process) partly addresses the Court’s concern about inferences of discrimination drawn from the product of the discretionary choices of multiple decision-makers. *McCleskey v. Kemp*, 481 U.S. 279, 279–80 (1987) (criticizing petitioner’s use of data from multiple decision-makers in the capital charging-and-sentencing process to draw inferences of racial discrimination); Sorensen & Wallace, *supra* note 103 (finding evidence of intentional racial discrimination in death charging by a prosecutor).

133. *See, e.g.*, Raymond Paternoster et al., *Justice by Geography and Race: The Administration of the Death Penalty in Maryland, 1978-1999*, 4 U. MD. L.J. RACE RELIGION GENDER & CLASS 1 (2004) (discussing variation in death charging-and-sentencing across counties for similarly situated defendants); Ashley Rupp, *Death Penalty Prosecutorial Charging Decisions and County Budgetary Restrictions: Is the Death Penalty Arbitrarily Applied Based on County Funding?*, 71 FORDHAM L. REV. 2735 (2003) (same).

Kentucky became the first state to enact a “Racial Justice Act.” The law, initially, permitted defendants to present statistical or other evidence suggesting that their race, the race of their victim(s), or both, played a significant part in prosecutor’s decision to seek the death sentence in their particular case. If a defendant prevailed on her claim, the prosecutor was required withdraw the capital charge. Kentucky Racial Justice Act, KY. REV. STAT ANN. § 532.300 (West 1998); *see also* David C. Baldus et al., *False Attacks on the Racial Justice Acts*, DES MOINES REG., Jun. 7, 1994, at A9 (describing unsuccessful efforts to enact Racial Justice Acts at the federal and state levels); Michael Mears, *Georgia Needs a Racial Justice Act*, DAILY REPORT, Sept. 25, 1998 (same).

assessing disproportionate (i.e., “excessive”) punishments at the case-level.¹³⁴ Whereas traditional studies focus on the average effect of case-level factors on outcomes, and treat variation in the impact of these factors across jurisdictions as a “nuisance” that must be taken into account but not closely examined, this Article highlights the importance of focusing on this type of variability in the capital charging process.¹³⁵ The utility of the model for analyzing actual capital charging decisions is demonstrated in the next section.

B. *Applying the Model to Georgia*

1. Georgia Capital Charging Data

I use the aforementioned model to diagnose Georgia’s capital charging decisions. Specifically, I use the model to analyze eight years of death penalty charging from Georgia (1993–2000) in an effort to assess the potential arbitrariness in death penalty charging decisions. This timeframe was selected for three important reasons. First, Georgia enacted its life without the possibility of parole (LWOP) statute in 1993.¹³⁶ The legislation was specifically designed as a sentencing alternative in capital murder trials, therefore potentially having a substantial impact on prosecutorial, judicial, and jury discretion. Second, Georgia created an agency tasked with collecting data on capital charging-and-sentencing decisions throughout the state in October 1992, so comprehensive data is only available after that point.¹³⁷ Finally, the year 2000 was selected as the cut-off in order to allow sufficient time for all of the cases to advance from the charging phase through the (initial) penalty phase.

134. See *supra* Part I; see also *Solem v. Helm*, 463 U.S. 277, 284–90 (1983) (describing the framework reviewing courts must adopt in assessing the constitutionality of a punishment under the Eighth Amendment); *Enmund v. Florida*, 458 U.S. 782, 798–801 (1982) (same).

135. Raymond J. Carroll, *Variances Are not Always Nuisance Parameters*, 59 *BIOMETRICS* 211, 211 (2003) (emphasizing the importance of modeling the variance structure in statistical analysis); Donald Hedeker et al., *Modeling Between- and Within-Subject Variance in Ecological Momentary Assessment (EMA) Data Using Mixed-Effects Location Scale Models*, 31 *STAT. MED.* 3328, 3328 (2012) (arguing the importance of investigating why subjects differ in variability rather than just their average level).

136. 1993 Ga. Laws 569, § 4; GA. CODE ANN. § 17-10-30.1 (West 1993) (repealed 2009).

137. MICHAEL MEARS, *THE DEATH PENALTY IN GEORGIA: A MODERN HISTORY, 1970-2000* (1999).

I collected and compiled these data from five different sources: the Georgia Bureau of Investigation (GBI),¹³⁸ the Georgia Department of Corrections (GDC),¹³⁹ the Office of the Georgia Capital Defender (GCD),¹⁴⁰ the Clerk's Office of the Georgia Supreme Court (CO),¹⁴¹ and the Atlanta Journal-Constitution (AJC).¹⁴² These various sources were consulted in order to obtain as much relevant case-level data on death-eligible defendants. From these sources, I coded facts about the defendant, codefendant(s), victim(s), judge, prosecutor, defense counsel, and the crime. The data consist of the entire population of homicide cases, and not a mere sample, so statistical inference based on sample statistics (e.g. *p*-values, significance tests, confidence intervals, etc.) does not apply in the conventional sense; instead, attention is given to the direction and magnitude of the statistical parameters and quantities of interest derived from these parameters.¹⁴³

138. The GBI collects data on all homicides known to the police in the state, disaggregated by year, month, and county of occurrence. Information in the GBI data include, among other things, the age, race/ethnicity, and gender of the victim(s) and alleged offender(s), the circumstances under which the homicide took place (e.g., robbery, burglary, etc.), the relationship between the victim and the offender, and the weapon used in the homicide. The major shortcomings of these data are the high rate of missing information and the limited information about each homicide. GA. BUREAU OF INVESTIGATION, <https://gbi.georgia.gov/> (last visited Mar. 27, 2017).

139. The GDC compiles data on every defendant convicted of a criminal homicide, which includes both murder and manslaughter. Included in the GDC data is information on offender demographic characteristics, criminal history, employment status at time of the offense, alcohol/drug use and abuse history, and family background. The GDC does not keep information on victims. GA. DEP'T OF CORR., <http://www.dcor.state.ga.us/> (last visited Mar. 27, 2017).

140. The GCD is a trial resource center for capital defense attorneys in Georgia. The attorneys from GCD serve as lead/co-counsel or consultants on the vast majority of capital cases throughout the state. The office keeps basic information on every capital case in which it is involved, including but not limited to demographic characteristics of defendants, victims, attorneys, and judges; defendant criminal and mental health history, crime-specific information; and method of disposition. *Office of the Capital Defender*, GA. PUB. DEF. COUNCIL, <http://www.gapubdef.org/index.php/divisions/office-of-the-capital-defender> (last visited Mar. 27, 2017).

141. Under Georgia law, all notices of intent to seek the death penalty must be filed with the Clerk's Office. The clerk maintains a list of all death notices filed, recording the name of the defendant, the date the notice was filed, the county in which the notice was filed, and the name(s) of the prosecutor filing the notice. The Clerk's Office keeps case files from the local county courts for all cases resulting in a death sentence. These files are used by the Georgia Supreme Court when reviewing death sentences on direct appeal. *Clerk's Office*, SUP. CT. OF GA., <http://www.gasupreme.us/court-information/clerk-office/> (last visited Mar. 27, 2017).

142. The AJC is Georgia's flagship daily newspaper. The newspaper collected information about death-eligible cases throughout the state as part of an investigative article on Georgia's capital punishment system. These data were made publicly available. Bill Rankin et al., *High Court Botched Death Reviews*, ATLANTA J. & CONST., Sept. 26, 2007, at A1.

143. Accord Adrian E. Raftery, *Bayesian Model Selection in Social Research*, 25 SOC. METHODOLOGY 111 (1995). See generally Darrell Steffensmeier & Stephen Demuth, *Ethnicity and Judges' Sentencing Decisions: Hispanic-Black-White Comparisons*, 39 CRIMINOLOGY 145,

I created a master list of all potentially capital cases from which prosecutors could identify and select defendants for the death penalty. Georgia's death penalty statute lists eleven that qualify a crime for the death penalty,¹⁴⁴ and the list of death-eligible defendants is comprised of all persons seventeen years of age or older who were convicted of murder and had at least one of the eleven special elements present.¹⁴⁵ These data allow for the determination of which defendants were factually eligible for the death penalty and which defendants received a formal death notice.¹⁴⁶

During the period under investigation (1993–2000), there were 1,238 cases resulting in a murder conviction that were eligible for the death penalty under Georgia's capital statute.¹⁴⁷ Prosecutors filed a notice of intent to seek the death penalty in 400 cases and 54 defendants ultimately received the death penalty. Of the 395 capitally charged cases in which the method of disposition is known, 59% were ultimately resolved by plea and 41% were resolved by trial. Cases that were technically death-eligible under the Georgia statute but were not charged with a capital crime were disposed by plea 39% of the time and by trial 61% of the time.

Georgia's Administrative Office of the Courts (AOC) organizes the state's 159 counties into 49 superior court judicial circuits.¹⁴⁸ Cases from the various counties were grouped at the judicial circuit level because, in Georgia, there is one district attorney per judicial circuit. While large counties comprise a single judicial circuit, many smaller counties are grouped together to form a judicial circuit. As a result, a single prosecutor may be responsible for charging and plea bargain decisions for several counties in her judicial circuit.¹⁴⁹ Also, if a judicial circuit consists of multiple counties, trial judges rotate throughout these counties. Treating counties that share a single judicial

160 (2001) (explaining that sample-based significance tests are inappropriate when analyzing the entire population of cases).

144. See GA. CODE ANN. § 17-10-30 (West 2012); *supra* Part A.

145. In *Roper v. Simmons*, 543 U.S. 551, 575 (2005), the Court ruled that the death penalty was unconstitutional for defendants who were juveniles at the time they committed their crimes. Prior to *Roper*, Georgia permitted the death penalty for defendants ages seventeen and older. The data examined in this Article focus on the pre-*Roper* period. Slightly under 3% of offenders in the data eligible for the death penalty, but for their age.

146. See *infra* Table 1 for descriptive statistics of the variables employed in this study.

147. The data include twenty-eight instances where a defendant was charged with the death penalty but the defendant was either acquitted of the capital crime or convicted of a non-homicide offense. The current analysis is limited to cases resulting in a murder conviction as a proxy for strength of evidence.

148. *Your Guide to the Georgia Courts*, JUD. COUNCIL GA., <http://www.georgiacourts.org/content/your-guide-georgia-courts> (last visited Mar. 27, 2017). The number of judicial circuits was 46 during the time period of the study.

149. See *infra* Table 2.

circuit as if they were independent ignores the similarities they share in the administration of capital punishment resulting from shared decision-makers.¹⁵⁰ Figure 1 displays a map of Georgia, divided into counties and judicial circuits.¹⁵¹

Relevant Variables. Prior research suggests the primary legally legitimate determinants of capital charging decisions are the: (1) death eligibility factors enumerated in the capital statute, (2) criminal charges concurrent with the homicide, (3) defendant's prior criminal history, and (4) relationship between the defendant and victim(s).¹⁵² Professor John Donohue's recent examination of Connecticut's death penalty system, which formed the centerpiece of litigation over its constitutionality, included nearly an identical set of variables that are included in my analysis.¹⁵³ Methodologically rigorous examinations of discretionary choices of prosecutors, judges, and juries have included as many as 230 explanatory variables and as few as a dozen.¹⁵⁴ The explanatory power of these models do not significantly vary—whether including a few hundred variables, or just a few handfuls. Moreover, the actual parameter estimates (i.e., regression coefficients) for many of these factors are statistically indistinguishable across the study designs, although the complexity of the models varied considerably.¹⁵⁵ Due to the current study's focus on capital charging decisions, rather than capital sentencing

150. Isaac Unah & John Charles Boger, *Race, Politics, and the Process of Capital Punishment in the South* 17, Presentation at the Annual Meeting of the American Political Science Association, Boston, Mass., Sept. 30-Oct. 3, 2002. Death penalty cases are extremely rare events, so aggregating county-level data to the judicial circuit level offers the additional benefit of more cases per contextual unit. Better statistical estimates of the relationships occurring at both the case- and contextual-level can be obtained without altering the dependence structure of the cases due to their clustering. *See also* RAUDENBUSH & BRYK, *supra* note 123.

151. The map also shows that Georgia's forty-nine judicial circuits are organized into Judicial Administrative Districts. These districts were created to provide regional court administration to the superior courts of Georgia. The districts were created along Georgia Congressional District lines and each district is served by an administrative judge and district court administrator selected by the superior court judges and senior judge in each particular district. The primary function of the administrative judge is to assist chief judges in preparing, presenting, and managing local court budgets. These judges, however, do not exert any influence on the capital charging-and-sentencing process in the judicial circuits that comprise the judicial district.

152. DAVID C. BALDUS & GEORGE WOODWORTH, RACE DISCRIMINATION IN AMERICA'S CAPITAL PUNISHMENT SYSTEM SINCE *FURMAN V. GEORGIA* (1972) (1997) (report to the ABA Section of Individual Rights and Responsibilities reviewing empirical literature on capital charging-and-sentencing decisions); David C. Baldus & George Woodworth, *Race Discrimination in the Administration of the Death Penalty: An Overview of the Empirical Evidence with Special Emphasis on the Post-1990 Research*, 39 CRIM. L. BULL. 194 (2003) (describing the social scientific research literature on capital charging-and-sentencing decisions).

153. Donohue, *supra* note 117, at 646.

154. Baldus & Woodworth, *supra* note 152.

155. *Id.*

decisions, a more limited range of information is available to prosecutors at the time of the charging decisions. As a result, an analytical model predicting capital sentencing decisions would necessarily include a more comprehensive set of variables stemming from the fact that prosecutors and defense counsel have access to a much wider range of information at this stage of the adjudicatory process.¹⁵⁶ This intuition is supported by the fact that research suggests that a limited number of factors are predictive of capital charging.

An additional constraint on the complexity of statistical models of death penalty decision-making is the relatively small number of potential capital cases. This even holds true for studies that examine the entire population of death eligible defendants under a period of study. For example, recent studies of capital punishment systems in Colorado ($N=539$),¹⁵⁷ Connecticut ($N=205$),¹⁵⁸ Maryland ($N=1,041$),¹⁵⁹ North Carolina ($N=151$),¹⁶⁰ Texas ($N=504$),¹⁶¹ and Washington ($N=266$)¹⁶² contain an insufficient number of cases to reasonably permit the simultaneous examination of the hundreds of theoretically relevant factors impacting capital charging-and-sentencing decisions.¹⁶³ As a result, analysts have developed streamlined models that include a few dozen explanatory variables, at most. I employ a model specification that includes forty case-level variables indexing the heinousness of the crime and the culpability of the defendant (see Table 1). These

156. See, e.g., RAYMOND PATERNOSTER & ROBERT BRAME, AN EMPIRICAL ANALYSIS OF MARYLAND'S DEATH SENTENCING SYSTEM WITH RESPECT TO THE INFLUENCE OF RACE AND LEGAL JURISDICTION (2003), http://www.aclu-md.org/uploaded_files/0000/0377/md_death_penalty_race_study.pdf (using different subsets of variables to model the death noticing and death sentencing decisions).

157. Justin Marceau et al., *Death Eligibility in Colorado: Many Are Called, Few Are Chosen*, 84 U. COLO. L. REV. 1069, 1106–07 (2013).

158. Donohue, *supra* note 117, at 646.

159. Paternoster & Brame, *supra* note 126, at 984.

160. Unah, *supra* note 126, at 164.

161. Scott Phillips, *Continued Racial Disparities in the Capital of Capital Punishment: The Rosenthal Era*, 50 HOUS. L. REV. 131, 133 (2012).

162. Katherine Beckett & Heather Evans, *Race, Death, and Justice: Capital Sentencing in Washington State, 1981-2014*, 6 COLUM. J. RACE & L. 77, 99 (2016).

163. A study of South Carolina death charging decisions identified 2,227 non-negligent homicide cases occurring between 1993 and 1997, and discovered 130 cases where a death penalty notice was actually filed. Songer & Unah, *supra* note 126, at 185–86. Although researchers investigated a significantly large population of homicides, they did not differentiate murder cases from non-negligent (i.e., voluntary) manslaughter cases. While it may be the case that non-negligent manslaughter cases contain special circumstances enumerated in the South Carolina death penalty statute, these cases cannot be deemed “death eligible” because only homicides committed with either expressed or implied malice are potentially subject to the death penalty.

variables are organized into three categories: *crime-related factors* (statutorily defined aggravating factors, type of murder weapon, motive for killing, strength of evidence,¹⁶⁴ and jurisdiction where killing occurred), *defendant-related factors* (number of defendants, defendant's race/ethnicity, sex, age, level of education, employment status, marital status, number of children, military service, history of drug use, psychiatric status, IQ Score, troubled family history, prior felony conviction, county of residence, and trigger-person status), and *victim-related factors* (number of victims, victims' race/ethnicity, sex, age, and prior relationship with defendant). Important inculpatory variables include the total number of statutorily defined aggravating circumstances present in the case, defendant's contemporary convictions and prior criminal history, money or sex-related motive, the number of victims, the relationship between the defendant and the victim(s), and the age of victim. Potentially mitigating factors include the defendant's age, marital status, educational background, employment history, troubled family history, military service, history of drug and alcohol use/abuse, psychiatric status, IQ, and religious affiliation.¹⁶⁵ The race/ethnicity of the defendant and victim(s), which are included in the model, are clearly legally impermissible factors.¹⁶⁶ The Georgia dataset includes more information than the forty variables included in the final model specifications, but I limited the variables under consideration to those case-level factors that were most legally relevant and/or predictive of death charging.

As explained, *supra*, the model includes forty key explanatory variables; however this actually understates the comprehensive of the variables examined in the study, when compared to prior research, because I employ a conservative counting method in order to reduce the number of model parameters that must be estimated from the data.¹⁶⁷ For example, in terms of inculpatory/aggravation evidence, I have information on the presence or absence of the eleven statutorily defined special circumstances enumerated in Georgia's capital statute, but rather than count them separately, I combined

164. Consistent with prior research, we limit our analysis to cases that ultimately resulted in a conviction for murder as a proxy for the strength of evidence in the case. BALDUS ET AL., *supra* note 53, at 40–42, 477.

165. Admittedly, many of these factors can cut in either direction, so they may be deemed as aggravating depending on the situation, but the direction of the effect in any individual case is immaterial for the purposes of the analysis. The overall effect is estimated from the data and, therefore, reflects the manner in which prosecutors, on average, treat these factors.

166. The fact that a homicide involved a defendant and victim(s) from different racial/ethnic groups does not imply that the homicide was racially motivated. Twenty-eight percent of the death eligible murders in the data were “interracial.” An examination of a random sample of interracial homicide cases did not suggest that there was a pattern of racially-motivated killings.

167. See *infra* note 168–70 and accompanying text.

them into a single variable that indexes the total number of statutory aggravating circumstances present in the case. Similarly, with respect to exculpatory/mitigation evidence, for example, I have information on the presence or absence of five types of “troubled family background” factors. I combined these factors into a single variable, capturing the total number of problematic family features occurring in a defendant’s background. This is an extremely important point to highlight because studies purporting to include 100-plus or 200-plus variables in their analysis are employing a different counting method,¹⁶⁸ but even these studies report final results from the stable models that only include a handful or a few dozen variables.¹⁶⁹ The fatal flaw in including so many variables is the well-known “curse of dimensionality.”¹⁷⁰

The model also includes the jurisdiction of every capital case in the dataset. Evidence of regional disparities in death penalty charging and sentencing has been characterized as proof of arbitrariness, bias, or both,¹⁷¹ but that debate need not be resolved here. It suffices that, for the purposes of these analyses, any sizable association between location and the likelihood a defendant receives a death notice that cannot be accounted for by case-level factors undermines the reliability and rationality of the system, so at a

168. By way of comparison, Weiss and colleagues disaggregate a defendants’ and victims’ ages into, respectively, four and five variables. This results in nine variables capturing age, whereas I use a total of two variables. Robert E. Weiss et al., *Death Penalty Charging in Los Angeles County: An Illustrative Data Analysis Using Skeptical Priors*, 28 SOC. METHODS & RES. 91, 94–95 (1999). Similarly, Weiss et al. use fifteen variables to capture the various defendant-victim racial combinations; whereas my model includes three variables. *Id.*

169. See, e.g., BALDUS ET AL., *supra* note 53, at 42–46 (examining models as large as 230 variables, but settling on a final model with 43 variables); PATERNOSTER & BRAME, *supra* note 156, at 26–28 (finding little difference in the predictive ability of models with as many as 176 variables and a few as a dozen); ISSAC UNAH & JOHN CHARLES BOGER, RACE AND THE DEATH PENALTY IN NORTH CAROLINA: AN EMPIRICAL ANALYSIS, 1993–1997 (2001), <http://www.unc.edu/~jcboger/NCDeathPenaltyReport2001.pdf> (examining models with upwards of 100 variables, but finding similar predictive ability with much a simpler model that included 36 variables); Weiss et al., *supra* note 168, at 105, 109–10 (constructing a database of 600 variables, but including less than 20 variables in any single model specification).

170. The curse of dimensionality refers to the problem arising from fitting complex models to data. Regression analysis becomes more difficult as the number of estimated relationships (i.e., parameters) increases relative to the number of data points. When models contain a large number of variables, it may not be possible to estimate the parameters in the model because there is not sufficient variation on all of those variables across observations. RICHARD E. BELLMAN, ADAPTIVE CONTROL PROCESSES: A GUIDED TOUR 94 (1961).

171. See also *Glossip v. Gross*, 135 S. Ct. 2726, 2761–62 (2015) (Breyer, J., dissenting) (noting that geography, like race and gender, impermissibly affect the application of the death penalty); cf. Paternoster et al., *supra* note 133, at 28–30 (characterizing geographic differences in the administration of the death penalty as bias); Smith, *supra* note 110, at 252–57 (describing inter-county variation in death penalty decision-making as being suggestive of arbitrariness).

minimum, it is indicative of arbitrariness.¹⁷² Professor Richard Berk has remarked: “[I]f the concern is about illegitimate factors affecting capital cases, the impact of location needs to be studied in much greater depth. As now measured, a county or city is just a proxy for processes that are not analyzed.”¹⁷³

2. Arbitrariness as Unreliability

An important advantage of MLMs over prior approaches to analyzing capital charging data is the ability of the model to assess the level of consistency in prosecutorial charging decisions.¹⁷⁴ MLMs permit the calculation of the *Intra-class Correlation Coefficient Type-1*, ICC_1 , which measures the degree of stability (i.e., reliability) of death charging decisions within judicial circuits.¹⁷⁵ In other words, the ICC_1 “can also be thought of as the correlation among units within the same group”¹⁷⁶ or “as an estimate of the extent to which raters are interchangeable—that is, the extent to which one rater from a group may represent all the raters within the group.”¹⁷⁷ $ICC_1 = \psi / (\psi + \theta)$, where ψ is the between-circuit variance in capital charging and θ is the within-circuit variance, and the statistic ranges from 0 to 1 (or 0% to 100%).¹⁷⁸ When explanatory variables are included in the model, the within and between-circuit variances are “residual variances” because of the effects of those variables are removed from the variances, permitted an assessment of the degree of variability that is unaccounted for by the explanatory variables. The ICC_1 can also be interpreted as the proportion of the total variation attributable to variation between clusters.¹⁷⁹ When the ICC_1 is large, a single case from a circuit is likely to provide a reliable estimate of the other cases (i.e., the group average), and therefore one can infer a strong

172. See *supra* Part B. Maintaining the analytical clarity between arbitrariness and bias assists in systematically evaluating capital charging systems, but does not negate the relationship between the two concepts.

173. Berk et al., *supra* note 96, at 387.

174. Katherine J. Klein & Steve W.J. Kozlowski, *From Micro to Meso: Critical Steps in Conceptualizing and Conducting Multilevel Research*, 3 ORGANIZATIONAL RES. METHODS 211, 224–25 (2000).

175. Klaus Larsen & Juan Merlo, *Appropriate Assessment of Neighborhood Effects on Individual Health: Integrating Random and Fixed Effects in Multilevel Logistic Regression*, 161 AM. J. EPIDEMIOLOGY 81, 82–83 (2005).

176. GELMAN & HILL, *supra* note 123, at 448.

177. Klein & Kozlowski, *supra* note 174, at 224.

178. The ICC is sometimes denoted as *lambda* (λ) or *rho* (ρ) because it is both a correlation coefficient and a reliability measure. See also *infra* Methodological Appendix.

179. GELMAN & HILL, *supra* note 123, at 448.

dependency (i.e., consistency) across cases in the circuit.¹⁸⁰ On the other hand, when the ICC_1 is small, multiple cases are necessary to provide a reliable estimate of the group average, and therefore charging decisions for individual cases can be viewed as inconsistent (or independent).¹⁸¹ Scholars disagree about the threshold the ICC_1 statistic must meet in order for a group of ratings to be deemed consistent. Arguably, thresholds for reliability should be problem-specific and researchers should avoid adopting any universal cut-off criterion.¹⁸² A general rule of thumb, however, is that an ICC_1 value of .7 (or 70%) is indicative of a reliable system.¹⁸³

The unadjusted model yields an ICC_1 of .14 (or 14%), indicating considerable inconsistency in capital charging behavior across prosecutors. It must be emphasized, however, that this statistic does not take into account potentially important factual differences in cases, so it is not extremely helpful in understanding capital charging dynamics, in and of itself. We expect factually dissimilar cases to be treated differently by prosecutors both within and between circuits. Nevertheless, the unadjusted ICC_1 provides a baseline to which comparisons can be made when legally legitimate case characteristics are added to the model (i.e., the adjusted model). The adjusted model, including 40 case-level predictors, produced an ICC_1 of .19 (or 19%). In other words, 81% of the variability in death noticing for factually similar cases is attributable to within-circuit dynamics.¹⁸⁴ The increase in ICC_1 from 14% to 19% suggests slightly better within-circuit consistency in the handling of cases once legitimate case characteristics are taken into account. The within and between-circuit variance components proportions will always sum to 1 (or 100%), so the increase in ICC_1 based on the adjusted model also means that factually similar cases are treated less consistently across circuits than factually dissimilar cases.¹⁸⁵ Nonetheless, the overall reliability in charging behavior based on cases that are factually similar is extremely low.

180. *Id.* at 258.

181. *Id.*

182. Charles E. Lance et al., *The Sources of Four Commonly Reported Cutoff Criteria: What Did They Really Say?*, 9 ORGANIZATIONAL RES. METHODS 202, 205–07 (2006) (discussing the purported origins of the cut-off criterion for reliability measures).

183. The greater between-group variance relative to within-group variance, the larger the value of the ICC_1 statistic. *Id.*; Klein & Kozlowski, *supra* note 174, at 225.

184. Klein & Kozlowski, *supra* note 174, at 224 (describing the ICC_1 statistic as a measure of the influence of unit membership).

185. Although most of the unexplained variability in charging can be attributed to inconsistencies in the manner in which different prosecutors handle similarly situated defendants, (i.e., within-circuit variability), the between-circuit dynamics are still meaningful: a rule of thumb is that an ICC_1 over 5% is substantively meaningful and that inter-circuit based inconsistency cannot be ignored. TOM A.B. SNIJDERS & ROEL J. BOSKER, MULTILEVEL ANALYSIS: AN INTRODUCTION TO BASIC AND ADVANCED MULTILEVEL MODELING 38 (1999).

MLMs also permit the calculation of the relative consistency of charging decisions across defendant and victim racial/ethnic groups.¹⁸⁶ Specifically, I evaluate whether inconsistency in prosecutorial discretion appears to be a function of race/ethnicity.¹⁸⁷ Recall that the ICC₁ for the adjusted model of the entire sample is 19%, suggesting a very low, but non-trivial amount of within-circuit consistency. The ICC₁ for Caucasian-defendant and non-Caucasian-defendant cases is 21% and 14%, respectively. For Caucasian and non-Caucasian victim cases, the ICC₁ values are, respectively, 20% and 14%. Although the differences between the subgroups in terms of consistency are not especially stark, these results provided evidence that the relative reliability in charging behavior is related to legally impermissible factors, and therefore indicative of an unacceptable risk of arbitrariness.

In addition to determining the level of consistency in charging behavior across cases within a circuit, I evaluate the reliability of the estimated circuit-level effects in Georgia. I use a variant of the ICC₁, called the *Intra-Class Correlation Coefficient Type-2* (ICC₂), to determine whether circuits can be meaningfully differentiated in terms of their death noticing behavior—that is, are the observed differences between circuits sufficiently pronounced to suggest that the inter-circuit variability is an important feature of capital charging dynamics.¹⁸⁸ $ICC_2 = \psi / (\psi + \theta / \bar{n}_j)$, where ψ is the between-circuit variance, θ is the within-circuit variance, and \bar{n}_j is the average number of death eligible cases per circuit.¹⁸⁹ The intuition behind the measure is that circuit-specific effects based on circuits that, on average, contain many cases are more stable and useful measures of circuit-level properties than circuit-specific effects based on circuits that have, on average, only a few cases.¹⁹⁰ Similar to the ICC₁, the range for ICC₂ is from 0 to 1 (or 0% to 100%). ICC₂ will always be larger than ICC₁ for the same model, with higher ICC₂ scores indicating significant between-circuit variability.¹⁹¹ The ICC₂ for the

With data drawn from a sample, one could examine whether the variance was statistically distinguishable from zero. Such an examination is unnecessary because the data analyzed in this study include the entire population of death eligible cases in Georgia during the time period under investigation, and not a mere sample. *See supra* note 142 and accompanying text. If our data were drawn from some “hyper-population,” the results still suggest that the variation in capital charging is statistically significant.

186. *See supra* Part 2.

187. Hedeker et al., *supra* note 135.

188. Klein & Kozlowski, *supra* note 162, at 225; RAUDENBUSH & BRYK, *supra* note 123, at 111.

189. RAUDENBUSH & BRYK, *supra* note 123, at 111.

190. Klein & Kozlowski, *supra* note 162, at 225.

191. Large group sizes generally result in more stable mean scores, therefore it is possible to have high ICC₂ values and low ICC₁ values. RAUDENBUSH & BRYK, *supra* note 123, at 72.

unadjusted and adjusted models are, respectively, 75% and 80%—both well above the .7 (or 70%) threshold.¹⁹² These statistics provide rather strong evidence that the circuit-level charging practices are highly variable (i.e., inconsistent) and the inclusion of legally relevant case-level predictors have a negligible impact on increasing consistency across circuits.

This circuit-level reliability can also be depicted graphically to provide a more intuitive presentation of the results. Circuit-specific effects were calculated for the unadjusted and adjusted models, and these circuit-specific effects were plotted.¹⁹³ The graphical depiction provides two key pieces of information: (1) the magnitude of each circuit-specific effect and (2) the overall variability of the circuit-specific effects throughout the Georgia. Viewed collectively, the graph permits a visual assessment of the “institutional performance” of the individual circuits—that is, which circuits appeared to be “rogue” jurisdictions with respect to their charging behavior—and the level randomness or idiosyncrasy at the jurisdictional level. Figures 2 and 3 display unadjusted variation in death noticing across jurisdictions. That is, these two figures display inter-circuit differences in the probability that a death-eligible homicide will be charged capitally. Figure 2 depicts the information spatially on a map of Georgia. The legend located to the left of the map displays a color grid (in grayscale) corresponding to the magnitude of the probability of receiving a death notice for each circuit. The black horizontal line in Figure 3 represents the statewide probability of a death-eligible defendant receiving a death notice (.33) and the circles denote the probability of a death notice for each judicial circuit.¹⁹⁴ The probability of a death-eligible defendant receiving a death notice ranges from approximately .06 (Atlanta Circuit) to .62 (Ocmulgee Circuit).

Mean Absolute Deviation. I also calculated the mean absolute deviation (MAD), which is an estimate of the spread of the circuit-specific effects and is calculated by subtracting the mean of a distribution of circuit-effects (i.e., the statewide average) from each of the absolute values of the circuit-specific effects and then taking the mean of the resulting scores: $\sum |\beta_{0j} - \gamma_{00}| / J$, where β_{0j} is the circuit-level predicted probability of a death notice, γ_{00} is the statewide predicted probability, and J is the total number of circuits.¹⁹⁵

192. See *supra* note 183 and accompanying text.

193. For a description of the calculation of circuit-specific effects, see *infra* Methodological Appendix.

194. See *infra* Figure 3. Because this is the statewide probability of a death notice, based on all of the death-eligible cases in the state during the time period, this is the probability that the “average” case receives a death notice.

195. The MAD is an estimate of the spread of ratings and is calculated by subtracting the mean of a distribution of ratings from each of the absolute values of the ratings and then taking

The MAD for the unadjusted model is .11 (or 11%). In other words, the “typical” circuit death-noticing probability differs from the state-wide probability by 11 percentage points. Both MAD and ICC₂ statistics measure interrater reliability, but do so in slightly different ways. The MAD assesses how different, on average, are the circuits from the state-wide average; whereas the ICC₂ captures whether circuits can be reliably differentiated in terms of their charging patterns.¹⁹⁶

Figures 4 and 5 display the adjusted variation in the probability of receiving a death notice across Georgia’s judicial circuits. The figures reveal that cases which are factually similar along several key dimensions, including the overall level of aggravation according to Georgia’s capital statute,¹⁹⁷ are still handled very differently across Georgia’s judicial circuits with respect their probability of being formally noticed for the death penalty. The expected probability of a case receiving a death notice ranges from .11 to .59. The MAD for the adjusted model is .08 (or 8%), so the inclusion of the 40 case-level factors decreases the average circuit deviation from by the state-wide average by about 3 percentage points.¹⁹⁸ Figure 6 overlays Figures 2 and 4 and indicates the relatively small impact individual-level case characteristics have on explaining circuit-level inconsistency in death-noticing in Georgia.

Median Odds Ratio. A third useful measure for assessing between-circuit variability is the median odds ratio (MOR). The MOR quantifies the variation between circuits by comparing two charging decisions in factually similar cases from two randomly chosen circuits. The MOR is the average ratio between the cases of higher propensity with the cases of lower propensity. $MOR = \exp[\sqrt{2 \times \psi} \times \Phi^{-1}(0.75)]$, where ψ is the between-circuit variance, $\Phi(\cdot)$ is the cumulative distribution function of the normal distribution, and $\exp[\cdot]$ is the exponential function.¹⁹⁹ The MOR encapsulates the increased risk that would occur if a particular case moved from one context to another. For the unadjusted model, the MOR describes the extent to which the outcome depends on context. When covariates are included, the MOR is a measure of the variation between outcomes across circuits that is not explained by the explanatory variables. The MOR will always be greater than or equal to 1. If the MOR is 1, then there is no variation between circuits.

the mean of the resulting scores. $\sum |\beta_{0j} - \gamma_{00}|/J$, where β_{0j} is the circuit-level predicted probability of a death notice, γ_{00} is the statewide predicted probability, and J is the total number of circuits. VOGT, *supra* note 86, at 190; *see also infra* Methodological Appendix.

196. *See supra* note 191 and accompanying text.

197. GA. CODE ANN. § 17-10-30 (West 2012).

198. The median absolute deviation—which is more resistant to extreme circuit values—for the unadjusted and adjusted models are, respectively, .09 and .07.

199. Larsen & Merlo, *supra* note 175.

The MOR and the ICC are complementary, but not equivalent, measures.²⁰⁰ The MOR for the unadjusted and adjusted models are, respectively, 2.03 and 2.32. Both of these statistics indicate that, on average, a case moving from a lower probability death charging circuit to the higher probability death charging circuit is a least twice as likely to be noticed for the death penalty. Of greater significance, however, is that factually similar cases are treated more dissimilarly across circuits. The difference is slight, but supports the earlier findings from the ICC₁ analysis—the addition of relevant explanatory variables increase between-circuit inconsistency.²⁰¹

These results provide rather clear evidence of legally unjustifiable inconsistency in charging behavior. While some variation across prosecutors and judicial circuits is to be expected—and perhaps even valued—the fact that the inconsistency in charging practices, at both the individual and circuit levels is so severe and largely unaffected by important case-level characteristics should give even the staunchest death penalty retentionists pause.

3. Arbitrariness as Invalidity

Another benefit of the MLM approach is the ability to describe how well case-level factors explain the observed variability in capital charging decisions. The *Coefficient of Determination* (R^2) describes the reduction in the proportion of residual variance based on the model of interest.²⁰² In other words, it measures the improvement in the predicative ability of the adjusted model compared to the unadjusted model.²⁰³ The R^2 statistic can be viewed as a proxy for the “rational connection between the facts found and the choice

200. Some analysts believe the MOR is more interpretable than the ICC₁ because there is a parallel between the coefficients of the explanatory variables, β , and the coefficient of the between-circuit random effect, ζ_j . See, e.g., Larsen & Merlo, *supra* note 175; Germán Rodríguez & Irma Elo, *Intra-Class Correlation in Random Effects Models for Binary Data*, 3 STATA J. 32, 43 (2003).

201. See *supra* note 185 and accompanying text.

202. To be sure, the R^2 measure has been the subject of intense debate and abuses of the statistic are commonplace; nonetheless it is still one of the most widely used model summary statistics in quantitative research. The statistic is used in the present study to assess the ability of the model, which is primarily derived from Georgia’s capital statute, to explain variability in capital charging behavior. The model does not purport to fully describe the process under investigation, rather it is used to explore whether death penalty charging decisions may be deemed sufficiently rational and predictable.

203. See *infra* METHODOLOGICAL APPENDIX.

made”—that is, it is a measure of validity of the decision-marking.²⁰⁴ As emphasized earlier,²⁰⁵ validity (i. e., R^2) and reliability (i. e., ICC₁, ICC₂, and MAD) statistics capture distinct aspects of capital charging. A reliability analysis concerns the consistency/consensus of capital charging, and therefore relates to the interchangeability of prosecutors (and circuits). A validity analysis, on the other hand, addresses the accuracy/truthfulness of capital charging. Consistency in capital charging does not necessarily imply accuracy; however, inconsistency in capital charging will unavoidably undermine its accuracy.

Recall that the adjusted model takes into account the 40 case-level variables, whereas the unadjusted model does not include any predictors. The overall R^2 statistic for the adjusted model is .40. Unlike traditional regression models, however, MLMs also allow the R^2 statistic to be disaggregated into case-level (“within”) and circuit-level (“between”) components.²⁰⁶ The R^2 statistics for case-level and judicial circuit-level variance components are 0.44 and 0.21, respectively, for the adjusted model. In other words, the covariates in the model explain, approximately, 44% of the variability within circuits and 21% of the variability across circuits.²⁰⁷ Location in a particular circuit accounts for approximately 19% of the variability in death charging, so approximately 4% (.21 x .19 = .039) of the total variability in death noticing can be explained by inter-circuit differences in case-level factors. This disaggregation of the variability in death noticing decisions into within- and between-circuit components reveals that case-level explanatory variables do a slightly better job of explaining within-circuit variability than a model that ignored the hierarchical structure of the data would suggest (44% versus 41%).²⁰⁸

In addition to the overall explanatory power of the model based on legally relevant case-level characteristics, another indicator of the rationality of capital charging is the lack of association between legally illegitimate

204. See *Motor Vehicle Mfrs. Ass’n v. State Farm Mut. Auto. Ins.*, 463 U.S. 29, 43 (1983) (quoting *Burlington Truck Lines v. United States*, 371 U.S. 156, 168 (1962)) (defining arbitrary and capricious action in the context of administrative law); accord *Burlington Truck Lines*, 371 U.S. at 168.

205. See *supra* Part I.B.

206. RAUDENBUSH & BRYK, *supra* note 123, at 79–85.

207. Location in a particular jurisdiction accounts for approximately 19% of the variability in death charging, see *infra*, so approximately 4% (.21 x .19 = .039) of the total variability can be explained by inter-jurisdictional differences in case-level factors.

208. RAUDENBUSH & BRYK, *supra* note 123, at 109–10 (“The estimates of the proportion of variance explained from a hierarchical analysis may be quite different from those generated in conventional level-1 or level-2 analyses and may lead to different conclusions.”).

considerations and the predictability of charging decisions.²⁰⁹ Specifically, the explanatory power of the case-level characteristics should not vary according to defendants' and victims' race/ethnicity. I examined the R^2 statistics for four sub-models: Caucasian defendant, non-Caucasian defendant, Caucasian victim, non-Caucasian victim.²¹⁰ The R^2 for the entire sample was 43%; whereas the R^2 was 59% for Caucasian defendant cases, 42% for non-Caucasian defendant cases, 51% for the Caucasian-victim cases, 37% for non-Caucasian victim cases.²¹¹

I examined a "trimmed" model that only includes the number statutorily defined aggravating circumstances, defendant's criminal history, number of victims and defendants, and the relationship between the defendant and victim. This model captures the "legal core" of the case that should, theoretically, drive capital charging.²¹² The trimmed model explains 29% of the overall variance, 31% of the variance within jurisdictions and 12% of the variance between jurisdictions. Stated differently, legally legitimate case factors that purportedly guide discretion in the capital punishment process account for less than one-third of the variation in prosecutors' charging decisions.

The R^2 statistic may not be especially intuitive when analyzing models with a dichotomous variable, as is the case in these analyses. By construction, the total variance, and how much of that variance is explained by the model, is based on an underlying continuous latent variable,²¹³ so the R^2 statistic relates to a transformation of the dependent variable rather than the actual variable.²¹⁴ A more interpretable measure for assessing the validity of capital charging decisions is Tjur's "D" Statistic.²¹⁵ Tjur's D , also called the "coefficient of discrimination," compares the predicted probability of observing an outcome when the outcome is actually observed to the predicted

209. See *supra* Part II.B.3.

210. In order to compare R^2 across sub-groups, the statistic had to be re-scaled in order to take into account the different total variance of the intercept only model. See JOOP J. HOX, MULTILEVEL ANALYSIS: TECHNIQUES AND APPLICATIONS (2d ed. 2010).

211. There is, admittedly, significant overlap between Caucasian-defendant cases and Caucasian-victim cases because the vast majority of those cases are intra-racial, so the higher R^2 statistics in those sub-models are likely a function of both of those factors.

212. See DONALD J. BLACK, SOCIOLOGICAL JUSTICE 20 (1989) (describing the legal core of a case as "the rules in the face of the evidence . . . that can be meaningfully analyzed in the jurisprudential tradition.").

213. For a detailed description of the R^2 statistic, see *infra* Methodological Appendix.

214. The transformation is the natural logarithm of the odds of observing the event (i.e., formal charging of a death notice). See *infra* Methodological Appendix.

215. Tue Tjur, *Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination*, 63 AM. STATISTICIAN 366, 369–70 (2009).

probability of observing an outcome when the outcome is not observed: $D = \hat{\mathbb{P}}(y = 1 | y = 1) - \hat{\mathbb{P}}(y = 1 | y = 0)$, where \mathbb{P} is the probability operator, the first term of the right-hand side of the equation is defined as the true positive rate, and second term is the false positive rate.²¹⁶ Tjur's D has a range from 0 to 1 (or 0% to 100%). The larger the Tjur's D statistic, the more rational the decision-making process is because the predicted probability of a positive outcome will increase for cases with a positive outcome and decrease for cases with a negative outcome.²¹⁷ In other words, the rationality of the decision-making process is directly proportional to its ability to minimize both false positive and false negatives. Tjur's D for the entire sample was 39% for the full model and 30% for the trimmed model.²¹⁸ Unfortunately, Tjur's D cannot be disaggregated into within and between-circuit components like R^2 . Tjur's D was 42% for Caucasian defendant cases, 38% for non-Caucasian defendant cases, 43% for the Caucasian-victim cases, 29% for non-Caucasian victim cases.²¹⁹

These results suggest that cases involving Caucasian defendants and victims are handled more rationally, and therefore are less arbitrary when compared to the cases involving non-Caucasian defendants and victims or the entire population of cases. According to the R^2 measure, the difference in validity of capital charging decisions based on the defendant's race/ethnicity and victim's race/ethnicity was, respectively, 17 and 14 percentage points. Calculations based on Tjur's D yielded similar results: a 5 percentage difference based on the defendants' race/ethnicity and 14 percentage point difference according to victims' race/ethnicity.

216. *Id.* at 369.

217. *Id.* at 369–70.

218. The accuracy of a model can also be assessed by the improvement it provides in classifying the cases over simple chance. JACOB COHEN ET AL., *APPLIED MULTIPLE REGRESSION/CORRELATION ANALYSIS FOR THE BEHAVIORAL SCIENCES* 516 (3d ed. 2003). In the Georgia data, prosecutors sought the death penalty in 33% of death-eligible cases. If one were to predict the likelihood that a case would receive the death penalty without knowing anything else about the case, one could classify all of the cases as not being noticed for the death penalty and be correct 67% of the time. The usefulness of the model in explaining capital charging, then, could be measured assessing the improvement in classification when the explanatory variables are included in the model. The full model predicts the correct response approximately 82% of the time, so the model improves classification over pure chance by 15 percentage points. The trimmed model improves classification by 5.4 percentage points.

The major drawback of this model is that a threshold must be chosen, *ex ante*, in order to classify cases. The default cut-off value is .5, so cases with a predicted probability greater than or equal to .5 will be classified as positives. As a consequence, the predictive power of the model greatly depends on the cut-off value the analyst chooses. The R^2 and Tjur's D statistics avoid this problem.

219. Tjur, *supra* note 215.

Some scholars have challenged using the predictability of death penalty charging and sentencing decisions as a measure of the arbitrariness of the capital punishment system, noting the unpredictability may not only stem from capricious behavior, but also imperfections in the data and underlying model.²²⁰ While it is true that the criminal justice system may not lend itself to highly accurate statistical modeling, the Court's heightened reliability requirement under its capital punishment jurisprudence demands a higher standard than would be typically expected of non-capital criminal justice decision-making.²²¹ Moreover, dozens of methodologically rigorous studies of capital charging-and-sentencing decisions employing different model specifications and examining different time periods and jurisdictions have been unable to predict the discretionary choices of prosecutors and jurors with much accuracy.²²² As Professor Berk and colleagues have explained, "It is difficult to imagine that a few covariates exist that if included as predictors would lead to clear and justified distinctions between defendants who are charged with a capital crime and defendants who are not . . . if idiosyncrasies associated with the case, the defendant, or the adjudication process seem to determine a substantial part of the outcome, the adjudication process is suspect . . ." ²²³ Irrespective of the shortcomings inherent in data and statistical models, when the statutorily defined culpability factors predict capital charging decisions only slightly better than chance alone,²²⁴ then the death penalty cannot be functioning in a rational manner. Also recall that several of the Justices in *Furman* expressed concern that the arbitrary administration of the death penalty invited disparate treatment based on race/ethnicity.²²⁵ The following section discusses the relationship between race/ethnicity and capital charging decisions in the context of excessiveness/disproportionality.

220. *E.g.*, STEPHEN P. KLEIN ET AL., RACE AND THE DECISION TO SEEK THE DEATH PENALTY IN FEDERAL CASES 40 (2006).

221. *See generally* *Thompson v. Oklahoma*, 487 U.S. 815, 856 (1988) (O'Connor, J., concurring) ("Under the Eighth Amendment, the death penalty has been treated differently from all other punishments."); *Woodson v. North Carolina*, 428 U.S. 280, 305 (1976) ("Because of that qualitative difference [between the death penalty and other severe punishments], there is a corresponding difference in the need for reliability in the determination that death is the appropriate punishment in a specific case.").

222. Berk et al., *supra* note 96, at 386.

223. *Id.* at 387.

224. *See supra* note 218 and accompanying text.

225. *See supra* Part I.A. The defendants in the consolidated cases comprising *Furman*—William Furman (Georgia), Lucius Jackson (Georgia), and Elmer Branch (Texas)—were all African American. *Id.* at 252–53. Moreover, Furman's homicide appeared to be accidental and Jackson and Branch were sentenced to death for non-homicidal rape. *Id.*

4. Arbitrariness as Disproportionality

My prior analyses have revealed that race/ethnicity, of both the defendant and the victim, is related to irrationality and inconsistency in capital charging decisions;²²⁶ however those analyses did not estimate the direct impact of race/ethnicity on capital charging decisions, rather they examined whether the degree of irrationality and inconsistency in charging practices varied by race/ethnicity.

I conducted additional analysis to determine whether race/ethnicity had a direct effect on the likelihood that a defendant received a formal death notice, as well as whether this impact varied across circuits. When the race/ethnicity of the defendant or the victim has a direct impact on the probability of a defendant being formally charged with the death penalty, then the system produces racially/ethnically disproportionate outcomes. In other words, a system that either imposes a penalty or confers a benefit based upon membership (as a defendant or victim) in a particular racial/ethnic group, then by definition, that system is producing (potential) punishments that are not “graduated and proportioned to the offense.”²²⁷ And, as emphasized earlier, it is unnecessary to posit that racial/ethnic disparities are motivated by intentional bias because compelling evidence of the effects of unconscious/implicit bias in capital decision-making may still support a constitutional challenge under the Eighth Amendment.²²⁸ In other words, a racial disparities claim under the Eighth Amendment’s Cruel and Unusual Punishment Clause is doctrinally and analytically distinct from a racial disparities claim under the Fourteenth Amendment’s Equal Protection Clause.²²⁹ My focus, here, is on the Eighth Amendment framework and the type of evidence that is illustrative of a violation of that specific constitutional provision.

The results indicate that both the race/ethnicity of the defendant and victim influence the probability that a defendant is charged with the death penalty.

226. Recall that arbitrary and capricious government action has been defined as the reliance on factors that legislative bodies have not intended the government actor to consider, or the failure of the government actor to consider an “important aspect of the problem.” *Motor Vehicle Mfrs. Ass’n v. State Farm Mut. Auto. Ins.*, 463 U.S. 29, 43 (1983).

227. *See supra* Part I.A.

228. *See supra* Part II.A.

229. *See supra* Part I. This doctrinal and analytical distinction is underscored by the proportionality review provision in Georgia’s death penalty statute that was viewed favorably by the Court in *Gregg*. The statute required the reviewing court to assess “whether the sentence [] was imposed under the influence of “passion, prejudice, any other arbitrary factor” in one subsection, and “whether the sentence [] is excessive or disproportionate to the penalty imposed in similar cases, considering both the crime and the defendant” in a separate sub-section. *Gregg*, *Gregg*, 428 U.S. at 153–54.

At a minimum, this suggests that there is a high risk of arbitrariness irrespective of discriminatory intent. The probability of a death notice being filed against a defendant increased by 7 percentage points if the defendant was Caucasian and 15 percentage points if the victim was Caucasian, all else equal. The race-of-defendant effect should be interpreted with caution, however, because the overwhelming majority of homicides involving Caucasian defendants are intra-racial (92%); whereas homicides committed by non-Caucasian defendants are less racially-homogeneous (71%). Forty-five percent of death-eligible cases involved Caucasian victims, but 65% of cases actually noticed for the death penalty involved Caucasian victims. Similarly, Caucasian defendants comprised 25% of death eligible cases, but 37% of cases receiving a death notice. These results are consistent with prior research—cases involving Caucasian victims are much more likely to be charged with the death penalty, all else equal.²³⁰

It is not clear whether these results demonstrate intentional racial/ethnic bias because, by definition, the statistical models represent simplifications of the underlying data-generating process and there may be unobservable factors that account for the observed racial/ethnic differences in capital charging.²³¹ It is worth noting that nearly every methodologically rigorous examination of the death penalty has uncovered similar results: cases involving white victims are the most likely to be noticed for the death penalty and sentenced to death, all else equal.²³² And it is especially interesting (and informative) that methodologically rigorous studies are more likely to discover racial/ethnic disparities than studies that are less methodologically rigorous in the very same jurisdictions.²³³ However, as explained, it suffices that these results strongly suggest that charging behavior is arbitrary via its relationship to a legally arbitrary factor: race/ethnicity.²³⁴

An important advantage of MLMs is their ability to examine the variability in the effects of race/ethnicity—for both the defendant and the victim—across judicial circuits. Prior research on capital charging decisions that was attentive to jurisdictional-level variability was limited by models that precluded the empirical examination of heterogeneous case-level effects, so

230. See, e.g., Baldus & Woodworth, *supra* note 152, at 1273.

231. George E.P. Box, *Science and Statistics*, 71 J. AM. STAT. ASS'N 791, 792 (1976) (explaining that statistical models are, invariably, under-inclusive).

232. See, e.g., Baldus & Woodworth, *supra* note 152; Carol S. Steiker & Jordan M. Steiker, *The American Death Penalty and the (In)Visibility of Race*, 82 U. CHI. L. REV. 243 (2015) (summarizing the empirical literature on race and capital punishment).

233. David C. Baldus et al., *Racial Discrimination and the Death Penalty in the Post-Furman Era: An Empirical and Legal Overview, with Recent Findings from Philadelphia*, 83 CORNELL L. REV. 1638, 1661–62 (1998).

234. See *supra* Part I.A.

this question has not been adequately explored in the literature. As previously noted, evidence of racial/ethnic disparities as proof of arbitrariness was presented to the court in *McCleskey* and rejected as insufficient.²³⁵ What evidence, then, would be sufficiently compelling to deem capital charging unconstitutionally arbitrary because of pronounced racial/ethnic disparities? Simple answers to this question are elusive, but exploring variability in the effect of defendants' and victims' race/ethnicity on capital charging decisions across the state may be particularly illuminating because it is capable of providing a statewide baseline against which all jurisdictions can be assessed.²³⁶

Figures 7 and 8 display the variable effects of the race/ethnicity of the defendant across circuits. Figures 9 and 10 provide similar information for race-of-victim effects. For Figures 8 and 10, the vertical axis is the magnitude of the effect and the horizontal axis lists the circuit. These figures reveal that the influence of these aforementioned case-level characteristics can vary considerably across jurisdictions, indicating significant inconsistency in the manner in which prosecutors treat race/ethnicity in their charging decisions, and therefore a strong indication of racially disproportionate decision-making. In concrete terms, the race-of-defendant effect ranged from -.12 to .31 (the statewide average was .07). This means that the probability that case involving a Caucasian defendant receives a death notice, relative to a case with a non-Caucasian defendant, ranges from 12 percentage points lower or 31 percentage points higher, all else equal. Figure 7 displays the spatial distribution of this effect on a map of Georgia. Limiting the analysis to circuits that filed at least ten death notices produces a range from -.08 to .31, and .16 to .29 for circuits with at least 20 death notices filed.

The race-of-victim effect ranged from -.07 to .41 (the statewide average was .15), with a similar interpretation as the aforementioned race-of-defendant effect. Figure 9 displays the spatial distribution and Figure 10 reports the race-of-victim effect for each circuit. The range was -.07 to .41 for circuits with at least ten death noticed filed and .16 to .40 for circuits filing at least 20 death notices. Clearly, circuits with race/ethnicity effects that are markedly different from the statewide average create cause for concern about disproportionality.²³⁷

235. See *supra* note 76 and accompanying text.

236. Graham, *supra* note 112.

237. The direct race/ethnicity-effect estimates of arbitrariness (via disproportionality) reported in this section should be viewed as conservative because results obtained from circuits with only a small number of cases were weighted towards the statewide average. These "shrunk" estimates (also called empirical Bayes estimates) exhibit less overall variability than what would be obtained using completely unpooled data, but permit more reasonable inferences

5. Summary

My analyses of the Georgia data provide considerable evidence of an arbitrary and racially disproportionate capital charging process. The inability of legally legitimate case characteristics to meaningfully improve consistency, rationality, and racial/ethnic proportionality—along with the strong association of race/ethnicity with inconsistency and irrationality—is especially alarming considering that Georgia’s death penalty system has been subject to serious scrutiny for more than forty years. The information obtained about Georgia’s capital charging process is extremely valuable in its own right, but its usefulness extends beyond extracting patterns of behavior from past decisions—the model can also assist us in making assessments about future potentially capital cases. The next section describes exactly how this can be done.

C. *Model Predictions for Case-Specific Outcomes*

The parameter estimates obtained from the empirical model described in the previous section can be used to predict unobserved (yet observable) case outcomes.²³⁸ After analyzing the death penalty data within the MLM framework, one can obtain educated guesses about a defendant’s probability of receiving a death notice even though the defendant was not included the previously analyzed data.²³⁹ One of the key advantages of prediction in the framework I have adopted is that it also permits contextually dependent predictions. In other words, the analyst can make counterfactual predictions based on the location of the actual case, other locations throughout the state, and predictions based on the statewide baseline.²⁴⁰ These predictions can provide greater understanding of the relative culpability of the case at issue, vis-à-vis all other death-eligible cases previously processed in the system.

from the data. RAUDENBUSH & BRYK, *supra* note 123, at 254–69 (discussing advantages of empirical Bayes estimation for hierarchical data). The completely unpooled estimates for race-of-defendant range from -.21 to .41, whereas the shrinkage estimates range from -.12 to .31. For the race-of-victim estimates, the un-pooled estimates range is -.24 to .48 compared to the shrinkage estimate range of -.07 to .41.

238. The terms *prediction* and *forecast* are often used interchangeably in the statistics literature. Technically speaking, forecasts only involve out-of-sample predictions, whereas predictions may include either in-sample or out-of-sample predictions. *See generally* Gary King & Langche Zeng, *The Dangers of Extreme Counterfactuals*, 14 POL. ANALYSIS 131 (2006). For the purposes of this discussion, I will use the term *prediction* to refer to both in-sample and out-of-sample predictions.

239. *See, e.g.*, David Afshartous & Jan de Leeuw, *Prediction in Multilevel Models*, 30 J. EDUC. & BEHAV. STAT. 109 (2005).

240. *Id.*

By way of illustration, assume a prosecutor is considering pursuing a capital charge against a defendant.²⁴¹ A prediction of the probability the case would receive a formal death notice can be obtained by inserting the values of the variables specified in the model for that particular case.²⁴² Imagine the prosecutor presents a case representative of a “typical” death-eligible case in the state based on the data: the defendant is a African American male, 27 years old, two contemporaneous felony charges (in addition to the murder charge), one prior felony conviction, used a firearm, monetary motive for the killing, non-gang related, has at least one child, employed at the time of the killing, unmarried, without a high school diploma or equivalent, no military service, history of alcohol and drug use, 1.3 (out of 5) on the troubled family background index, an IQ of 100, no or minimal psychiatric impairment, and resided in the same county where the killing occurred.²⁴³ The sole victim was an African American male, 37 years old, and unacquainted with the defendant prior to the murder. The prosecutor also alleges two statutorily-defined aggravating circumstances. Based on the statistical model, there is a 33 percent chance that a defendant would be noticed for the death penalty.

It is important to keep in mind that this prediction is also based on the defendant being tried in the “typical” jurisdiction in the state. When we take into account differences in death-noticing behavior for similar cases across jurisdictions in the state, the expected likelihood of a death notice for this defendant may be as large as 59 percent or a low as 12 percent.²⁴⁴ Figures 4

241. A detailed discussion of the statistical theory underlying prediction in multilevel models is beyond the scope of this paper, but thorough descriptions are readily available. *See id.*

242. Prediction uncertainty can also be quantified, so both the expected value and a range in which the expected value may fall can be calculated. Gary King et al., *Making the Most of Statistical Analyses: Improving Interpretation and Presentation*, 44 AM. J. POL. SCI. 341, 348–49 (2000) (discussing sources of uncertainty in data simulation).

243. The consideration of factors such as a defendant’s marital status, number of children, employment status, etc., may be legally relevant for the purposes of the defendant’s mitigation case, so it is important that these factors are expressly modeled. *See* Stephen P. Garvey, *Aggravation and Mitigation in Capital Cases: What Do Jurors Think?*, 98 COLUM. L. REV. 1538 (1998) (discussing jurors’ assessment of mitigation evidence in capital trials); Robert J. Smith et al., *The Failure of Mitigation?*, 65 HASTINGS L.J. 1221 (2014).

244. Prediction using both hierarchical and non-hierarchical data structures depends on the proper estimation of the model parameters. In the case of multilevel models, the fixed effects (β) and variance components (ψ and θ) are assumed to be “true” when calculating expected outcomes. Admittedly, all statistical models are simplifications of much more complex dynamics and, therefore, no model is ever true. Box, *supra* note 231 (“Since all models are wrong, the scientist cannot obtain a ‘correct’ one by excessive elaboration.”). Nonetheless, the closer the proposed model is to the true model, the more accurate the parameter estimates and, by extension, the more plausible the predictions derived from those estimates. The number of cases in a group (n_j), as well as the number of groups (J), may profoundly impact the quality of the parameter

and 5 depict the jurisdictional variability in probability of receiving a death notice based on the aforementioned case characteristics. This predicted probability provides a rough sense of the death-worthiness of a case based on prosecutors' patterns of actual charging decisions throughout the state. To be clear, this information cannot indicate whether the case *should* be authorized for the death penalty; it merely identifies how we would expect the case to be handled based on prior death penalty noticing activity.

Now imagine a prosecutor presenting a similar case that differs only with respect to the number of statutory aggravating circumstances present in the case, which is now four. In this case, the predicted probability that the same defendant is noticed for the death penalty increases to 57 percent.²⁴⁵ Again, this is the expected probability of a death notice in the typical jurisdiction, and this expected value would range from 27 percent to 81 percent depending on the jurisdiction.

Keeping with this example, now suppose that two cases are factually similar, except for the race of the victim. The model reveals that a case involving a Caucasian victim has an expected probability of receiving a death notice that is 15 percentage points greater than a case with a non-Caucasian victim. Assuming that the 15 percentage point difference in the expected probability is not attributable to any other legitimate case characteristics not captured in the statistical model, this evidence may assist in the assessment of the defendant's culpability—especially in borderline cases. Returning to our example of the typical case in Georgia, which has an African American victim, the expected probability of 33% of receiving a death notice (i.e., one out of three cases). A factually identical case, with the exception of having a Caucasian victim, would have an expected probability of 48% of being noticed for death.

This admittedly simplified example assumes that the race-of-victim effect is constant throughout the state, but as explained, the data reveal this assumption is unwarranted, and the race-of-victim may be stronger or weaker depending on the jurisdiction. In some jurisdictions the race-of-victim effect may be significantly stronger,²⁴⁶ while in a handful of others, white-victim

estimates, so one must be careful when drawing inferences from the statistical model. Afshartous & de Leeuw, *supra* note 239.

245. This example assumes a linear effect of the number of statutory aggravating factors on the probability that a case receives a death notice. It is possible that the actual effect is nonlinear, in which case the predicted probability could be larger or smaller than fifty-seven percent. *See, e.g.,* SIMON N. WOOD, GENERALIZED ADDITIVE MODELS: AN INTRODUCTION WITH R 316–24 (2006) (describing multilevel models that relax the linearity assumption for the effects of predictor variables).

246. The effect of the race-of-victim coefficient ranges from -.07 to .40.

cases are less likely to be noticed for death.²⁴⁷ In jurisdictions where the race-of-victim effect is much larger, in absolute value, than the statewide average, an inference of an unacceptably high risk of arbitrariness in death charging in that particular jurisdiction may be justified.

The real world, of course, is more complicated because cases typically differ along multiple important dimensions. A crucial shortcoming of existing practices is the inability of appellate courts to develop general culpability measures to identify comparable cases irrespective of factual differences.²⁴⁸ The calculation of the expected probability of receiving a death notice becomes extremely useful, perhaps essential, when analysts cannot identify a sufficient number of similarly situated defendants.²⁴⁹ This expected probability constitutes a “charging propensity score” and cases can be organized into various groupings/tiers based on similar propensity scores.²⁵⁰ The predictions can be “normalized” with respect to race/ethnicity and geography, by either adding or subtracting the influence of those factors from the prediction.²⁵¹ This approach permits comparisons based upon the similarly predicted scores, rather than only focusing on cases that have similar facts.²⁵²

247. The race-of-victim effect resulting in a lower probability that a defendant was charged with the death penalty in Caucasian-victim cases was only present in jurisdictions that sought the death penalty on fewer than ten occasions.

248. *See infra* Part III.

249. An additional advantage of this approach is that the systematic collection of data and uniform coding of variables will greatly assist courts in identifying cases that are factually similar among relevant dimensions. Cases need not necessarily have the exact same values on these key variables, and the analyst can identify a range of permissible values that would satisfy the query. *See, e.g.*, Stefano M. Iacus et al., *Causal Inference Without Balance Checking: Coarsened Exact Matching*, POL. ANALYSIS 1 (2012) (describing a method for identifying similarly situated units in social science research).

250. Paul R. Rosenbaum & Donald B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 BIOMETRIKA 41, 41 (1983) (defining a propensity score as the conditional probability of assignment into a group given a set of observed covariates).

251. The choice to add or subtract the influence of race/ethnicity would depend on which group is chosen as the baseline.

252. To improve the accuracy of the model predictions, it is also important that prior cases that were deemed excessive on appeal and reversed must be excluded from the analysis. BALDUS ET AL., *supra* note 53, at 282 (noting flaws in proportionality review systems); Rankin et al., *supra* note 142 (discovering the Georgia Supreme Court’s routine use of overturned cases when conducting its mandatory proportionality review).

D. *Generalizability of the Findings and Exportability of the Model*

1. Generalizability

There is strong evidence of arbitrariness in capital charging decisions in Georgia during the time period investigated in this study. Admittedly, the model provides a simplification of a more complex decision-making process, but consistent with prior research, it captures the essential features of the process—i.e., relevant aggravation and mitigation evidence. The current analysis focused solely on eight years of capital charging data in Georgia, but there are two compelling reasons to believe that the findings may be generalizable to other states. First, following the Court’s approval of Georgia’s capital punishment statute in *Gregg*, many states developed statutes that were very similar to Georgia’s statute. As noted, Georgia’s statute was closely modeled after the ALI’s model death penalty statute.²⁵³ The demonstrated failure of Georgia’s statute to eliminate or sufficiently reduce arbitrariness and bias in capital charging seriously brings into the question the ability of similarly structured statutes to accomplish that goal. Second, the empirical results obtained in the current study are very similar to the findings reported in other states. For example, the marginal effect of the race-of-victim on the probability of a death notice was 15 percentage points. Recent studies of from Connecticut and Maryland report similar marginal effects—respectively, .20 and .13.²⁵⁴ This provides some evidence that similar dynamics are occurring across jurisdictions.

2. Exportability

The usefulness of the diagnostic tool developed in this Article extends beyond its applicability to Georgia. This is important because developing legally cognizable claims about arbitrariness in capital charging will require empirically grounded research that is reproducible.²⁵⁵ The common structure of capital statutes, along with the Court’s governing capital punishment jurisprudence, allows the model to be applied in all capital jurisdictions with

253. *See supra* Part I.A

254. Donohue, *supra* note 117, at 50; PATERNOSTER & BRAME, *supra* note 156, at 53.

255. Smith, *supra* note 110 (positing that the systematic collection and analysis of data showing regional variation in capital charging is essential for challenges to capital punishment systems); David Zuckerman, Building a Capital Arbitrariness Claim from the Ground up: A “How To” Primer Based on the Pennsylvania Experience, presented at Capital Punishment Training Conference (Aug. 1997) (emphasizing the need for social science research in developing arbitrariness challenges to capital punishment systems).

only minor modifications. The key doctrinally and empirically relevant factors influencing capital charging-and-sentencing decisions have been well-documented. Prosecutors, themselves, routinely identify the very same factors that researchers include in their predictive models.²⁵⁶ The model I develop in this Article is capable of incorporating additional legally relevant (as well as empirically relevant) factors necessary to properly understanding capital charging processes. All current capital statutes enumerate aggravating circumstances; some expressly identify mitigating factors while others do not.²⁵⁷ The presence or absence of statutorily defined mitigating factors was made largely irrelevant by the Court's ruling in *Lockett v. Ohio* because courts are not allowed to place many restrictions on the types of mitigation evidence defendants can present at trial.²⁵⁸

Identifying jurisdictional variability in the expected probability of receiving a death notice and the influence of case-level factors is straightforward in my model. Death charging decisions are made at the county-level (or similar sub-state unit level) and information about the locality of capital prosecutions is readily available. My model did not attempt to explain variability across the judicial circuits in Georgia using circuit-level factors (e.g., social and economic variables), but these relationships have been explored elsewhere.²⁵⁹ I did discover, however, that case-level variables explain only a very small proportion of the between-circuit variability in capital charging.

The use of the model is dependent on sufficient data from which to draw reasonable inferences about capital charging practices. While it is true that data will be more available in some states than in others, data limitations need not be an insurmountable obstacle. Preliminary results, based on models lacking sufficiently detailed case-level information, may still be informative. To the extent that these models provide tentative evidence of arbitrariness, a strong claim can be made to local and state governments to systematize the collection of relevant data, or at least impose a temporary moratorium on

256. See generally Ehrhard, *supra* note 116 (discussing factors that prosecutors list as relevant and irrelevant to their capital charging decisions); Ehrhard-Dietzel, *supra* note 116 (same); PATERNOSTER & BRAME, *supra* note 156 (same).

257. See Jeffrey L. Kirchmeier, *Casting a Wider Net: Another Decade of Legislative Expansion of the Death Penalty in the United States*, 34 PEPP. L. REV. 1 (2006) (describing capital statutes in every state).

258. See *supra* Part I.

259. Sherod Thaxton, *The Social Geometry of Death: Social Structure and Capital Punishment in Georgia, 1993–2000* (2009) (unpublished Ph.D. dissertation, Emory University) (on file with author).

capital charging until their death penalty systems can be explored in greater depth.²⁶⁰

E. Summary

The utility of social scientific methods and evidence in both understanding the administration of capital punishment and remedying persistent problems is beyond serious dispute in the scholarly community,²⁶¹ yet many courts and legislatures remain resistant to fully translating this body of knowledge into meaningful death penalty doctrines and reform efforts.²⁶² In order to gain traction with courts and legislatures, social/behavioral science models of death penalty decision-making must be packaged in such a way that emphasizes value in their adoption, relative simplicity of their implementation, and ease in their justification to the public. In the next section, I briefly describe how my diagnostic model can be used to improve accuracy and consistency in capital charging, and thereby impose greater discipline on the use of the death penalty. I do not attempt to articulate a fully developed proposal, rather my aim is to lay a foundation upon which a more elaborate model can be built. Key to laying this foundation, I argue, are meaningful mechanisms to discourage inadequate charging screening by prosecutors that legislatures are likely to seriously consider.²⁶³

260. See, e.g., Jeffrey L. Kirchmeier, *Another Place Beyond Here: The Death Penalty Moratorium Movement in the United States*, 73 U. COLO. L. REV. 1 (2002) (discussing death penalty moratoriums across the nation in response to concerns over the improper administration of capital punishment); Charles S. Lanier & James R. Acker, *Capital Punishment, the Moratorium Movement, and Empirical Questions: Looking Beyond Innocence, Race, and Bad Lawyering in Death Penalty Cases*, 10 PSYCHOL. PUB. POLICY. & L. 577 (2004).

261. See David C. Baldus, *Keynote Address: The Death Penalty Dialogue Between Law and Social Science*, 70 IND. L.J. 1033 (1995) (discussing the influence of social science on nearly every aspect of the administration of capital punishment); Michael L. Radelet & Marian J. Borg, *The Changing Nature of Death Penalty Debates*, 26 ANN. REV. SOCIOLOGY 43 (2000).

262. See generally James R. Acker, *A Different Agenda: The Supreme Court, Empirical Research Evidence, and Capital Punishment Decisions, 1986–1989*, 27 LAW & SOC'Y REV. 65 (1993) (noting that the U.S. Supreme Court justices were much more likely to discredit social scientific evidence and announce principles detached from the evidence of the actual administration of capital punishment); Baldus et al., *supra* note 133 (documenting legislators' resistance to statistical evidence of racial bias in the administration of the death penalty).

263. Liebman, *supra* note 7, at 333 (arguing that most death penalty reforms are unlikely to be adopted by courts or legislatures because they fail to provide adequate incentives).

III. POLICY IMPLICATIONS

The majority of proposed and enacted death penalty reforms offer the appearance of greater regularity, but do very little to reduce arbitrariness and bias.²⁶⁴ Professor James Liebman aptly notes that death penalty reforms must avoid becoming mere “window-dressing” for change that further disserve capital defendants by legitimating a broken system and reducing incentives for adopting subsequent meaningful reforms.²⁶⁵ Combating arbitrariness in front-end charging decisions is an indispensable step in any effort to improve the rationality and transparency of the capital punishment process.²⁶⁶ Numerous proposals aimed at either explicitly restricting the breadth of prosecutorial discretion or better illuminating and policing prosecutors’ discretionary choices already exist.²⁶⁷ Detailed descriptions and critiques of these existing proposals can be found elsewhere.²⁶⁸ The purpose of this section is to sketch a foundation for a general framework for death penalty reform that can be both effective in reducing arbitrariness, capriciousness, and excessiveness, but also sufficiently attractive to legislators to have an honest chance at being implemented. My framework draws inspiration from some of these existing proposals; nonetheless, there are multiple important points of departure that may hold the promise of offering unique improvements over these other models.

264. Carol S. Steiker & Jordan M. Steiker, *Should Abolitionists Support Legislative “Reform” of the Death Penalty?*, 63 OHIO ST. L.J. 417, 417–21 (2002) (describing the tendency of criminal justice officials to adopt “toothless” reforms).

265. Liebman, *supra* note 7, at 333–34; *see also* Charles J. Ogletree Jr., *Black Man’s Burden: Race and the Death Penalty in America*, 81 OR. L. REV. 15 (2002) (arguing that the Court’s procedural regulation of the death penalty has further entrenched an inconsistent and racially discriminatory system).

266. *See supra* Part I.

267. *See, e.g.*, Oren Bar-Gill & Omri Ben-Shahar, *The Prisoners’ (Plea Bargain) Dilemma*, 1 J. LEGAL ANALYSIS 737 (2009); Rachel E. Barkow, *Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law*, 61 STAN. L. REV. 869 (2009); Stephanos Bibas, *Regulating the Plea-Bargaining Market: From Caveat Emptor to Consumer Protection*, 99 CALIF. L. REV. 1117 (2011); Josh Bowers, *Punishing the Innocent*, 156 U. PA. L. REV. 1117 (2008); Lucian E. Dervan, *The Surprising Lessons from Plea Bargaining in the Shadow of Terror*, 27 GA. ST. U. L. REV. 239 (2011); Russell M. Gold, *Promoting Democracy in Prosecution*, 86 WASH. L. REV. 69 (2011); Roger A. Fairfax, Jr., *Delegation of the Criminal Prosecution Function to Private Actors*, 43 U.C. DAVIS L. REV. 411 (2009); Oren Gazal-Ayal, *Partial Ban on Plea Bargains*, 27 CARDOZO L. REV. 2295 (2006); Joseph L. Hoffman et al., *Plea Bargaining in the Shadow of Death*, 69 FORDHAM L. REV. 2313 (2001); Liebman, *supra* note 5; Robert L. Misner, *Recasting Prosecutorial Discretion*, 86 J. CRIM. L. & CRIMINOLOGY 717 (1996); F. Thomas Schornhorst, *Preliminary Screening of Prosecutorial Access to Death Qualified Juries: A Missing Constitutional Link*, 62 IND. L.J. 295 (1986–87); William J. Stuntz, *Plea Bargaining and Criminal Law’s Disappearing Shadow*, 117 HARV. L. REV. 2548 (2004).

268. *See, e.g.*, Liebman & Clarke, *supra* note 129; Steiker & Steiker, *supra* note 264.

My key intervention involves front-end proportionality review of death charging heavily guided by social scientific inquiry into arbitrariness and disproportionality. This intervention would likely need to be accompanied by meaningful financial, administrative, and reputational disincentives for foreseeable charging errors because useful reforms must raise the costs of making foreseeable errors in capital charging.²⁶⁹ Current capital charging-and-sentencing systems make trial error virtually costless to prosecutors—and to a lesser extent the capital defense bar²⁷⁰—because proper feedback mechanisms from appellate and post-conviction stages to trial actors are nearly non-existent.²⁷¹ By adequately discouraging poor screening decisions with respect to guilt and punishment, not only can high error correction costs be substantially reduced,²⁷² but also the greater expense associated with pre-trial and trial stages of capital cases. In the interest of space, however, I only focus on the front-end proportionality review component in this section. The necessary details for the second component are fleshed-out elsewhere.²⁷³

269. Misner, *supra* note 267, at 719 (“The current flaw in the evolving power of the prosecutor is the failure to force her to face the full costs of prosecutorial decisions.”); William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 572 (2001) (“[U]nless the trial system imposes costs on them for making mistakes, they will make too many. Broader liability rules are a way of evading the adjudication system, and therefore of making mistakes cheaper.”).

270. Due to the fact that the pool of capital charged cases is so large, the capital defense bar tends to neglect capital pre-trial, trial, and direct appellate proceedings, and focuses on procedural issues that are winnable in habeas proceedings rather than the substantive issues that are in play at the earlier stages. *See* Liebman et al., *supra* note 114, at 2076. This is particularly true because condemned inmates’ access to state-compensated attorneys shrinks, so it becomes easier for outside attorneys to intervene during the post-conviction stages. *Id.*; *see also* Murray v. Giarratano, 492 U.S. 1, 10 (1989) (no right to counsel in state-post conviction proceedings); Ross v. Moffitt, 417 U.S. 600, 611–12 (1974) (no right to counsel in certiorari proceedings after direct appeal); *cf.* Emily Garcia Uhrig, *A Case for a Constitutional Right to Counsel in Habeas Corpus*, 60 HASTINGS L.J. 541, 541 (2009) (commenting on the absence of a constitutional guarantee of counsel in post-conviction proceedings and arguing that the liberty interests of condemned inmates improperly sentenced to death require such a guarantee).

271. Liebman, *supra* note 7, at 324–25.

272. Bowers, *supra* note 267, at 1179 (suggesting that systemic miscarriages of justice are not due to the practice of plea bargaining, but result from overreaching in cases that technically meet the statutory requirements for the highest possible charge, but fall outside of systemic and communal norms); Gazal-Ayal, *supra* note 267, at 2306 (arguing that poor charge screening decisions are responsible for miscarriages of justice, and not plea bargaining itself).

273. *See* Thaxton, *supra* note 32 (describing a system of financial, administrative, and reputational disincentives for poor charge screening in potentially capital cases).

A. *Designing Front-End Proportionality Review*

The feature of post-*Furman* capital statutes approved in *Gregg* holding the most promise for eliminating, or at least significantly reducing, arbitrariness, bias, and disproportionality in the administration of capital punishment was meaningful appellate review of all death sentences imposed at trial.²⁷⁴ The *Gregg* Court focused much attention on the alleged narrowing function that the newly drafted capital statutes performed, but even at the time of the Court's ruling, the aggravating circumstances enumerated in the capital statutes encompassed a very wide range of murders. Thus from the very beginning, it was highly unlikely that those circumstances, in and of themselves, could sufficiently narrow the death-eligible pool to make the administration of capital punishment less arbitrary or discriminatory. This was particularly true at the time of *Gregg* because the death penalty was still permissible for a wide range of non-homicide offenses (e.g., rape, kidnapping, armed robbery, and arson). If the revised statutes were to have any reformatory force, then their impact would have been primarily—if not exclusively—through appellate review of death sentences.

Critics of the statutes approved by the Court in *Gregg* voiced two key objections. First, they argued that the new laws were incapable of ensuring the constitutionally permissible administration of capital punishment required under *Furman*.²⁷⁵ The statutes merely shifted the unbridled discretion of the pre-*Furman* era statutes to other parts of the process—namely, prosecutorial charging, plea-bargaining, and executive clemency.²⁷⁶ Chief Justice Burger commented during oral argument in *Gregg* that front-end (charging) and back-end (clemency) discretion were inevitable components of any capital scheme and outside of the effective control of legislatures.²⁷⁷ According to the Chief Justice, if the critics' arguments were taken to their logical conclusion, no statute would meet the standards that critics of the current schemes advocated, no matter how narrowly death eligibility was defined.²⁷⁸ But the Chief Justice appears to have missed the critics' underlying point. The revised statutes were insufficiently expansive to implement the heightened reliability and accuracy standards required under *Furman*'s "death is different" logic. Potential abuses of executive branch power needed to be monitored and, when appropriate, remedied by

274. See *supra* Part I.A.

275. BANNER, *supra* note 53, at 273.

276. *Id.*

277. *Id.*

278. BANNER, *supra* note 53, at 273 ("Since there is always an initial discretion on the part of the prosecutor, and . . . at the far end a power of clemency by an executive," he pointed out, "then no statutes can meet [your] standards.")

the courts. It was within the legislature's power to craft capital statutes that imposed greater justificatory and evidentiary burdens on prosecutors when pursuing the death penalty, thus impacting the front-end discretionary process. Regulating clemency decisions, admittedly, would be more difficult, but the relative infrequency and heightened transparency of these decisions made the practice less of a concern to these critics.

The second major objection to the revised statutes was the Court's unsupported belief that the "heightened procedural regulation approach" to capital punishment would satisfy the Eighth Amendment. Even assuming, *arguendo*, that the revised capital statutes could accomplish what they purported, neither the states defending their statutes nor any independent party provided evidence that their new regimes were non-arbitrary and unbiased. For example, Georgia's statute was closely modeled on "an untested innovation in 1962" by the American Law Institute (ALI).²⁷⁹ Neither the ALI, Georgia, nor the Court had any reliable evidence as to whether the guided-discretion statutes presented before the Court in *Gregg* and its companion cases were capable of eliminating or sufficiently reducing the rampant arbitrariness (and potential bias) in death penalty charging-and-sentencing practices that the Court deemed violated the Eighth Amendment.²⁸⁰ In the nearly forty years since *Gregg*, the Court has deemphasized *Furman*'s strong concerns about actual outcomes of death penalty cases, refused to test whether *Gregg*'s assumption that the guided-discretion statutes would result in accurately and consistently imposed death sentences,²⁸¹ and ignored social science evidence on the operation of the death penalty.²⁸² Numerous scholars have attributed the Court's reluctance to embrace the social scientific evidence of the realities of capital charging-and-

279. AM. LAW INST., REPORT OF THE COUNCIL TO THE MEMBERSHIP OF THE AMERICAN LAW INSTITUTE ON THE MATTER OF THE DEATH PENALTY 4, Annex A 1-3 (2009).

280. Statistical evidence of racial bias in the administration of the death penalty was presented to the Court in *Furman*, but this evidence did not form the basis for the Court's ruling. See *Furman v. Georgia*, 408 U.S. 238, 239-40 (1972). Only Justices Douglas and Marshall cited statistical evidence regarding racial bias in their opinions. *Id.* at 250-51 (Douglas, J., concurring); *Id.* at 348-56 (Marshall, J., concurring). See generally HAINES, *supra* note 53.

Interestingly, the Court, in *Gregg*, also criticized the petitioners' challenges to the revised capital statutes for failing to provide evidence that those schemes did not (or could not) satisfy the constitutional mandate of *Furman*. See HANEY, *supra* note 33, at 12-13.

281. Smith, *supra* note 110, at 249. It would have been possible for the Court to periodically grant review for a group of cases and determine whether the state's statute was effectively distinguishing the worst-of-the-worst cases from those cases that were technically death-eligible but, nonetheless, received a sentence less than death. *Id.*

282. *McCleskey v. Kemp* clearly highlighted the fact that the promise of *Gregg* had not been fulfilled—at least in Georgia. *McCleskey v. Kemp*, 481 U.S. 279, 301 (1987). See generally BALDUS ET AL., *supra* note 53, at 306-93.

sentencing practices to its lack of expertise in evaluating statistical evidence of arbitrariness, bias, and excessiveness.²⁸³ As Professors Carol Steiker and Jordan Steiker note, the Court's avoidance of direct engagement in the statistical evidence provided by petitioners and respondents in capital cases was routine and "many of the justices may have felt that their personal legitimacy as jurists was threatened in cases involving statistical proof."²⁸⁴ These two criticisms—the absence of constraints on front-end/back-end decision-makers in the revised capital statutes and the Court's unwillingness to acknowledge and respond to applicable social science—foreshadowed problems for the next forty-plus years.

But the appellate court's review could, in theory, correct errors of inadequate charge screening by identifying factors in the cases it reviewed that warranted a punishment less than the death penalty, irrespective of the defendant's eligibility under the governing statute. The appellate court was not limited to reviewing cases for trial error, so it was free to engage in a more comprehensive assessment of the totality of circumstances in each case. General standards by which defendants and criminals could be assessed was a necessary component to guide this evaluation. Comparative proportionality review, which entailed a systematic inquiry into similar and dissimilar cases, appeared to provide the vehicle through which these culpability assessments could be made. Arbitrariness and bias, on a systemic level, could be reduced by rigorous (dis)proportionality assessments at the individual level. Individual punishments that were appropriately calibrated based on the disciplined consideration of legitimate defendant and crime factors could increase overall consistency and rationality.

The ability of comparative proportionality review to accomplish this daunting, yet doable, task was cut short by the Court a mere eight years after *Gregg* in *Pulley v. Harris*²⁸⁵—a case that has been viewed by many analysts

283. Steiker & Steiker, *supra* note 232, at 279–82.

284. *Id.* at 282. According to Professor Scott Sundby, Justice Powell's aversion to quantitative reasoning was not evident across all cases. Scott E. Sundby, *The Loss of Constitutional Faith: McCleskey v. Kemp and the Dark Side of Procedure*, 10 OHIO ST. J. CRIM. L. 5, 13 n.42 (2012). In a series of antitrust opinions, Justice Powell employed sophisticated economic reasoning guided by mathematical models. *Id.* Although econometric (i.e., statistical) analyses differs from economic reasoning, Professor Sundby noted that "one might wonder whether Powell's comfort in employing sophisticated economic analysis reflects that a person's 'numberphobia' to some extent tracks one's ideological priors. Complex theories look helpful and clear when they lead to conclusions that are congenial with our views, but appear confusing and incomplete when leading to conclusions less favorable to our predispositions." *Id.* See generally THOMAS S. KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* 97–98 (2d ed. 1970) (explaining that disconfirming facts vary in their importance depending on the point in which the data are observed).

285. *Pulley v. Harris*, 465 U.S. 37 (1984).

as especially counterproductive to the stated goals articulated in *Furman* and *Gregg*.²⁸⁶ In *Pulley*, the constitutionality of California's death penalty statute was challenged, in part, because the California Supreme Court refused to conduct comparative review of the defendant's case with sentences imposed in similar capital cases to determine whether the defendant's death sentence was proportionate.²⁸⁷ The Ninth Circuit Court of Appeals had previously ruled that comparative proportionality review was constitutionality required,²⁸⁸ but the Court reversed and held that comparative proportionality review was not an indispensable feature of constitutional death penalty statutes.²⁸⁹ The Court did not completely negate the possibility that comparative proportionality review might be required for a particular death penalty statute, but it reasoned that California's statute was not "so lacking in other checks on arbitrariness that it would not pass constitutional muster without comparative proportionality review."²⁹⁰ The Court also acknowledged that "any capital sentencing scheme may occasionally produce aberrational outcomes [, but] [s]uch inconsistencies are a far cry from the major systemic defects identified in *Furman*."²⁹¹ While the Court could have been correct, in principle, it failed to reference concrete evidence supporting its assertion that those major pre-*Furman* defects were, in fact, relics of the past. In fact, there was a growing body of evidence to the contrary, and Justice Brennan highlighted this discrepancy, among others, in his dissenting opinion.²⁹² Also troubling to Brennan was the Court's refusal to consider

286. See, e.g., Timothy V. Kaufman-Osborn, *Capital Punishment, Proportionality Review, and Claims of Fairness (with Lessons from Washington State)*, 79 WASH. L. REV. 775, 775 (2004) (discussing the debate surrounding the constitutionality and effectiveness of comparative proportionality review and finding that "[t]he failure of comparative proportionality review furnishes yet another reason for concluding that capital punishment cannot be conducted in a way that comports with claims of fairness"); Barry Latzer, *The Failure of Comparative Proportionality Review of Capital Cases (with Lessons from New Jersey)*, 64 ALB. L. REV. 1161, 1164 (2001) ("Comparative review, deconstitutionalized by *Pulley*, should be abolished and replaced by more traditional proportionality review of capital cases, what I will call 'inherent' or 'retributive' proportionality review.").

287. *Pulley*, 465 U.S. at 39–40.

288. *Harris v. Pulley*, 692 F.2d 1189, 1192 (9th Cir. 1982).

289. *Pulley*, 465 U.S. at 45 ("Examination of our 1976 cases makes clear that they do not establish proportionality review as a constitutional requirement.").

290. *Id.* at 38.

291. *Id.* at 54.

292. *Id.* at 59–74 (Brennan, J., dissenting). Studies highlighted by Justice Brennan in his dissenting opinion, in draft or published form, included: David C. Baldus, Charles Pulaski, & George Woodworth, *Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience*, 74 J. CRIM. L. & CRIMINOLOGY 661 (1983); William J. Bowers & Glenn L. Pierce, *Arbitrariness and Discrimination Under Post-Furman Capital Statutes*, 26 CRIME & DELINQ. 563 (1980); Linda A. Foley & Richard S. Powell, *The Discretion of Prosecutors, Judges, and Juries*

whether “comparative proportionality review should be required in order to ensure that the irrational, arbitrary, and capricious imposition of the death penalty invalidated by *Furman* does not still exist.”²⁹³

The implications of the Court’s ruling in *Pulley* were far-reaching. Several states that adopted Georgia’s model of proportionality review either explicitly repealed the applicable provisions or expressed views that the process was unnecessary.²⁹⁴ Detailed assessments of comparative proportionality review practices in states retaining the procedural safeguard uncovered a common characteristic: “an apparent inability or unwillingness to monitor their capital-sentencing in a sufficiently consistent, comprehensive, and principled manner to identify excessive or discriminatory sentences when they occur.”²⁹⁵ The National Center for State Courts (NCSC) Project on Comparative Proportionality Review in Death Sentence Cases released a report explaining that the only conceivable way to achieve an effective proportionality review process was for reviewing courts to: (a) identify both life- and death-sentence cases comparable to the case being reviewed, (b) determine the proportion of cases resulting in a death sentence, and (c) make a legal judgment as to whether the relative frequency

in *Capital Cases*, 7 CRIM. JUST. REV. 16 (1982); Samuel R. Gross & Robert Mauro, *Patterns of Death: An Analysis of Racial Disparities in Capital Sentencing and Homicide Victimization*, 37 STAN. L. REV. 27 (1984); Michael L. Radelet & Glenn L. Pierce, *Race and Prosecutorial Discretion in Homicide Cases*, 19 L. & SOC’Y REV. 587 (1985); Marc Riedel, *Discrimination in the Imposition of the Death Penalty: A Comparison of the Characteristics of Offenders Sentenced Pre-Furman and Post-Furman*, 49 TEMP. L. Q. 261 (1976); Hans Zeisel, *Race Bias in the Administration of the Death Penalty: The Florida Experience*, 95 HARV. L. REV. 456 (1981).

293. *Pulley*, 465 U.S. at 74 (Brennan, J., dissenting).

294. BALDUS ET AL., *supra* note 53, at 280, 290.

295. *Id.* at 280; see also Rhonda G. Hartman, *Critiquing Pennsylvania’s Comparative Proportionality Review in Capital Cases*, 52 U. PITT. L. REV. 871, 872 (1991) (noting that Pennsylvania’s comparative proportionality review protocols are inconsistent with the principles of fairness and uniformity); Kaufman-Osborn, *supra* note 286 (identifying fatal shortcomings in Washington State’s comparative proportionality review system); Rankin et al., *supra* note 142 (describing problems with Georgia’s system of comparative proportionality review); Donald H. Wallace & Jonathan R. Sorensen, *Comparative Proportionality Review: A Nationwide Examination of Reversed Death Sentences*, 22 AM. J. CRIM. JUST. 13 (1997) (conducting a systematic review of comparative proportionality review nationwide and discovering that appellate courts routinely failed to conduct meaningful review of death sentences); Donald H. Wallace & Jonathan R. Sorensen, *Missouri Proportionality Review: An Assessment of a State Supreme Court’s Procedures in Capital Cases*, 8 NOTRE DAME J.L. ETHICS & PUB. POL’Y 281, 310–13 (1994) (explaining that the Missouri Supreme Court employs an enfeebled proportionality review process that exacerbates a system that is not operating to minimize the possibility of arbitrary and capricious sentencing). *But see* Ken Driggs, “*The Most Aggravated and Least Mitigated Murders*”: *Capital Proportionality Review in Florida*, 11 ST. THOMAS L. REV. 207, 275 (1999) (praising Florida’s proportionality review as meaningful and superior to most other jurisdictions, including neighboring Georgia).

of death sentences within a group is insufficiently large to warrant affirming the sentence.²⁹⁶ Currently, no court appears to be conducting the type of comparative proportionality review consistent with the NCSC model.²⁹⁷

The diagnostic model I developed in Part II facilitates front-end, empirically informed, comparative review of death penalty charging decisions. This front-end review can be used to compliment, and not replace, back-end review. Moreover, my diagnostic model need not be the only, or even the most important, feature of the front-end review, but it would be an indispensable component of this review process because of its social scientific rigor and replicability. The use of my proposed model would also avoid three persistent shortcomings of current back-end proportionality review practices: (1) the failure to develop general measures of culpability that enable courts to identify comparable cases irrespective of factual differences; (2) failure to make available to the parties in litigation the information on the cases considered by the court in its proportionality review; and (3) the inability (or unwillingness) of courts to identify evidence of racial/ethnic discrimination in the imposition of death sentences.²⁹⁸

First, my statistical model provides a general measure of culpability, based on the relationship between numerous case-level factors and actual capital charging outcomes. As such, it avoids the problems of comparison methods that require the identification of factually identical (or at least very similar) cases.²⁹⁹ Second, the data and methods used for the assessment of the capital charging system and future individual cases would be available to all parties involved in the litigation. This facilitates an open inquiry into the data and model used to draw inferences about general patterns of charging behavior. Third, and finally, the model allows a more nuanced understanding of racial/ethnic disparities in capital charging. By identifying a statewide “baseline race/ethnicity effect” and the variability of the effect across sub-state units, one can draw more reliable inferences about the influence of race/ethnicity in capital charging as an indicator of arbitrariness. Even if it is

296. BALDUS ET AL., *supra* note 53, at 281–91.

297. *Id.* at 282 (“Our investigation indicates that not a single state court has explicitly raised the question of whether it should adopt a precedent-seeking rather than a frequency approach to proportionality review.”).

298. *Id.* at 286.

299. I develop a general measure of culpability, based on the observed relationships between numerous case-level factors and actual capital charging outcomes. The culpability measure is comprised of a weighted scale of the explanatory variables and each case is given a culpability score based on a summation of the specific values of the explanatory variables for the case, multiplied by the empirically derived weight for that specific variable. Therefore, even when cases are not factually identical, they can be compared based on their overall empirically derived culpability score. BALDUS ET AL., *supra* note 53, at 286.

the case that the overall race/ethnicity effect (for defendants or victims) appears to be negligible or an artifact of the model misspecification of deficiencies with the data, it is much less likely that race/ethnicity effects that are two or three orders of magnitude greater than the statewide average are merely the product of potential shortcomings of the model and data.

Front-end comparative proportionality review may hold genuine promise for disciplining capital charging. It removes much of the mystery that has continued to plague the back-end process—namely, inadequate or inappropriate comparisons and lack of transparency. The process would entail some start-up costs, but the intellectual and financial resources necessary to get the project operational should be relatively minimal. In fact, there are numerous examples of commissions tasked with gathering and analyzing data on the operation of the death penalty in their respective jurisdictions. At the federal level, attorneys at the Department of Justice (DOJ) collect data on all potential federal death penalty cases, and front-end charge screening is performed by a committee with input from attorneys for both parties.³⁰⁰ Two state legislatures, Kentucky and North Carolina, even enacted legislation requiring not only the collection and analysis of capital punishment litigation data, but also providing legal causes of action for defendants raising certain claims, such as racial/ethnic discrimination in charging or sentencing, with the statistical evidence.³⁰¹

Legislators and prosecutors in other jurisdictions have been aware of the extensive empirical literature documenting arbitrariness and bias for decades, often in their own counties and states, but this evidence has had little impact, if any, on general charging patterns. The material and psychological benefits of the death penalty for elected officials and frontline prosecutors may simply override any commitment to truly rationalizing the process at a systemic level. Social scientific evidence identifying problems with the administration of capital punishment has never been self-implementing. Actually disciplining prosecutorial behavior requires more. As I mentioned above, genuine feedback mechanisms from appellate and post-conviction stages to prosecutors that force them (or their counties) to internalize the costs of their

300. U.S. DEP'T OF JUSTICE, THE FEDERAL DEATH PENALTY SYSTEM: SUPPLEMENTARY DATA, ANALYSIS AND REVISED PROTOCOLS FOR CAPITAL CASE REVIEW (2001), <https://www.justice.gov/archive/dag/pubdoc/deathpenaltystudy.htm>.

301. Kentucky Racial Justice Act, KY. REV. STAT. ANN. § 532.300 (West 1998); North Carolina Racial Justice Act of 2009, N.C. GEN. STAT. § 15A-210 (2010) *repealed by* 2013 N.C. Sess. Laws 2013-154, § 5(a). Under the North Carolina statute, for example, a capital defendant can have his or her sentence reduced to life imprisonment without parole if there is evidence proving “that race was a significant factor in decisions to seek or impose the sentence of death in the county, the prosecutorial district, the judicial division, or the State at the time the death sentence was sought or imposed.” *Id.*

mistakes are essential. While my diagnostic model is useful in identifying potential charging errors, the insights gained from it may not become fully realized unless embedded in a larger framework that forces prosecutors to confront the consequences of foreseeably poor choices in capital charging.³⁰² Again, the development and description of this larger framework is beyond the scope of this project.³⁰³

B. Summary

The purpose of this section was to provide a fairly rough sketch of how a rigorous social scientific model of capital charging can be incorporated into a more robust front-end comparative proportionality review process to incentivize more carefully and empirically informed capital charge screening. I have only described what I believe are some (not necessarily all) of the indispensable features of this front-end reform, but hopefully, this can provide a springboard for future research; nonetheless, two important caveats are in order. First, future proposals that are likely to gain traction with courts, legislatures, and the general public are unlikely to constitute radical departures from the corpus of proposals articulated by other scholars also concerned with reducing or eliminating both arbitrariness and bias in the capital charging-and-sentencing process. This conventionality is understandable, in many respects, given the nature of the doctrinal, political, and structural constraints confronting death penalty reformers, either abolitionist or retentionists. Any feasible proposal must be developed with full awareness of the (im)practicalities accompanying the administration of the government's ultimate sanction.³⁰⁴

Second, death penalty reformers should recognize that there are important, yet often underappreciated, limitations to *technocratic* thought—that is, the widely held view among lawyers and legal scholars that problems of law can

302. Liebman, *supra* note 7, at 324–25.

303. Thaxton, *supra* note 32.

304. See generally FROM LYNCH MOBS TO THE KILLING STATE: RACE AND THE DEATH PENALTY IN AMERICA (Charles J. Ogletree Jr. & Austin Sarat eds., 2006); ANDREW WELSH-HUGGINS, NO WINNERS HERE TONIGHT: RACE, POLITICS, AND GEOGRAPHY IN ONE OF THE COUNTRY'S BUSIEST DEATH PENALTY STATES (2009); Stephen B. Bright, *In Defense of Life: Enforcing the Bill of Rights on Behalf of Poor, Minority and Disadvantaged Persons Facing the Death Penalty*, 57 MO. L. REV. 849 (1992); Stephen B. Bright, *The Politics of Crime and the Death Penalty: Not "Soft on Crime," But Hard on the Bill of Rights*, 39 ST. LOUIS U. L.J. 479 (1995); Stephen B. Bright & Patrick J. Keenan, *Judges and the Politics of Death: Deciding Between the Bill of Rights and the Next Election in Capital Cases*, 75 B.U. L. REV. 759 (1995); Ogletree Jr., *supra* note 265; Steiker & Steiker, *supra* note 264.

be reduced to problems of technique.³⁰⁵ It remains an open question whether the capital punishment process is amenable to the types of “fixes” that have been successfully implemented by the federal government to substantially reduce (to varying degrees) the arbitrariness and racial/ethnic bias in previously problematic areas such as voting, housing, employment, and public education.³⁰⁶ Although broad consensus exists among death penalty analysts concerning the sources of arbitrariness and bias,³⁰⁷ there is considerable disagreement over whether any set of proposed remedies will produce the desired result.³⁰⁸ According to Carol Steiker and Jordan Steiker, “[t]he body of doctrine produced by the Court is enormously complex and its applicability to specific cases [is] difficult to discern; yet, it remains unresponsive to the central animating concerns that inspired the Court to embark on its regulatory regime in the first place. Indeed, most surprisingly, the overall effect of [forty]-odd years of doctrinal head-banging has been to substantially reproduce the pre-*Furman* world of capital sentencing.”³⁰⁹ All proposals aimed at repairing the “broken system” of capital punishment must fully acknowledge that inherent limitations may exist with respect to altering doctrines, policies, and practices in order to achieve the even-handed administration of the death penalty. Establishing “super due process” for criminal defendants on the one hand, and encouraging “tighter” monitoring and regulation of charging and plea-bargaining practices in district attorney offices on the other hand, may only provide minor fixes to longstanding problems of the improper influence of race/ethnicity, gender, social class, and

305. Donald J. Black, *The Boundaries of Legal Sociology*, 81 YALE L.J. 1086, 1090–91 (1972).

306. Howe, *supra* note 101, at 2085–94; Ogletree Jr., *supra* note 265, at 34–38.

307. These factors include overly-broad capital statutes, inadequate legal representation, decentralized decision-makers, broad prosecutorial and sentencer discretion, and overly restrictive merits review in state and federal habeas proceedings.

308. *See also* Howe, *supra* note 101, at 2124–27 (suggesting that even an increase in the due process protections, (“super due process”) for capital defendants is unlikely to significantly reduce or eliminate racial bias in the capital punishment process); Smith et al., *supra* note 243, at 1224–25 (demonstrating that the Court’s mitigation □ facilitating doctrines have largely failed to benefit capital defendants with compelling evidence for mercy). *Compare, e.g.*, Baldus et al., *supra* note 20, at 361–64 (refuting the claim that discrimination in the imposition of the death penalty is inevitable and impossible to prevent), *with, e.g.*, William J. Bowers & Wanda D. Foglia, *Still Singularly Agonizing: Law’s Failure to Purge Arbitrariness from Capital Sentencing*, 39 CRIM. L. BULL. 51, 54 (2003) (demonstrating that constitutionally mandated requirements to guide jury discretion and eliminate arbitrariness in sentencing are not working).

309. Carol S. Steiker & Jordan M. Steiker, *Sober Second Thoughts: Reflections on Two Decades of Constitutional Regulation of Capital Punishment*, 109 HARV. L. REV. 355, 359 (1995).

geography. The reality may be that it is unlikely that more “tinkering” is all that is needed to satisfy the still unfilled promise of *Furman*.³¹⁰

IV. CONCLUSION

Following a series of landmark rulings in the late-1970s and early-1980s,³¹¹ the U.S. Supreme Court appeared to craft a capital punishment jurisprudence that unambiguously mandated that frontline criminal justice officials reserve the death penalty for the worst crimes and worst criminals.³¹² The Court announced the arbitrary and discriminatory administration of capital punishment would no longer be tolerated. The guided-discretion statutes that emerged in the aftermath of *Furman* offered the promise of constraining the hyper-discretion that existed in the pre-*Furman* era, but the Court’s faith in these doctrines to effectively guard against caprice and bias in the capital charging-and-sentencing was misplaced.³¹³ Much to the dismay of the various current and former U.S. Supreme Court Justices and cautiously optimistic death reformers,³¹⁴ rigorous empirical research on the capital charging-and-sentencing process almost unequivocally reveals that states have failed to purge the process of the arbitrariness and bias that the Court believed to be particularly rampant pre-*Furman*.³¹⁵ Prosecutors’ charging

310. See *Callins v. Collins*, 510 U.S. 1141, 1145 (1994) (Blackmun, J., dissenting) (stating “[f]rom this day forward, I no longer shall tinker with the machinery of death”).

311. See *Enmund v. Florida*, 458 U.S. 782 (1982) (death sentence unconstitutional when defendants are neither killers nor had the intention to kill); *Godfrey v. Georgia*, 446 U.S. 420 (1980) (vaguely defined aggravating factors are unconstitutional); *Presnell v. Georgia*, 439 U.S. 14 (1978) (death sentence must be based on a statutorily defined factor); *Lockett v. Ohio*, 438 U.S. 586 (1978) (statutory restrictions on mitigation evidence is unconstitutional); *Eberheart v. Georgia*, 433 U.S. 917 (1977) (death penalty is unconstitutional for non-homicidal kidnapping); *Coker v. Georgia*, 433 U.S. 584 (1977) (death penalty for non-homicidal rape is unconstitutional); *Gregg v. Georgia*, 428 U.S. 153 (1976); *Furman v. Georgia*, 408 U.S. 238 (1972).

312. In the seven years immediately following the Court’s decision in *Gregg*, it ruled in favor of 14 out of 15 death-sentenced inmates whose appeals were fully heard. HAINES, *supra* note 53.

313. BALDUS ET AL., *supra* note 53, at 398–99.

314. Multiple U.S. Supreme Court Justices have made statements in recent years either expressly condemning the practice of capital punishment, or raising serious concerns as to its fair administration. Moreover, several Justices who voted to uphold the constitutionality of the death penalty while on the Court publically criticized the death penalty after retiring from the Court. See *supra* note 5 and accompanying text.

315. BALDUS & WOODWORTH, *supra* note 152; Baldus & Woodworth, *supra* note 152; Richard C. Dieter, *Twenty Years of Capital Punishment: A Re-Evaluation*, DEATH PENALTY INFO. CTR. (June 1996), <http://www.deathpenaltyinfo.org/twenty-years-of-capital-punishment>; Steiker & Steiker, *supra* note 309.

decisions remain highly arbitrary both within and across jurisdictions.³¹⁶ Effective death penalty reform, if possible, must begin with the gatekeepers of the system.

METHODOLOGICAL APPENDIX

The multilevel models (MLMs) utilized in this Article offer four key improvements over prior research investigating capital charging dynamics.³¹⁷ These models (1) correctly take into account the non-independence of cases nested in the same judicial circuit;³¹⁸ (2) allow for the partitioning of variation in the case-level charging decisions into within- and between-circuit effects;³¹⁹ (3) provide better estimates of the effects of case-level explanatory

316. In his concurring opinion in *Furman*, Justice Stewart famously wrote that “death sentences are cruel and unusual in the same way that being struck by lightning is cruel and unusual.” *Furman*, 408 U.S. at 309 (Stewart, J., concurring). More recently, Washington Supreme Court Justice Charles Johnson reiterated Justice Stewart’s concern in his dissenting opinion: “Reviewing the history of this court’s proportionality review reveals how the administration of capital cases defies any rational analysis. . . . These cases exemplify the arbitrariness with which the penalty of death is exacted. They are symptoms of a system where statutory comparability defies rational explanation. The death penalty is like lightening, randomly striking some defendants and not others. . . . No rational explanation exists to explain why some individuals escape the penalty of death and others do not.” *State v. Cross*, 132 P.3d 80, 109–10, 115 (Wash. 2006) (Johnson, J., dissenting).

From 2006 to 2013, 355 condemned inmates were executed and 261 individuals were killed by lightning in the United States—a ratio of approximately 1.4 executions for every lightning strike death. Richard C. Dieter, *Facts About the Death Penalty*, DEATH PENALTY INFO. CTR., <http://www.deathpenaltyinfo.org/documents/FactSheet.pdf> (last updated Jan. 27, 2017); *U.S. Lightning Fatalities 2006-2016*, NAT’L WEATHER SERV., <http://www.lightningsafety.noaa.gov/fatalities.shtml> (last visited Jan. 29, 2017).

317. See Sherod Thaxton, *Un-Gregg-Ulated: Capital Charging and the Missing Mandate of Gregg v. Georgia*, 11 DUKE J. CONST. L. & PUB. POL’Y 145, 166 (2016) (describing the virtues of multilevel modeling to investigate the constitutionality of death penalty dynamics).

318. Inferences drawn from analytical frameworks that do not explicitly account for the fact that death eligible cases are nested in different jurisdictions are often misleading because relationships measured at one level of analysis (e.g., between cases) do not necessarily hold at another level of analysis (e.g., between circuits). Interpreting associations at the higher level as pertaining to the lower level is known as an *ecological fallacy*. The opposite of the ecological fallacy is an *atomistic fallacy*, and this occurs when one draws inferences about the relationships between group-level variables based on information about individual-level relationships. These fallacies are problems of *inference*, not of measurement. DOUGLAS A. LUKE, MULTILEVEL MODELING 5–6 (2004).

319. Multilevel models have a complex error structure because the total variability in individual outcomes is comprised of two components: the within-cluster variance and the between-cluster variance. Decomposing the random part of the multilevel model into unit-specific and cluster-specific effects allows the analyst to determine how much variability in the outcome can be attributed to each level. ANDERS SKRONDAL & SOPHIA RABE-HESKETH, GENERALIZED

variables by combining information on both the with- and between-circuit effects of those variables,³²⁰ and (4) produce sensible (and statistically defensible) calculations of circuit-specific effects that facilitate the assessment and ranking of the institutional performance of the circuits.³²¹

Prior studies have *recognized* that between-circuit processes may be an important source of variability in capital charging and sentencing, yet they have not properly incorporated this information in their analyses.³²² The unfortunate result of this oversight has been an incomplete description and understanding of the operation of capital punishment systems. Rather than taking inter-jurisdictional variability as an important object of study, prior research has treated inter-jurisdictional variability as a nuisance that needed to be minimized or corrected in an effort to properly analyze intra-jurisdictional variability.³²³

The traditional approach adjusts for differences across jurisdictions through the use of “fixed effects”—that is, the estimation of a set of jurisdiction-specific regression coefficients intended to capture differences between jurisdictions for similar cases.³²⁴ But the fixed effects modeling framework is ill-suited for the investigation of death penalty charging data for at least four reasons.³²⁵ First, information from jurisdictions that either

LATENT VARIABLE MODELING: MULTILEVEL, LONGITUDINAL, AND STRUCTURAL EQUATION MODELS 51 (2004).

320. The case-level estimates of explanatory variables in MLMs is the weighted average of the within- and between-circuit effects. GELMAN & HILL, *supra* note 123, at 478 (explaining the use of pooling factors for weighting individual coefficients in a multilevel model).

321. Estimates of the variability of the within- and between-circuit influences on the outcome, net of the case-level explanatory factors, are used to predict the circuit-specific effects. SKRONDAL & RABE-HESKETH, *supra* note 319, at 225; Duncan et al., *supra* note 27.

322. The importance of separating variability into within- and between-circuit effects was brilliantly underscored by two statisticians who famously remarked, “One statistician’s error term is another’s career!” DONALD HEDEKER & ROBERT D. GIBBONS, LONGITUDINAL DATA ANALYSIS 56 (2006).

323. Carroll, *supra* note 135, at 211–15 (noting the strong tendency of statistical analyses to treat variability in effects as uninteresting rather than an object of study).

324. Even studies focusing on geographical differences fail to employ the MLM framework, which is specifically designed to permit a more nuanced assessment of jurisdictional variability. See, e.g., Katherine Barnes et al., *Place Matters (Most): An Empirical Study of Prosecutorial Decision-Making in Death-Eligible Cases*, 51 ARIZ. L. REV. 305 (2009); Donohue, *supra* note 117; Paternoster et al., *supra* note 133; Songer & Unah, *supra* note 126; David Weisburd & Joseph Naus, *Report to Special Master David Baime: Re Systemic Proportionality Review*, in REPORT TO THE SUPREME COURT SYSTEMIC PROPORTIONALITY REVIEW PROJECT 67 (2001). Although these studies show that significant differences remain between jurisdictions even after taking account numerous case-level explanatory variables, the analytical framework they employ cannot provide answers to questions that I specifically address in my analyses by using the MLM framework.

325. *Compare* PAUL D. ALLISON, FIXED EFFECTS REGRESSION MODELS 2 (2009) (explaining that prior studies have preferred the fixed effects approach, in part, because it can take into account

have a single case (called *singleton clusters*) or from jurisdictions where all cases received the same outcome (i.e., no death-eligible case was noticed for the death penalty or all death eligible cases were noticed for the death penalty) must be discarded.³²⁶ MLMs, on the other hand, utilize information from circuits with a single observation and from circuits containing cases that are identical with respect to their death noticing outcomes. Although the circuit-level effect will not be precisely estimated, the information from the single case contributes to the estimation of the coefficients and variance parameters of the individual and circuit-level regressions.

Second, under the fixed-effects framework, cases must exhibit substantial within-circuit variation along multiple case-level characteristics in order to reliably explore case-level dynamics. If a substantial proportion of the variation in the case-level characteristics is between-circuit and not within-within circuit, then the traditional approach will give imprecise estimates because the case-level estimates only deal with a small subsection of the variance of the case-level characteristic. As two scholars have recently noted, “[I]n controlling out context, [fixed effects] models effectively cut out much of what is going on—goings-on that are usually of interest to the researcher, the reader and the policy maker. . . . and offer overly simplistic and impoverished results that can lead to misleading interpretations.”³²⁷ By explicitly modeling circuit-level heterogeneity in capital charging, MLMs can sensibly incorporate information about within- and between-circuit variability in the effects of explanatory factors, and therefore provide reasonable answers about the general effect of a variable even when a very large proportion of variability in explanatory variable is between-circuit.³²⁸

Third, fixed effects models require the data to contain a moderate to large number of cases in each jurisdiction in order to provide an accurate measure of the jurisdiction-specific effect. This results from the fact that the

any jurisdiction-level unobserved effects on case outcomes that may be potentially correlated with case-level explanatory variables), *with* Bell & Jones, *supra* note 127, at 134 (“[W]e take the strong, and rather heterodox, view that there are few, if any, occasions in which [fixed effects] modeling is preferable to [MLM].”) (alterations added).

326. The fixed effects framework removes the between-jurisdiction variability from the model, so all that remains to be examined is within-jurisdiction variability. As a consequence, singleton clusters or jurisdictions in which all death-eligible cases either receive or do not receive a death notice are dropped from the analysis. There can be no within-jurisdiction variation with a single case or multiple cases that are identical across the variables examined in the model. A. COLIN CAMERON & PRAVIN K. TRIVEDI, *MICROECONOMETRICS: METHODS AND APPLICATIONS* 796–97 (2005).

327. Bell & Jones, *supra* note 127, at 134.

328. The variance of the parameter estimate will also be impacted by the weighing because of the uncertainty around the effect of any particular variable related to the (dis)similarity of the jurisdictions. RAUDENBUSH & BRYK, *supra* note 123, at 47.

framework treats each jurisdiction separately from the other jurisdictions in order to capture the within-jurisdiction variability. As one scholar aptly noted, the fixed effects approach is “amnesic” because every new cluster is treated like a new world and does not benefit from the information about other clusters.³²⁹ Whereas fixed effects models are hindered by their naïve treatment of circuit-specific effects that give evidence from small circuits undue influence, MLMs avoid this problem by weighing small circuits that exhibit high within-circuit variability (i.e., unreliable information) towards the overall average, thereby minimizing the influence of these small circuits on the determination of cluster-specific effects.³³⁰

Lastly, the fixed effects approach does not permit inferences about the between-jurisdiction variation, including whether the variability is substantively meaningful. This is especially true when the outcome variable is binary, such as the decision to seek the death penalty against a defendant, because the even naïve estimates of the circuit-specific effects are precluded.³³¹ All of these aforementioned problems originate from the inability of the fixed effects framework to explicitly model context and heterogeneity, and therefore simultaneously consider within- and between-jurisdiction variability. Bell and Jones have argued that the fixed effect technique is used too readily as a default option without a full understanding of what is being estimated and what is being lost by doing so.³³²

The intuition behind MLMs is the estimation of the mean and variance of the *distribution* of the circuit-specific effects—but not the actual circuit-specific effects—via the imposition of a modest constraint on the variability of the between-circuit effects: a probability model that assumes the circuit-specific effects arise from deviations from a typical circuit in the state. The constraint allows the model to utilize all of the information from available

329. RICHARD MCELREATH, STATISTICAL RETHINKING: A BAYESIAN COURSE WITH EXAMPLES IN R AND STAN 355 (2016).

330. *Id.*

331. CAMERON & TRIVEDI, *supra* note 326, at 796 (explaining that fixed effects models for binary outcomes, such as noticing decisions, cannot produce estimates of the circuit-specific effects). Some scholars have employed a linear probability model (LPM) to estimate the cluster-specific fixed-effects, called “unconditional fixed effects,” on a binary outcome. *E.g.*, David S. Abrams & Albert H. Yoon, *The Luck of the Draw: Using Random Case Assignment to Investigate Attorney Ability*, 74 U. CHI. L. REV. 1145, 1168 (2007). Not only is the LPM subject to the same shortcomings as the traditional fixed effects models for a non-binary outcome, but the unconditional effects are also inconsistent (i.e., they fail to converge to the true effect as the sample size increases. *See* Ethan Katz, *Bias in Conditional and Unconditional Fixed Effects Logit Estimation*, 9 POL. ANALYSIS 379, 380 (2001). For a discussion of the additional shortcomings for the LPM as it pertains to the analysis of capital charging data, *see infra* note 336 and accompanying text.

332. Bell & Jones, *supra* note 127, at 134.

cases to provide better estimates of case-level effects and circuit-level variability. MLMs offer a sensible compromise between within- and between-circuit effects because the effect of a case-level factor is neither purely a between-circuit effect (because the case-level factor can vary across cases within jurisdictions) nor purely within-circuit effects (because the case-level factor may be constant across cases within a particular circuit).³³³

MLMs can be written in two parts: a Level-1 model and a Level-2 model. The Level-1 model is, essentially, a series of sub-models for each Level-2 unit (e.g., judicial circuit). The unit of analysis is the death-eligible murder case and the sample size for each regression is number of death-eligible cases for each particular judicial circuit. This model captures variability in death noticing among cases within the judicial circuit. Formally, the Level-1 model can be written as: $\Pr(y_{ij} = 1) = g^{-1}(\beta_{0j} + \beta_k X_{kij} + \epsilon_{ij})$, where the subscripts i and j index the i th defendant and j th judicial circuit, respectively, y_{ij} is a binary outcome indicating “1” if the defendant is noticed for the death penalty and “0” if otherwise, $\Pr(y_{ij} = 1)$ is the probability that the defendant i th in circuit j th is noticed for the death penalty, β_{0j} (*beta*) is the circuit-level probability that a defendant is noticed for the death penalty (conditional on all explanatory variables being equal to zero), X_k are k explanatory variables with β regression coefficients, and ϵ_{ij} (*epsilon*) are Level-1 errors (case-level the deviation from the expected probability for the i th defendant in the j th judicial circuit).³³⁴ In order to meaningfully interpret β_{0j} , explanatory variables, X_k , are centered at their overall (i.e., statewide) average values, $\bar{X}_k = 0$, so the intercept is the probability of a “typical” case in Georgia receiving the death notice in the j th circuit.

The unit of analysis for the Level-2 model is the judicial circuit, not the individual death-eligible cases, and the outcome variable is the circuit-specific probability (β_{0j}). Formally, the Level-2 model is: $\beta_{0j} = \gamma_{00} + \zeta_{0j}$,

333. The multilevel estimate, β_k^{ML} , is calculated as: $\lambda_j \beta_j^B + (1 - \lambda_j) \times \beta_k^W$; where $\lambda_j = \psi_j / [\psi_j + (\theta/n_j)]$, n_j is the sample size of judicial circuit, j , ψ is the between-circuit variance, θ is the variance within circuits, β_k^B is the between-circuit effect, and β_k^W is the within-circuit effect. GELMAN & HILL, *supra* note 123, at 477–78.

334. Here $g^{-1}(\cdot)$ is the *inverse link function* (also called the logistic function), so $\Pr(y_{ij} = 1) = \frac{\exp(\beta_{0j} + \beta_k X_{kij} + \epsilon_{ij})}{1 + \exp(\beta_{0j} + \beta_k X_{kij} + \epsilon_{ij})}$. Technically speaking, there is no Level-1 residual error (ϵ_{ij}) in this mathematical expression, but ϵ_{ij} appears in another equivalent formulation: $y_{ij}^* = \beta_{0j} + \beta_k X_{kij} + \epsilon_{ij}$, where y_{ij}^* represents the propensity to notice a case for the death penalty, such that $y_{ij} = 1$ if $y_{ij}^* > 0$ and $y_{ij} = 0$ if $y_{ij}^* \leq 0$. The equivalence of the two equations can be shown: $\Pr(y_{ij} = 1) = \Pr(\beta_{0j} + \beta_k X_{kij} + \epsilon_{ij} > 0)$. Nevertheless, I include ϵ_{ij} in the prior equation in order to make the interpretation of the regression coefficients more intuitive and relationship of the variance components more apparent in the MLM framework.

where β_{0j} is the same as described above, γ_{00} is the probability of a death notice for the typical circuit (i.e., the statewide average across circuits, not cases), and ζ_{0j} (*zeta*) is a circuit-specific deviation from the statewide average.³³⁵ The total variance of the circuit-specific intercepts, $Var(\beta_{0j}) = Var(\zeta_{0j}) + Var(\epsilon_{ij}) = \psi + \theta$, where ψ (*psi*) is the between-circuit variance, $Var(\zeta_j)$, and θ (*theta*) is the within-circuit variance, $Var(\epsilon_{ij})$.³³⁶ When case-level explanatory variables, X_k , are included in the model, ψ and θ are *residual variances*—i.e., variability left unexplained after taking into account the explanatory variables. The ζ 's are not model parameters, but are quantities of interest predicted from the estimated parameters ($\hat{\beta}$, $\hat{\psi}$, and $\hat{\theta}$) which are treated as known. The ζ 's are crucial for making inferences for the circuits in the data (e.g., assessment of institutional performance) and can be used to compare the various circuits in terms of their punitiveness (or leniency) with respect to death noticing behavior because ζ 's are residual deviations (i.e., the deviations take into account the case-level characteristics included in the model). Under the assumption that the key legal features of the death noticing process have been included in the model (see Table 1) or have been proxied by other variables included in the model, these two

335. The Level-2 model cannot be estimated on its own because the random intercept, β_{0j} , is not observed. Instead, the Level-2 model must be substituted in the Level-1 model to obtain a reduced form model for the observed responses: $y_{ij}^* = \gamma_{00} + \zeta_{0j} + \beta_k X_{kij} + \epsilon_{ij}$, where $(\gamma_{00} + \zeta_{0j}) = \beta_{0j}$.

336. For the logistic regression model, θ has a fixed variance that is specified, *a priori*, by the logistic distribution: $\theta = \frac{\pi^2}{3} \approx 3.29$. The use of a linear probability model (LPM), which treats a binary outcome variable as continuous, to examine clustered data will give misleading results because θ will be incorrectly estimated from the data, and therefore all inferences based on those statistics will be unreliable.

The LPM suffers from two additional limitations that makes it ill-suited for the current project. First, the LPM assumes that the relationship between the explanatory variables and the binary outcome variable is linear, which is an unrealistic assumption for this project because the explanatory variables attempt to index a defendant's culpability level. For example, it is improbable that the impact of an increase in the number of victims in a homicide case on the probability that a defendant receives the death penalty is the same when the number increases from one to two as it would be from five to six. The logistic regression model explicitly takes this nonlinearity into account to properly estimate the relationships between explanatory variables and the probability of receiving a death notice.

Second, with respect to predicting the probability that a case is noticed for the death penalty, particularly cases not included in the estimation sample, the LPM is much more likely give probabilities that are less than "0" and greater than "1". These out-of-range predictions are caused, in part, by the erroneous assumption of a linear relationship between the explanatory variables and the binary outcome variable. Although it is possible to round the predictions up or down to obtain probabilities bounded at zero and one, the out-of-range predictions are strong evidence that data do not meet the assumptions of the model.

variance components can be used to measure different aspects of arbitrariness in death noticing decision-making, such as those detailed in Parts 2, 3, and 4.

Unreliability Measures. Part 2 described four measures of unreliability/inconsistency in capital charging based on the aforementioned variance components: within-circuit unreliability (ICC_1),³³⁷ between-circuit unreliability (ICC_2),³³⁸ the mean absolute deviation (MAD) of the circuits from the statewide average,³³⁹ and the median odds ratio (MOR).³⁴⁰ The (un)reliability of capital charging behavior was also explored for subsets of cases that varied according to the race of the defendant and race of the victim. As noted, it is important to examine whether cases differ in variability along legally impermissible dimensions and not just the average level of an outcome when assessing the level of arbitrariness impacting legal decision-making.³⁴¹ This approach, referred to as a “heterogeneous variance analysis,” is useful for exploring whether certain classes of cases appear to be handled more haphazardly than others.

Invalidity Measures. The invalidity/irrationality analysis discussed in Part 3 was based on the *coefficient of determination* (R^2) and the *coefficient of discrimination* (Tjur’s D). Both of these statistics measure how well the specified model predicts the outcome. The R^2 statistic quantifies the proportional reduction in prediction error variance comparing the model without covariates (the “null” or “unadjusted” model) with the model of interest containing all relevant predictors (the “adjusted” model).³⁴² As noted, the error variance has a complex structure in the multilevel context: $Var(\zeta_j + \varepsilon_{ij}) = \psi + \theta$. The coefficient of determination for two-level models is the proportional reduction in the estimated total error variance comparing the unadjusted model without covariates with adjusted model: $R^2 = [(\hat{\psi}_0 + \hat{\theta}_0) - (\hat{\psi}_1 + \hat{\theta}_1)] / (\hat{\psi}_0 + \hat{\theta}_0)$, where $\hat{\psi}_0$ and $\hat{\theta}_0$ are estimates for the

337. See Larsen & Merlo, *supra* note 175, at 82.

338. See RAUDENBUSH & BRYK, *supra* note 123.

339. See *supra* note 195 and accompanying text. The mean absolute deviation (MAD) uses the circuit-specific deviation to calculate the average difference of the circuit-level probabilities from the state-wide probability for a factual similar case.

340. See *supra* note 199. The median odds ratio relies on the between-circuit residual variance to quantify the variation between circuits by comparing two charging decisions in factually similar cases from two randomly chosen clusters. The MOR is the average ratio between the cases of higher propensity with the cases of lower propensity.

341. Hedeker et al., *supra* note 135.

342. Analysts disagree as to whether the coefficient of determination is an accurate measure of “model fit” because the R^2 will tend to be small when the “true” model has a large residual variance, and therefore it would be erroneous to interpret a small R^2 as indicating model misspecification. See Gary King, *How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science*, 30 AM. J. POL. SCI. 666, 675 (1986).

unadjusted model, and $\hat{\psi}_1$ and $\hat{\theta}_1$ are the estimates for the adjusted model.³⁴³ The proportional reduction in each of the variance components can be evaluated: $R_1^2 = (\hat{\theta}_0 - \hat{\theta}_1)/\hat{\theta}_0$ and $R_2^2 = (\hat{\psi}_0 - \hat{\psi}_1)/\hat{\psi}_0$, where R_1^2 and R_2^2 are, respectively, the proportional reduction in the within- and between-circuit residual variances.³⁴⁴ Similar to the aforementioned unreliability analyses, I disaggregate the measure of the rationality of death penalty charging by race of the defendant and race of the victim and provide an assessment of whether prosecutors death charging decisions appear to be more or less rational/valid depending on the particular subclass of case.

Unlike models analyzing continuous predictors, the Level-1 residual variance in the logistic regression model, θ , is fixed so it cannot decrease when adding other variables to the model.³⁴⁵ As a consequence, the variance estimates of the random effects become larger when explanatory variables are included and, therefore, lack a straightforward interpretation.³⁴⁶ The random effects must be rescaled to permit the calculation of the R^2 statistics for the different model specifications.³⁴⁷ The rescaling procedure includes: (a) calculating the total variance of the unadjusted model; (b) calculating the total variance of adjusted model; (c) calculating the scale correction factor (i.e., the square root of the ratio of the variances of unadjusted model to the adjusted model); and (d) rescaling the random effects by using the scale correction factor.³⁴⁸

A complimentary measure to the R^2 statistic was also calculated: Tjur's D . This statistic assesses the fit of a model on observed data by comparing the predicted probability that the prosecutor filed a death penalty notice when a death notice was actually filed to the predicted probability that a death notice was filed when there was actually no noticed filed by the prosecution. The intuition underlying the Tjur's D is that the rationality of death notice decision-making is directly proportionality to its ability to minimize both false positive and false negatives. When the predictive model, which includes many of the legally (and empirically) relevant factors purportedly driving death penalty charging, fails to differentiate cases with an acceptable degree accuracy, then the rationality/validity/accuracy of the decision-making is

343. SNIJDERS & BOSKER, *supra* note 185, at 110–19.

344. RAUDENBUSH & BRYK, *supra* note 123, at 68–98.

345. HOX, *supra* note 210, at 133–39.

346. There is no direct analog to the coefficient of determination for logistic regression, but several “pseudo- R^2 ” measures have been developed. *See, e.g.*, J. SCOTT LONG, REGRESSION MODELS FOR CATEGORICAL AND LIMITED DEPENDENT VARIABLES 102–08 (1997) (describing goodness-of-fit measures for binary regression models).

347. HOX, *supra* note 210.

348. *Id.*

highly questionable. Tjur's D often provides similar answers as R^2 , but is often deemed to be a more interpretable measure when the decision is binary, as is the case with death charging decisions (i.e., yes/no).

Disproportionality Measures. Part 4 presented results from the racial disproportionality component of the inquiry. This analysis relied on another key feature of MLMs—the ability to measure the variability of the effects of case-level explanatory factors across judicial circuits. These models, sometimes called “random coefficient” or “heterogeneous effects” models, capture differences in case-level effects across circuits: $\eta_{ij} = (\beta_0 + \zeta_{0j}) + (\beta_k + \zeta_{kj})$, where β_0 and ζ_{0j} were defined earlier, β_k is the average effect of the case-level explanatory variable for the state, and ζ_{kj} is the judicial circuit-specific deviation from the average effect of β_k .³⁴⁹ For the purposes of the current study, the coefficients of interest were the race-of-defendant and race-of-victim effects.³⁵⁰ Recall from the example described in the paper, the statewide average race-of-victim effect, β_k , was a 15 percentage point increase in the probability of receiving a death notice when the victim was Caucasian, compared to when the victim was African American, all else equal. The jurisdiction-specific deviations for the race-of-victim effect, ζ_{kj} , were as large as 26 percentage points (Ocumlgee).

349. SKRONDAL & RABE-HESKETH, *supra* note 319, at 50.

350. *Id.*

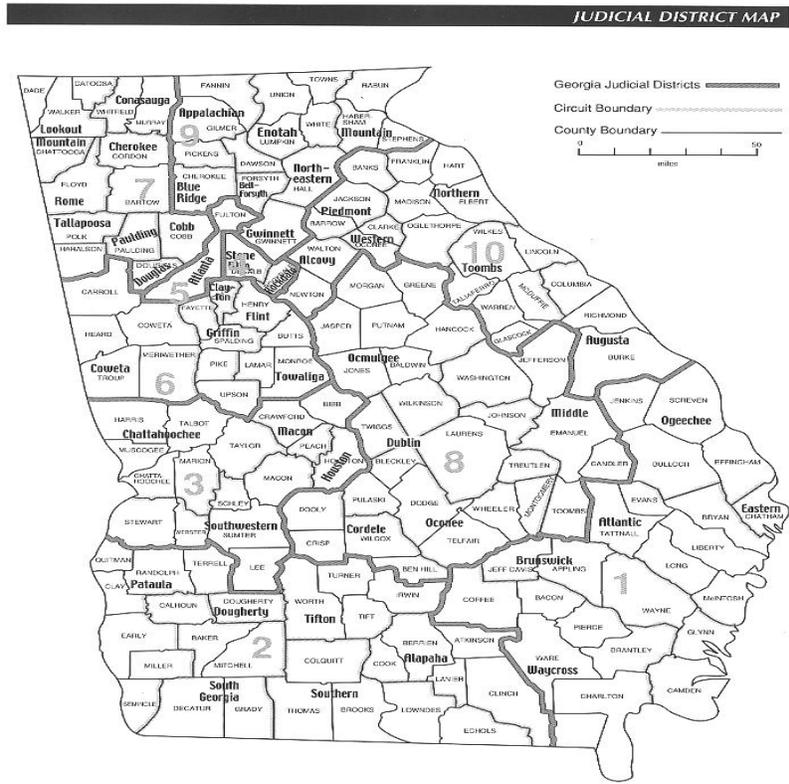


FIGURE 1: MAP OF GEORGIA

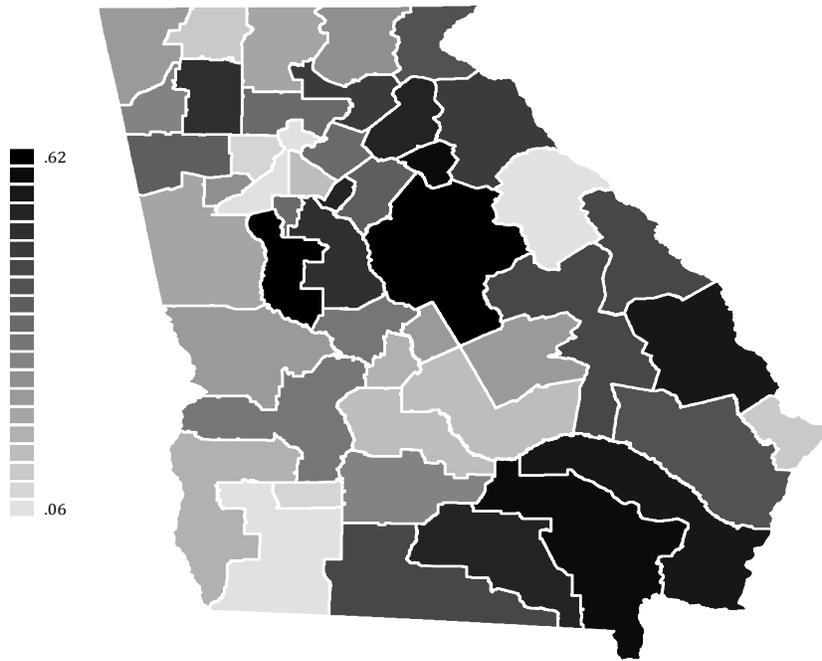


FIGURE 2: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (UNADJUSTED)

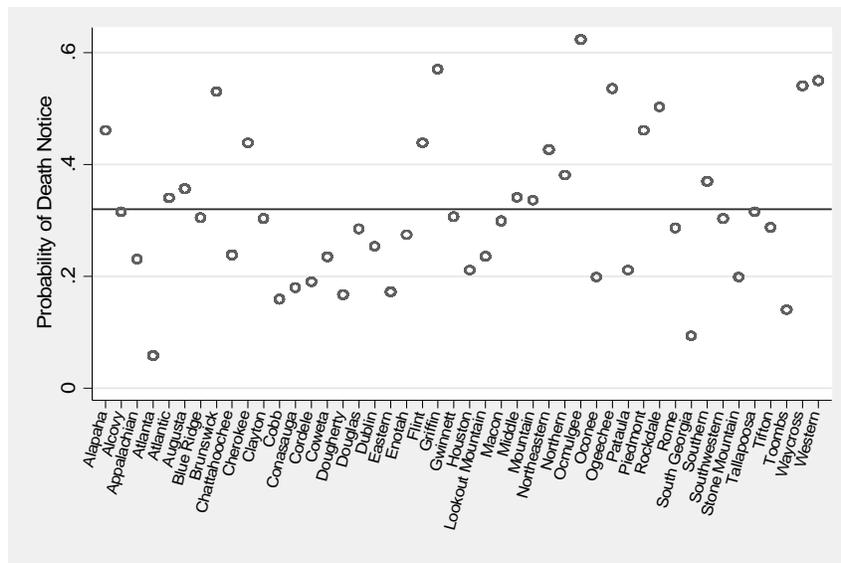


FIGURE 3: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (UNADJUSTED)

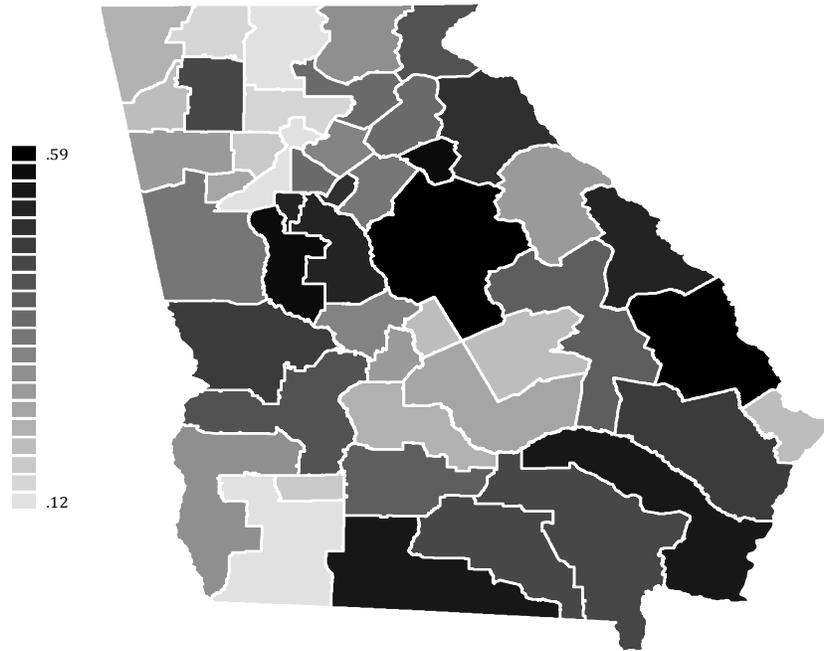


FIGURE 4: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (ADJUSTED)

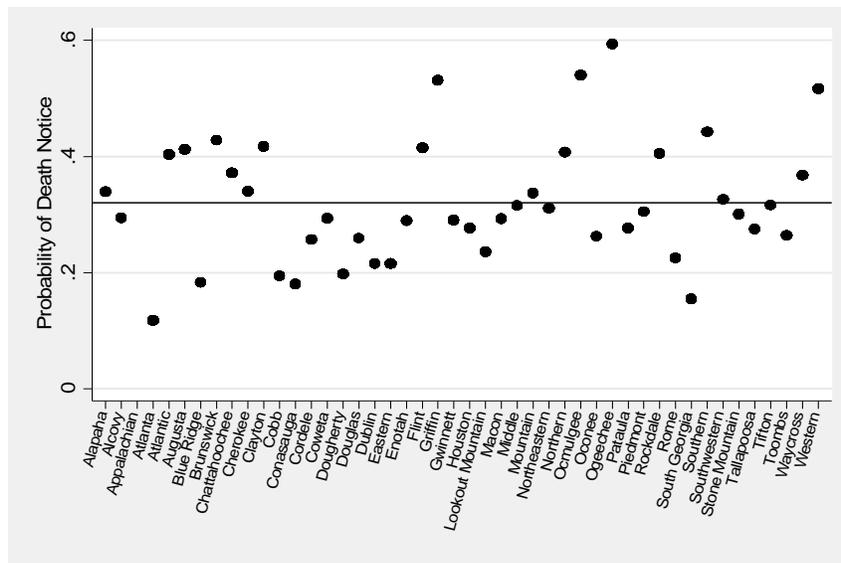


FIGURE 5: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (ADJUSTED)

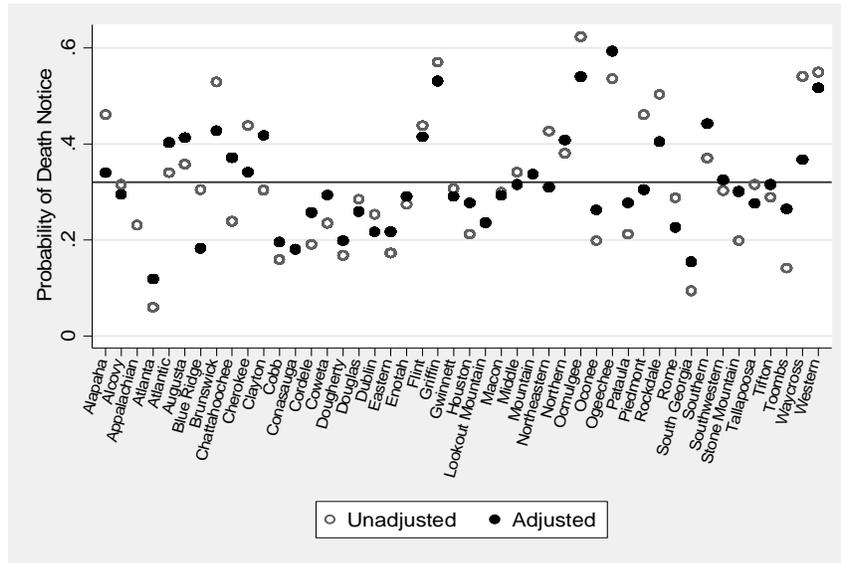


FIGURE 6: JURISDICTIONAL VARIABILITY IN THE PROBABILITY OF RECEIVING A DEATH NOTICE (UNADJUSTED & ADJUSTED ESTIMATES)

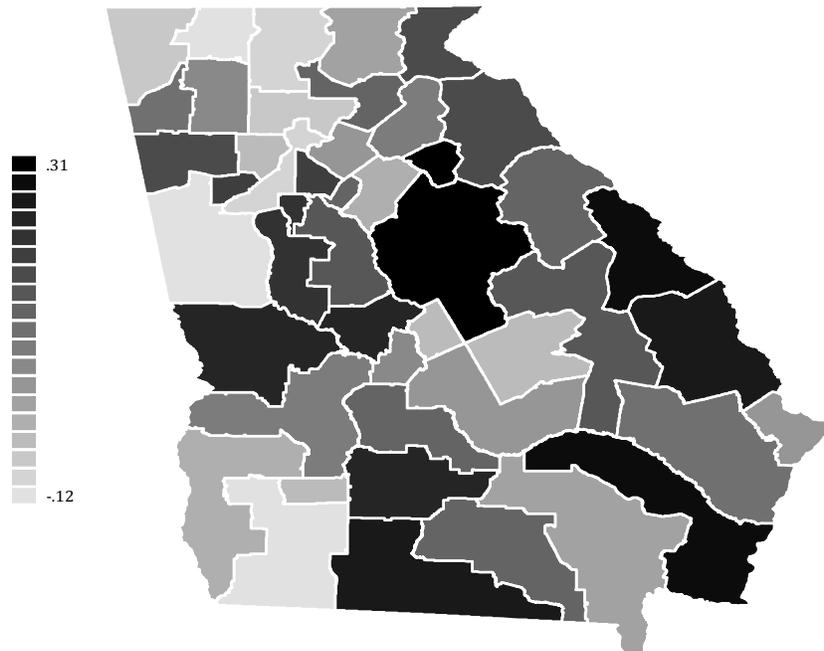


FIGURE 7: JURISDICTIONAL VARIABILITY IN THE RACE OF DEFENDANT EFFECT (CAUCASIAN)

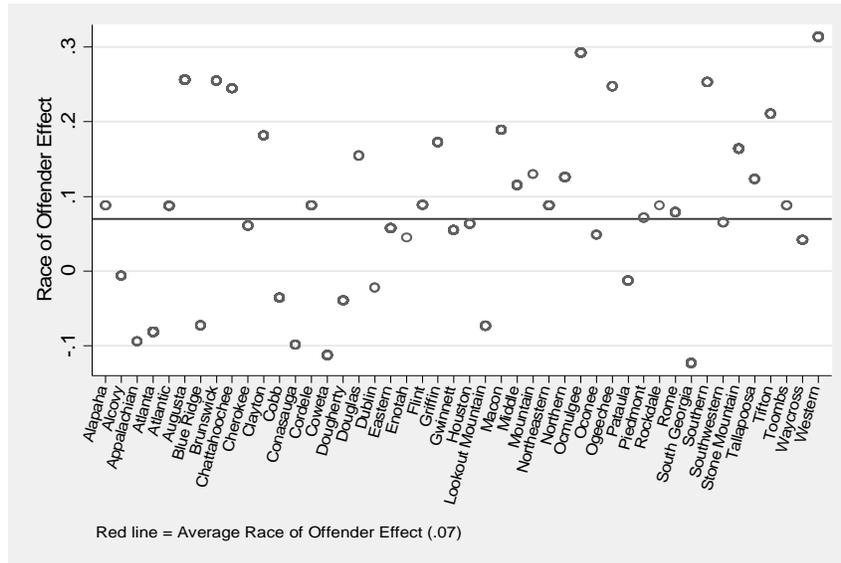


FIGURE 8: JURISDICTIONAL VARIABILITY IN THE RACE OF DEFENDANT EFFECT (CAUCASIAN)

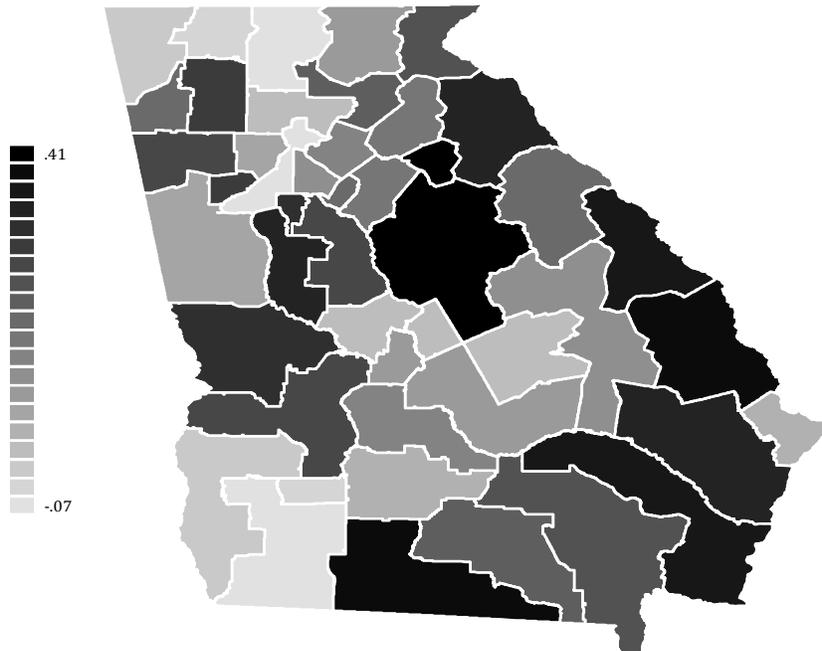


FIGURE 9: JURISDICTIONAL VARIABILITY IN THE RACE OF VICTIM EFFECT (CAUCASIAN)

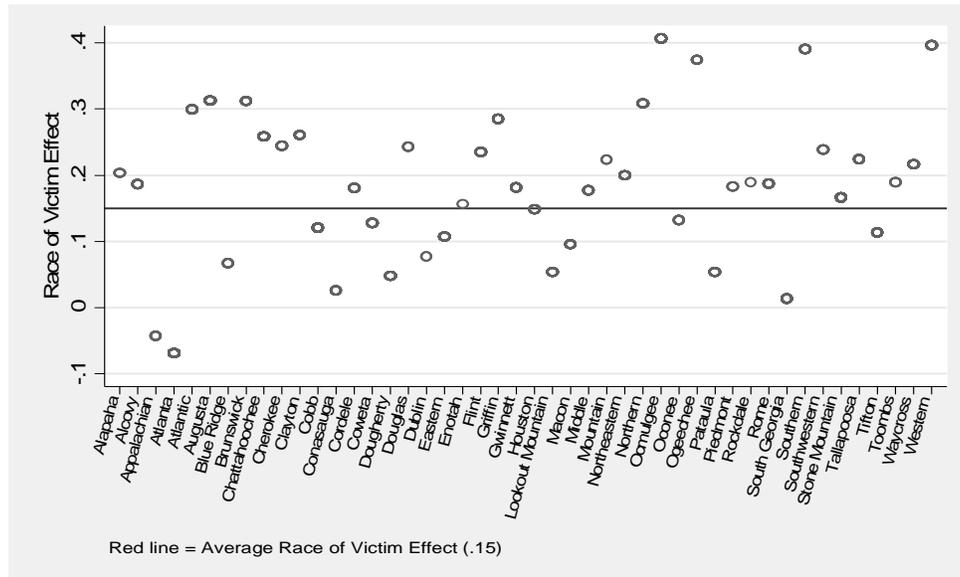


FIGURE 10: JURISDICTIONAL VARIABILITY IN THE RACE OF VICTIM EFFECT (CAUCASIAN)

TABLE 1: SUMMARY STATISTICS

Variables	Mean/ Proportion	Std. Dev.	Min	Max
DP Notice Filed	0.322	--	0	1
Total Statutory Aggs	2.276	1.091	1	7
Year of Offense	--	--	1993	2000
# of Defs	1.793	1.109	1	7
Def White	0.248	--	0	1
Def Black	0.728	--	0	1
Def Latino	0.018	--	0	1
Def Male	0.946	--	0	1
Def Age	27.15	9.935	17	69
Def Contemp Fels	1.724	1.602	0	9
Def Prior Fels	0.514	1.332	0	10
Def # of Children	0.583	--	0	1
Def Employed	0.562	--	0	1
Def Married	0.179	--	0	1
Def HS Grad	0.262	--	0	1
Def Military Service	0.084	--	0	1
Def Drug Use	0.506	--	0	1
Def Psych Status ³⁵¹	1.219	0.508	1	4
Def IQ (Culture Fair)	110	14.833	50	151
Def WRAT ³⁵²	8.089	3.494	1	13
Def Fam History ³⁵³	1.298	1.224	0	5
Monetary Motive	0.577	--	0	1
Sex-Crime Motive	0.053	--	0	1
Gang Related Motive	0.003	--	0	1
Def is "Trigger Person"	0.853	--	0	1
Firearm Homicide	0.644	--	0	1
# of Vics	1.185	0.504	1	6
Vic White	0.448	--	0	1
Vic Black	0.497	--	0	1
Vic Latino	0.034	--	0	1
Vic. Asian/PI	0.021	--	0	1
Vic Female	0.368	--	0	1
Vic Age	36.72	18.20	0	97
Vic Stranger	0.350	--	0	1
Interracial Crime	0.283	--	0	1
County	--	--	1	159
Judicial Circuit	--	--	1	46
Total Cases	1,238			

351. Defendant's psychiatric status: no impairment, minimal impairment, serious impairment, and severe impairment.

352. Wide Range Achievement Test (reading, math, and spelling).

353. Summary measure of how many risk factors for criminality were present in the defendant's family environment during childhood: alcohol/drug abuse, emotional/psychological abuse, physical abuse, family criminality, and "broken home."

TABLE 2: DEATH NOTICES BY JUDICIAL CIRCUIT (1993-2000)

Judicial Circuit	Death Notices	Percent of Total Death Notices
Alapaha	3	0.8
Alcovy	6	1.5
Appalachian	1	0.3
Atlanta	21	5.3
Atlantic	11	2.8
Augusta	29	7.2
Blue Ridge	3	0.8
Brunswick	16	4.0
Chattahoochee	16	4.0
Cherokee	9	2.3
Clayton	19	4.8
Cobb	11	2.8
Conasauga	1	0.3
Cordele	3	0.8
Coweta	6	1.5
Dougherty	7	1.8
Douglas	3	0.8
Dublin	1	0.3
Eastern	10	2.5
Flint	9	2.3
Griffin	17	4.3
Gwinnett	13	3.3
Houston	2	0.5
Lookout Mountain	3	0.8
Macon	8	2.0
Middle	8	2.0
Mountain	2	0.5
Northeastern	11	2.8
Northern	9	2.3
Ocmulgee	26	6.5
Oconee	2	0.5
Ogeechee	10	2.5
Pataula	3	0.8
Paulding	5	1.3
Rockdale	4	1.0
Rome	5	1.3
South Georgia	2	0.5
Southern	12	3.0
Southwestern	4	1.0
Stone Mountain	25	6.3
Tallapoosa	6	1.5
Tifton	8	2.0
Waycross	12	3.0
Western	18	4.5
<i>Total Death Notices: 400</i>		
<i>Percent of all judicial circuits filing a death notice: 96%</i>		

Dr. Sherod Thaxton,
Leveraging Death

LEVERAGING DEATH

SHEROD THAXTON*

Empirical research addressing the use of the death penalty as leverage in plea negotiations is virtually nonexistent. This is particularly surprising given the fact that both plea bargaining and capital punishment have been the focus of much scholarly attention. The U.S. Supreme Court has explicitly approved guilty pleas induced out of fear of the death penalty, yet the impact of the threat of the death penalty on the likelihood of parties reaching a plea agreement is far from obvious. On the one hand, prosecutors, defense attorneys, and defendants may have especially strong incentives to plea bargain in death-eligible cases. On the other hand, many of these advantages may be offset by forces pushing against compromise on both sides of the aisle precisely because the death penalty is an option, so the role the death penalty is playing in plea negotiations in the aggregate remains ambiguous. To date, the only empirical study to explore this issue concluded that the threat of capital punishment does not impact the likelihood of reaching a plea agreement. Unfortunately the study suffers from several limitations that may have ultimately masked any true effect that the death penalty has on plea-bargaining rates. This Article reexamines this question using an originally constructed data set of recent capital charging-and-sentencing decisions in Georgia (1993–2000) that is able to avoid many of the shortcomings of the sparse prior research. The results provide strong evidence that the threat of the death penalty has a robust causal effect on the likelihood of a plea agreement—the threat of the death penalty increases the probability of a plea agreement by approximately 20 to 25 percentage points across various model

* Acting Professor of Law, UCLA School of Law. Former Staff Attorney, Capital Habeas Unit, Office of the Federal Defender for the Eastern District of California. I would like to thank Daniel Abebe, Anya Bernstein, Alexander Boni-Saenz, Bernard Harcourt, William Hubbard, Aziz Huq, Richard McAdams, Tracey Meares, Thomas Miles, Jennifer Nou, Andres Sawicki, Julia Simon-Kerr, and participants of The University of Chicago Law School Faculty Workshop for useful comments on drafts of this Article. Naturally, all remaining errors are my own.

specifications. Not only is this finding important in its own right by illuminating capital defendants' behavioral response to the death penalty, it also has meaningful implications for other purported benefits of plea bargaining in the capital context. The paper briefly considers one of the most commonly identified benefits of plea bargaining—cost reduction—and concludes that the death penalty fails to deter sufficient numbers of murder defendants from opting for trial to offset the significant expense of a capital case and subsequent appeals.

TABLE OF CONTENTS

INTRODUCTION.....	477
I. PLEA BARGAINING IN THE SHADOW OF DEATH	484
A. Bargaining incentives	484
B. Bargaining disincentives.....	487
II. PRIOR RESEARCH.....	492
III. GEORGIA'S DEATH PENALTY	498
A. History and Description of the Modern Statute.....	498
B. Life Without the Possibility of Parole	502
C. Capital Case Progression	503
IV. DATA	505
V. EMPIRICAL STRATEGY.....	507
A. Designation of Treatment and Control Groups	507
B. Statistical Model	513
C. Missing Data	519
VI. RESULTS	521
A. Fixed-effects logit specifications.....	521
B. Sensitivity analyses.....	524
VII. FINANCIAL IMPLICATIONS AND DISCUSSION.....	540
VIII. CONCLUSION	549

INTRODUCTION

Plea bargaining¹ is a crucial feature of our criminal justice system, as approximately 95% of convictions that occur within a year of arrest are obtained by a guilty plea.² Despite its current centrality, however, “plea bargaining did not occur with any frequency until well into the nineteenth century,”³ and the Supreme Court did not specifically address its constitutionality until long after it was common practice in the criminal justice system.⁴ In *Brady v. United States*,⁵ the Court reasoned that plea bargaining benefited both sides of the adversarial system and was “inherent in the criminal law and its administration.”⁶ The following year, in *Santobello v. New York*,⁷ the Court defended the practice of plea bargaining, calling it “an essential component of the administration of justice”⁸ that was to be encouraged as long as it was properly administered (i.e., as long as pleas were intelligent and voluntary). Several years later in *Bordenkircher v. Hayes*,⁹ the Court endorsed prosecutorial threats of stiffer penalties when defendants refuse to accept a plea offer.¹⁰ In fact, since its formal

¹ Plea bargaining is “[a] negotiated agreement between a prosecutor and a criminal defendant whereby the defendant pleads guilty to a lesser offense or to one of multiple charges in exchange for some concession by the prosecutor, usu. a more lenient sentence or a dismissal of the other charges.” BLACK’S LAW DICTIONARY 1270 (9th ed. 2009). Plea bargaining may involve three areas of negotiation: charge bargaining (negotiation to plead guilty or no contest in exchange for the dismissal of some counts or reduction of the charge); sentence bargaining (negotiation to plead guilty or no contest in exchange for the prosecutor’s recommendation to the court for a lighter sentence); and fact bargaining (negotiation to stipulate to certain facts in exchange for an agreement not to introduce other facts into evidence). *Id.*

² See THOMAS H. COHEN & TRACEY KYCKELHAHN, BUREAU OF JUSTICE STATISTICS, U.S. DEP’T OF JUSTICE, FELONY DEFENDANTS IN LARGE URBAN COUNTIES, 2006, at 10 (2010); see also *Padilla v. Kentucky*, 130 S. Ct. 1473, 1485 n.13 (2010) (recognizing that only 5% of federal and state felony prosecutions are resolved by trial).

³ Albert W. Alschuler, *Plea Bargaining and Its History*, 79 COLUM. L. REV. 1, 5 (1979); see also *id.* at 19 (“It was only after the Civil War that cases of plea bargaining began to appear in American appellate court reports.”).

⁴ See John H. Langbein, *Torture and Plea Bargaining*, 46 U. CHI. L. REV. 3, 9 (1978) (“If you turn to the American Constitution in search of authority for plea bargaining, you will look in vain. Instead, you will find—in no less hallowed a place than the Bill of Rights—an opposite guarantee, a guarantee of trial.”).

⁵ 397 U.S. 742 (1970).

⁶ *Id.* at 751.

⁷ 404 U.S. 257 (1971).

⁸ *Id.* at 260.

⁹ 434 U.S. 357 (1978).

¹⁰ *Id.* at 363. Eight years before *Bordenkircher*, the Court ruled that the mere fact that a defendant pleads guilty to murder to avoid the death penalty does not make the plea involuntary, especially when the defendant is represented by competent counsel who can assess the weight of the evidence against the defendant. *North Carolina v. Alford*, 400 U.S.

endorsement of plea bargaining, the Court has been reluctant to regulate plea bargains, recently noting that “[h]indsight and second guesses are [] inappropriate . . . where a plea has been entered without a full trial”¹¹

Most judges support the system of plea bargaining because it allows them to alleviate the need to schedule and hold a trial on what are typically already overcrowded dockets.¹² Prosecutors desire both the reduced caseload and assurance of a conviction from plea bargaining. Plea negotiations also allow prosecutors to strengthen their cases against codefendants by offering certain defendants a plea arrangement in exchange for testimony against one or more codefendants. This practice assures prosecutors at least one conviction while also enhancing the chances of a subsequent conviction. Defendants are allowed to avoid a more serious

25, 31 (1970).

¹¹ *Premo v. Moore*, 131 S. Ct. 733, 745 (2011). Notable exceptions include the Court’s recent decisions in *Missouri v. Frye*, 132 S. Ct. 1399 (2012), *Lafler v. Cooper*, 132 S. Ct. 1376 (2012), and *Padilla v. Kentucky*, 130 S. Ct. 1473 (2010). In *Padilla*, the Court held that defense counsel’s incorrect advice to a client regarding the deportation risk associated with pleading guilty may violate the Sixth Amendment guarantee to effective assistance of counsel if it prejudices the client’s decision. 130 S. Ct. at 1478. The Court also rejected the argument that its ruling would ultimately open the “floodgates” of litigation by providing defendants with new causes of action. *Id.* at 1484–85. *But see* Derek Wikstrom, Note, “*No Logical Stopping-Point*: The Consequences of *Padilla v. Kentucky*’s Inevitable Expansion,” 106 NW. U. L. REV. 351, 374 (2012) (“Even if *Padilla* has an eventual logical stopping-point, that point comes after extension to many other collateral consequences that are as serious as deportation.”).

In early 2012, the Court decided *Cooper* and *Frye*, extending *Padilla*’s logic to nondeportation cases and holding that a defendant’s Sixth Amendment right to effective assistance of counsel may be violated when (1) defense counsel fails to properly inform the defendant of a beneficial plea agreement offered by the prosecution (*Frye*), or (2) defense counsel incorrectly advises the defendant on the state of the law, leading the defendant to reject a beneficial plea agreement (*Cooper*). *Frye*, 132 S. Ct. 1399; *Cooper*, 132 S. Ct. 1376. The Court recognized that defense counsel’s duty to inform a client of formal plea offers may be subject to exceptions. *Frye*, 132 S. Ct. at 1408. It also pointed out that claims raised under “ineffective advice” must satisfy three conditions: (1) the ineffective advice was the “but for” cause of the plea not being accepted by the defendant, (2) the trial court would have ultimately accepted the terms of the bargain, and (3) the conviction or sentence would have been lesser than those imposed after trial. *Cooper*, 132 S. Ct. at 1385; *see also* Justin F. Marceau, *Embracing a New Era of Ineffective Assistance of Counsel*, 14 U. PA. J. CONST. L. 1161, 1191–92 (2012) (analyzing the Court’s recent plea-bargaining decisions). The impact of these decisions on plea bargaining is debatable, however. As one federal trial judge has explained, most of the problems that occur in the plea-bargaining process do *not* result from ineffective assistance of counsel, but rather prosecutorial overconfidence in the face of questionable evidence and sources. Jed S. Rakoff, *Frye and Lafler: Bearers of Mixed Messages*, 122 YALE L.J. ONLINE 25, 26 (2012).

¹² *Blackledge v. Allison*, 431 U.S. 63, 71 (1977) (noting that plea bargaining permits “[j]udges and prosecutors [to] conserve vital and scarce resources”). *But see* Jo Dixon, *The Organizational Context of Criminal Sentencing*, 100 AM. J. SOC. 1157, 1177 (1995) (suggesting that the level of plea bargaining is high irrespective of caseload pressure).

charge or sentence and, if represented by private counsel, avoid the cost of a trial.¹³ So, on balance, the practice of plea bargaining is generally believed to be superior to trials due to reduced costs, improvements in the speed and efficiency of case processing, and increases in the certainty of convictions.¹⁴ The practice is not without its detractors, however, as legal academics and practitioners continue to debate its fairness and desirability.

Critics of the plea-bargaining system emphasize that it encourages prosecutors to “overcharge” at the start of the case in an effort to coerce defendants into accepting a plea, allows prosecutors to “cure” defects in their cases by avoiding trial, and encourages defendants to plead guilty to crimes that they did not commit.¹⁵ A defendant who agrees to a plea bargain may also be required to relinquish certain constitutional and statutory rights in exchange for a negotiated plea.¹⁶ Opponents of plea bargaining also suggest that the practice allows defendants to avoid the appropriate punishment for their crimes (as established by state legislative

¹³ See, e.g., Frank H. Easterbrook, *Criminal Procedure as a Market System*, 12 J. LEGAL STUD. 289, 297–98 (1983) (defending plea bargains as an element of a well-functioning market system and cautioning that efforts to restrict actors’ discretion in criminal procedure will have negative consequences); Frank H. Easterbrook, *Plea Bargaining as Compromise*, 101 YALE L.J. 1969 (1992); William M. Landes, *An Economic Analysis of the Courts*, 14 J.L. & ECON. 61, 66–69 (1971) (describing factors influencing the likelihood of parties reaching plea agreements); Robert E. Scott & William J. Stuntz, *Plea Bargaining as Contract*, 101 YALE L.J. 1909, 1918–34, 1967–68 (1992) (explaining the risks and entitlements parties trade in plea bargaining).

¹⁴ See, e.g., GEORGE FISHER, *PLEA BARGAINING’S TRIUMPH: A HISTORY OF PLEA BARGAINING IN AMERICA 200–01* (2003) (discussing the growth of plea bargaining in response to the operational goals of judges and prosecutors); Donald J. Newman, *Pleading Guilty for Considerations: A Study of Bargain Justice*, 46 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 780, 790 (1956) (describing plea bargaining “as a natural, expedient outgrowth of deficiencies in the administration of our ‘trial-by-combat’ theory of justice”); Edward A. Ruttenburg, *Plea Bargaining Analytically—The Nash Solution to the Landes Model*, 7 AM. J. CRIM. L. 323, 353 (1979) (“Plea bargaining should be accepted openly as a system which can accomplish the goals of justice as completely as can a pure trial system, while at the same time releasing resources to serve society in other areas of life.”).

¹⁵ See Langbein, *supra* note 4, at 3 (discussing the parallels between the coerciveness of current plea-bargaining practices and torture in medieval European courts); see also Oren Bar-Gill & Omri Ben-Shahar, *The Prisoners’ (Plea Bargain) Dilemma*, 1 J. LEGAL ANALYSIS 737, 769 (2009) (positing that plea bargaining may benefit individual defendants, but due to coordination problems, may not benefit defendants as a group); Oren Gazal-Ayal, *Partial Ban on Plea Bargains*, 27 CARDOZO L. REV. 2295, 2306 (2006) (arguing that prosecutors use negotiated pleas in cases with weak—possibly inadmissible—evidence and when specific defenses might be established in court); Stephen J. Schulhofer, *Plea Bargaining as Disaster*, 101 YALE L.J. 1979, 1980–91 (1992) (describing structural flaws—e.g., innocence and conflicts of interest—in the plea-bargaining system).

¹⁶ Stephanos Bibas, *Regulating the Plea-Bargaining Market: From Caveat Emptor to Consumer Protection*, 99 CALIF. L. REV. 1117, 1123 (2011) (discussing the rights defendants generally forfeit that could otherwise provide a cause of action for an appeal).

bodies) and that the practice heavily favors defendants with savvy lawyers, irrespective of the defendants' actual culpability.¹⁷ Also, claims of plea bargaining as a "necessary" and "inevitable" component of our criminal justice system have been challenged, as critics of plea bargaining point to jurisdictions that have experimented with partial or complete bans on the practice.¹⁸

The debate over the promises and pitfalls of plea bargaining is perhaps most contentious in the context of the death penalty,¹⁹ yet scholars have

¹⁷ WELSH S. WHITE, THE DEATH PENALTY IN THE NINETIES: AN EXAMINATION OF THE MODERN SYSTEM OF CAPITAL PUNISHMENT 56 (1991) [hereinafter WHITE, DEATH PENALTY] (explaining that many prosecutors disfavor litigating pretrial motions, and defense attorneys who engage in intensive pretrial motion work are likely to obtain favorable plea bargains for their clients); Albert W. Alschuler, *Plea Bargaining and the Death Penalty*, 58 DEPAUL L. REV. 671, 674 (2009) (suggesting that plea bargaining for a reduced punishment undermines both the symbolic and instrumental purposes of punishment); Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463, 2493 n.116 (2004) (arguing that plea bargaining tends to favor repeat offenders).

¹⁸ See Michael L. Rubinstein & Teresa J. White, *Alaska's Ban on Plea Bargaining*, 13 LAW & SOC'Y REV. 367, 367 (1979) (describing Alaska's attempt to ban plea bargaining for the vast majority of cases); Robert A. Weninger, *The Abolition of Plea Bargaining: A Case Study of El Paso County, Texas*, 35 UCLA L. REV. 265, 311–13 (1987) (explaining that efforts to abolish plea bargaining have achieved mixed results). But several scholars have noted that the incentives to plea bargain are so strong for all the primary actors in the criminal justice system that they simply find alternative ways to achieve the same end even when the practice has been expressly prohibited. See, e.g., Joseph L. Hoffmann et al., *Plea Bargaining in the Shadow of Death*, 69 FORDHAM L. REV. 2313, 2390 (2001).

¹⁹ See *infra* Part I. Compare, e.g., ROBERT M. BOHM, DEATHQUEST II: AN INTRODUCTION TO THE THEORY AND PRACTICE OF CAPITAL PUNISHMENT IN THE UNITED STATES 207–08 (2003) (commenting that the availability and use of plea bargaining can contribute to arbitrariness and discrimination in the implementation of the death penalty), WHITE, *supra* note 17, at 54 (conducting interviews with capital defense attorneys and concluding that "the likelihood of a plea bargain in a capital case will be dramatically affected by factors that have nothing to do with the nature of the crime or the strength of the evidence against the defendant"), Alschuler, *supra* note 17, at 674 (suggesting that plea bargaining "undermines the [death penalty's] most common rationale . . . some crimes are so horrible that they simply require it"), and Daniel Givelber, *The New Law of Murder*, 69 IND. L.J. 375, 410–11 (1994) (explaining that the existence of statutorily defined special circumstances for death eligibility strengthens the prosecutor's bargaining position and facilitates arbitrariness), with WELSH S. WHITE, LITIGATING IN THE SHADOW OF DEATH 145–171 (2006) [hereinafter WHITE, LITIGATING IN THE SHADOW OF DEATH] (explaining that plea bargaining is an important tool for defense counsel to save their clients' lives, but also recognizing how the practice leads to increased arbitrariness and potentially wrongful convictions), and Russell Stetler, *Commentary on Counsel's Duty to Seek and Negotiate a Disposition in Capital Cases (ABA Guideline 10.9.1)*, 31 HOFSTRA L. REV. 1157, 1157–58 (2003) (noting that, in 2003, the American Bar Association Guidelines recognized that attorneys representing a capital defendant have an obligation to seek negotiated pleas, and suggesting that the current number of death row inmates could have been drastically reduced if their defense attorneys worked more diligently to obtain a negotiated plea).

conducted very little research on the relationship between the death penalty and plea negotiations.²⁰ Instead, they have focused the bulk of their attention on the examination of the possible deterrent effect of the death penalty on potential murderers.²¹ This Article offers an empirical examination of the causal impact of the threat of the death penalty on the likelihood of parties reaching a plea agreement. This type of inquiry is particularly relevant because the effect of the death penalty on plea bargaining is theoretically ambiguous. The threat of the death penalty may induce defendants who may not have otherwise accepted a plea agreement to plead to avoid the risk of possible execution, so the overall number of cases proceeding to trial is reduced. An opposite effect is plausible as well: armed with the threat of the “ultimate penalty,” prosecutors may be less willing to offer capital defendants desirable plea bargains, if any bargain at all, so the number of cases going to trial may increase.²² It is also possible

²⁰ See Ilyana Kuziemko, *Does the Threat of the Death Penalty Affect Plea Bargaining in Murder Cases? Evidence from New York's 1995 Reinstatement of Capital Punishment*, 8 AM. L. & ECON. REV. 116, 141 (2006) (conducting the first study of the effect of the death penalty on plea bargaining); see also Susan Ehrhard, *Plea Bargaining and the Death Penalty: An Exploratory Study*, 29 JUST. SYS. J. 313, 315 (2008) (noting that there has been only one systematic examination of the impact of the death penalty on plea bargaining); Kent S. Scheidegger, *The Death Penalty and Plea Bargaining to Life Sentences 1–2* (Criminal Justice Legal Found., Working Paper No. 09-01, 2009) (same); cf. ANDREW WELSH-HUGGINS, NO WINNERS HERE TONIGHT: RACE, POLITICS, AND GEOGRAPHY IN ONE OF THE COUNTRY'S BUSIEST DEATH PENALTY STATES 96–100 (2009) (discussing the high prevalence of plea bargaining in capital cases in California, New York, Ohio, and at the federal level).

²¹ To be sure, social scientists—particularly those working in fields outside of economics—have conducted empirical research on various aspects of the death penalty for more than 75 years, and over the past 30 years, socio-legal research has figured prominently into the debate over the appropriateness and effectiveness of capital punishment in the United States. See David C. Baldus, *Keynote Address: The Death Penalty Dialogue Between Law and Social Science*, 70 IND. L.J. 1033, 1041 (1995); Michael L. Radelet & Marian J. Borg, *The Changing Nature of Death Penalty Debates*, 26 ANN. REV. SOC. 43, 43–44 (2000). But no other issue related to the death penalty has received more systematic attention from legal and academic communities. William C. Bailey & Ruth D. Peterson, *Murder, Capital Punishment, and Deterrence: A Review of the Literature*, in THE DEATH PENALTY IN AMERICA: CURRENT CONTROVERSIES 135, 135 (Hugo Adam Bedau ed., 1997).

²² See Kuziemko, *supra* note 20, at 117 (asserting that the death penalty may make prosecutors more aggressive or recalcitrant, so they may be less willing to strike deals with defendants); see also WHITE, LITIGATING IN THE SHADOW OF DEATH, *supra* note 19, at 157–58 (presenting the problem that some defendants prefer execution rather than life imprisonment); Ehrhard, *supra* note 20, at 318 (highlighting that many prosecutors only make “take it or leave it” offers of life without the possibility of parole in death-eligible cases); James S. Liebman, *Opting for Real Death Penalty Reform*, 63 OHIO ST. L.J. 315, 322–24 (2002) (showing that the availability of the death penalty can trump the usual pressure for prosecutors and defendants to reasonably compromise).

Both chief prosecutors and rank-and-file prosecutors may anticipate huge career payoffs from pursuing a capital case even when the defendant ultimately receives a sentence

that the death penalty will have no impact on the likelihood that a defendant accepts a plea bargain and will only impact the terms of the bargain.²³ As two scholars have recently noted, “opposing hypotheses about the effect of the death penalty on prosecutorial discretion have never been rigorously tested.”²⁴

Examining the impact of capital punishment on plea bargaining is important for several reasons. First, it helps inform our understanding of how sentencing law influences plea bargaining. Given the centrality of pleas for the disposition of criminal cases, studying how sentencing structure impacts the incentives of prosecutors and defendants in plea negotiations is important to our understanding of the criminal justice system.

Second, the use of the death penalty as leverage in plea negotiations raises important legal and ethical issues. Defendants are typically required to waive important constitutional rights as a condition of the plea agreement.²⁵ Many of these rights are considered crucial to the accurate determination of guilt and punishment at trial, so the absence of these protections may undermine our confidence in that determination. This may be of particular concern in the capital context because although defendants who plead guilty avoid the risk of execution, they still receive very lengthy

less than death or the capital sentence is reversed on appeal. *See, e.g.*, WHITE, DEATH PENALTY, *supra* note 17, at 55 (arguing that capital defense lawyers routinely remark that upcoming prosecutorial elections are the most important factor in predicting the likelihood of a plea bargain in a capital case, and savvy defense attorneys attempt to delay trial until after an election in order to increase the likelihood that a prosecutor will offer a plea); Jonathan E. Gradess & Andrew L. B. Davies, *The Cost of the Death Penalty in America: Directions for Future Research*, in THE FUTURE OF AMERICA’S DEATH PENALTY: AN AGENDA FOR THE NEXT GENERATION OF CAPITAL PUNISHMENT RESEARCH 397, 409 (Charles S. Lanier et al. eds., 2009) (“Prosecutors are often hawkish about the death penalty during election campaigns, increasing the probability that they will press for it in office and use it when available.”); Liebman, *supra* note 22, at 324–25 (arguing that trial error is virtually costless to prosecutors). *But see* Ronald F. Wright, *How Prosecutor Elections Fail Us*, 6 OHIO ST. J. CRIM. L. 581, 606 (2009) (suggesting that campaigning prosecutors invoke the death penalty infrequently). Alaska, Connecticut, Delaware, New Jersey, and Rhode Island are the only states that do not popularly elect their district attorneys. John A. Horowitz, Note, *Prosecutorial Discretion and the Death Penalty: Creating a Committee to Decide Whether to Seek the Death Penalty*, 65 FORDHAM L. REV. 2571, 2575 n.24 (1997).

²³ Kuziemko, *supra* note 20, at 140. *See generally* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984) (explaining why parties fail to reach settlement agreements when such agreements are apparently mutually beneficial).

²⁴ Gradess & Davies, *supra* note 22, at 409.

²⁵ Hoffmann et al., *supra* note 18, at 2317–30 (discussing the U.S. Supreme Court’s heightened concern over statutory schemes that “needlessly burden” the exercise of constitutional rights).

sentences—typically life imprisonment. The threat of the death penalty has also induced innocent defendants to plead guilty (and even falsely implicate others) to avoid execution.²⁶

Finally, the study of the plea-negotiation process in the capital context permits us to gain a better understanding of the financial and administrative costs of capital punishment to states and the federal government. Capital trials are extremely expensive and they rarely reduce prison costs because of the infrequency of executions and the added expense associated with housing inmates on death row.²⁷ The threat of capital punishment may result in substantial savings, however, if the threat of execution deters sufficient numbers of individuals from pursuing trial.

Part I of the Article discusses many of the unique aspects of plea bargaining in the capital context and how these factors cut both in favor and against successful plea negotiations. Part II describes the limited empirical research on the impact of the threat of the death penalty on plea bargaining and how certain weaknesses of that particular work undermine our ability to draw firm conclusions about the relationship between capital punishment and plea bargaining. This section also explains how the limitations of this prior research are avoided in the study conducted in this Article using an originally constructed data set on capital charging-and-sentencing decisions in Georgia. Part III briefly outlines Georgia's modern capital punishment and life without the possibility of parole (LWOP) statutes, Part IV describes the data collected on Georgia's capital charging-and-sentencing system that are used to analyze the impact of the death penalty on plea negotiations, and Part V discusses the empirical strategy employed to analyze the data. Part VI reports the results from various model specifications. According to these findings, my conservative estimate is that the threat of the death penalty increases the likelihood of reaching a plea agreement by approximately 20 percentage points. In practical terms, the death penalty increases the plea-bargaining rate from approximately 40% to 60%. In other words, the threat of capital punishment deters roughly two out of every ten death-noticed defendants from pursuing a trial. Part VII discusses the implications of the findings for the administrative and financial costs of

²⁶ See, e.g., Samuel R. Gross et al., *Exonerations in the United States 1989 Through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523, 544–46 (2005) (describing cases in which innocent defendants pleaded guilty to murder, and even falsely implicated others, in order to avoid the death penalty); Paul Hammel, 'Beatrice 6' Cleared; '100 Percent Innocent,' OMAHA WORLD-HERALD, Jan. 27, 2009, at B1 (discussing five exonerated convicted murderers who falsely pleaded guilty after being threatened with the death penalty).

²⁷ John K. Roman et al., *Reassessing the Cost of the Death Penalty Using Quasi-Experimental Methods: Evidence from Maryland*, 11 AM. L. & ECON. REV. 530, 551–53, 571 (2009); see also *infra* Part VII (detailing financial and administrative costs of capital cases).

the death penalty. Based on the high costs associated with litigating a single capital trial and the rather modest ability of the death penalty to deter defendants from pursuing trial, capital punishment does not appear to be a cost-justified bargaining chip.

I. PLEA BARGAINING IN THE SHADOW OF DEATH

A. BARGAINING INCENTIVES

In addition to the perceived increase in plea-bargaining leverage resulting from the severity of the punishment, prosecutors typically enjoy huge advantages by merely seeking the death sentence. First, it enables the government to empanel a “death-qualified” jury. The primary purpose of “death qualification” during voir dire is to remove jurors who unequivocally oppose the death penalty or, conversely, who believe that the death sentence is required in every homicide case.²⁸ Although the Supreme Court has upheld the constitutionality of death-qualified juries in the face of the claim that these juries placed defendants at an unfair risk of conviction,²⁹ available evidence strongly suggests that death-qualified juries are much more likely to convict than non-death-qualified juries.³⁰ The death-qualification process also permits prosecutors to identify jurors with mildly skeptical views of the death penalty (or the criminal justice system in general) who may survive exclusion for cause, and subsequently exercise their peremptory challenges to remove them from the jury.³¹

²⁸ See, e.g., *Wainwright v. Witt*, 469 U.S. 412, 424 (1985) (explaining that the standard for determining when the Court may exclude a prospective juror because of his views on capital punishment is whether “the juror’s views would ‘prevent or substantially impair the performance of his duties as a juror in accordance with his instructions and his oath’” (footnote omitted)).

²⁹ *Witherspoon v. Illinois*, 391 U.S. 510, 517–18 (1968) (“We simply cannot conclude . . . that the exclusion of jurors opposed to capital punishment results in an unrepresentative jury on the issue of guilt or substantially increases the risk of conviction.”).

³⁰ See generally CRAIG HANEY, *DEATH BY DESIGN: CAPITAL PUNISHMENT AS A SOCIAL PSYCHOLOGICAL SYSTEM* 118–21 (2005) (describing the conviction proneness of death-qualified juries); William J. Bowers, *The Capital Jury: Is It Tilted Toward Death?*, 79 *JUDICATURE* 220, 222–23 (1996) (same); Robert Fitzgerald & Phoebe C. Ellsworth, *Due Process vs. Crime Control: Death Qualification and Jury Attitudes*, 8 *LAW & HUM. BEHAV.* 31, 48 (1984) (same); Walter E. Oberer, *Does Disqualification of Jurors for Scruples Against Capital Punishment Constitute Denial of Fair Trial on Issue of Guilt?*, 39 *TEX. L. REV.* 545 (1961) (same); Robert L. Young, *Guilty Until Proven Innocent: Conviction Orientation, Racial Attitudes, and Support for Capital Punishment*, 25 *DEVIANT BEHAV.* 151, 155 (2004) (same).

³¹ James S. Liebman, *The Overproduction of Death*, 100 *COLUM. L. REV.* 2030, 2097 (2000) (explaining that voir dire in capital cases allows prosecutors to “jettison[] the segment of the jury pool that is most likely to be skeptical of informer, police, and forensic testimony and to take seriously the beyond a reasonable doubt standard”); F. Thomas Schornhorst,

Second, seeking the death penalty substantially increases the defense's burden with very little increase, if any, in the government's burden by vastly expanding the defense attorney's role and the requisite skill set and financial resources.³² The overwhelming majority of criminal defense attorneys, both in private practice and employed by the government, are routine negotiators of deals that permit their clients to avoid trial. Representation of a capitally charged client now possibly requires defense counsel to become serious investigators of a horrific crime (or at least coordinators of complex investigations), competent consumers (and communicators) of mental health and forensic science, and experts on complicated (and constantly evolving) specialized constitutional and statutory law.³³ Prosecutors are well aware that many highly skilled trial lawyers will simply refuse to represent a capital defendant, and evidence suggests that capital defendants represented by court-appointed counsel are more likely to receive the death sentence and more likely to have their appeals denied.³⁴ According to one scholar:

Preliminary Screening of Prosecutorial Access to Death Qualified Juries: A Missing Constitutional Link, 62 IND. L.J. 295, 325 (1987) (arguing for prescreening of capital cases for probable cause of the existence of "death-eligible" special circumstances prior to death qualification of a jury); Young, *supra* note 30, at 151 (explaining that death-qualified jurors are more likely to prefer convicting the innocent over acquitting the guilty).

³² Liebman, *supra* note 31, at 2097–98; Liebman, *supra* note 22, at 322; see also Stephen B. Bright, *Counsel for the Poor: The Death Sentence Not for the Worst Crime but for the Worst Lawyer*, 103 YALE L.J. 1835, 1844 (1994) (discussing interrelated reasons for the poor quality of representation in capital cases).

³³ Liebman, *supra* note 22, at 322–23; see also Bibas, *supra* note 16, at 1141 ("Good defense lawyers must know, for example, whether a defendant's small children, ill health, apology, cooperation, or restitution can lower his sentence."); Richard G. Dudley, Jr. & Pamela Blume Leonard, *Getting It Right: Life History Investigation as the Foundation for a Reliable Mental Health Assessment*, 36 HOFSTRA L. REV. 963, 988 (2008) (emphasizing the importance and accompanying complexity of thorough mental health investigation for mitigation in capital cases). See generally Stephen B. Bright, *Death by Lottery—Procedural Bar of Constitutional Claims in Capital Cases Due to Inadequate Representation of Indigent Defendants*, 92 W. VA. L. REV. 679 (1990); Stephen B. Bright, *In Defense of Life: Enforcing the Bill of Rights on Behalf of Poor, Minority and Disadvantaged Persons Facing the Death Penalty*, 57 MO. L. REV. 849, 857–62 (1992) (providing examples of gross ineffective assistance of defense counsel in capital cases).

³⁴ See James C. Beck & Robert Shumsky, *A Comparison of Retained and Appointed Counsel in Cases of Capital Murder*, 21 LAW & HUM. BEHAV. 525, 538 (1997); see also RAYMOND PATERNOSTER ET AL., *THE DEATH PENALTY: AMERICA'S EXPERIENCE WITH CAPITAL PUNISHMENT* 237 (2008) (reporting that court-appointed lawyers representing capitally charged clients in the most active death penalty jurisdictions were significantly more likely to have been professionally disciplined prior to the appointment); Jules Epstein, *Death-Worthiness and Prosecutorial Discretion in Capital Case Charging*, 19 TEMP. POL. & CIV. RTS. L. REV. 389, 400–01 (2010) (suggesting a direct consequence of overly broad capital statutes, and the resulting sky-rocketing capital docket, is the difficulty of finding competent counsel for representation); Liebman, *supra* note 22, at 322 (noting that, when

“There are many small communities that do not have surgeons. But that does not mean that we allow chiropractors to do brain surgery in those communities.” We do, however, let “chiropractors” with law degrees perform the equivalent of brain surgery in capital cases and, predictably, the “patient” often dies.³⁵

Third, prosecutors generally understand that defense attorneys will adopt minimalist, risk-averse pretrial and litigation strategies and practices, focusing primarily on penalty-phase investigation and preparation and on preserving “credibility” at the penalty phase. Consequently, the risk of an acquittal even in cases with genuine evidentiary problems regarding guilt is particularly low.³⁶ In fact, a national study of jurors who served on capital trials revealed that jurors were more likely to vote for the death sentence when defense counsel’s guilt-phase and penalty-phase presentations were logically inconsistent.³⁷ The Supreme Court has also recognized that “[a]ttorneys representing capital defendants face daunting challenges in developing trial strategies [for guilt and penalty phases]”³⁸ and “must strive at the guilt phase to avoid a counterproductive course.”³⁹ Finally, prosecutors recognize that judges at both the trial and appellate levels—especially elected judges—are less likely to require capital trials to strictly follow reliable procedure out of fear that prosecutors will publicly blame judges for losses based on “legal technicalities.”⁴⁰

taking into consideration the vast responsibilities associated with defending a capital defendant, the number of qualified attorneys in any jurisdiction dwindles to a handful or completely disappears); Ronald J. Tabak & J. Mark Lane, *The Execution of Injustice: A Cost and Lack-of-Benefit Analysis of the Death Penalty*, 23 LOY. L.A. L. REV. 59, 70 (1989) (reporting that 90% of inmates on death row could not afford private counsel).

³⁵ Vivian Berger, *The Chiropractor as Brain Surgeon: Defense Lawyering in Capital Cases*, 18 N.Y.U. REV. L. & SOC. CHANGE 245, 254 (1990–1991) (citations omitted).

³⁶ Ehrhard, *supra* note 20, at 318 (describing capital defense attorneys’ concerns about compromised credibility at the penalty phase after aggressively pursuing an innocence defense).

³⁷ Scott E. Sundby, *The Capital Jury and Absolution: The Intersection of Trial Strategy, Remorse, and the Death Penalty*, 83 CORNELL L. REV. 1557, 1589–91 (1998).

³⁸ *Florida v. Nixon*, 543 U.S. 175, 191 (2004) (discussing the potential tensions between guilt and penalty phase strategies).

³⁹ *Id.*

⁴⁰ See Stephen B. Bright & Patrick J. Keenan, *Judges and the Politics of Death: Deciding Between the Bill of Rights and the Next Election in Capital Cases*, 75 B.U. L. REV. 759, 834–35 (1995); Richard R. W. Brooks & Steven Raphael, *Life Terms or Death Sentences: The Uneasy Relationship Between Judicial Elections and Capital Punishment*, 92 J. CRIM. L. & CRIMINOLOGY 609, 638–39 (2002); Andrew Gelman et al., *A Broken System: The Persistent Patterns of Reversals of Death Sentences in the United States*, 1 J. EMPIRICAL LEGAL STUD. 209, 260 (2004); Liebman, *supra* note 31, at 2111–14; Ashley Rupp, *Death Penalty Prosecutorial Charging Decisions and County Budgetary Restrictions: Is the Death Penalty Arbitrarily Applied Based on County Funding?*, 71 FORDHAM L. REV. 2735, 2736–37 (2003) (discussing the tremendous political and community pressure on judges in capital cases).

B. BARGAINING DISINCENTIVES

The aforementioned advantages that prosecutors enjoy in seeking the death penalty would appear to increase the likelihood of a plea bargain; however, these advantages may be offset by forces pushing against compromise on both sides of the aisle. Both systematic and anecdotal evidence suggests that plea negotiations are most likely to occur when (1) both parties perceive a similar expected probability of conviction, (2) the expected trial penalty is relatively small (i.e., low-severity offenses), (3) the settlement discount offered by the government is substantial, (4) the defendant's risk aversion is high, (5) pretrial detention and court delays substantially increase opportunity costs for defendants, and (6) the parties (particularly the defendant) will directly incur the full costs of litigation.⁴¹ But many of these features may be missing or substantially weakened in the majority of capital cases.

The central point of contention in a capital trial is not the expected probability of conviction—approximately 90% of all murder trials result in conviction;⁴² rather, it is the likelihood of a death sentence (or a sentence greater than the statutory minimum) at the penalty phase,⁴³ and perhaps even the expected likelihood of ultimately being executed.⁴⁴ Juries wield enormous discretion in deciding whether to impose the death sentence, which is inherently a subjective enterprise.⁴⁵ Rigorous examinations of jury

⁴¹ See, e.g., Gazal-Ayal, *supra* note 15, at 2299 (explaining that prosecutors obtain guilty pleas when they can offer substantial concessions); Landes, *supra* note 13, at 99 (presenting evidence that the likelihood of a plea agreement depends, *inter alia*, on the severity of the crime, differences in the expected probability of conviction, resources available to the parties, and parties' risk aversion).

⁴² COHEN & KYCKELHAHN, *supra* note 2, at 11; THOMAS H. COHEN & BRIAN A. REAVES, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, FELONY DEFENDANTS IN LARGE URBAN COUNTIES, 2002, at 24 (2006); BRIAN A. REAVES, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, FELONY DEFENDANTS IN LARGE URBAN COUNTIES, 1994, at 24 (1998).

⁴³ FRANKLIN E. ZIMRING, THE CONTRADICTIONS OF AMERICAN CAPITAL PUNISHMENT 53–55 (2003) (describing the penalty phase of a capital trial as a “status competition” between the defendant and the victim's family); Jon Kyl et al., *On the Wings of Their Angels: The Scott Campbell, Stephanie Roper, Wendy Preston, Louarna Gillis, and Nila Lynn Crime Victims' Rights Act*, 9 LEWIS & CLARK L. REV. 581, 607–08 (2005) (describing the proliferation of victims' rights legislation and victim-impact statements during the penalty phase of capital trials).

⁴⁴ NANCY J. KING & JOSEPH L. HOFFMANN, HABEAS FOR THE TWENTY-FIRST CENTURY: USES, ABUSES, AND THE FUTURE OF THE GREAT WRIT 147 (2011) (explaining that capital cases are thirty-five times more likely to be granted relief upon federal habeas review than noncapital cases); James S. Liebman et al., *Capital Attrition: Error Rates in Capital Cases, 1973–1995*, 78 TEX. L. REV. 1839, 1850, 1852 (2000) (noting that 68% of death sentences were overturned on appeal from 1973 to 1995 and 82% of those defendants ultimately avoided the death penalty at resentencing).

⁴⁵ *Zant v. Stephens*, 462 U.S. 862, 891 (1983) (holding that aggravating circumstances

sentencing patterns are unable to identify any meaningful (i.e., legally legitimate) defendant or crime characteristics that consistently distinguish cases that receive the death penalty from cases that do not.⁴⁶ This strongly suggests that prosecutors and defendants in capital cases may significantly differ in their assessments of the expected trial sentence. As a result, prosecutors will only have imperfect information about a defendant's reservation price (i.e., the maximum plea sentence that the defendant would accept to avoid a trial), and might inadvertently make a plea offer exceeding that reservation price.⁴⁷ While prosecutors may generally know the probability that a case would result in a plea bargain, they cannot accurately predict the actual result in specific cases.⁴⁸ Nationally, only one-third of capital trials result in a death sentence,⁴⁹ so a capital defendant may require a larger discount than what a prosecutor is willing to offer. This is

are only required to narrow death eligibility and not to channel jurors' discretion at the penalty phase, and nonstatutory aggravating circumstances may be considered by the jury as well); *accord* Lee v. State, 365 S.E.2d 99, 105 (Ga. 1988), *cert. denied*, 488 U.S. 879 (1988); *see also* Nance v. State, 623 S.E.2d 470, 473 (Ga. 2005) (stating that there is no requirement under the Georgia capital statute that nonstatutory aggravating circumstances be proven beyond a reasonable doubt), *cert. denied*, 549 U.S. 868 (2006); Chelsea Creo Sharon, *The "Most Deserving" of Death: The Narrowing Requirement and the Proliferation of Aggravating Factors in Capital Sentencing Statutes*, 46 HARV. C.R.-C.L. L. REV. 223, 245 (2011) (describing the U.S. Supreme Court's refusal to place greater restrictions on the factors juries may consider when deciding whether to impose the death sentence).

⁴⁶ See Richard Berk et al., *Statistical Difficulties in Determining the Role of Race in Capital Cases: A Re-Analysis of Data from the State of Maryland*, 21 J. QUANTITATIVE CRIMINOLOGY 365, 387 (2005) [hereinafter Berk et al., *Statistical Difficulties*] ("It is difficult to imagine that a few covariates exist that if included as predictors would lead to clear and justified distinctions between defendants who are charged with a capital crime and defendants who are not; likewise for death sentences. . . . [I]f idiosyncrasies associated with the case, the defendant, or the adjudication process seem to determine a substantial part of the outcome, the adjudication process is suspect whether race is important or not."); *see also* BARRY NAKELL & KENNETH A. HARDY, *THE ARBITRARINESS OF THE DEATH PENALTY* 120, 151 (1987) (noting that legally legitimate characteristics in capital cases only explain a small portion of the variation in charging-and-sentencing decisions); Richard A. Berk et al., *Chance and the Death Penalty*, 27 LAW & SOC'Y REV. 89, 107-09 (1993) [hereinafter Berk et al., *Chance and the Death Penalty*] (same); William J. Bowers & Wanda D. Foglia, *Still Singularly Agonizing: Law's Failure to Purge Arbitrariness from Capital Sentencing*, 39 CRIM. L. BULL. 51, 84 (2003) (demonstrating that constitutionally mandated requirements to guide jury discretion and eliminate arbitrariness in sentencing are not working); Deon Brock et al., *Arbitrariness in the Imposition of Death Sentences in Texas: An Analysis of Four Counties by Offense Seriousness, Race of Victim, and Race of Offender*, 28 AM. J. CRIM. L. 43, 70 (2000) (reporting evidence of significant within-jurisdiction arbitrariness in the imposition of the death penalty after taking into account offense seriousness).

⁴⁷ Bar-Gill & Ben-Shahar, *supra* note 15, at 757.

⁴⁸ Bibas, *supra* note 17, at 2467; Gazal-Ayal, *supra* note 15, at 2321.

⁴⁹ RICHARD C. DIETER, *DEATH PENALTY INFO. CTR., SMART ON CRIME: RECONSIDERING THE DEATH PENALTY IN A TIME OF ECONOMIC CRISIS* 14 (2009).

important because prosecutors are concerned with more than just avoiding trials. They also care about fairness and reputation, and this may lead prosecutors to refuse to adjust plea offers in a manner that would avoid trial, even when prosecutors accurately assess the defendant's reservation price.⁵⁰ Pressure from victims' family and the community may also factor into the minimum plea a prosecutor is willing to offer.⁵¹

Statutory minima in capital cases—typically life with or without the possibility of parole—also preclude prosecutors from offering substantial concessions (at least in the minds of many defendants).⁵² Prosecutors are also less willing to drop or reduce homicide charges because such actions are politically costly, so they often feel bound to dedicate resources to trying homicide cases even with evidentiary problems.⁵³ Even plea offers of a life sentence *with* the possibility of parole may not be considered a meaningful concession relative to life *without* the possibility of parole in the current climate in which determinate-sentencing and truth-in-sentencing laws typically require inmates sentenced to life imprisonment to serve at least twenty-five years, and parole boards are increasingly reluctant to release inmates serving life sentences.⁵⁴ As one scholar has explained, plea

⁵⁰ WILLIAM J. STUNTZ, *THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE* 258 (2011) (“A prosecutor who becomes known as a pushover will be taken advantage of, not once but many times . . . once [a] threat [is] made, it ha[s] to be carried out.”); Josh Bowers, *Punishing the Innocent*, 156 U. PA. L. REV. 1117, 1152–53 (2008).

⁵¹ See, e.g., Kyl et al., *supra* note 43, at 621 (advocating increased victim and victim family involvement in plea negotiations, including the ability to reopen a plea or sentence when the accused has pleaded to a reduced offense).

⁵² See, e.g., WHITE, *LITIGATING IN THE SHADOW OF DEATH*, *supra* note 19, at 158 (describing the “free me or fry me” attitude of many capital defendants); Robert L. Misner, *Recasting Prosecutorial Discretion*, 86 J. CRIM. L. & CRIMINOLOGY 717, 742 (1996) (explaining how the prevalence of statutory minima has dramatically increased the importance of charge bargaining relative to sentence bargaining for serious offenses); see also Lucian E. Dervan, *The Surprising Lessons from Plea Bargaining in the Shadow of Terror*, 27 GA. ST. U. L. REV. 239, 245 (2011) (suggesting that the ability to offer substantial sentence reductions is key to prosecutors' success in the use of plea bargaining).

⁵³ Gazal-Ayal, *supra* note 15, at 2306; Daniel C. Richman & William J. Stuntz, *Al Capone's Revenge: An Essay on the Political Economy of Pretextual Prosecution*, 105 COLUM. L. REV. 583, 600–05 (2005) (arguing that state prosecutors are more often politically obliged to prosecute a defendant for the offense she is suspected of committing, and thus cannot drop the case or offer a charge bargain).

⁵⁴ MARC MAUER ET AL., *THE SENTENCING PROJECT, THE MEANING OF “LIFE”: LONG PRISON SENTENCES IN CONTEXT* 12 (2004) (discussing the steady increase in time actually served for life with the possibility of parole sentences across the nation); Bibas, *supra* note 16, at 1141 (noting trend in truth-in-sentencing laws and parole board practices); Ehrhard, *supra* note 20, at 316 (explaining that the distinction between life with parole and without parole sentences is disappearing in current political climate because parole boards are resistant to granting early release); Press Release, Ga. State Bd. Pardons & Paroles, *More Violent-Crime Lifers Die in Prison Than Are Parole [sic]* (June 1, 1998), *available at*

bargaining draws “sustenance . . . from the *principles* of the indeterminate sentence,”⁵⁵ so statutory minima and de facto determinate sentences (via extremely inactive parole boards) severely restrict the magnitude of a plea discount.

Many capital defendants also suffer from substantial cognitive and emotional deficits and biases, and consequently are more likely to be risk-seeking (or risk-neutral at best) and less likely to accept favorable plea bargains.⁵⁶ As a result, defense counsel in capital cases must be particularly skillful with potential mental health issues affecting their clients.⁵⁷ In fact, a recent study of federal habeas corpus actions discovered that client mental health issues continue to provide a strong basis for relief.⁵⁸ Capital defendants are also more likely to be highly skeptical of their defense attorneys during plea negotiations, especially in situations when the defendant is a racial or ethnic minority and defense counsel is not.⁵⁹

Pretrial detention, court delays, and litigation costs are much less of a concern for capital defendants as well. Defendants charged with murder are the least likely to be released pending trial (19%), have the highest bail amounts (i.e., represent the largest percentage of defendants with bail exceeding \$50,000), and are decreasingly (over the past twenty years) likely to be granted pretrial release.⁶⁰ In many jurisdictions, defendants charged with murder are “non-bondable” either through statute or practice (e.g.,

http://www.pap.state.ga.us/opencms/export/sites/default/press_releases/1998_press_releases/news_0005.html (refuting popular misconception that “straight lifers” are released and reporting that twenty-one “straight lifers” died in prison while serving their sentences in the previous twelve months).

⁵⁵ FISHER, *supra* note 14, at 127.

⁵⁶ See, e.g., WHITE, LITIGATING IN THE SHADOW OF DEATH, *supra* note 19, at 162 (explaining that capital defendants with mental health problems pose a significant obstacle in plea negotiations and often express a “free me or fry me” position to their defense counsel); ALSCHULER, *supra* note 17, at 674 (explaining that defendants are executed “for the crime of being [] optimist[s] . . . [the] inability to think 100 yards in front of [themselves]”); BIBAS, *supra* note 17, at 2467 (highlighting that the combination of poor lawyering and irrational behavior can lead some defendants to reject bargains they should otherwise accept).

⁵⁷ Liebman, *supra* note 22, at 322.

⁵⁸ KING & HOFFMANN, *supra* note 44, at 151 (examining habeas corpus petitions after the enactment of the Anti-Terrorism and Effective Death Penalty Act (AEDPA)).

⁵⁹ See, e.g., WHITE, LITIGATING IN THE SHADOW OF DEATH, *supra* note 19, at 156–57 (describing how defendants’ mistrust of defense counsel in capital cases is often compounded by racial or cultural differences); cf. ALBERT W. ALSCHULER, *The Defense Attorney’s Role in Plea Bargaining*, 84 YALE L.J. 1179, 1180–82 (1975) (claiming that plea bargaining is destructive to attorney–client relationships).

⁶⁰ THOMAS H. COHEN & BRIAN A. REAVES, BUREAU OF JUSTICE STATISTICS, U.S. DEP’T OF JUSTICE, PRETRIAL RELEASE OF FELONY DEFENDANTS IN STATE COURTS 3, 6 (2007); BRIAN A. REAVES & JACOB PEREZ, BUREAU OF JUSTICE STATISTICS, U.S. DEP’T OF JUSTICE, PRETRIAL RELEASE OF FELONY DEFENDANTS, 1992, at 2 (1994).

defendants lack the requisite financial resources to post bail),⁶¹ defense counsel has a strategic incentive to delay trial absent any strong claims of actual innocence,⁶² and the costs associated with defending a capital trial make it virtually impossible for defendants to finance their own defense,⁶³ so the state must almost invariably declare them indigent and appoint counsel at the government's expense.⁶⁴ Local prosecutors also avoid the burden of incurring the full cost of capital trials—particularly the substantial costs of error correction at the appellate level.⁶⁵

Finally, there are a variety of additional reasons why prosecutors and defendants may not reach plea agreements in capital cases. The highly publicized nature of capital cases, relative to noncapital cases, can increase the reluctance of defendants to admit their participation in the crime in open court.⁶⁶ Complicating matters is the fact that many prosecutors and judges (and even some capital statutes) do not permit defendants to enter Alford or *nolo contendere* pleas in capital murder cases, but do allow such pleas in noncapital murder cases.⁶⁷ Prosecutors may also seek the death penalty

⁶¹ COHEN & REAVES, *supra* note 60, at 3–6.

⁶² BARRY LATZER & JAMES N.G. CAUTHEN, JUSTICE DELAYED? TIME CONSUMPTION IN CAPITAL APPEALS: A MULTISTATE STUDY 30 (2007) (remarking on the extraordinary amount of time some defense attorneys take to file a notice of appeal in capital cases); Michael E. Tigar, *Judges, Lawyers and the Penalty of Death*, 23 LOY. L.A. L. REV. 147, 148 (1989) (discussing judges' condemnation of defense counsel's conduct in capital cases, including delay tactics).

⁶³ See *infra* Part VII for a discussion of the specific costs associated with capital trials.

⁶⁴ See, e.g., AM. BAR ASS'N, EVALUATING FAIRNESS AND ACCURACY IN STATE DEATH PENALTY SYSTEMS: THE GEORGIA DEATH PENALTY ASSESSMENT REPORT 143 (2006), available at <http://www.americanbar.org/content/dam/aba/migrated/moratorium/assessment/project/georgia/report.authcheckdam.pdf> (noting that appointment of counsel is required for a defendant indicted for a capital felony if she can establish that she is indigent). To be declared indigent, a person must “lack[] sufficient income or other resources to employ a qualified lawyer to defend him or her without undue hardship on the individual or his or her dependants [sic].” *Id.* at 143 n.73 (citations omitted) (internal quotation marks omitted); cf. Alschuler, *supra* note 17, at 677 (discussing how defense counsel in capital cases often threaten or attempt to make the trial as expensive as possible in order to improve their bargaining positions).

⁶⁵ Arthur L. Alarcón & Paula M. Mitchell, *Executing the Will of the Voters?: A Roadmap to Mend or End the California Legislature's Multi-Billion-Dollar Death Penalty Debacle*, 44 LOY. L.A. L. REV. S41, S88 (2011) (reporting the significant costs associated with federal habeas review of state death sentences in California); Liebman, *supra* note 22, at 325 (explaining that state and federal courts incur the financial burden of correcting trial mistakes); Misner, *supra* note 52, at 719–20 (arguing that our current system is seriously flawed because prosecutors are not required to take into account finite criminal justice resources when making charging-and-sentencing decisions).

⁶⁶ WHITE, LITIGATING IN THE SHADOW OF DEATH, *supra* note 19, at 157.

⁶⁷ An Alford plea is similar to a plea of *nolo contendere* where a defendant pleads guilty to a crime but does not admit guilt. *Alford v. North Carolina*, 400 U.S. 25, 25–26, 37 (1970)

against a particular defendant in an attempt to extract information that may have otherwise been impossible to obtain—e.g., the names of possible accomplices and additional victims—by inducing the defendant’s cooperation in exchange for a sentence reduction. If the defendant refuses to cooperate, the prosecutor may be less likely to offer a favorable plea.⁶⁸ Such cooperation may not be forthcoming in capital cases both because of the high-profile nature of the cases and the severity of the potential penalties for possible codefendants implicated by the defendants.

Due to the influence of these competing incentives on plea-bargaining dynamics in the capital context, coupled with the lower likelihood of pleas in murder cases in general, it is far from obvious what role capital punishment is playing in plea negotiations in the aggregate. The empirical analysis of the influence of the threat of the death penalty on parties’ propensity to reach a plea agreement assists in informing this debate.⁶⁹

II. PRIOR RESEARCH

To date, only one study has attempted to systematically examine the impact of the death penalty on plea bargaining.⁷⁰ Professor Ilyana

(explaining that “[a]n individual accused of crime may voluntarily, knowingly, and understandingly consent to the imposition of a prison sentence even if he is unwilling or unable to admit his participation in the acts constituting the crime” and there is no “material difference between a plea that refuses to admit commission of the criminal act and a plea containing a protestation of innocence when . . . a defendant intelligently concludes that his interests require entry of a guilty plea and the record before the judge contains strong evidence of actual guilt”); see also *Henderson v. Morgan*, 426 U.S. 637, 644–45 (1976) (explaining the constitutional prerequisites for Alford pleas). Georgia’s capital statutes prohibit Alford pleas in capital cases. GA. CODE ANN. § 17-7-95(a) (West 2003).

⁶⁸ See, e.g., David Garland, “Symbolic” and “Instrumental” Aspects of Capital Punishment, in *THE FUTURE OF AMERICA’S DEATH PENALTY*, *supra* note 22, at 421, 437.

⁶⁹ The focus of the empirical analyses in this Article is whether a plea agreement was obtained, not the specific dynamics of the negotiations involved in obtaining a plea. In cases that were ultimately disposed by trial, the data used in these analyses cannot distinguish whether plea negotiations were attempted and, if so, why the negotiation failed. Nevertheless, it is reasonable to assume that in the vast majority of cases, *defendants* rejected the plea offer from the prosecutor because the offer was deemed unacceptable. Although no systematic research on this topic has been conducted, experts estimate between 50% and 75% of inmates on death row rejected plea offers that would have avoided the death sentence. WHITE, *LITIGATING IN THE SHADOW OF DEATH*, *supra* note 19, at 145–46; Alschuler, *supra* note 17, at 671; Stetler, *supra* note 19, at 1157.

According to a recent empirical study of federal postconviction review of capital cases, 3% of defendants sentenced to death plead guilty at the conviction phase without obtaining a sentence bargain from the prosecutor. KING & HOFFMANN, *supra* note 44, at 147. Some defendants may perceive additional benefits from accepting a plea apart from what the government is willing or able to offer, such as leniency from the judge or jury during sentencing. See Dervan, *supra* note 52, at 259.

⁷⁰ Kuziemko, *supra* note 20, at 141 (“This work is, to the best of my knowledge, the first

Kuziemko's study of the impact of the death penalty on plea bargaining in New York State is the first study to examine whether the threat of the death penalty deters capital defendants from taking their cases to trial.⁷¹ After the reinstatement of the death penalty in 1995, several district attorneys from across New York's sixty-two counties publicly announced their refusal to seek the death penalty in death-eligible murder cases (i.e., first-degree murder cases).⁷² The study used the reinstatement of the death penalty in 1995, coupled with these refusals to seek the death penalty, as a "natural experiment to estimate the effect of the death penalty on plea bargaining."⁷³ The study identifies defendants who were arrested for first- or second-degree murder as the "treatment group"—that is, the group of defendants potentially susceptible to the death penalty—and defendants who were arrested for burglary, forcible rape, or armed robbery as the "control group," and compares these groups before and after the policy change.⁷⁴ The study concludes that the threat of the death penalty in New York did *not* increase defendants' overall propensity to plead guilty, but the death penalty did lead defendants to accept plea bargains with harsher terms.⁷⁵

The approach taken in the study was sensible considering the available data on New York's capital charging-and-sentencing process, but several deficiencies with the data likely undermine the substantive conclusions. The first problem is the jurisdiction selected for the study: New York State. While it is understandable that New York was chosen because of the "natural experiment" conditions presented by the reinstatement of the death

to examine the effect of the death penalty on plea bargaining."); *see also* Ehrhard, *supra* note 20, at 315; Scheidegger, *supra* note 20, at 2–3.

⁷¹ *See generally* Kuziemko, *supra* note 20 (studying the effect of this threat on behavior).

⁷² *Id.* at 118, 121; *see also* Al Baker, *Effort to Reinstatement Death Penalty Law Is Stalled in Albany*, N.Y. TIMES, Nov. 18, 2004, at A1.

⁷³ Kuziemko, *supra* note 20, at 118.

⁷⁴ *Id.* at 118, 122. The capital statute enacted in 1995 also expanded the definition of first-degree murder in the state. Before the act, first-degree murder was limited to individuals who were convicted of willfully killing law enforcement officers, and second-degree murder was a catchall category that included a wide range of homicide offenses. In an attempt to take into account the expanded definition of first-degree murder, and therefore keep the treatment group consistent across time, the study aggregated first- and second-degree homicides. *Id.* at 120.

⁷⁵ *Id.* at 140–41. Recognizing the possibly limited generalizability of the New York findings, the study examines a national cross section of murder defendants in 1988 drawn from the thirty-three largest counties in the United States. Results from the national data also corroborate these findings. *Id.* at 135–40. Unfortunately, the national data suffer from important limitations that also limit generalizability (i.e., improper focus on large urban counties and improper identification of treatment and control groups). Stephanie Hindson et al., *Race, Gender, Region and Death Sentencing in Colorado, 1980–1999*, 77 U. COLO. L. REV. 549, 570–77 (2006) (discussing county variation in use of the death penalty within a state).

penalty, New York has a rather inactive death penalty when compared to many other jurisdictions with capital statutes. For example, from 1995 through 2000, district attorneys in New York issued only thirty-six death notices, an average of six notices per year,⁷⁶ although New York averaged nearly 500 first- and second-degree murder arrests each year.⁷⁷ Furthermore, by the year 2000, only six people sat on death row in New York, and no executions have taken place since the death penalty was reinstated.⁷⁸

The second problem is that the study does not directly examine the impact the death penalty has on the bargaining process in cases that are actually noticed for the death penalty. The study posits that the effect of the death penalty may not be limited to defendants who are actually noticed for the death penalty because “its specter may have encouraged some defendants to secure plea bargains after the [district attorney] merely mentioned a death sentence as a possibility but before he actually issued a death notice.”⁷⁹ This assumption is problematic because it misrepresents actual capital charging dynamics. District attorneys do not deem all first-degree murders as worthy of the death penalty.⁸⁰ While the fact that a defendant is charged with first-degree murder under New York’s statute is sufficient to permit the prosecutor to seek the death sentence, the prosecutor’s mere mention of the possibility of seeking the death penalty against a defendant is unlikely to be viewed as a credible threat because the death penalty is used so infrequently.⁸¹ Even in cases that are technically

⁷⁶ Kuziemko, *supra* note 20, at 121.

⁷⁷ *Id.*

⁷⁸ *Id.* (stating the last execution in New York occurred in 1963). Three additional factors may limit the study’s generalizability: (1) the limited scope of New York’s death penalty (i.e., felony murder is ineligible for the death penalty); (2) New York’s provision allowing defendants the right to plead guilty and automatically receive life imprisonment; and (3) the dramatic increase in the quality of defense counsel available to capital defendants resulting from the creation of the state-initiated Capital Defender Office, which was part of the capital punishment statute. *Id.* at 135 & n.14. Furthermore, New York’s highest court invalidated the death penalty statute in 2004, just four years after the time frame the study examines. *See People v. LaValle*, 817 N.E.2d 341 (N.Y. 2004).

⁷⁹ Kuziemko, *supra* note 20, at 133.

⁸⁰ *See* Raymond Paternoster et al., *Justice by Geography and Race: The Administration of the Death Penalty in Maryland, 1978–1999*, 4 U. MD. L.J. RACE RELIGION GENDER & CLASS 1, 17–18 (2004) (explaining that different prosecutors in different offices, or even within the same office, may not evaluate a case as death eligible, even when the case satisfies all of the statutory criteria for death eligibility); Glenn L. Pierce & Michael L. Radelet, *Race, Region, and Death Sentencing in Illinois, 1988–1997*, 81 OR. L. REV. 39, 41, 46 (2002) (describing the same phenomenon).

⁸¹ This observation is underscored by the fact that prosecutors only formally sought the death penalty thirty-six times over a six-year period. *See, e.g., Bar-Gill & Ben-Shahar, supra* note 15, at 738 (underscoring the importance of “credible threats” by prosecutors in

death eligible, it is more plausible that the defendant would wait until the prosecutor actually carried out her threat to seek the death penalty before agreeing to a plea bargain. In fact, defense counsel, who is typically a repeat player with the prosecutor, is likely to advise her client as to whether the prosecutor's threat is credible.⁸²

The study also implausibly assumes, without evidence, that district attorneys (or defense counsel) actually mention the possibility of the death sentence in all cases where an indictment for first-degree murder is obtained. This is especially unlikely because capital defendants are permitted to accept plea bargains allowing them to avoid the death sentence after their cases are noticed for the death penalty, but they do not enjoy a constitutional right to withdraw guilty pleas in capital cases when the pleas were made voluntarily and knowingly.⁸³ Nearly all death penalty jurisdictions (including New York) permit capital defendants to accept plea bargains up until the penalty phase of their capital trials.⁸⁴ As mentioned *supra*, the majority of the thirty-six death notices in the study were withdrawn after plea bargains were reached. Therefore, if the assumption that the mere possibility of a death notice influences defendants' plea calculus is incorrect, then the study's treatment group includes defendants who are not impacted by the possible threat of the death penalty and the treatment effect will be biased towards zero because of systematic measurement error. Stated differently, the treatment group will actually contain individuals who should be either in the control group or entirely excluded from the analysis.

The New York State study defends this approach by characterizing its estimates as "intent to treat" (ITT) measures,⁸⁵ rather than the conventional estimates of "average treatment effect" (ATE) or average treatment effect

the plea-bargaining context).

⁸² STUNTZ, *supra* note 50, at 258 ("Plea bargaining is what academics call a 'repeat-play' game; the same lawyers negotiate pleas again and again.").

⁸³ Compare *Santobello v. New York*, 404 U.S. 257, 260 (1971) (explaining that a defendant does not have a constitutional right to withdraw a guilty plea in all circumstances), with *Fair v. State*, 268 S.E.2d 316, 323–24 (Ga. 1980) (holding that a defendant has no constitutional right to withdraw a permissible guilty plea in a capital case and that motions to withdraw a plea must be assessed on a case-by-case basis).

⁸⁴ DAVID C. BALDUS ET AL., *EQUAL JUSTICE AND THE DEATH PENALTY: A LEGAL AND EMPIRICAL ANALYSIS* 23 (1990) (discussing bifurcated death penalty trials post-*Furman*); Kuziemko, *supra* note 20, at 119–20 (explaining that capital trials are conducted in two parts: in the first phase, the court is only concerned with the question of guilt; if the defendant is convicted of capital murder in the first phase, the same jury proceeds to the sentencing phase of the trial to determine the appropriate punishment); see also *infra* Part III.A (discussing Georgia's capital charging-and-sentencing process and noting that defendants are allowed to accept pleas up until the penalty phase).

⁸⁵ Kuziemko, *supra* note 20, at 133.

on the treated (ATT).⁸⁶ “The ITT is one of a number of possible parameters of interest and may not always be of greatest scientific or policy relevance.”⁸⁷ It measures the effect of treatment assignment, but not the effect of the treatment itself.⁸⁸ But it is questionable that the study’s estimates can be accurately characterized as “intent to treat” estimates because of the reasons previously stated: (1) prosecutors do not deem all death-eligible cases as “death worthy,” (2) not all death-eligible defendants (or their counsel) believe their cases are at risk for the death penalty, and (3) prosecutors do not “threaten” the death penalty in all (or even the vast majority) of death-eligible cases. Furthermore, the study does not alleviate this concern because it does not distinguish cases in which a threat was ever mentioned from the larger group of defendants indicted for first-degree murder.⁸⁹ Problems of systematic measurement error notwithstanding,⁹⁰ scholars have repeatedly remarked that the ATT is both easier to identify and likely to be more theoretically informative because it describes the impact of the treatment only among the units who are actually exposed to it.⁹¹ Formally, assuming some selection on observables, $ATT = E[Y_{(1)} - Y_{(0)} | X, T = 1]$; where $Y_{(1)}$ and $Y_{(0)}$ denote the two potential outcomes under treatment and control conditions, respectively, X indicates observable

⁸⁶ Michael E. Sobel, *Causal Inference in Randomized and Non-Randomized Studies: The Definition, Identification, and Estimation of Causal Parameters*, in THE SAGE HANDBOOK OF QUANTITATIVE METHODS IN PSYCHOLOGY 3, 7–8 (Roger E. Millsap & Alberto Maydeu-Olivares eds., 2009) (defining ITT, ATE, and ATT causal effects).

⁸⁷ *Id.* at 3, 7. The ITT, ATE, and ATT address different causal questions: (1) ITT measures the average effect of the treatment per person offered the treatment, irrespective of how many treatment group members actually received it; (2) ATE measures the average effect of the treatment if all members in the treatment group actually received the treatment; and (3) ATT measures the average treatment effect per person receiving the treatment. Lisa A. Gennetian et al., *Constructing Instrumental Variables from Experimental Data to Explore How Treatments Produce Effects*, in LEARNING MORE FROM SOCIAL EXPERIMENTS: EVOLVING ANALYTIC APPROACHES 75, 86–87 (Howard S. Bloom ed., 2006).

⁸⁸ Sobel, *supra* note 87, at 7. It is also important to emphasize that the intent to treat effect “is commonly featured in connection with randomized clinical trials,” in order to justify the assumption that the treatment effect is identifiable. *Id.*

⁸⁹ *But see* Els Goetghebeur & Tom Loeys, *Beyond Intention to Treat*, 24 EPIDEMIOLOGIC REVS. 85, 85 (2002) (arguing that the “upside” of noncompliance is that it more closely resembles the heterogeneous population of future treatment groups).

⁹⁰ *See, e.g.*, Goetghebeur & Loeys, *supra* note 89, at 89 (“The more we seek to tailor possibly dynamic treatments to individual characteristics . . . the more imperative it becomes to acknowledge treatment actually received as an important source of variation in treatment effect.”).

⁹¹ *See* STEPHEN MORGAN & CHRISTOPHER WINSHIP, COUNTERFACTUALS AND CAUSAL INFERENCE 43 (2007) (“[T]he average treatment effect among the treated is a theoretically important quantity”); Markus Gangl, *Causal Inference in Sociological Research*, 36 ANN. REV. SOC. 21, 24 (2010) (same).

covariates in the model, and T is an indicator variable for treatment assignment.⁹²

The New York State study should be applauded for highlighting this gap in the literature and making an important empirical contribution to the existing debate, although the limitations of the data partly undermine the reliability and generalizability of the conclusions that were reached.⁹³ A significant improvement on that seminal study would entail several modifications. First, it would examine a jurisdiction that more actively pursues the death penalty with respect to both sentencing individuals to death and carrying out executions.⁹⁴ Second, the study would directly examine actual treatment effects and not merely “intent to treat” effects. In fact, prior research has appropriately defined the “treatment effect” of the death penalty as the government’s “decision to file a death notice that formally announces [the] state’s intention to seek a death sentence.”⁹⁵ Finally, it would control for a wider range of case factors relevant to charging and plea-bargaining decisions.⁹⁶ The present study incorporates all of these improvements by analyzing a rich data set from Georgia, which is better suited to test this hypothesis.

⁹² MORGAN & WINSHIP, *supra* note 91, at 42; *see also* Donald B. Rubin, *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies*, 66 J. EDUC. PSYCHOL. 688, 689–90 (1974).

⁹³ *See infra* Part V.

⁹⁴ It is important to make a distinction between those jurisdictions that actively pursue the death penalty but infrequently execute individuals and those that actively pursue the death penalty and execute defendants. California, for example, has nearly double the death row population of any other state (721 death row inmates as of January 1, 2011), but rarely executes inmates who have been sentenced to death. CAL. COMM’N ON THE FAIR ADMIN. OF JUSTICE, REPORT AND RECOMMENDATIONS ON THE ADMINISTRATION OF THE DEATH PENALTY IN CALIFORNIA 21–27 (2008); DEBORAH FINS, NAACP LEGAL DEF. & EDUC. FUND, DEATH ROW U.S.A. 34, 39–45 (2011).

⁹⁵ *See* Roman et al., *supra* note 27, at 533.

⁹⁶ The New York study examines the following case-level factors: defendant’s sex, race/ethnicity, age, number of prior convictions, county of arrest, original charge, and plea charge. Kuziemko, *supra* note 20, at 129. The New York study fails to explore victim characteristics, contemporaneous convictions, and specific case factors identified in the statute that make a case a death-eligible crime. The supplemental national cross-section analysis includes both offender and victim demographic information, but lacks legally relevant, case-specific information outside of the number of prior convictions. *Id.* at 136. Taking into account more case information also permits the analyst to determine whether the assumptions of the “natural experiment” actually hold—that is, whether the treatment and control groups are truly indistinguishable except for the intervention. Richard A. Berk, *Randomized Experiments as the Bronze Standard*, 1 J. EXPERIMENTAL CRIMINOLOGY 417, 421, 428 (2005).

III. GEORGIA'S DEATH PENALTY

Prosecutors in Georgia have aggressively sought the death penalty since the practice first began in the colonies in 1608—Georgia ranks fifth in executions carried out since that time (1,002).⁹⁷ Georgia also ranks seventh in the nation with respect to total executions since the death penalty was reinstated (52) and ninth in terms of its death row population (102).⁹⁸ At the time of this writing, the most recent execution carried out in Georgia was that of Andrew Cook on February 21, 2013.⁹⁹ Georgia has also been the most influential state in shaping national death penalty policy in the modern era of capital punishment.¹⁰⁰ No less than seventeen cases originating in Georgia have set legal precedent with respect to the administration of capital punishment, including *Furman v. Georgia*¹⁰¹ and *Gregg v. Georgia*,¹⁰² which, respectively, were responsible for placing and lifting the moratorium on executions in the United States in the 1970s.¹⁰³ The following section provides a brief history of Georgia's current capital statute.

A. HISTORY AND DESCRIPTION OF THE MODERN STATUTE

On June 29, 1972, in *Furman v. Georgia*,¹⁰⁴ the U.S. Supreme Court invalidated Georgia's death penalty statute, ruling that the lack of

⁹⁷ M. WATT ESPY & JOHN ORTIZ SMYKLA, EXECUTIONS IN THE UNITED STATES, 1608–2002: THE ESPY FILE 2 (ICPSR 4th ed. 2004); DEATH PENALTY INFO. CTR., FACTS ABOUT THE DEATH PENALTY (2012), available at <http://www.deathpenaltyinfo.org/FactSheet.pdf> (providing information about executions in Georgia from 2003 through the present).

⁹⁸ DEATH PENALTY INFO. CTR., *supra* note 97, at 2–3.

⁹⁹ *Id.*; Rhonda Cook, *Executed Man Makes Apology*, ATLANTA J.-CONST., Feb. 23, 2013, at B1.

¹⁰⁰ BALDUS ET AL., *supra* note 84, at 3.

¹⁰¹ 408 U.S. 238, 239–40 (1972) (per curiam) (holding that unguided-discretion death penalty statutes are unconstitutional).

¹⁰² 428 U.S. 153, 207 (1976) (holding that guided-discretion capital statutes are constitutionally permissible).

¹⁰³ Rhonda Cook, *Georgia Cases Have Set Legal Precedent*, ATLANTA J.-CONST., June 2, 1996, at C4. Other significant cases include *McCleskey v. Kemp*, 481 U.S. 279, 319 (1987) (finding statistical evidence of system-wide discrimination irrelevant; the defendant must show discrimination in the case at hand); *Francis v. Franklin*, 471 U.S. 307, 324–25 (1985) (holding that capital defendants enjoy the presumption of innocence and do not have the burden of proof in capital cases); *Zant v. Stephens*, 462 U.S. 862, 891 (1983) (holding that death penalty may be imposed as long as the jury finds at least one valid aggravating circumstance, and nonstatutory aggravating circumstances may also be considered by the jury when deciding whether to sentence a defendant to death); *Eberheart v. Georgia*, 433 U.S. 917, 917 (1977) (declaring death penalty judgment for nonhomicidal kidnapping with bodily injury unconstitutional); and *Coker v. Georgia*, 433 U.S. 584, 592 (1977) (declaring death penalty for nonhomicidal rape of an adult unconstitutional).

¹⁰⁴ 408 U.S. 238.

sentencing guidance for capital jurors was a violation of the Eighth Amendment, and subsequently invalidated all capital statutes that lacked such standards.¹⁰⁵ The Georgia General Assembly quickly drafted new death penalty legislation the following January and the bill was signed into law by then-Governor Jimmy Carter on March 28, 1973.¹⁰⁶ The legislation was soon codified; it provided for eleven separate instances where the death penalty could be imposed on someone convicted of a capital offense.¹⁰⁷ Less than a year later, Troy Leon Gregg was convicted of murder and armed robbery and sentenced to death under Georgia's new death penalty statute. The U.S. Supreme Court would eventually grant certiorari in Gregg's¹⁰⁸ case (consolidated with two other death penalty cases from Florida¹⁰⁹ and Texas¹¹⁰) and ultimately rule that Georgia's new death penalty statute was constitutionally acceptable.¹¹¹ The Court's decision officially ended the nation's four-year moratorium on the death penalty.¹¹²

¹⁰⁵ *Id.* During the previous year, the Supreme Court ruled by a six-to-three vote that neither the absence of sentencing guidelines nor single-verdict procedures violated the Fourteenth Amendment. *McGautha v. California*, 402 U.S. 183, 196–208, 210–13 (1971). The *Furman* ruling, which was decided by a five-to-four vote, was particularly surprising because the composition of the Court had not changed from the *McGautha* ruling. HERBERT H. HAINES, *AGAINST CAPITAL PUNISHMENT: THE ANTI-DEATH PENALTY MOVEMENT IN AMERICA, 1972–1994*, at 37–39 (1996).

The immediate effect of *Furman* was that approximately 558 death row inmates had their sentences commuted to life imprisonment. Although there was widespread speculation by death penalty proponents that many of these inmates would kill again once released from prison, subsequent research revealed that only one of the 239 *Furman*-commuted inmates released from prison committed a second murder in the fifteen years following the *Furman* decision. James W. Marquart & Jonathan R. Sorensen, *A National Study of the Furman-Commuted Inmates: Assessing the Threat to Society from Capital Offenders*, 23 *LOY. L.A. L. REV.* 5, 23–24 (1989).

¹⁰⁶ MICHAEL MEARS, *THE DEATH PENALTY IN GEORGIA: A MODERN HISTORY, 1970–2000*, at 14–41 (1999) (providing a detailed history of the modern death penalty in Georgia). On December 8, 1972, less than six months after *Furman*, Florida became the first state to officially restore capital punishment when Governor Reubin Askew signed new death penalty legislation into law. HAINES, *supra* note 105, at 45.

¹⁰⁷ 1973 Ga. Laws 163–65, § 3 (originally codified at GA. CODE ANN. § 27-2534.1 (1973); current version at GA. CODE ANN. § 17-10-30 (West 2003)).

¹⁰⁸ *Gregg v. Georgia*, 428 U.S. 153 (1976).

¹⁰⁹ *Proffitt v. Florida*, 428 U.S. 242 (1976).

¹¹⁰ *Jurek v. Texas*, 428 U.S. 262 (1976).

¹¹¹ *Gregg*, 428 U.S. at 169, 195 (holding, seven-to-two, that the death penalty for murder did not by itself violate the Eighth Amendment and all three of the capital statutes contained sufficient procedural reforms to warrant them constitutional under *Furman*); *see also* HAINES, *supra* note 105, at 52–54. Interestingly, the court had no evidence suggesting that the new statutes eliminated arbitrariness and bias in capital sentencing; rather, the court based its decision on whether the procedural reforms enacted in each statute were capable of producing outcomes different from those produced under the pre-*Furman* statutes.

¹¹² MEARS, *supra* note 106, at 65–69.

As a result of the decision, thirty-four states, the federal government, and the U.S. Armed Forces currently permit the death penalty in their jurisdictions.¹¹³

As noted *supra*, Georgia's new death penalty was originally enacted in 1973 and enumerated eleven separate instances where the death penalty could be imposed on someone convicted of a capital offense:

- (a) The death penalty may be imposed for the offenses of aircraft hijacking or treason, in any case.
- (b) In all cases of other offenses for which the death penalty may be authorized, the judge shall consider, or he shall include in his instructions to the jury for it to consider, any mitigating circumstances or aggravating circumstances otherwise authorized by law and any of the following statutory aggravating circumstances which may be supported by the evidence:
 - (1) The offense of murder, rape, armed robbery, or kidnapping was committed by a person with a prior record of conviction for a capital felony, *or the offense of murder was committed by a person who has a substantial history of serious assaultive criminal convictions.*¹¹⁴
 - (2) The offense of murder, rape, armed robbery, or kidnapping was committed while the offender was engaged in the commission of another capital felony, or aggravated battery, or the offense of murder was committed while the offender was engaged in the commission of burglary or arson in the first degree.
 - (3) The offender by his act of murder, armed robbery, or kidnapping knowingly created a great risk of death to more than one person in a public place by means of a weapon or device which would normally be hazardous to the lives of more than one person.
 - (4) The offender committed the offense of murder for himself or another, for the purpose of receiving money or any other thing of monetary value.
 - (5) The murder of a judicial officer, former judicial officer, district attorney or solicitor or former district attorney or solicitor during or because of the exercise of his official duty.
 - (6) The offender caused or directed another to commit murder or committed murder as an agent or employee of another person.
 - (7) The offense of murder, rape, armed robbery, or kidnapping was outrageously or wantonly vile, horrible or inhuman in that it involved torture, depravity of mind, or an aggravated battery to the victim.
 - (8) The offense of murder was committed against any peace officer, corrections employee or fireman while engaged in the performance of his official duties.
 - (9) The offense of murder was committed by a person in, or who has escaped from, the lawful custody of a peace officer or place of lawful confinement.

¹¹³ DEATH PENALTY INFO. CTR., *supra* note 97.

¹¹⁴ The italicized portion of the death penalty statute was subsequently declared unconstitutionally vague. *Arnold v. State*, 224 S.E.2d 386, 392 (Ga. 1976).

(10) The murder was committed for the purpose of avoiding, interfering with, or preventing a lawful arrest or custody in a place of lawful confinement, of himself or another.

(c) The statutory instructions as determined by the trial judge to be warranted by the evidence shall be given in charge and in writing to the jury for its deliberation. The jury, if its verdict be a recommendation of death, shall designate in writing, signed by the foreman of the jury, the aggravating circumstance or circumstances which it found beyond a reasonable doubt. In non-jury cases the judge shall make such designation. Except in cases of treason or aircraft hijacking, unless at least one of the statutory aggravating circumstances enumerated in Code section 27-2434.1(b) is so found, the death penalty shall not be imposed.¹¹⁵

With very few changes, Georgia's death penalty legislation has remained in place since Governor Jimmy Carter first signed it into law;¹¹⁶ however, there were several changes mandated by subsequent U.S. Supreme Court rulings. The year after the Court officially reinstated Georgia's death penalty in *Gregg*, it invalidated the death penalty for defendants convicted of non-homicidal rape and kidnapping with bodily injury in, respectively, *Coker v. Georgia* and *Eberheart v. Georgia*.¹¹⁷ Georgia's juvenile death penalty was also invalidated following the Court's decision in *Roper v. Simmons*,¹¹⁸ which forbade the death penalty for defendants who were under the age of eighteen during the commission of their crime. The Georgia statute had permitted the death penalty for defendants who were seventeen at the time of their crime.¹¹⁹

¹¹⁵ 1973 Ga. Laws 163–65, § 3 (emphasis added) (originally codified at GA. CODE ANN. § 27-2534.1 (1973); current version at GA. CODE ANN. § 17-10-30 (West 2003)). Throughout the remainder of this Article, the specific elements of the capital statute listed in subsection (b) of Georgia's capital statute will be referred to as *B1*, *B2*, *B3*, etc.

¹¹⁶ In 1996 and 1997, there were two unsuccessful proposals to lower the age of eligibility for the death penalty to sixteen. There were also two attempts to add an additional aggravating circumstance that would allow the death penalty in the event a person was convicted of the rape of a child under the age of twelve; however, this legislation was also unsuccessful. MEARS, *supra* note 106, at 46. In 2006, an additional element, *B11*, was added: "The offense of murder, rape, or kidnapping was committed by a person previously convicted of rape, aggravated sodomy, aggravated child molestation, or aggravated sexual battery." GA. CODE ANN. § 17-10-30(b)(11) (West Supp. 2012).

¹¹⁷ See *Eberheart v. Georgia*, 433 U.S. 917 (1977); *Coker v. Georgia*, 433 U.S. 584 (1977). The current Georgia statute permits the death penalty for murder (malice or felony), aircraft hijacking, and treason. See *Collins v. State*, 236 S.E.2d 759, 762 (Ga. 1977) (Jordan, J., concurring) ("Of course the crimes of treason and aircraft hijacking, along with murder, remain capital felonies . . .").

¹¹⁸ 543 U.S. 551 (2005).

¹¹⁹ DEATH PENALTY INFO. CTR., *supra* note 97. For a critical treatment of the U.S. Supreme Court's rulings in *Coker* and *Roper*, see generally Adam S. Goldstone, *The Death Penalty: How America's Highest Court Is Narrowing Its Application*, 4 CRIM. L. BRIEF 23 (2009) (arguing that the *Coker* and *Roper* decisions are examples of judicial activism and inappropriately limit the application of the death penalty).

B. LIFE WITHOUT THE POSSIBILITY OF PAROLE

Georgia's life sentence without the possibility of parole statute was enacted in May 1993 and allowed juries to deny parole to defendants convicted of certain high felonies, including murder.¹²⁰ Presently, all thirty-four states that authorize the death penalty have enacted similar legislation.¹²¹ Georgia's LWOP legislation may have significantly altered the administration of capital punishment in Georgia by restricting LWOP to murder cases in which the prosecution has filed notices of intent to seek the death penalty.¹²² As a result, prosecutors may seek the death sentence in cases they do not believe are deserving of the death penalty, but they do believe warrant LWOP.¹²³ For example, in 2001, Devonia "Eddie" Inman was convicted in Adel, Georgia, of the murder of Donna Brown and sentenced to life without the possibility of parole.¹²⁴ In commenting on the verdict and sentence, Alapaha Judicial Circuit District Attorney Bob Ellis remarked, "Had we not sought the death penalty, we could have not gotten life without parole."¹²⁵ Ellis further explained that by seeking the death penalty, he gave the jury the opportunity to deny parole to Inman.¹²⁶

Even in jurisdictions that do not restrict the LWOP sentencing option to death cases, it is unlikely that a defendant charged with a death-eligible homicide offense would agree to plea to an LWOP sentence without the threat of a death sentence at trial. Absent the risk of a death sentence at trial, the defendant could do no worse at trial than the LWOP plea offered by the prosecutor. The defendant would be better off taking her chances,

¹²⁰ 1993 Ga. Laws 1656–57, § 4 (codified at GA. CODE ANN. § 17-10-30.1 (repealed 2009)).

¹²¹ DEATH PENALTY INFO. CTR., *supra* note 97.

¹²² In 2009, after two failed attempts, the Georgia General Assembly enacted legislation permitting the imposition of life without the possibility of parole in murder cases, independent of a death penalty prosecution. 2009 Ga. Laws 227, § 10; *see also* H.R. 142, § 17, 151st Gen. Assemb., Reg. Sess. (Ga. 2011) (codifying the 2009 law at GA. CODE ANN. § 17-10-16.1 (West Supp. 2012)).

¹²³ *See, e.g.,* Tony Perry, *Drug Lord Avoids Death Penalty with Plea Deal: Arellano Felix Pleads Guilty to Charges that Will Put Him in Prison for Life Without the Possibility of Parole*, L.A. TIMES, Sept. 18, 2007, at B1 (discussing defendant agreeing to plea to life without the possibility of parole in exchange for withdrawal of death penalty); Gene Johnson, *Strategy Changing on Death Penalty*, NEWS TRIB. (Tacoma, Wash.) (July 30, 2007, 1:00 AM), <http://www.thenewstribune.com/2007/07/30/121534/strategy-changing-on-death-penalty.html> (quoting a prosecutor explaining that the threat of the death penalty is the only leverage available in some cases).

¹²⁴ Peter Failor, *Man Gets Life for 1998 Adel Murder*, VALDOSTA DAILY TIMES, June 28, 2001, at A1.

¹²⁵ *Id.*

¹²⁶ *Id.*

however slim, at trial for the possibility of receiving a lesser sentence.¹²⁷ This is particularly true in jurisdictions such as Georgia that permit judges and juries to impose a life with the possibility of parole sentence for anyone convicted of murder or felony murder.¹²⁸ In Georgia, the defendant and prosecutor may enter into a plea agreement at any time up until the jury renders its sentence in the penalty phase. The following section briefly outlines the major stages of a case from indictment through the penalty phase that differentiate capital from noncapital cases.

C. CAPITAL CASE PROGRESSION¹²⁹

Indictment Through Arraignment. Georgia is an indictment jurisdiction, so a grand jury is required to decide formally whether there is probable cause to believe that the accused has committed the specified crime.¹³⁰ Following the indictment, the accused may be eligible for the appointment of counsel.¹³¹ If deemed eligible for appointed counsel and the charge is a capital felony, two attorneys must be appointed before the accused is called upon to plea to the charges, which generally occurs at the arraignment. Prior to arraignment, a pretrial conference is held and the prosecuting attorney must announce whether she intends to seek the death penalty and then file a notice of intent with the clerk of the superior court.¹³²

¹²⁷ Ehrhard, *supra* note 20, at 313 (summarizing statements from prosecutors explaining that the death penalty is often the only leverage they have in plea negotiations in murder cases). Indeed, the likelihood of receiving a straight life sentence at trial in a capital murder case in Georgia does not appear to be particularly slim. From 1993 to 2000, 31% of capital cases disposed by trial resulted in straight life sentences, whereas 36% received LWOP and 32% received the death sentence.

¹²⁸ GA. CODE ANN. § 17-10-30 (West 2003). In economics parlance, there is no “price” or “penalty” associated with a defendant invoking her constitutional right to trial under the Georgia regime in place at the time of this study if the prosecutor only offers LWOP in a noncapital case. To be sure, individuals who finance their own defenses will incur those costs, but the vast majority of murder defendants are represented by court-appointed counsel. *See generally* Beck & Shumsky, *supra* note 34, at 525; Tabak & Lane, *supra* note 34, at 59.

¹²⁹ In the interest of space, only the most relevant stages of the “typical” progression of a Georgia death penalty case through automatic appeal are described. The qualifier “typical” is used because there are numerous factors that may cause a case to deviate from this abbreviated description.

¹³⁰ GA. CODE ANN. §§ 15-12-60 *et seq.*

¹³¹ Upon a showing of indigence, an individual indicted for a capital felony is eligible for appointed counsel. Pursuant to the Georgia Indigent Defense Act of 2003 (GIDA), an indigent is “[a] person charged with a . . . [crime] punishable by imprisonment who earns less than 100 percent of the federal poverty guidelines unless there is evidence that the person has other resources that might reasonably be used to employ a lawyer without undue hardship on the person or his or her dependents.” GA. CODE ANN. § 17-12-2 (West 2003 & Supp. 2012); *see also* 2003 Ga. Laws 192–217, § 1.

¹³² UNIF. APP. R. IIC(1). The specific aggravating circumstances the government intends

The superior court must then transmit the notice to the clerk of the Supreme Court of Georgia.¹³³ During the arraignment, the court must read the indictment and ask the defendant to plead to the capital felony and any lesser-included offenses charged. The defendant is allowed to plead guilty, not guilty, or mentally incompetent to stand trial.¹³⁴

Capital Trial. The court must empanel forty-two prospective jurors from whom the state and defense must select a total of twelve jurors and one or more alternative jurors, if deemed necessary by the judge.¹³⁵ All capital cases are heard before the Georgia Superior Court¹³⁶ and conducted in two phases: the conviction phase (also commonly referred to as the guilt/innocence phase) and, if the defendant is found guilty of a capital felony, the penalty phase. Immediately prior to the conviction phase, the court must conduct a conference with the state, defense counsel, and the defendant to resolve several matters, including, *inter alia*, any last-minute motions, stipulations, and objections to defense counsel.¹³⁷

In situations where the defendant is found guilty of capital murder at the conclusion of the conviction phase, the case proceeds to the penalty phase (i.e., a presentencing hearing) where both the prosecutor and defense counsel may present witnesses and evidence regarding the statutory aggravating circumstances, as well as nonstatutory aggravating and mitigating circumstances.¹³⁸ The jury may sentence the defendant to death only if they find one or more statutory aggravating circumstances beyond a reasonable doubt, but a death sentence is never required.¹³⁹ Following a conviction for a capital felony and a sentence of death, the defendant may challenge her conviction or death sentence by: (1) filing a motion for a new trial with the superior court, or (2) filing a direct appeal with the Georgia

to prove at trial need not be included in the indictment. *See generally* Lewis v. State, 620 S.E.2d 778 (Ga. 2005).

¹³³ UNIF. APP. R. IIC(1).

¹³⁴ A defendant indicted for a capital felony may not plead *nolo contendere*. GA. CODE ANN. § 17-7-95(a).

¹³⁵ *Id.* §§ 15-12-160, -168.

¹³⁶ GA. CONST. art. 6, § 4.

¹³⁷ UNIF. APP. R. IIIA(1).

¹³⁸ GA. CODE ANN. § 17-10-2(c); *see also* Zant v. Stephens, 462 U.S. 862, 891 (1983) (holding that nonstatutory aggravating circumstances may be considered by the jury during the presentence hearing).

¹³⁹ GA. CODE ANN. § 17-10-31.1(c). Prior to *Ring v. Arizona*, 536 U.S. 584 (2002), which prohibited judge-imposed death sentences, judges in Georgia were permitted to sentence a capital defendant to death if the defendant waived her right to a jury at the penalty phase. This situation, however, never occurred in the years under investigation for this study.

Supreme Court.¹⁴⁰ If the defendant does not initiate any sort of review, the case will automatically be appealed to the Georgia Supreme Court within ten days of the filing of the trial transcript by the court reporter of the Georgia Superior Court.¹⁴¹ This automatic review will occur even if the defendant does not wish to appeal her conviction or sentence.¹⁴²

It is worth emphasizing that it is very common for multiple years to elapse between the initial arrest and the trial in death penalty cases, so both parties have ample time to negotiate a plea agreement. The data analyzed for this study¹⁴³ reveal that the average time between arrest and sentencing for defendants noticed for the death penalty and opting for trial was 31.9 months. The average time for defendants noticed for the death penalty but ultimately pleading at some point before the penalty phase was approximately 24.6 months. For defendants noticed for the death penalty, irrespective of whether they pleaded or took their cases to trial, the average amount of time between arrest and sentencing was 27.9 months. By comparison, the average time between arrest and sentencing for death-eligible defendants not facing the death penalty but ultimately convicted of murder was 17.4 months for those opting for trial and 13.5 months for those who pleaded (and 15.9 months irrespective of method of disposition). The specific data collected in Georgia and analyzed in this Article are discussed in the next section.

IV. DATA

The current study analyzes eight years of death penalty charging-and-sentencing data from Georgia (1993–2000) and addresses each of the aforementioned modifications in an effort to assess more accurately the impact of the death penalty on the plea-bargaining process.¹⁴⁴ The data used for these analyses were collected from the Georgia Bureau of Investigation (GBI), the Georgia Department of Corrections (GDC), the Office of the Georgia Capital Defender (GCD), the Clerk's Office of the

¹⁴⁰ GA. CODE ANN. § 17-10-35; UNIF. APP. R. IVA(1)–(2).

¹⁴¹ GA. CODE ANN. § 17-10-35.

¹⁴² *Id.*

¹⁴³ *See infra* Part IV.

¹⁴⁴ This particular time frame was selected for two important reasons. First, Georgia's life without the possibility of parole legislation was enacted in 1993. The legislation was specifically designed as a sentencing alternative in capital murder trials, therefore potentially having a significant impact on prosecutorial discretion. GA. CODE ANN. § 17-10-30.1 (repealed 2009); 1993 Ga. Laws 1656–57, § 4. Second, in October 1992, the Georgia General Assembly, along with the Georgia Supreme Court, established a statewide agency to actively monitor all death penalty cases in Georgia's 159 counties. MEARS, *supra* note 106, at 4. For a discussion of additional advantages of examining Georgia's capital charging-and-sentencing process, see BALDUS ET AL., *supra* note 84, at 3.

Georgia Supreme Court (CO), the *Atlanta Journal-Constitution* (AJC),¹⁴⁵ and the U.S. Census Bureau.¹⁴⁶ These data contain detailed information on each homicide case in Georgia with respect to the defendant, codefendant(s), victim(s), judge, prosecutor, defense counsel, and the crime. As mentioned *supra*,¹⁴⁷ the Georgia death penalty statute lists eleven elements making a crime eligible for the death penalty.¹⁴⁸ These data allow for the determination of which defendants are actually eligible for the death penalty in Georgia. Information obtained concerning prosecutors' actual decisions to seek the death penalty in each case permit the examination of a *genuine* treatment effect.¹⁴⁹ The more nuanced data also permit the inclusion of a much richer set of statistical controls than were considered in prior research.

During the period under investigation (1993–2000), prosecutors filed a notice of intent to seek the death penalty in 400 cases and fifty-four defendants ultimately received the death penalty.¹⁵⁰ Of the 395 capitally charged cases in which the method of disposition is known, 59% (234) were resolved by plea and 41% (161) were resolved by trial. With respect to cases that were technically death eligible under the Georgia statute but in which the prosecutor declined to seek the death penalty, 39% (350) were disposed by plea and 61% (551) disposed by trial. Of the 724 cases resulting in a murder conviction that were ineligible for the death penalty, 30% (212) were disposed of by plea and 70% (505) by trial.¹⁵¹ The plea rate for cases noticed for the death penalty ranged from 38.7% (in 1998) to

¹⁴⁵ Bill Rankin et al., *A Matter of Life or Death: An AJC Special Report: High Court Botched Death Reviews*, ATLANTA J.-CONST., Sept. 26, 2007, at A1 (discussing a collection of data on 2,328 murder convictions in Georgia between 1995 and 2004); Raymond Paternoster, *The Death Penalty in Georgia, 1995–2004* (Sept. 17, 2007) (unpublished report) (on file with the Journal of Criminal Law and Criminology).

¹⁴⁶ See *infra* Appendix A.

¹⁴⁷ *Supra* Part III.A.

¹⁴⁸ The original statute permitted the death penalty for the crimes of murder, rape, armed robbery, or kidnapping with bodily injury, but rape, armed robbery, and kidnapping with bodily injury were removed as death-eligible offenses following the Supreme Court's decisions in *Coker v. Georgia*, 433 U.S. 584, 584 (1977) (declaring the death penalty for rape unconstitutional), and *Eberheart v. Georgia*, 433 U.S. 917, 917 (1977) (holding that the death penalty for armed robbery was unconstitutional). See *supra* note 103.

¹⁴⁹ See Gangl, *supra* note 91, at 24.

¹⁵⁰ These 400 death notices and the fifty-four death sentences are with respect to incident dates, not disposition dates. Seven of these sentences (12.9%) were imposed after 2000.

¹⁵¹ Forty-five percent (584) of all death-eligible cases were disposed by plea. Information concerning the method of disposition was missing for seventeen cases in the sample (0.8%). A total of 1,628 non-capitally-charged murder convictions were obtained in the period under study—34% were disposed by plea. In these data, there are twenty-eight cases in which the defendant was initially noticed for the death penalty but was ultimately acquitted, had charges dismissed, or was convicted of a lesser offense.

75% (in 1999). The plea rate for death-eligible cases that were disposed by plea in which the defendant did not face the death penalty ranged from 31% (in 1998) to 53% (in 1993).

V. EMPIRICAL STRATEGY

A. DESIGNATION OF TREATMENT AND CONTROL GROUPS

The “treatment group” for the purposes of the study consists of defendants who were noticed for the death penalty. The “control group” was comprised of defendants eligible for the death penalty, but against whom the prosecutor chose not to seek the death penalty.¹⁵² The difficulty in defining the control group is specifying what qualifies as a “death-eligible” case. Some analysts suggest that any homicide committed in Georgia could be death eligible resulting from the state’s felony-murder statute and the *B7* statutory aggravating circumstance.¹⁵³ Recall that *B7* reads: “The offense of murder, rape, armed robbery, or kidnapping was outrageously or wantonly vile, horrible, or inhuman in that it involved torture, depravity of mind, or an aggravated battery to the victim.”¹⁵⁴ Due to this ambiguity, two different approaches were adopted to identify death-eligible defendants.

The first approach categorizes defendants as eligible for the death penalty based on the presence of at least one of the special aggravating circumstances listed in Georgia’s capital statute. The presence of these specific factors in each case was assessed in two ways. First, data from a study conducted by the *Atlanta Journal-Constitution* on 2,328 murder convictions obtained between 1995 and 2004 were used to determine the number of aggravating circumstances present in the 967 murder convictions with incident dates from 1995 through 2000 in that study.¹⁵⁵

¹⁵² Roman et al., *supra* note 27, at 533 (defining treatment and control groups in death penalty studies in a similar fashion).

¹⁵³ Kuziemko, *supra* note 20, at 137 n.15; Kathryn W. Riley, *The Death Penalty in Georgia: An Aggravating Circumstance*, 30 AM. U. L. REV. 835, 853–54 (1981) (explaining that the vagueness and overbreadth of the *B7* circumstance is in conflict with the narrowing requirement articulated in *Furman* and *Gregg*); Richard A. Rosen, *The “Especially Heinous” Aggravating Circumstance in Capital Cases—The Standardless Standard*, 64 N.C. L. REV. 941, 945 (1986) (arguing that the “especially heinous” aggravating factor is overinclusive, has been applied inconsistently, and fails to guide prosecutorial discretion).

¹⁵⁴ GA. CODE ANN. § 17-10-30(b)(7) (West 2003); *see also supra* Part III.A for a full description of Georgia’s capital statute. In *Godfrey v. Georgia*, the Court ruled the *B7* special circumstance was not unconstitutionally vague on its face. 446 U.S. 420, 420 (1980).

¹⁵⁵ *See Rankin et al., supra* note 145; Paternoster, *supra* note 145.

Table 1
Death Notices in Georgia by County (1993–2000)

County	Death Notices	% of Total Death Notices	County	Death Notices	% of Total Death Notices
Appling	5	1.3	Henry	5	1.3
Baldwin	5	1.3	Houston	2	0.5
Barrow	2	0.5	Irwin	1	0.3
Bartow	9	2.3	Jackson	3	0.8
Ben Hill	1	0.3	Jeff Davis	3	0.8
Bibb	8	2.0	Jefferson	1	0.3
Bleckley	1	0.3	Jenkins	1	0.3
Brooks	1	0.3	Jones	2	0.5
Bryan	1	0.3	Laurens	1	0.3
Bulloch	5	1.3	Lee	1	0.3
Burke	5	1.3	Liberty	4	1.0
Butts	2	0.5	Long	5	1.3
Camden	1	0.3	Lowndes	7	1.8
Candler	1	0.3	Macon	1	0.3
Carroll	3	0.8	Madison	1	0.3
Catoosa	1	0.3	Marion	2	0.5
Charlton	1	0.3	Meriwether	1	0.3
Chatham	10	2.5	Mitchell	2	0.5
Cherokee	2	0.5	Monroe	2	0.5
Clarke	11	2.8	Morgan	9	2.3
Clayton	19	4.8	Muscogee	15	3.8
Cobb	11	2.8	Newton	3	0.8
Coffee	3	0.8	Oconee	7	1.8
Colquitt	1	0.3	Oglethorpe	1	0.3
Columbia	4	1.0	Paulding	2	0.5
Cook	3	0.8	Pierce	1	0.3
Crawford	1	0.3	Pike	3	0.8
Crisp	2	0.5	Polk	4	1.0
DeKalb	25	6.3	Putnam	8	2.0
Dodge	1	0.3	Rabun	1	0.3
Dougherty	7	1.8	Randolph	1	0.3
Douglas	3	0.8	Richmond	20	5.0
Early	1	0.3	Rockdale	4	1.0
Effingham	2	0.5	Screven	2	0.5
Elbert	2	0.5	Spalding	8	2.0
Emanuel	1	0.3	Stephens	1	0.3
Fayette	4	1.0	Sumter	1	0.3
Floyd	5	1.3	Terrell	1	0.3
Forsyth	1	0.3	Thomas	2	0.5
Fulton	21	5.3	Tift	7	1.8
Gilmer	1	0.3	Toombs	3	0.8
Glynn	5	1.3	Troup	2	0.5
Greene	1	0.3	Upson	2	0.5
Gwinnett	13	3.3	Walker	2	0.5
Hall	11	2.8	Walton	3	0.8
Hancock	1	0.3	Ware	7	1.8
Harris	1	0.3	Washington	2	0.5
Hart	6	1.5	Wayne	1	0.3
			Whitfield	1	0.3

Total Death Notices: 400

Percent of all counties filing a death notice: 61%

Table 2
Death Notices in Georgia by Judicial Circuit (1993–2000)

Judicial Circuit	Death Notices	% of Total Death Notices
Alapaha	3	0.8
Alcovy	6	1.5
Appalachian	1	0.3
Atlanta	21	5.3
Atlantic	11	2.8
Augusta	29	7.2
Blue Ridge	3	0.8
Brunswick	16	4.0
Chattahoochee	16	4.0
Cherokee	9	2.3
Clayton	19	4.8
Cobb	11	2.8
Conasauga	1	0.3
Cordele	3	0.8
Coweta	6	1.5
Dougherty	7	1.8
Douglas	3	0.8
Dublin	1	0.3
Eastern	10	2.5
Flint	9	2.3
Griffin	17	4.3
Gwinnett	13	3.3
Houston	2	0.5
Lookout Mountain	3	0.8
Macon	8	2.0
Middle	8	2.0
Mountain	2	0.5
Northeastern	11	2.8
Northern	9	2.3
Ocmulgee	26	6.5
Oconee	2	0.5
Ogeechee	10	2.5
Pataula	3	0.8
Paulding	5	1.3
Rockdale	4	1.0
Rome	5	1.3
South Georgia	2	0.5
Southern	12	3.0
Southwestern	4	1.0
Stone Mountain	25	6.3
Tallapoosa	6	1.5
Tifton	8	2.0
Waycross	12	3.0
Western	18	4.5

Total Death Notices: 400

The second manner in which the presence of statutorily defined elements was determined was based on the presence of *B1*, *B2*, or *B4* special circumstances from inmate records from the Georgia Department of Corrections and the Georgia Bureau of Investigation. Recall that under *B1*, a defendant is eligible for the death penalty when “[t]he offense of murder, rape, armed robbery, or kidnapping was committed by a person with a prior record of conviction for a capital felony.”¹⁵⁶ Under Georgia’s statute, capital felonies are defined as murder, rape, armed robbery, or kidnapping.¹⁵⁷ A capital offense refers to statutorily defined capital offenses, not necessarily death-eligible offenses.¹⁵⁸ The predicate offenses for which the death penalty can be imposed are murder (malice or felony), aircraft hijacking, and treason.¹⁵⁹ According to the *B2* statutory aggravating circumstance, a defendant is eligible for the death penalty when “[t]he offense of murder, rape, armed robbery, or kidnapping was committed while the offender was engaged in the commission of another capital felony or aggravated battery, or the offense of murder was committed while the offender was engaged in the commission of burglary or arson in the first degree.”¹⁶⁰ Defendants who were convicted of contemporaneously committing these crimes—or had one of these crimes initially listed in their arrest reports—were categorized as death eligible.¹⁶¹ Defendants convicted

¹⁵⁶ GA. CODE ANN. § 17-10-30(b)(1); *see also supra* Part III.A.

¹⁵⁷ GA. CODE ANN. § 17-10-30(b)(1).

¹⁵⁸ *Merrow v. State*, 601 S.E.2d 428, 431 (Ga. Ct. App. 2004) (“A capital offense . . . refers to offenses defined by statute as capital offenses, not necessarily offenses for which the state could or actually does seek the death penalty.” (quoting *White v. State*, 414 S.E.2d 296, 297 (Ga. Ct. App. 1991))); *accord* *Peek v. State*, 238 S.E.2d 12, 20 (Ga. 1977) (explaining that nonhomicide “capital offenses” listed in the death penalty statute qualify as “capital felonies” for purposes of applying the aggravating circumstance provision of the capital statute).

¹⁵⁹ *Collins v. State*, 236 S.E.2d 759, 762 (Ga. 1977) (Jordan, J., concurring) (“Of course the crimes of treason and aircraft hijacking, along with murder, remain capital felonies . . .”).

¹⁶⁰ GA. CODE ANN. § 17-10-30(b)(2).

¹⁶¹ The major limitation of this measure is that the Georgia Department of Corrections does not technically distinguish between offenses committed during the actual commission of the murder and offenses the defendant was simply convicted of during the same trial as the murder. As a check, several cases were randomly selected from the Department of Corrections website that lists the separate offenses (if a multiple-offense case) by the incident date. In the vast majority of these cases, the murder and other felony were committed on the same day. This suggests that, for most cases, the *B2* measure is valid for determining death eligibility. I also cross-referenced the conviction data from the Georgia Department of Corrections with the arrest data from the Georgia Bureau of Investigation in order to determine whether the contemporary felony present at time of arrest was consistent with the conviction data for lesser felonies. This additional check supports the assertion that the murder and the other felony conviction were truly contemporaneous.

of murdering multiple victims were also categorized as death eligible because multiple victim homicides satisfy the requirements of the *B2* statutory aggravating circumstance.¹⁶²

Death eligibility based on the presence of the *B4* circumstance—“offender committed the offense of murder for himself or another, for the purpose of receiving money or any other thing of monetary value”—was determined by a contemporaneous conviction for an economically motivated crime (i.e., robbery and theft). Admittedly, the three factors employed to determine death-eligible defendants are narrow and ignore a host of other factors listed in the statute, but it is worth noting that prior research strongly suggests that the presence of multiple victims and a contemporaneous felony are “the most commonly used factors in death sentence cases, and thus account for a high proportion of death eligible cases.”¹⁶³ Consequently, these two types of aggravating circumstances are also the most common factors used by judges and jurors to justify death sentences, as well as the strongest predictors of a death sentence, even after holding other legally relevant factors constant.¹⁶⁴ Perhaps more

¹⁶² The presence of multiple victims or a contemporaneous felony is commonly employed by researchers to identify death-eligible cases. See generally SAMUEL R. GROSS & ROBERT MAURO, *DEATH AND DISCRIMINATION: RACIAL DISPARITIES IN CAPITAL SENTENCING* (1989); Pierce & Radelet, *supra* note 80, at 72–91 (2002); Glenn L. Pierce & Michael L. Radelet, *The Impact of Legally Inappropriate Factors on Death Sentencing for California Homicides, 1990–1999*, 46 SANTA CLARA L. REV. 1, 21–25 (2005). Not only are these two types of aggravating circumstances the most common set of aggravating circumstances used by prosecutors, jurors, and judges to justify death sentences, but the number of victims is consistently one of the strongest predictors of a death sentence, even after holding other legally relevant factors constant. Pierce & Radelet, *supra* note 80, at 72–91; Steven F. Shatz & Nina Rivkind, *The California Death Penalty Scheme: Requiem for Furman?*, 72 N.Y.U. L. REV. 1283, 1328–32 (1997).

To be sure, the multiple victim measure is imperfect because it is possible that a defendant murdered multiple victims on the same day but in unrelated situations. The vast majority of cases, however, are single victim (87.3%), so this measure, in and of itself, does not influence the categorization of most defendants. Perhaps more importantly, the *B1* and *B2* criteria are not mutually exclusive—in fact, they share considerable overlap. It is possible, even likely, then, that a defendant’s categorization as death eligible will be valid on one or more of these measures.

¹⁶³ Compare Pierce & Radelet, *supra* note 80, at 66, with Shatz & Rivkind, *supra* note 162, at 1329 (“[T]he felony murder special circumstances play the predominant role in defining death-eligibility.”).

¹⁶⁴ Pierce & Radelet, *supra* note 80, at 61 (noting that juries were most likely to impose the death sentence in cases involving multiple victims); Shatz & Rivkind, *supra* note 162, at 1329–30 (explaining three-quarters of death-sentenced cases involved a felony-murder circumstance); see also *Gregg v. Georgia*, 428 U.S. 153, 225 (1976) (White, J., concurring) (“[T]he standards by which [prosecutors] decide whether to charge a capital felony will be the same as those by which the jury will decide the questions of guilt and sentence.”); William J. Bowers & Glenn L. Pierce, *Arbitrariness and Discrimination Under Post-Furman*

importantly, data on death eligibility from the more detailed study conducted by the *Atlanta Journal-Constitution* overlap with nearly 80% of the cases in the current sample, so a more “fine-grained” assessment of statutorily defined elements is available for the vast majority of cases.¹⁶⁵

Death eligibility was also limited to situations in which the defendant was ultimately convicted of murder. Obviously prosecutors seek the death penalty against defendants prior to obtaining a guilty plea or guilty verdict at trial, but limiting the pool of death-eligible defendants to those who are actually convicted of murder serves as a proxy for strength of evidence. Other scholars have employed this limiting strategy when examining capital charging-and-sentencing processes.¹⁶⁶ Perhaps of equal significance is the fact that Georgia’s murder statute does not include “degrees” of murder like many other states. Instead, the statute specifies that individuals may be charged with malice murder (intentional) or felony murder (unintentional, but during the commission of *any* other felony),¹⁶⁷ and only these two types of murder may be death eligible, depending on the presence of at least one statutory aggravating circumstance. “Lesser” degrees of homicide are categorized as voluntary and involuntary manslaughter.¹⁶⁸ It is very unlikely that prosecutors will offer a charge bargain from murder to manslaughter when the available evidence permits a conviction for murder, which requires a mandatory minimum life sentence.¹⁶⁹ Similarly,

Capital Statutes, 26 CRIME & DELINQ. 563, 627 (1980) (discovering that the vast majority of death sentences imposed in Georgia and Florida involved cases with a felony circumstance).

¹⁶⁵ The “crude” measure of death eligibility based on the presence of *B1*, *B2*, or *B4* special circumstances failed to classify 36% of cases identified by the AJC study (and 16% of cases noticed for the death penalty). Thirty-nine percent of the cases misclassified were technically eligible for death based upon the presence of the *B7* circumstance.

¹⁶⁶ See, e.g., BALDUS ET AL., *supra* note 84, at 40–42, 477 n.72 (discussing the use of murder conviction as a proxy for strength of evidence); accord David C. Baldus et al., *Racial Discrimination and the Death Penalty in the Post-Furman Era: An Empirical and Legal Overview, with Recent Findings from Philadelphia*, 83 CORNELL L. REV. 1638, 1668–70 (1998); Paternoster et al., *supra* note 80.

¹⁶⁷ GA. CODE ANN. § 16-5-1(a) (West 2003) (malice murder); *id.* § 16-5-1(c) (felony murder).

¹⁶⁸ Voluntary manslaughter carries a maximum sentence of twenty years and involuntary manslaughter carries a maximum sentence of ten years. GA. CODE ANN. § 16-5-2 (voluntary manslaughter); *id.* § 16-5-3(a) (involuntary manslaughter).

¹⁶⁹ COHEN & KYCKELHAHN, *supra* note 2, at 10–11 (noting that, regardless of method of adjudication, the vast majority of defendants initially charged with murder are ultimately convicted of murder). See William J. Stuntz, *Plea Bargaining and Criminal Law’s Disappearing Shadow*, 117 HARV. L. REV. 2548, 2563 (2004) (commenting that prosecutors generally pursue every murder case they can, which is why the acquittal rate in murder cases is higher than for other violent felonies); see also Bowers, *supra* note 50, at 1153 (same); Gazal-Ayal, *supra* note 15, at 2320 (explaining that dropping or reducing murder charges can be politically costly for prosecutors and this is part of the reason they dedicate resources

prosecutors are not likely to seek the death penalty in cases in which they believe a conviction for manslaughter is warranted, even in the presence of aggravating circumstances that would make the case eligible for the death penalty if a murder conviction were obtained.¹⁷⁰

Focusing exclusively on defendants subsequently convicted of murder suffers from the drawback of excluding death-noticed defendants who are not ultimately convicted of murder. In these data, twenty-eight individuals (7% of death-noticed defendants) were initially noticed for the death penalty, but later were acquitted, had their charges dismissed, or were convicted of an offense other than murder. Nevertheless, a closer inspection of the data supports the previously stated intuition that prosecutors do not typically offer charge bargains in exchange for guilty pleas. Of the twelve cases noticed for the death penalty that resulted in a plea bargain for a charge other than murder, only two were for manslaughter.¹⁷¹ Thus, the more cautious approach adopted in these analyses (i.e., underinclusion) in an attempt to most accurately compare “apples to apples” would appear to outweigh its disadvantages.

The second approach utilized to identify death-eligible defendants was to define *all* cases that ultimately resulted in a murder conviction as death eligible. Because some scholars claim that capital statutes like Georgia’s permit any homicide to be deemed death eligible,¹⁷² this very broad definition of death eligibility was employed in order to determine whether the results are robust to the specific criteria used to identify the control group.

B. STATISTICAL MODEL

There are, essentially, two reasons why death-noticing and plea-bargaining decisions might be related. First, the decisions may be causally related.¹⁷³ Second, death noticing and plea bargaining may be related

to trying even weak murder cases).

¹⁷⁰ See Liebman, *supra* note 31, at 2097–98 (charging a case capitally increases the chances of winning, but it also increases the embarrassment and publicity of losing); William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 570 (2001) (explaining that defeats at trial for prosecutors are so vivid because they are rare, so prosecutors are less likely to pursue cases that are unwinnable).

¹⁷¹ The remaining plea bargains to nonhomicide charges were for aggravated assault, armed robbery, burglary, concealing the death of another, and kidnapping.

¹⁷² See, e.g., John Blume & Theodore Eisenberg, *Judicial Politics, Death Penalty Appeals, and Case Selection: An Empirical Study*, 72 S. CAL. L. REV. 465 (1999); Rosen, *supra* note 153.

¹⁷³ Death noticing typically precedes plea negotiations. In fact, it is common for prosecutors to announce the intent to seek the death penalty before having obtained an indictment or meeting with the defendant. It remains plausible, however, that some

because they are caused by other shared factors (i.e., “third-variables”) that may or may not be observed, and if these factors are taken into account, the relationship between death noticing and plea bargaining disappears; that is, the relationship between death noticing and plea bargaining is not independent of those other factors.¹⁷⁴

In an effort to determine the impact of the death penalty on the decision to go to trial, a conditional fixed-effects logit model is estimated according to the following equation:

$$\Pr(\text{Trial}_{ict} = 1) = \frac{\exp(\alpha_c + \gamma_k X_{ict} + \beta \times \text{DPNotice}_{ict} + \zeta_{ict})}{1 + \exp(\alpha_c + \gamma_k X_{ict} + \beta \times \text{DPNotice}_{ict} + \zeta_{ict})} \quad [1]$$

where i indexes the defendant, c indexes the judicial circuit, and t indexes the year.¹⁷⁵ In all of the analyses employed, X_{ict} is a matrix of case characteristics, including, *inter alia*, $t - 1$ incident-year dummies, and DPNotice is a dummy (binary) variable indicating whether the prosecutor sought the death penalty against the defendant. The model makes the following assumptions: (a) conditional on X_{ict} and DPNotice, Trial_{ict} is an independent Bernoulli random variable¹⁷⁶ with probability given by Equation [1]; (b) $\Pr(\text{Trial}_{ict} = 1)$ depends on X_{ict} and DPNotice through the logistic function; (c) $\Pr(\text{Trial}_{ict} = 1)$ is governed by parameters γ and β ,¹⁷⁷ a

prosecutors elect to file a death notice following an initial failed attempt to obtain a plea bargain, but ultimately withdraw the death notice after a defendant agrees to a negotiated plea. This potential complication, known as simultaneous causation, is addressed more fully in Part VI.B.

¹⁷⁴ For a detailed discussion of the key requirements of causal inference, see RICHARD A. BERK, *REGRESSION ANALYSIS: A CONSTRUCTIVE CRITIQUE* 82–83 (2004). It is important to emphasize that researchers need not control for every conceivable variable possibly influencing plea bargaining. Candidate variables must meet three conditions: (1) correlated with the key causal variable (i.e., death-noticing decision); (2) affects the outcome variable (i.e., plea-bargaining decision); and (3) causally prior to the key causal variable. Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 78 (2002). If one of these three conditions is absent, then controlling for the rival variable is not only unnecessary when examining the causal impact of the key variable of interest, but it may also lead to incorrect inferences if the variable is a consequence of the key causal variable (i.e., “post-treatment” bias). *Id.* at 79–80; see also GARY KING ET AL., *DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH* 78 (1994) (controlling for a consequence of the cause produces the incorrect causal effect).

¹⁷⁵ WILLIAM H. GREENE, *ECONOMETRIC ANALYSIS* 839 (4th ed. 2000).

¹⁷⁶ A random variable, Y , that can only take on two values, 0 and 1, with $\Pr(Y=1) = p$ is a Bernoulli random variable with parameter p . This variable has a mean of p and a variance of $p(1 - p)$. MORRIS H. DEGROOT & MARK J. SCHERVISH, *PROBABILITY AND STATISTICS* 276 (4th ed. 2012).

¹⁷⁷ More precisely, the γ and β coefficients represent the expected change in the probability of trial corresponding to changes in each predictor in the model (i.e., X and

unit-specific (e.g., jurisdiction-specific) parameter α_c , and (d) ζ is a vector of residual error terms with a mean of zero and variance of $\pi^2/3$.¹⁷⁸

The fixed-effects specification is particularly advisable with these data because the models control for unobserved heterogeneity across judicial circuits and years.¹⁷⁹ Failing to account for these fixed effects can result in omitted variable bias and lead to inconsistent estimates of a hypothesized causal effect.¹⁸⁰ Georgia's Administrative Office of the Courts (AOC)

DPNotice, respectively). Gelman et al., *supra* note 40, at 239. There are a total of $k + 1$ parameters estimated in the model.

¹⁷⁸ GREENE, *supra* note 175, at 70. I employ the conditional fixed-effects estimator, *id.* at 839, due to the bias introduced by using the unconditional fixed effects (indicator set) when cluster sizes are relatively small. See Tom Coupé, *Bias in Conditional and Unconditional Fixed Effects Logit Estimation: A Correction*, 13 POL. ANALYSIS 292, 295 (2005); Ethan Katz, *Bias in Conditional and Unconditional Fixed Effects Logit Estimation*, 9 POL. ANALYSIS 379, 384 (2001). An acceptable probit alternative specification does not exist because there is no sufficient statistic that allows the fixed effects to be conditioned out of the likelihood function. William Greene, *The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects*, 7 ECONOMETRICS J. 98, 102–03 (2004). I also reanalyzed the data with the unconditional estimator and obtained similar results.

Alternative specifications were analyzed employing a random-effects estimator (i.e., random intercept models). ANDERS SKRONDAL & SOPHIA RABE-HESKETH, *GENERALIZED LATENT VARIABLE MODELING: MULTILEVEL, LONGITUDINAL, AND STRUCTURAL EQUATION MODELS* 49–50 (2004). Random effects models rely on the strong assumption that the unobserved cluster-specific influences are uncorrelated with individual-level case characteristics, so fixed-effects models are preferred when that assumption is unrealistic. GREENE, *supra* note 175, at 576–77. It is possible, however, to allow the random effect (i.e., intercepts) to be correlated with the individual-level variables by creating an aggregated measure of the individual variables for each group and including that measure as a predictor in the model. Any covariance between an individual-level predictor and a group-level random effect must operate through the covariance between the group-level average of the individual-level predictor and that random effect; therefore, inclusion of the group-level average of the predictor as a covariate in the model will eliminate any confounding between the individual-level predictor and omitted variables at the group level. STEPHEN W. RAUDENBUSH & ANTHONY S. BRYK, *HIERARCHICAL LINEAR MODELS: APPLICATIONS AND DATA ANALYSIS METHODS* 261–62 (2d ed. 2002); see also ANDREW GELMAN & JENNIFER HILL, *DATA ANALYSIS USING REGRESSION AND MULTILEVEL/HIERARCHICAL MODELS* 506 (2007). Results (not reported) from these models were very similar to those obtained from the fixed-effects specifications.

¹⁷⁹ The variance of the jurisdiction-level effects tells us the extent to which there is variability among jurisdictions in the data beyond that explained by the other regression predictors. Similarly, the year effects represent unexplained variation among years. Gelman et al., *supra* note 40, at 238 n.84; see also Thomas R. Ten Have et al., *Deviations from the Population-Averaged Versus Cluster-Specific Relationship for Clustered Binary Data*, 13 STAT. METHODS MED. RES. 3, 9 (2004) (“In the binary response case, conditional likelihood estimation is the only approach that is less susceptible [to confounding of treatment effect due to cluster-level unobserved heterogeneity], because it conditions out all cluster-level information that may confound within-cluster effects.”).

¹⁸⁰ GREENE, *supra* note 175, at 839–40.

organizes the state's 159 counties into forty-nine superior court judicial circuits.¹⁸¹ As a result, county-level data are nested in the judicial circuits for the analyses conducted in this paper.¹⁸² The distribution of death notices by county and judicial circuit are presented in, respectively, Tables 1 and 2. The specific variables used, coding conventions adopted, and summary statistics can be found in Appendix A and Tables 3, 4, and 5.

Although the model controls for average differences across jurisdictions and average differences across years, it does not take into account omitted covariates that are case specific.¹⁸³ Prosecutors do not randomly select cases for the death penalty, so the estimation of the true impact of capital punishment on trials may be biased and inconsistent if the death-noticing decision is correlated with some other unobserved variable(s) that also has a causal impact on the decision to go to trial (i.e., endogeneity bias). Relatedly, if a death-noticing decision is also influenced by a prior failed plea negotiation (i.e., simultaneous causality), death noticing is also endogenous because it will be correlated with an unobserved variable influencing both decisions: preliminary plea negotiation. The fixed-effects logit model discussed earlier implicitly assumes that all factors simultaneously influencing the death noticing and trial decisions are included in the model via the covariates and circuit and year fixed effects. This may be a questionable assumption, so several alternative models that explicitly account for confounding omitted variables at the case level are examined in Part VI.B.

¹⁸¹ JUDICIAL COUNCIL OF GA., YOUR GUIDE TO THE GEORGIA COURTS (2003).

¹⁸² This is necessary for two important reasons. First, in Georgia, there is one district attorney per judicial circuit. While large counties comprise a single judicial circuit, many smaller counties are grouped together to form a single judicial circuit. As a result, a single prosecutor may be responsible for charging and plea-bargain decisions for several counties in her judicial circuit. Also, if a judicial circuit consists of multiple counties, trial judges rotate throughout these counties in the circuit. Treating counties that share a single judicial circuit as if they were independent ignores the similarities they share in the administration of capital punishment resulting from shared decisionmakers. Second, death penalty cases are extremely rare events, so aggregating county-level data to the judicial-circuit level allows one to observe more cases per contextual unit and better statistically estimate relationships occurring at both the case- and contextual-level without altering the dependence structure of the cases due to their clustering. RAUDENBUSH & BRYK, *supra* note 178, at 45 (noting that aggregating data at the highest level of nesting preserves the dependence structure of the units); Gary King & Langche Zeng, *Logistic Regression in Rare Events Data*, 9 POL. ANALYSIS 137, 163 (2001) (discussing the difficulties associated with analyzing rare events in binary data).

¹⁸³ When the death-noticing decision is correlated with the case-specific error term rather than just the judicial-circuit or year effect, a simultaneous equation approach is necessary. John Antonakis et al., *On Making Causal Claims: A Review and Recommendations*, 21 LEADERSHIP Q. 1086, 1092 (2010).

Table 3
Summary Statistics (DE Subsample)

Variable Name	N	Mean	Std. Dev.	Min	Max
Plea/Verdict (Verdict)	852	0.54	0.50	0	1
DP Notice	852	0.32	0.47	0	1
# of Stat Aggs	852	2.28	1.09	1	7
# of Offenders	852	1.80	1.11	1	7
Offender Black	852	0.73	0.45	0	1
Offender White	852	0.25	0.43	0	1
Offender Sex (Male)	852	0.95	0.23	0	1
Offender Age (Yrs)	852	27.14	9.95	17	69
Contemp. Felony	852	1.72	1.60	0	9
Prior Felony	852	0.51	1.33	0	10
Offender HS Grad	852	0.26	0.44	0	1
Offender Employed	852	0.56	0.50	0	1
Offender Married	852	0.18	0.38	0	1
# of Children	852	0.58	0.49	0	1
# of Victims	852	1.18	0.50	1	6
Victim Black	852	0.50	0.50	0	1
Victim White	852	0.44	0.50	0	1
Victim Sex (Female)	852	0.36	0.48	0	1
Victim Age	852	36.66	18.17	0	99
Victim Stranger	852	0.35	0.48	0	1
Interracial Homicide	852	0.28	0.45	0	1

Table 4
Summary Statistics (MC Subsample)

Variable Name	N	Mean	Std. Dev.	Min	Max
Plea/Verdict (Verdict)	1238	0.58	0.49	0	1
DP Notice	1238	0.22	0.42	0	1
# of Stat Aggs	1238	1.57	1.39	0	7
# of Offenders	1238	1.65	1.05	1	7
Offender Black	1238	0.72	0.45	0	1
Offender White	1238	0.26	0.44	0	1
Offender Sex (Male)	1238	0.93	0.25	0	1
Offender Age (Yrs)	1238	28.07	10.50	17	74
Contemp. Felony	1238	1.43	1.49	0	9
Prior Felony	1238	0.46	1.23	0	10
Offender HS Grad	1238	0.28	0.45	0	1
Offender Employed	1238	0.57	0.50	0	1
Offender Married	1238	0.18	0.38	0	1
# of Children	1238	0.60	0.49	0	1
# of Victims	1238	1.13	0.43	1	6
Victim Black	1238	0.56	0.50	0	1
Victim White	1238	0.40	0.49	0	1
Victim Sex (Female)	1238	0.35	0.48	0	1
Victim Age	1238	34.58	17.22	0	99
Victim Stranger	1238	0.29	0.45	0	1
Interracial Homicide	1238	0.22	0.41	0	1
DP Eligible	1238	0.69	0.46	0	1

Table 5
Summary Statistics (Death-Noticed Subsample)

Variable Name	N	Mean	Std. Dev.	Min	Max
Plea/Verdict (Verdict)	395	0.41	0.49	0	1
# of Stat Aggs	391	2.88	1.23	1	7
# of Offenders	400	1.89	1.14	1	7
Offender Black	400	0.59	0.49	0	1
Offender White	400	0.37	0.48	0	1
Offender Sex (Male)	400	0.94	0.23	0	1
Offender Age (Yrs)	400	27.29	9.28	17	66
Contemp. Felony	400	2.11	1.85	0	9
Prior Felony	400	0.52	1.42	0	10
Offender HS Grad	326	0.27	0.44	0	1
Offender Employed	330	0.62	0.49	0	1
Offender Married	344	0.19	0.39	0	1
# of Children	322	0.60	0.49	0	1
# of Victims	400	1.37	0.75	1	6
Victim Black	400	0.30	0.46	0	1
Victim White	397	0.66	0.47	0	1
Victim Sex (Female)	399	0.50	0.50	0	1
Victim Age	400	38.17	19.45	0	90
Victim Stranger	385	0.45	0.50	0	1
Interracial Homicide	397	0.35	0.48	0	1

C. MISSING DATA

An additional complication with analyzing official homicide records is incomplete information (i.e., missing data).¹⁸⁴ The vast majority of statistical analyses must be performed on a full data matrix; therefore, the common practice among social scientists is to perform *casewise* deletion by eliminating observations that have missing data on one or more variables.¹⁸⁵ Casewise deletion is problematic because it (1) potentially forces researchers to discard much useful information about the relationships

¹⁸⁴ See generally Wendy C. Regoeczi & Marc Riedel, *The Application of Missing Data Estimation Models to the Problem of Unknown Victim/Offender Relationships in Homicide Cases*, 19 J. QUANTITATIVE CRIMINOLOGY 155 (2003) (suggesting ways to deal with missing data problems when analyzing official homicide data).

¹⁸⁵ Roderick J. A. Little, *Regression with Missing X's: A Review*, 87 J. AM. STAT. ASS'N 1227, 1228 (1992).

between variables, (2) results in inefficient parameter estimates due to a reduction of sample size, and (3) may bias parameter estimates if the data are not missing completely at random (i.e., if the missing data are not a random subset of the overall population).¹⁸⁶ There is missing data on at least one variable in approximately one-third of the observations in the Georgia data; however, when it is possible to predict the probability that a variable is missing information for an observation (using information from other covariates in the data), the most appropriate strategy is to attempt to predict those missing values.¹⁸⁷ This is particularly true when examining the capital punishment process because death penalty cases in Georgia (and elsewhere) are very rare occurrences, so it is crucial to retain as much information as possible.

Over the past two decades, quantitative methodologists have developed several approaches to “guess” the values of missing data by using information about the association of the variable of interest with other variables in the data. A regression-based multiple imputation approach is employed in these analyses, which provides a significant improvement over simple imputation methods and traditional single imputation strategies.¹⁸⁸ In brief, Stef van Buuren and colleagues’ “fully conditional specification” (FCS) approach is used because it offers the greatest flexibility in creating multivariate imputation models by allowing for specialized methods that are impractical under the other approaches.¹⁸⁹ The FCS approach imputes the data on a variable-by-variable basis by specifying an imputation model for each variable, thereby allowing the analyst to preserve unique features of the data such as bounds, skip patterns, interactions, and bracketed responses, and to incorporate appropriate constraints between different variables in order to avoid logical inconsistencies in the imputed data.¹⁹⁰

A somewhat simplified description of the algorithm is that observed data are used to impute missing values and incorporate estimation

¹⁸⁶ *Id.* at 1227.

¹⁸⁷ *Id.* at 1230, 1233.

¹⁸⁸ *Id.* at 1234–35.

¹⁸⁹ Stef van Buuren, *Multiple Imputation of Discrete and Continuous Data by Fully Conditional Specification*, 16 *STAT. METHODS MED. RES.* 219, 219 (2007).

¹⁹⁰ *Id.* at 219, 222. The statistical properties of FCS are not fully understood, but simulation studies suggest that FCS performs well in a variety of applications. S. van Buuren et al., *Fully Conditional Specification in Multivariate Imputation*, 76 *J. STAT. COMPUTATION & SIMULATION* 1049, 1061 (2006); Trivellore E. Raghunathan et al., *A Multivariate Technique for Multiply Imputing Missing Values Using A Sequence of Regression Models*, 27 *SURV. METHODOLOGY* 85, 92–93 (2001). More importantly, when there are missing variables that follow a mixture of distributions (e.g., continuous, ordinal, categorical), FCS is the only sensible parametric approach. Van Buuren et al., *supra*, at 1061.

uncertainty (resulting from analyzing a finite number of observations) and fundamental uncertainty (resulting from unmodeled variation in the dependent variable and represented by the stochastic component of the model) in their prediction of plausible values. For the present study, this process was repeated five times to create five complete data sets, with each data set containing different plausible values for missing variables to account for the uncertainty surrounding the imputations. After these data sets were created, a complete-case analysis was repeated on each data set. The overall point estimate of each parameter was obtained by averaging across the five separate point estimates for that particular parameter.¹⁹¹ The variance of the point estimate was computed by averaging across the five estimated variances from within each completed data set, plus the sample variance in the point estimate across the data sets (multiplied by a factor that corrects for bias because the number of imputed data sets is finite).¹⁹²

The next section presents results from models using both casewise deletion and multiple imputation strategies.

VI. RESULTS

A. FIXED-EFFECTS LOGIT SPECIFICATIONS

Table 6 displays results from four different specifications.¹⁹³ Across all four models, defendants noticed for the death penalty were significantly less likely to opt for trial (i.e., significantly more likely to accept a plea).¹⁹⁴ Models 1 and 3 analyze the impact of the death penalty on the probability of going to trial for defendants classified as death eligible according to the

¹⁹¹ Little, *supra* note 185, at 1235.

¹⁹² *Id.*

¹⁹³ Binary regression models in this Article report Cragg and Uhler's pseudo- R^2 statistic, defined as:

$$\frac{1 - \exp(2 \times [LL_{\text{null}} - LL_{\text{full}}] / N)}{1 - \exp(2 \times [LL_{\text{null}} - LL_{\text{max}}] / N)}$$

where LL_0 is the log-likelihood for the null model (i.e., constant-only model), LL_1 is the full regression model, LL_{max} is the maximum possible likelihood (i.e., perfect fit), and N is the sample size. See John G. Cragg & Russell S. Uhler, *The Demand for Automobiles*, 3 CAN. J. ECON. 386, 400 n.20 (1970). The Cragg and Uhler pseudo- R^2 statistic is most analogous to the traditional R^2 statistic used in ordinary least squares (OLS) regressions because, unlike most other pseudo- R^2 statistics, it is "normed" so the upward bound approaches unity. See J. SCOTT LONG, REGRESSION MODELS FOR CATEGORICAL AND LIMITED DEPENDENT VARIABLES 106 (1997).

¹⁹⁴ Recall that the models estimate the probability of a defendant taking her case to trial, so a negative sign on the DPNotice coefficient indicates that defendants noticed for the death penalty are *less* likely to have their cases resolved by trial, which is equivalent to being *more* likely to have their cases resolved by plea agreement.

first criteria discussed: eligibility based on the presence of statutorily defined elements (hereinafter, “DE”). Models 2 and 4 analyze the impact of the death penalty on defendants classified as death eligible based on the second criteria: murder conviction (hereinafter, “MC”). All specifications include judicial-circuit and year fixed effects, as well as controls for the number of codefendants; defendant’s race/ethnicity, gender, and age; defendant’s employment status at time of arrest; defendant’s marital status; number of statutory aggravating factors; contemporaneous felony conviction; prior felony conviction; whether defendant graduated from high school; number of children defendant has; number of victims; victim race/ethnicity, gender, and age; relationship between the victim and offender; and whether the homicide was interracial.¹⁹⁵

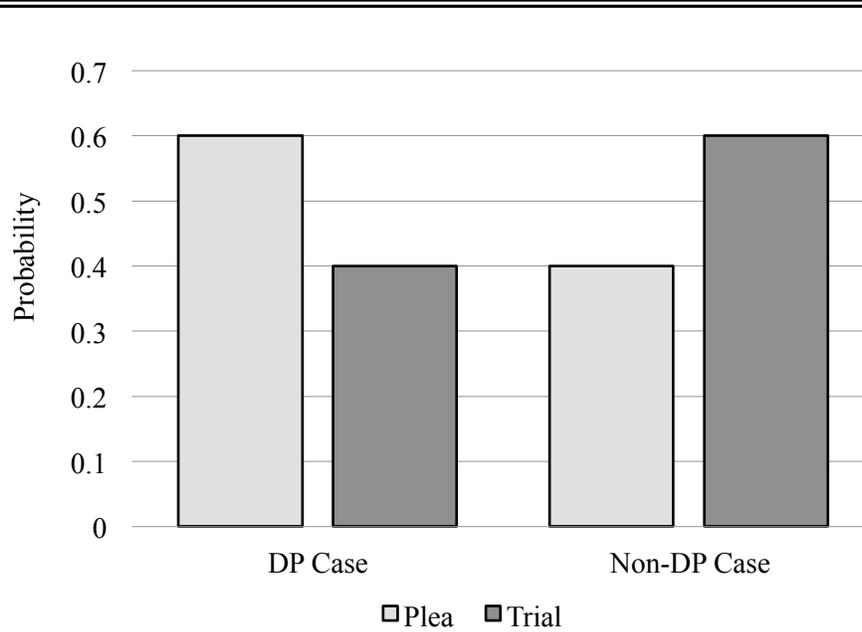
The natural coefficients from the logistic regression model, “logit coefficients or log odds,” lack an intuitive interpretation, so marginal effects are presented. The marginal effect represents the change in the probability of a case being disposed by trial, holding all other variables constant.¹⁹⁶ Model 1 (DE) and Model 2 (MC) reveal that being noticed for the death penalty reduces the probability of a defendant taking her case to trial by, respectively, .23 and .22. For Model 1, this means the probability of death-noticed defendants going to trial is .37, whereas the probability of defendants not noticed for death going to trial is .60, all else equal (see Figure 1). For Model 2, the probabilities are .41 and .63, respectively.

¹⁹⁵ See *infra* Appendix A for a detailed description of these variables and Tables 3, 4, and 5 for summary statistics. Appendix B presents the marginal effects for all of the covariates in the fixed-effects logit model. Model 1 lists the point estimates for covariates predicting the probability of a case being disposed by trial. Model 2 displays the effects of the same covariates on the probability a case is noticed for the death penalty.

¹⁹⁶ LONG, *supra* note 193, at 71–74. The conditional fixed-effects model does not provide estimates of the judicial-circuit fixed effects, a_c , which are needed to compute marginal effects. Coupé, *supra* note 178, at 292. Thus, marginal effects were obtained using the unconditional fixed-effects logit model. The conditional and unconditional fixed-effects estimates are essentially the same when cluster sizes average at least eight, and the average cluster sizes are twenty and twenty-seven in, respectively, the DE and MC models. *Id.* The marginal effects obtained from the unconditional fixed-effects linear model were nearly identical, suggesting that the results are quite robust to various model specifications. See *infra* Part VI.B.

Population-average effects (sometimes referred to as marginal effects) were also obtained, and these estimates were similar to the conditional and unconditional fixed-effects estimates. The unconditional fixed-effects estimates represent the difference in the probability of trial, depending on being noticed for the death penalty, for the *same* defendant. The population-average effect, on the other hand, represents the difference in probability of trial of the *average* defendant noticed for the death penalty versus the *average* defendant not noticed for the death penalty—that is, the estimates do not control unobserved circuit-level effects. RAUDENBUSH & BRYK, *supra* note 178, at 304, 334.

Figure 1
Method of Disposition



As noted, *supra*, approximately one-third of the cases in the data have missing information on at least one variable. Specifically, the DE models lose 31.2% of the cases and the MC models lose 35.9% of the cases. Table 6 presents results from the multiple imputation models. As with the casewise deletion results previously reported, the coefficient for DPNotice is statistically significant across all specifications, although coefficients are slightly smaller with respect to their absolute value. In the Model 3 (DE), being noticed for the death penalty decreases the probability of trial by .18, and by .17 in Model 3 (MC). The baseline probabilities for the DE (Model 3) and MC (Model 4) specifications are very similar to the casewise deletion models (.60 and .64, respectively).

Table 6
Marginal Effect of Death Penalty on Trial (Fixed-Effects Logit Models)

	Model 1	Model 2	Model 3	Model 4
DP Notice	-0.234*** (0.038)	-0.218*** (0.038)	-0.184*** (0.032)	-0.168*** (0.031)
Pr(Trial)	0.603	0.628	0.603	0.635
N	847	1238	1236	1932
Fixed-Effects	Y	Y	Y	Y
Year Dummies	Y	Y	Y	Y
R-Squared	0.32	0.28		

*p<.05; **p<.01; ***p<.001

Standard errors adjusted for clustering on judicial circuit.

Models 1 & 3: Death Eligible Subsample.

Models 2 & 4: Murder Conviction Subsample.

Models 3 & 4: Multiple Imputation Estimates.

Pr(Trial) = Probability of trial without a death noticed filed.

All specifications include controls for the number of codefendants, victims, statutory aggravating factors, contemporaneous felony convictions, prior felony convictions, and children of defendant; the race/ethnicity, gender, and age of defendant and victim(s); defendant's employment status at time of arrest, marital status, high school graduation status; offender/victim relationship; and whether the homicide was interracial.

B. SENSITIVITY ANALYSES

Table 7 presents the linear (unconditional) fixed-effects estimates for the DE and MC models using casewise deletion (Models 1 and 2) and multiple imputation (Models 3 and 4). The classic fixed-effects linear model takes the form:

$$E(\text{Trial}) = \Pr(\text{Trial} = 1) = \alpha_c + \gamma X + \beta \times \text{DPNotice} + \zeta \quad [2]$$

where α , γ , β , ζ , X , and DPNotice are defined in Equation [1], but ζ are now identically and independently normally distributed: $\zeta \sim N(0, \sigma^2)$.¹⁹⁷ Linear

¹⁹⁷ See GREENE, *supra* note 175, at 560. Some subscripts have been omitted for

regression models applied to binary dependent variables are commonly referred to as linear probability models.¹⁹⁸ These models are generally deemed inappropriate for binary data because of heteroscedasticity, nonnormality, nonsensical predictions, and functional form misspecification.¹⁹⁹ These results are presented simply as a robustness check for the conditional fixed-effects logit estimates presented in Table 6. Unconditional fixed-effects models are most appropriate in the linear regression context, so the purpose of showing these estimates is to check that the DPNotice coefficients are similar in terms of direction, magnitude, and statistical significance.²⁰⁰ The interpretation of coefficients from the linear probability is similar to the linear regression model, so according to Model 1 (DE and casewise deletion), for example, being noticed for the death penalty decreases the probability of going to trial by .22 from a baseline probability of .62, holding all other variables constant.

simplicity.

¹⁹⁸ LONG, *supra* note 193, at 35.

¹⁹⁹ *Id.* at 38–40.

²⁰⁰ GREENE, *supra* note 175, and accompanying text; *accord* David S. Abrams & Albert H. Yoon, *The Luck of the Draw: Using Random Case Assignment to Investigate Attorney Ability*, 74 U. CHI. L. REV. 1145, 1168 (2007) (employing a linear model for a binary outcome, rather than a logit or probit model, in order to estimate unconditional fixed effects).

Table 7
Marginal Effects of Death Penalty on Trial (Fixed-Effects LPMs)

	Model 1	Model 2	Model 3	Model 4
DP Notice	-0.217*** (0.048)	-0.212*** (0.044)	-0.186*** (0.036)	-0.173*** (0.034)
Pr(Trial)	0.618	0.645	0.622	0.658
N	852	1238	1238	1932
Fixed-Effects	Y	Y	Y	Y
Year Dummies	Y	Y	Y	Y
R-Squared	0.18	0.16		

*p<.05; **p<.01; ***p<.001

Standard errors adjusted for clustering on judicial circuit.

Models 1 & 3: Death Eligible Subsample.

Models 2 & 4: Murder Conviction Subsample.

Models 3 & 4: Multiple Imputation Estimates.

Pr(Trial) = Probability of trial without a death noticed filed.

All specifications include controls for the number of co-defendants, victims, statutory aggravating factors, contemporaneous felony convictions, prior felony convictions, and children of defendant; the race/ethnicity, gender, and age of defendant and victim(s); defendant's employment status at time of arrest, marital status, high school graduation status; offender/victim relationship; and whether the homicide was interracial.

As noted *supra*, another concern with the models estimated in this study is possible endogeneity bias²⁰¹ resulting from either case-level

²⁰¹ Endogeneity occurs when “the values our explanatory variables take on are sometimes a consequence, rather than a cause, of our dependent variable.” KING ET AL., *supra* note 174, at 185.

The level of concern over endogeneity bias in observational studies varies across disciplines—e.g., econometricians tend to be much more concerned about endogeneity than, for example, epidemiologists, biostatisticians, psychologists, statisticians, and other social scientists. See Antonakis et al., *supra* note 183, at 1100 (remarking that attention to remedying possible endogeneity bias “has not had a big impact on other social science disciplines including psychology and management research”); Robert Gibbons, *What Is Economic Sociology and Should Any Economists Care?*, 19 J. ECON. PERSP. 3, 6 (2005);

omitted variable bias or possible reverse causality.²⁰² Due to the fact that in nonexperimental research, predictor and outcome variables may covary because of factors outside the control (and knowledge) of the researcher,²⁰³ standard regression techniques will result in biased and inconsistent estimators when unobserved factors affecting the response are correlated with unobserved factors affecting the causal variable of interest.²⁰⁴ While the problems of omitted variable bias and reverse causality may be theoretically distinct, they result in the same source of bias—correlation between the causal variable and the unobserved factors affecting the response variable²⁰⁵—so similar corrective approaches can be used to address both forms of this potential bias. Three alternative approaches were employed to examine the robustness of the previously reported results.

First, a nonrecursive simultaneous-equation model in which a dependent variable indicating selection into a treatment group (i.e., DPNotice) appears as an explanatory variable in a substantive equation predicting the outcome (i.e., Trial).²⁰⁶ “[M]ultiple equation models have

James J. Heckman, *The Scientific Model of Causality*, 35 SOC. METHODOLOGY 1, 5 (2005) (noting that epidemiological and statistical models often fail to take into account simultaneous causality and other sources of randomness generating unobservables in their models); S. Rabe-Hesketh & A. Skrondal, *Parameterization of Multivariate Random Effects Models for Categorical Data*, 57 BIOMETRICS 1256, 1256 (2001) (explaining that econometricians have given greater attention to identification problems than biostatisticians).

²⁰² KING ET AL., *supra* note 174, at 185. Measurement error is a third source of endogeneity bias, but is not of particular concern in these analyses.

²⁰³ *Id.* at 186.

²⁰⁴ James J. Heckman, *Dummy Endogenous Variables in a Simultaneous Equation System*, 46 ECONOMETRICA 931, 931 (1978).

²⁰⁵ KING ET AL., *supra* note 174, at 185.

²⁰⁶ Simultaneous equation models can be divided into two major types: recursive and nonrecursive. A nonrecursive model occurs when there are reciprocal relationships (i.e., feedback loops) between the outcome variables of two or more equations in the system or at least some of the disturbances are correlated. PAMELA M. PAXTON ET AL., NONRECURSIVE MODELS: ENDOGENEITY, RECIPROCAL RELATIONSHIPS, AND FEEDBACK LOOPS 13 (2011); *accord* DAVID KAPLAN, STRUCTURAL EQUATION MODELING: FOUNDATIONS AND EXTENSIONS 16–17 (2000) (noting that nonrecursive models have non-zero off-diagonal elements in the residual variance–covariance matrix); Rex B. Kline, *Reverse Arrow Dynamics: Formative Measurement and Feedback Loops*, in STRUCTURAL EQUATION MODELING: A SECOND COURSE 43, 56 (Gregory R. Hancock & Ralph O. Mueller eds., 2006) (“Nonrecursive models have feedback loops or disturbance covariances for endogenous variables with direct effects between them.”). Most econometricians, however, refer to models with correlated disturbances as recursive models if there are no feedback loops present. *E.g.*, GREENE, *supra* note 174, at 659 (explaining that a model is recursive when the matrix of coefficients of the endogenous variables is triangular); *accord* Joachim Wilde, *Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors*, 69 ECON. LETTERS 309, 310 (2000). For the purposes of these analyses, the simultaneous models are labeled nonrecursive to maintain consistency with the larger structural equation modeling literature.

been the key tools for many researchers [to] study complicated cause-and-effect relationships. . . . The regression equations are explicitly meant to represent the mechanisms by which causes have their effects.”²⁰⁷ This model is typically referred to as an *endogenous bivariate probit* or *endogenous switching model* due to the fact that DPNotice is a binary variable and the observational units (i.e., defendants) are allocated to a specific regime (i.e., death noticed/non-death noticed) depending on the value of this decision variable.²⁰⁸ The model is estimated from the following equation:

$$\pi_{jk} = \Phi_2[d_2(\gamma_2 X_2 + \beta_2 \times \text{DPNotice} + \lambda \xi + \delta_2) + d_1(\gamma_1 X_1 + \xi + \delta_1)], \quad [3]$$

where π_{jk} can represent four different joint probabilities, depending on the values of the Trial and DPNotice variables (j indexes the binary outcome for Trial and k indexes the binary outcome for DPNotice).²⁰⁹ So, for example, $\pi_{11} = \text{Pr}(\text{Trial} = 1, \text{DPNotice} = 1)$, $\pi_{10} = \text{Pr}(\text{Trial} = 1, \text{DPNotice} = 0)$, etc., and d_1 and d_2 are signs variables, being equal to 1 or -1 depending on whether the observed binary outcome equals 1 or 0.²¹⁰ The parameters γ

Labeling the simultaneous models nonrecursive also underscores the fact that the models take into account possible endogeneity bias resulting from reciprocal causation.

Bivariate logistic models were also estimated and produced nearly identical results. Bivariate probit models are generally preferred in the literature because the various extant multivariate logistic distributions have properties such as restrictions on possible values of correlation coefficients and asymmetric nonelliptical distributions that make such a direct approach less practical, and convergence problems are common. Barry C. Arnold, *Multivariate Logistic Distributions*, in HANDBOOK OF THE LOGISTIC DISTRIBUTION 237, 244–45 (N. Balakrishnan ed., 1992); Murray D. Smith & Peter G. Moffatt, *Fisher’s Information on the Correlation Coefficient in Bivariate Logistic Models*, 41 AUSTL. & N.Z. J. STAT. 315, 317–19 (1999).

²⁰⁷ BERK, *supra* note 174, at 190; accord ERIC A. HANUSHEK & JOHN E. JACKSON, STATISTICAL METHODS FOR SOCIAL SCIENTISTS 227 (1977) (“[S]tructural [i.e., multiple] equations represent the theoretical model hypothesized to underlie the observed data; this is the causal structure assumed to generate the data.”).

²⁰⁸ SKRONDAL & RABE-HESKETH, *supra* note 178, at 437. The endogenous switching model with a binary outcome is also known as the “multivariate probit model with structural shift.” Heckman, *supra* note 204, at 932.

²⁰⁹ Note that some subscripts have been omitted to simplify the presentation. The standard representation of simultaneous equation models lists β as the effect of endogenous variables and γ as the effect of exogenous variables. See PAXTON ET AL., *supra* note 206, at 4.

²¹⁰ Alfonso Miranda & Sophia Rabe-Hesketh, *Maximum Likelihood Estimation of Endogenous Switching and Sample Selection Models for Binary, Ordinal, and Count Variables*, 6 STATA J. 285, 288 (2006); accord Lorenzo Cappellari & Stephen P. Jenkins, *Calculation of Multivariate Normal Probabilities by Simulation, with Applications to Maximum Simulated Likelihood Estimation*, 6 STATA J. 156, 166 (2006). Technically, $d_m = 2y_m - 1$, where m indexes the particular equation ($m = 1, 2$). So when $y = 1$, $d_m = 1$ and when

and β , as well as the variables X and $DPNotice$, are defined the same as in Equation [1],²¹¹ δ is the error term that is unique to each equation, and Φ_2 is the bivariate normal cumulative distribution function (CDF).²¹² ξ (ξ) is a second-order latent variable—i.e., a latent variable whose indicators are themselves latent variables²¹³—and “merely represents the combined effect of all unobserved covariates.”²¹⁴ The inclusion of latent variables “in statistical models is a common way of taking unobserved heterogeneity into account.”²¹⁵ λ (λ) is a factor loading, representing the covariance between the disturbances (i.e., the covariance between the omitted variables after the influence of the included factors are accounted for). Due to the standardized parameterization of the model—i.e., the variables ξ and δ are normally distributed with mean equal to zero and variance equal to one—the covariance (λ) is equal to the correlation, ρ

$y = 0, d_m = -1$. *Id.*

²¹¹ GREENE, *supra* note 175, at 852–56 (noting that the endogenous nature of an explanatory variable can be ignored in formulating the log-likelihood in the bivariate probit model). Including observed endogenous variables in a system of probit equations yields likelihoods whose maximization generates consistent parameter estimates. G.S. MADDALA, LIMITED-DEPENDENT AND QUALITATIVE VARIABLES IN ECONOMETRICS 122–23 (1983).

²¹² GREENE, *supra* note 175, at 849–52, 854.

²¹³ The latent indicators are the “first-order” factors and “may be found to satisfy a factor analytic model themselves.” KENNETH A. BOLLEN, STRUCTURAL EQUATIONS WITH LATENT VARIABLES 313 (1989) (“Less widely appreciated is that more general and abstract latent variables may determine the ‘first-order’ latent variables.”); SKRONDAL & RABE-HESKETH, *supra* note 178, at 18 (“[L]atent variables pervade modern statistics and . . . are used to represent widely different phenomena such as true variables measured with error, hypothetical constructs, unobserved heterogeneity, missing data, counterfactuals and latent responses underlying categorical variables.”); David W. Gerbing & James C. Anderson, *On the Meaning of Within-Factor Correlated Measurement Errors*, 11 J. CONSUMER RES. 572, 574 (1984); Anders Skrondal & Sophia Rabe-Hesketh, *Latent Variable Modelling: A Survey*, 34 SCANDINAVIAN J. STAT. 712, 712 (2007) (“[L]atent variables are referred to by different names in different parts of statistics, examples including ‘random effects’, ‘common factors’, ‘latent classes’, ‘underlying variables’ and ‘frailties’.”).

²¹⁴ SKRONDAL & RABE-HESKETH, *supra* note 178, at 9; Heckman, *supra* note 204, at 935 (“[T]he error term in each equation consists of the sum of continuous and discrete random variables that are correlated.”).

Second-order factor models have at least three distinct advantages: (1) permit the testing of whether the hypothesized factor accounts for the relationships between the first-order factors (i.e., the residuals); (2) impose a structure on the pattern of the covariance between the first-order factors; and (3) separate the variance due to specific factors from measurement error. Fang Fang Chen et al., *Testing Measurement Invariance of Second-Order Factor Models*, 12 STRUCTURAL EQUATION MODELING 471, 473 (2005).

²¹⁵ SKRONDAL & RABE-HESKETH, *supra* note 178, at 9; *see also* Francesca Francavilla et al., *Mothers’ Employment and Their Children’s Schooling: A Joint Multilevel Analysis for India*, 41 WORLD DEV. 183, 186 (2012) (“Systems of random effects [i.e., latent variables] equations have been used to deal with endogenous covariates In such cases the outcome of an equation appears as a covariate in another equation.”).

(ρ).²¹⁶ The bivariate probit model assumes that unobserved factors influencing the treatment variable (i.e., death penalty notice) and the outcome variable (i.e., case disposed by trial) manifest themselves in the correlation of the error terms of the two equations.²¹⁷

In contrast to Equation [1], the model now represents a system of equations, so the numbered subscripts refer to the different equations (e.g., X_1 and X_2 index the explanatory variables for, respectively, the death-noticing and trial-decision equations), where the main response (i.e., Trial) and the switching dummy (i.e., DPNotice) are nested (i.e., clustered) within cases.²¹⁸ This is easily shown by writing the two equations separately:

$$\Pr(\text{DPNotice} = 1) = \Phi(\gamma_1 X_1 + \zeta_1) \quad [4a]$$

$$\Pr(\text{Trial} = 1) = \Phi(\gamma_2 X_2 + \beta_2 \times \text{DPNotice} + \zeta_2). \quad [4b]$$

The relationship between the errors across the two equations can be observed with the following equations:

$$\zeta_1 = \xi + \delta_1 \quad [5a]$$

$$\zeta_2 = \lambda\xi + \delta_2, \quad [5b]$$

where the error in each equation consists of a part that is unique to that

²¹⁶ Miranda & Rabe-Hesketh, *supra* note 210, at 287–88.

²¹⁷ Robert C. Luskin, *Estimating and Interpreting Correlations Between Disturbances and Residual Path Coefficients in Nonrecursive (and Recursive) Causal Models*, 22 AM. J. POL. SCI. 444, 450 (1978) (“Stated simply, the correlation between the disturbances of the [] structural equations expresses the extent to which those equations fail to recognize major causes of their dependent variables that are either the same or correlated.”). These disturbances represent effects of random influences or omitted covariates that are case-specific and immediate, whereas the fixed-effects models represent the effects of omitted influences that remain constant within a particular jurisdiction or year. GREENE, *supra* note 175, at 852–56 (explaining that the key advantage of the bivariate probit model is its ability to explicitly control for unobservable confounding factors); SKRONDAL & RABE-HESKETH, *supra* note 178, at 87. This approach shares similarities with propensity score adjustment, which has been used in prior research to estimate the causal impact of filing a notice to seek the death penalty on associated costs. Roman et al., *supra* note 27, at 556 (noting that propensity score models are a viable solution to modeling selection bias in models with binary treatments); see also Paul R. Rosenbaum & Donald B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 BIOMETRIKA 41, 55 (1983). In fact, sensitivity analyses reveal that the estimates from propensity score models were very similar. See *infra* Part VI.B.

²¹⁸ Miranda & Rabe-Hesketh, *supra* note 210, at 288.

equation, δ , and a second part, ξ , that is common to both.²¹⁹ Each error term (ζ) now depends, in part, on the value of ξ , which in turn means that ζ_1 and ζ_2 will be related to one another.²²⁰

Substituting [5a] into [4a] and [5b] into [4b] gives the following equations:

$$\Pr(\text{DPNotice} = 1) = \Phi(\gamma_1 X_1 + \xi + \delta_1) \quad [6a]$$

$$\Pr(\text{Trial} = 1) = \Phi(\gamma_2 X_2 + \beta_2 \times \text{DPNotice} + \lambda \xi + \delta_2), \quad [6b]$$

and combining [6a] and [6b] results in Equation [3].²²¹ The relationship between the two decision processes may be more easily observed in the (causal) path diagram of the model depicted in Figure 2. The bivariate probit model takes into account any unobserved causes influencing both decision processes through ξ , including potential simultaneous/reverse causality,²²² so β_2 can be considered the unbiased causal effect of the threat of the death penalty (via the filing of a notice of intent to seek the death penalty) on the probability of a case being disposed by trial.²²³

²¹⁹ BOLLEN, *supra* note 213, at 314; SKRONDAL & RABE-HESKETH, *supra* note 178, at 91; Gerbing & Anderson, *supra* note 213, at 574–76.

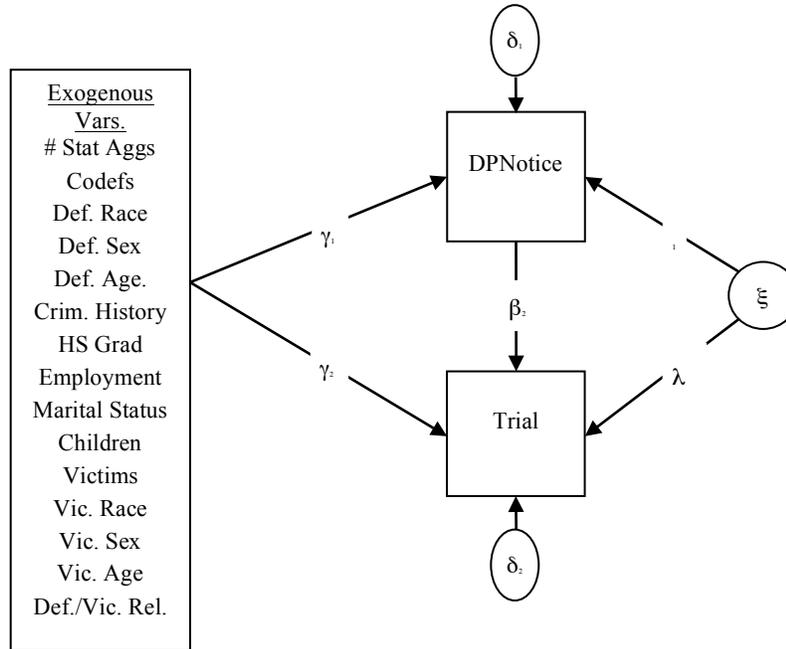
²²⁰ “We can induce dependence between responses by including factor structures [i.e., latent variables] in the linear predictor.” SKRONDAL & RABE-HESKETH, *supra* note 178, at 91.

²²¹ Technically, the combination of Equations [6a] and [6b] results in a specific instance of Equation [3]: when $\pi_{jk} = \pi_{jl}$, therefore, Equation [3] represents a more general formulation because it can estimate all four of the joint probabilities.

²²² The model depicted in Figure 2 does take into account both omitted variables and potential reverse causality, even though it does not estimate a causal relationship from the plea-decision variable to the death-noticing variable. Although the statistical literature is replete with incorrect examples of feedback arrows between two endogenous variables in a system of simultaneous equations, as well as correlated disturbances, such a formulation has been shown to be logically inconsistent in the bivariate probit context. The correlated disturbances properly adjust for reciprocal causation. See Peter Schmidt, *Constraints on the Parameters in Simultaneous Tobit and Probit Models*, in STRUCTURAL ANALYSIS OF DISCRETE DATA WITH ECONOMETRIC APPLICATIONS 422, 427 (Charles F. Manski & Daniel McFadden eds., 1981); see also MADDALA, *supra* note 211, at 117–18 (same).

²²³ BOLLEN, *supra* note 213, at 314 (explaining that second-order factors can eliminate bias resulting from correlated measurement errors). The model is fitted via maximum likelihood and the unobserved heterogeneity term, ξ , is integrated out using either ordinary Gauss–Hermite quadrature or adaptive quadrature. Miranda & Rabe-Hesketh, *supra* note 210, at 288.

Figure 2
Path Diagram of Model Predicting Trial



The two equations share identical explanatory variables except for DPNotice, which only appears in the trial equation. No exclusion restrictions are required to identify the model—the multi-equation probit model is identified as long as each equation contains one varying exogenous variable.²²⁴ In fact, it is unlikely that a valid “instrument” exists—that is, a variable that induces substantial variation in the endogenous covariate is independent of unmeasured confounders and has no direct effect on the outcome.²²⁵ Researchers have recognized that

²²⁴ SKRONDAL & RABE-HESKETH, *supra* note 178, at 438 (“Although beneficial for identification, the [exclusion] restrictions are not necessary for identification.”); Heckman, *supra* note 204, at 957 (remarking that “the restriction to exclusion restrictions is overly stringent” and that “[i]dentification through use of covariance restrictions is also permitted”); Giampiero Marra & Rosalba Radice, *Estimation of a Semiparametric Recursive Bivariate Probit Model in the Presence of Endogeneity*, 39 CAN. J. STAT. 259, 263 (2011) (noting that theoretical identification in the recursive bivariate probit context does not require the availability of any instrumental variables because the linear combination of the two equations does not contain the same variables as the original); Wilde, *supra* note 206, at 312 (exclusion restrictions are only necessary if there is no variation in exogenous regressors, and this is a rather weak assumption in economic applications).

²²⁵ This underlying identifying assumption of the instrumental variable approach is both

situations frequently arise in practice where identical explanatory variables influence selection and a subsequent outcome of interest, and the analysts may be required to base identification on distribution assumptions about the residuals alone.²²⁶ The assumption of joint normality of the residual terms in the bivariate probit model may be reasonable under certain weaker assumptions: (1) the selection equation and the equation of interest represent closely related decisions or goals; (2) the decisions have the same causes; and (3) the decisions occur within a short time frame or are close to each other geographically.²²⁷ These conditions would appear to hold rather well for the current analysis. Prior simulation studies also report that the bivariate probit model outperforms instrumental variable models in many applications and is rather robust to nonnormality of error terms, especially estimated covariate effects and variance of the random effects.²²⁸ Moreover, analysts have established that identification is less of a concern when causal effects, rather than structural parameters, are of primary interest.²²⁹ In fact, calculation of the treatment effect in the bivariate probit

very strong and *unverifiable*. MORGAN & WINSHIP, *supra* note 91, at 196–200 (detailing the shortcomings of traditional instrumental variable estimation and explaining how analysts are mistaken when believing the assumption is empirically testable).

²²⁶ See Anne E. Sartori, *An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions*, 11 POL. ANALYSIS 111–12 (2003); *cf.* GREENE, *supra* note 175, at 616 (“The case of identical regressors [across a system of equations] is quite common . . .”).

²²⁷ Sartori, *supra* note 226, at 112; *cf.* James J. Heckman et al., *Matching as an Econometric Evaluation Estimator*, 65 REV. ECON. STUD. 261, 264–65 (1998) (discussing bias in estimation of treatment effects resulting from geographic mismatch).

²²⁸ See, e.g., Jay Bhattacharya et al., *Estimating Probit Models with Self-Selected Treatments*, 25 STAT. MED. 389, 399–402 (2006); Marra & Radice, *supra* note 224, at 260 (“[I]t is well known, from both a theoretical and empirical point of view, that simultaneous likelihood estimation methods are superior to conventional two-stage instrumental variable procedures.”); Charles E. McCulloch & John M. Neuhaus, *Misspecifying the Shape of a Random Effects Distribution: Why Getting It Wrong May Not Matter*, 26 STAT. SCI. 388, 400 (2011) (“Theory and simulation studies indicate that most aspects of statistical inference are highly robust to this assumption [of normality for random effects] . . . including estimation of covariate effects, [and] estimation of the random effects variance . . .”); Gary Young et al., *Multivariate Probit Models for Conditional Claim-Types*, 44 INS.: MATHEMATICS & ECON. 214, 222 (2009).

²²⁹ Joshua D. Angrist, *Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice*, 19 J. BUS. & ECON. STAT. 2, 3–5 (2001) (endorsing the use of the potential outcomes/counterfactual framework instead of structural parameters, particularly in models examining limited dependent variables (LDV) with dummy endogenous variables and also remarking that identifying assumptions for structural parameters are largely unnecessary for causal inference in LDV models in light of the potential outcomes/counterfactual framework); Guido W. Imbens & Joshua D. Angrist, *Identification and Estimation of Local Average Treatment Effects*, 62 ECONOMETRICA 467, 475 (1994) (same).

model directly lends itself to the counterfactual/potential outcomes framework.²³⁰ A case “is only allocated to one of the regimes and never both, the responses in the regimes thus represent *potential outcomes*.”²³¹ As Nobel Prize-winning econometrician James Heckman explains, the “Neyman-Fisher-Cox-Rubin model of potential outcomes . . . is also the switching regression model of Quandt.”²³²

The average marginal effect of DPNotice on the likelihood of trial is the difference between two conditional probabilities: $\Pr(\text{Trial} = 1 | \text{DPNotice} = 1, X_1, X_2) - \Pr(\text{Trial} = 1 | \text{DPNotice} = 0, X_1, X_2)$. In words, the marginal effect is the probability that a case results in trial, given that a death penalty notice has been filed in that case, *minus* the probability that a case results in a trial, given that a death penalty notice has not been filed in that case.²³³ The results from the bivariate probit model support the earlier analyses (see Table 8). Being noticed for the death penalty decreases the probability of a case going to trial from .61 to .34 in the DE model. The correlation, *rho* (ρ), of the residuals across the two equations after controlling for the covariates is statistically significant ($\rho = -.559$, $p < .05$). In the MC model, being noticed for the death penalty decreases the probability of a case going to trial from .61 to .36. The correlation between the disturbances was statistically insignificant ($\rho = -.518$, $p < .05$).²³⁴ These effect sizes are somewhat larger than those obtained from the previous models, but it is important to emphasize that the bivariate probit models do *not* control for unobserved judicial-circuit-level factors, and this is likely to impact the causal estimates. The proportion of the variance in the residuals that is attributable to shared omitted variables is equal to the square of the

²³⁰ GREENE, *supra* note 175, at 853; Bhattacharya et al., *supra* note 228, at 400.

²³¹ SKRONDAL & RABE-HESKETH, *supra* note 178, at 437.

²³² James J. Heckman & Edward J. Vytlacil, *Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects*, 96 PROC. NAT'L ACAD. SCI. 4730, 4730 (1999) (citing Richard E. Quandt, *A New Approach to Estimating Switching Regressions*, 67 J. AM. STAT. ASS'N 306 (1972)).

²³³ GREENE, *supra* note 175, at 853. The conditional probabilities are obtained by: $\Pr(\text{Trial}=1 | \text{DPNotice}=1, X_1, X_2) = \Pr(\text{Trial}=1, \text{DPNotice}=1 | X_1, X_2) / (\Pr(\text{DPNotice}=1 | X_1))$ and $\Pr(\text{Trial}=1 | \text{DPNotice}=0, X_1, X_2) = \Pr(\text{Trial}=1, \text{DPNotice}=0 | X_1, X_2) / (\Pr(\text{DPNotice}=0 | X_1))$.

²³⁴ Even if the error terms from the two equations are not correlated in the overall population (as indicated by significant tests), they are correlated in the selected sample and can bias parameter estimates. See CHRISTOPHER H. ACHEN, *THE STATISTICAL ANALYSIS OF QUASI-EXPERIMENTS* 73–81 (1986); Steven Goodman, *A Dirty Dozen: Twelve P-Value Misconceptions*, 45 SEMINARS HEMATOLOGY 135, 136 (2008) (stating that the “effect best supported by the data from a given experiment is always the observed effect, regardless of its significance”). Nonetheless, the close correspondence between the single-equation results and the bivariate probit results strongly suggest that any bias in the estimates is minimal. Moreover, additional sensitivity checks reveal that the parameter estimates were not unduly influenced by endogeneity bias.

correlation coefficient, ρ , across the models.

Second, a semi-nonparametric version of the bivariate probit model (SNP) was analyzed to check the robustness of the previously estimated parametric bivariate model.²³⁵ The fully parametric bivariate probit model in Equation [3] assumes joint normality of residuals, and although simulation studies suggest the model is rather robust to nonnormality,²³⁶ bias in the causal estimates resulting from the violation of that assumption remains a possibility. The SNP model makes less restrictive assumptions about the distribution of unobservables, and therefore can handle a broader class of error distributions.²³⁷ The intuition behind the SNP approach is to approximate the unknown distributions of the residuals by Hermite polynomial expansions (i.e., the product of a squared polynomial and a normal density) and use the approximations to derive a pseudo-maximum likelihood estimator for the model parameters.²³⁸ The polynomial expansion can handle distributions with arbitrary skewness and kurtosis.²³⁹ The SNP models reveal that the threat of capital punishment reduces probability of trial by .249 (DE) and .258 (MC).²⁴⁰ These estimates are similar to those obtained via the classic bivariate probit model. The estimates are also somewhat larger than those obtained from the fixed-effects logit and linear probability models, however the SNP models also do not account for judicial-circuit-level fixed effects.

Finally, a propensity score-matching model is used to adjust for nonrandom selection into capital prosecution.²⁴¹ The intuition behind the model is that bias in estimates of treatment effects is reduced when the comparison of outcomes is performed using “treated” and “control” subjects who are as similar as possible along a large number of relevant dimensions.²⁴² Exact matching,²⁴³ or even coarsened exact matching,²⁴⁴ is

²³⁵ Giuseppe De Luca, *SNP and SML Estimation of Univariate and Bivariate Binary-Choice Models*, 8 STATA J. 190, 192 (2008); Marra & Radice, *supra* note 224, at 259–60.

²³⁶ See Bhattacharya et al., *supra* note 228; Young et al., *supra* note 228.

²³⁷ Siegfried Gabler et al., *Seminonparametric Estimation of Binary-Choice Models with an Application to Labor-Force Participation*, 11 J. BUS. & ECON. STAT. 61, 63 (1993).

²³⁸ Gabler et al., *supra* note 237, at 63.

²³⁹ De Luca, *supra* note 235, at 194; Mark B. Stewart, *Semi-Nonparametric Estimation of Extended Ordered Probit Models*, 4 STATA J. 27, 30 (2004).

²⁴⁰ DE model: $\rho = -.324, p < .10$; MC model: $\rho = -.123, p > .10$.

²⁴¹ See, e.g., Roman et al., *supra* note 27, at 556–58 (employing propensity score models to account for nonrandom selection into prosecution).

²⁴² Rosenbaum & Rubin, *supra* note 217, at 55.

²⁴³ Daniel E. Ho et al., *Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference*, 15 POL. ANALYSIS 199, 217 (2007).

²⁴⁴ Stefano M. Iacus et al., *Causal Inference Without Balance Checking: Coarsened Exact Matching*, 20 POL. ANALYSIS 1 (2012).

typically infeasible when the number of relevant variables is large, so propensity scores summarize pretreatment characteristics into a single-index variable.²⁴⁵ By definition, capital and noncapital cases with the same value of the propensity score have the same distribution of the full vector of regressors, so it is sufficient to match cases on their propensity score to obtain the same probability distribution of X for treated and nontreated match samples.²⁴⁶ Propensity score models rest on the strong, yet unverifiable, assumption that differences between cases are captured by their observable attributes (“conditional independence assumption”).²⁴⁷ This assumption is unlikely to hold perfectly with these data; nonetheless, propensity score models have been shown to reduce, but not eliminate, the bias generated by unobservable confounding factors. The extent to which this potential bias is reduced depends on the richness and quality of control variables used to compute the propensity scores. Thus, the models are estimated for the simple purpose of comparing their results to the previously estimated models in order to determine how sensitive the estimates are to model specifications. The propensity score (PS) is the conditional probability of a death-notice filing and can be expressed as:

$$\Pr(\text{DPNotice} = 1) = \frac{\exp(\alpha_c + \gamma X + \zeta)}{1 + \exp(\alpha_c + \gamma X + \zeta)}, \quad [7]$$

where α , γ , and ζ are defined in Equation [1].²⁴⁸ As is customary in the literature, the estimation sample is restricted to the region of common support: 20 cases were dropped from the DE model ($N = 832$) and 175 cases were dropped from the MC model ($N = 1063$).²⁴⁹ In practice, no two

²⁴⁵ Rosenbaum & Rubin, *supra* note 217.

²⁴⁶ The propensity scores satisfy the balance condition when observations with the same propensity score have the same distribution of observable (and hopefully unobservable) characteristics independent of treatment status—*i.e.*, for a given propensity score, exposure to the treatment is random so treatment and control units should be, on average, observationally identical. *Id.*; see also Sascha O. Becker & Andrea Ichino, *Estimation of Average Treatment Effects Based on Propensity Scores*, 2 STATA J. 358, 360 (2002) (describing the standard balancing algorithm).

²⁴⁷ Rosenbaum & Rubin, *supra* note 217, at 43.

²⁴⁸ Some subscripts have been omitted for simplicity.

²⁴⁹ The region of common support is the overlapping distribution of propensity scores for the treatment and the control groups. Roman et al., *supra* note 27, at 557 (explaining that “the average treatment effect is defined only in the region of common support”); see Gary King & Langche Zeng, *The Dangers of Extreme Counterfactuals*, 14 POL. ANALYSIS 131, 146–151 (2006) (noting that using data outside the region of common support induces some degree of model dependence and increases the risk of bias for almost any model chosen); see also Iacus et al., *supra* note 244, at 11 (same).

variables will share the exact same propensity score because it is a continuous variable, so a kernel-matching estimator is used.²⁵⁰ Following the matching of cases based on propensity scores, the average treatment effect can be estimated by simply taking the difference in the potential outcomes in the two counterfactual situations.²⁵¹ Again, the results from the propensity score models should be viewed with caution because of the likelihood of bias resulting from unobservable confounding factors. With this caveat in mind, it is worth noting that propensity score models corroborate results from the previously estimated models: a death notice decreased the probability of trial by .25 and .24 in, respectively, the DE and MC models.²⁵²

In summary, the current study was able to examine directly the causal impact of the death penalty on defendants' propensity to go to trial. Estimates from the various specifications suggest that the death penalty decreases the probability of a trial anywhere from .17 to .27, from a baseline probability of approximately .60, although estimates in the range of .17 to .23 are likely to be most accurate due to the consideration of unobserved heterogeneity across judicial circuits and years. As explained earlier this roughly equates to the death penalty deterring two out of every ten death-noticed defendants from pursuing a trial.

²⁵⁰ James J. Heckman et al., *Matching as an Econometric Evaluation Estimator*, 65 REV. ECON. STUD. 261, 271 (1998) (describing kernel-regression-based matching estimators). Matching estimators based on nearest neighbors, local linear regression, and Mahalanobis distances provided similar results. See also MORGAN & WINSHIP, *supra* note 91, at 107–16 (describing various matching estimators).

²⁵¹ Iacus et al., *supra* note 244, at 1; see also *supra* note 92 and accompanying text for discussion of the calculation of treatment effects under the counterfactual framework.

²⁵² Results not reported, but available from author upon request.

Table 8
Marginal Effect of Death Penalty on Trial (Multivariate Models)

	Model 1	Model 2	Model 3	Model 4
DP Notice	-0.268*** (0.048)	-0.265*** (0.041)	-0.246*** (0.047)	-0.258*** (0.034)
Pr(Trial)	0.607	0.606	0.684	0.622
N	852	1238	852	1238
Fixed-Effects	N	N	N	N
Year Dummies	Y	Y	Y	Y
R-Squared	N/A	N/A	N/A	N/A

*p<.05; **p<.01; ***p<.001

Standard errors adjusted for clustering on judicial circuit.

Models 1 & 3: Death Eligible Subsample.

Models 2 & 4: Murder Conviction Subsample.

Models 3 & 4: Semi-Parametric Model.

Pr(Trial) = Probability of trial without a death noticed filed.

All specifications include controls for the number of codefendants, victims, statutory aggravating factors, contemporaneous felony convictions, prior felony convictions, and children of defendant; the race/ethnicity, gender, and age of defendant and victim(s); defendant's employment status at time of arrest, marital status, high school graduation status; offender/victim relationship; and whether the homicide was interracial.

Defendants generally prefer charge bargains to sentencing bargains because a less serious charge is accompanied by a lower penalty (and, perhaps, less severe collateral consequences).²⁵³ Unfortunately, it was impossible to examine directly the impact of the threat of the death penalty on charge bargaining because the data used in this study do not contain information on whether the prosecutor modified her initial charge. As discussed in Part V.A, however, it is debatable whether much charge bargaining occurs when a defendant is initially charged with murder (and a prima facie case exists for the charge).²⁵⁴ Recall that Georgia's murder

²⁵³ Kuziemko, *supra* note 20, at 126.

²⁵⁴ COHEN & KYCKELHAHN, *supra* note 2, at 10–11 (explaining that the vast majority of

statute does not include “degrees” of murder like many other states. A defendant is either charged with murder (mandatory minimum life sentence), voluntary manslaughter (twenty-year maximum), or involuntary manslaughter (ten-year maximum).²⁵⁵ The statutory minimum for both manslaughter offenses is one year, and inmates convicted of manslaughter are generally eligible for parole after serving 65% of their sentences.²⁵⁶ The dramatic reduction in potential punishment can make it politically costly for a prosecutor to offer a charge bargain from murder to manslaughter simply to avoid trial—especially when he is faced with pressure from victims’ families and the electorate.²⁵⁷

Although prosecutors may lack the flexibility to offer charge or sentence bargains in the noncapital context,²⁵⁸ they still may present defendants with other incentives to plea. Such incentives might include: (1) dropping less serious offenses against the defendant, (2) dropping charges against codefendants,²⁵⁹ (3) keeping potentially embarrassing facts about the crime or defendant confidential, (4) asking the judge to impose multiple sentences concurrently rather than consecutively, (5) assisting with detention facility placement, or (6) agreeing to assist with parole board hearings. According to the Georgia data, approximately 40% of noncapital murder convictions were obtained via a plea bargain. This strongly suggests that noncapital murder defendants are willing to accept plea bargains for consideration other than a charge or sentence reduction. Charge bargaining in the capital context seems equally unlikely.²⁶⁰ That is, it is doubtful that prosecutors would seek the death penalty against

defendants initially charged with murder are ultimately convicted of murder, irrespective of method of adjudication).

²⁵⁵ GA. CODE ANN. § 16-5-2(b) (West 2003) (voluntary manslaughter); § 16-5-3(a) (involuntary manslaughter).

²⁵⁶ *Id.* §§ 16-5-2(b), 16-5-3(a).

²⁵⁷ See *supra* notes 51–53 and accompanying text. Both scholars and victims’ rights groups have advocated expanding victims’ involvement in plea negotiations, including the ability to reopen a plea or sentence when the accused has pleaded to a reduced offense. Kyl et al., *supra* note 43, at 621. Others have noted that state prosecutors are more restricted than federal prosecutors with respect to offering charge bargains due to actual and perceived political constraints and consequences. Gazal-Ayal, *supra* note 15, at 2306; Richman & Stuntz, *supra* note 53, at 600–05; see also WILLIAM S. MCFEELY, PROXIMITY TO DEATH 19, 80–82 (2000) (describing the intense pressure politicians and prosecutors face from the electorate to aggressively pursue the death penalty).

²⁵⁸ See, e.g., Ehrhard, *supra* note 20, at 322 (describing interviews conducted with prosecutors and defense attorneys who have litigated capital cases and reporting that both the prosecution and defense believe that LWOP is often necessary to get defendants to take an LS plea).

²⁵⁹ Prosecutors may also agree not to file charges against potential codefendants.

²⁶⁰ See *supra* Part V.A.

defendants absent a prima facie case for murder solely to obtain a plea for manslaughter.²⁶¹ The Georgia data, in fact, support this intuition: only two cases that were noticed for the death penalty during the period of the study resulted in a plea bargain for manslaughter.²⁶²

VII. FINANCIAL IMPLICATIONS AND DISCUSSION

The U.S. Supreme Court has reasoned that plea bargaining is “an essential component of the administration of justice” when properly conducted.²⁶³ Plea bargaining is purported to provide benefits in the form of reduced costs, increased efficiency, and certainty to defendants, prosecutors, defense counsel, judges, and victims.²⁶⁴ When plea bargaining in the “shadow of death,” incentives for parties to plea bargain may be significantly magnified, although some have suggested that disincentives to bargain may be overriding.²⁶⁵ Not only has the Court approved the use of the death penalty by prosecutors to secure guilty pleas,²⁶⁶ but prosecutors have openly stated that they (or other prosecutors with whom they are familiar) routinely use the death penalty as leverage in plea negotiations.²⁶⁷ Due to the absence of methodologically rigorous research on the subject, however, the actual impact of the threat of capital punishment on plea-bargaining dynamics has been largely speculative. This Article has demonstrated that, at least in Georgia, the death penalty does reduce the total number of cases proceeding to trial. Based on the magnitude of that effect, however, it does not appear that the threat of the death penalty deters enough murder trials to be cost-effective. This is especially relevant because the high price tag associated with pursuing the death penalty—coupled with the infrequency of death sentences and executions²⁶⁸—has

²⁶¹ See *supra* Part V.A.

²⁶² See *supra* Part V.A.

²⁶³ *Santobello v. New York*, 404 U.S. 257, 260 (1971).

²⁶⁴ See *supra* note 13 and accompanying text; see also Stephanos Bibas, *Harmonizing Substantive-Criminal-Law Values and Criminal Procedure: The Case of Alford and Nolo Contendere Pleas*, 88 CORNELL L. REV. 1361, 1367 (2003) (summarizing procedurally based arguments in favor of plea bargaining).

²⁶⁵ See *supra* Part I. See generally Ehrhard, *supra* note 20, at 313; Hoffmann et al., *supra* note 18, at 2390.

²⁶⁶ *Alford v. North Carolina*, 400 U.S. 25, 31 (1970); *Brady v. United States*, 397 U.S. 742 (1970).

²⁶⁷ See *supra* note 123 and accompanying text (describing instances of prosecutors admitting using the death penalty as leverage in plea negotiations); see also Ehrhard, *supra* note 20, at 319 (describing interviews with prosecutors who admitted that the death penalty is often used as a bargaining chip).

²⁶⁸ Clark Calhoun, Note, *Reviewing the Georgia Supreme Court's Efforts at Proportionality Review*, 39 GA. L. REV. 631, 632 (2005) (underscoring that less than 2% of homicide cases occurring in the modern era of the death penalty have resulted in a death

caused many state and local officials to rethink seriously the feasibility of maintaining the death penalty.²⁶⁹ These concerns have only been exacerbated in recent years due to the current financial crisis.

In 2009, at least eleven state legislatures considered bills to abolish the death penalty, citing associated costs as one of their primary concerns.²⁷⁰ Abolition bills were reintroduced in at least five of those states in 2011.²⁷¹ Georgia lawmakers introduced their own abolition bill in early 2012, also citing the financial burden of administering the death penalty.²⁷² Over the past five years, New Jersey, New Mexico, New York, Illinois, and Connecticut have abolished the death penalty,²⁷³ and the high cost of maintaining a properly functioning death penalty system figured prominently in all of those debates.²⁷⁴ A recent national study suggests that

sentence); Liebman et al., *supra* note 44, at 1844 (showing that 68% of death sentences handed down between 1973 and 1995 were reversed on appeal and less than 2% of death row inmates are executed in any given year).

²⁶⁹ Widespread cuts have been made (or proposed) to courthouse staff, attorney general offices, district attorney offices, and public defender offices. *See, e.g.*, Greg Bluestein, *State Budget Cuts Clog Criminal Justice System*, NBC NEWS (Oct. 26, 2011), http://www.nbcnews.com/id/45049812/ns/us_news-crime_and_courts/. The American Bar Association reported that most states have decreased court funding by 10% to 15%, including significantly scaling back indigent defense and collateral review. *Id.* Nationally, twenty-six states have been unable to fill judgeships for budgetary reasons and 14 states have reduced court staff. In San Francisco, for example, 40% (25 of 63) of the courtrooms have been closed, resulting in huge backlogs in both the civil and criminal dockets. *Id.* State and local governments are also forced to divert funding from hospitals and health care, police and public safety, education, and roads and infrastructure to pay for capital trials. *See* Abby Goodnough, *States Turning to Last Resorts in Budget Crisis*, N.Y. TIMES, June 22, 2009, at A1 (stating that in states across the nation, governors and legislators have recommended increasing taxes and fees, deepening spending cuts, and extending furloughs for government workers in the face of a \$121 billion budget gap).

²⁷⁰ These states included Colorado, Connecticut, Illinois, Kansas, Maryland, Montana, Nebraska, New Hampshire, New Mexico, Texas, and Washington. DIETER, *supra* note 49, at 14; *see also* James S. Liebman, *The New Death Penalty Debate: What's DNA Got to Do with It?*, 33 COLUM. HUM. RTS. L. REV. 527, 528 (2002) (noting that legislation to either abolish or moderate the use of the death penalty was considered in twenty-six death penalty jurisdictions in 2000–2001 and passed a committee vote in at least twelve states).

²⁷¹ *Death Penalty Abolished in Illinois*, PROJECT PRESS (Am. Bar Ass'n Death Penalty Representation Project), Spring 2011, at 1.

²⁷² S. 342, 151st Gen. Assemb., Reg. Sess. (Ga. 2012).

²⁷³ Ashby Jones & Steve Eder, *Costs Test Backing for Death Penalty: Some Former Supporters Say Capital Punishment Isn't Worth Huge Sums Spent on Drawn-Out Cases; Californians to Vote*, WALL ST. J. (Oct. 5, 2012), <http://online.wsj.com/article/SB10000872396390443493304578036792267666794.html>. Prior to 2007, when New Jersey and New York abolished capital punishment, it had been nearly twenty-three years since a state had officially abolished its death penalty. Massachusetts and Rhode Island both dismantled their capital statutes in 1984.

²⁷⁴ *Id.*

each additional capital trial causes an increase in county spending of more than \$2 million and these costs are borne primarily by increasing taxes.²⁷⁵ Such costs have forced many counties to seek help from state legislatures to create programs to diffuse death penalty costs across counties (even those that do not choose to use the death penalty).²⁷⁶ Total taxes and expenditures for capital trials from 1983 to 1999 were more than \$5.5 billion.²⁷⁷

Some of the most thorough examinations of costs associated with the death penalty over the past twenty years have expressly noted that the threat of the death penalty may actually produce financial savings by increasing capital charged defendants' propensity to accept a plea, thereby avoiding trial costs.²⁷⁸ Unfortunately, none of these studies were able to offer any definitive answers to this question because, based on available evidence, the impact of capital punishment on plea bargaining was too speculative.²⁷⁹ Studies in California,²⁸⁰ Indiana,²⁸¹ Kansas,²⁸² and North Carolina²⁸³ all

²⁷⁵ Katherine Baicker, *The Budgetary Repercussions of Capital Convictions*, 4 *ADVANCES ECON. ANALYSIS & POL'Y* 1, 10 (2004) (explaining that death penalty convictions have cost counties more than \$5.5 billion between 1982 and 1999).

²⁷⁶ See Richard Willing & Gary Fields, *Geography of the Death Penalty*, USA TODAY, Dec. 20, 1999, at 1A (reporting that "[f]ifteen counties account for nearly a third of all prisoners sentenced to death [in the United States] but only one-ninth of the population of the states with capital punishment").

Prosecutors from rural and suburban counties in Georgia account for a disproportionate number of death penalty cases. See *supra* note 75. District attorneys from ninety-two counties in Georgia (57.8%) have filed all of the 400 death penalty notices for homicides occurring between 1993 and 2000 (see Table 1). Three counties (1.9%)—Fulton (Atlanta), Dekalb (Decatur), and Chatham (Savannah)—account for nearly one-half of the reported homicides, but only 13.5% of death notices during this period. Particularly interesting is that nearly one-third of Georgia homicides (and 22% of murder convictions) occurred in Fulton County, although only 4.8% of death notices (and 1.9% of death sentences) came from Fulton County between 1993 and 2000. The seven counties with the highest death-noticing rates (i.e., percentage of murder convictions noticed for the death penalty)—Oconee (88%), Morgan (82%), Putnam (73%), Ware (70%), Appling (63%), Bartow (56%), and Lowndes (54%)—account for approximately 3% of the state's population and less than 4% of murder convictions (and less than 3% of total reported homicides), but 13% of death notices between 1993 and 2000. Collectively, these seven counties sought the death penalty in 68% (52 of 77) of homicide cases that ultimately resulted in a murder conviction.

²⁷⁷ Baicker, *supra* note 275, at 1321 (also estimating that each capital case costs approximately \$2.5 million to prosecute); see also *Public Policy Choices on Deterrence and the Death Penalty: A Critical Review of New Evidence: Hearing on H.B. 3834 Before the J. Comm. on the Judiciary*, 2005 Leg. (Mass. 2005) (statement of Jeffrey Fagan, Professor, Columbia University), available at <http://www.deathpenaltyinfo.org/MassTestimonyFagan.pdf> (summarizing studies of the financial cost of the death penalty).

²⁷⁸ DIETER, *supra* note 49, at 16–17 (citing studies that have recognized potential savings from capital statutes by increasing the likelihood of pleas).

²⁷⁹ *Id.*

²⁸⁰ CAL. COMM'N ON THE FAIR ADMIN. OF JUSTICE, *supra* note 94, at 80–81 (commenting

acknowledge that a truly comprehensive study of the cost of the death penalty in their respective jurisdictions would require a close examination of the causal impact of the death penalty on plea bargaining in death-eligible cases. Capital trials are extremely expensive, so numerous plea bargains in potentially capital trials may be required to offset the cost of a single capital trial. Capital cases are more expensive and time-consuming than comparable noncapital murder cases at every stage of the process: more time for pretrial preparation, more pretrial motions, more experts, more attorneys for the defense (and typically the government as well), more time to conduct voir dire, longer trials, longer jury deliberations, and more appeals that take longer.²⁸⁴ Georgia has yet to conduct a comprehensive study of the costs of its capital punishment system, but trial costs for the median capital case have been estimated as at least four times more expensive than the median noncapital murder trial.²⁸⁵ The elapsed time from arrest to murder conviction in capital cases in Georgia for the period

that, in some cases, the risk of the death penalty provides an incentive to plead to life without parole so removing the death penalty might result in more trials, but explaining that even if all murder cases eligible for life without parole went to trial, both trial and appellate costs would be considerably less expensive than capital cases, in large part because of the absence of a penalty phase and right to counsel for habeas petitions).

²⁸¹ KATHRYN JANEWAY, THE APPLICATION OF INDIANA'S CAPITAL SENTENCING LAW: FINDINGS OF THE INDIANA CRIMINAL LAW STUDY COMMISSION 120 (2002) (report prepared for the Office of the Governor and the Indiana General Assembly) (recognizing that plea bargaining also influences the cost of resolving a capital case, but also acknowledging that the impact of the death penalty on plea bargaining may not always be in both parties' best interests).

²⁸² LEGIS. DIV. OF POST AUDIT, STATE OF KAN., PERFORMANCE AUDIT REPORT: COSTS INCURRED FOR DEATH PENALTY CASES: A K-GOAL AUDIT OF THE DEPARTMENT OF CORRECTIONS 19, 32 (2003) (stating a potential benefit of the death penalty savings from avoiding trials, but acknowledging the instant study did not examine those potential savings).

²⁸³ Philip J. Cook & Donna B. Slawson, The Costs of Processing Murder Cases in North Carolina 31 (May 1993) (unpublished manuscript) (on file with the Terry Sanford Inst. of Pub. Policy, Duke Univ.) (noting that 21% of capital charged defendants plead guilty to murder and avoid trial, but concluding that "[s]ince we have no direct evidence on the effect of the death penalty option on the likelihood of trial, and since there are plausible arguments in both directions, we proceed on the assumption that there are neither more nor fewer trials as a result of the death penalty option").

²⁸⁴ See Roman et al., *supra* note 27, at 536 (explaining that case processing is more expensive at every stage of a capital case relative to a noncapital murder case); Robert L. Spangenberg & Elizabeth R. Walsh, *Capital Punishment or Life Imprisonment? Some Cost Considerations*, 23 LOY. L.A. L. REV. 45, 46-47 (1989) (conducting the first cost analysis of every stage of the capital charging-and-sentencing process and concluding that life imprisonment is a more economical alternative).

²⁸⁵ Stephen Gurr, *The High Cost of Death*, GAINESVILLE TIMES, Aug. 10, 2008, at 1D (reporting that defense expenses alone average about \$150,000, and juror and bailiff costs are seven times higher in capital cases).

under investigation in this study was nearly 1.8 times longer than the elapsed time from arrest to murder conviction in death-eligible noncapital murder cases (27.9 months versus 15.9 months). Interestingly, the average time between arrest and conviction via trial in death-eligible noncapital murder cases (17.4 months) was less than the time between arrest and plea in capital cases in Georgia (24.6 months).²⁸⁶

A recent study in New Jersey revealed that capital cases resulted in two to five times more pretrial motions filed, three to five times longer pretrial defense investigation, sixty-six times longer voir dire, and thirty more court days.²⁸⁷ Capital cases were also ten times as likely to proceed to trial, had twice as many lawyers (by statute), and resulted in longer and more complicated appeals.²⁸⁸ An examination of death penalty costs in Kansas discovered that, relative to noncapital murder cases, capital cases had investigation costs that were three times higher, trial costs that were sixteen times greater, and appeal costs that were twenty-one times more.²⁸⁹ The added expense associated with capital cases at every stage of the process has also been documented in the administration of the federal death penalty: attorney costs, expert costs, transcript costs, and out-of-court costs were significantly greater for capital cases.²⁹⁰ In general, capital cases were four times more expensive than noncapital cases for death-eligible defendants from 1989 to 1997, and 6.5 times more expensive from 1998 to 2004.²⁹¹

One of the few methodologically rigorous studies attempting to take both state and federal costs into account when comparing cost differentials between capital and noncapital death-eligible cases revealed that prosecuting a single capital case in Maryland adds *at minimum* \$1 million in total costs even after taking into account differential imprisonment

²⁸⁶ See *supra* Part III.C.

²⁸⁷ N.J. DEATH PENALTY STUDY COMM'N, NEW JERSEY DEATH PENALTY STUDY COMMISSION REPORT (2007).

²⁸⁸ *Id.*

²⁸⁹ LEGIS. DIV. OF POST AUDIT, *supra* note 282 (discovering that, relative to noncapital murder cases, capital cases had investigation costs that were three times higher, trial costs that were sixteen times greater, and appeal costs that were twenty-one times more).

²⁹⁰ JON B. GOULD & LISA GREENMAN, REPORT TO THE COMMITTEE ON DEFENDER SERVICES JUDICIAL CONFERENCE OF THE UNITED STATES: UPDATE ON THE COST AND QUALITY OF DEFENSE REPRESENTATION IN FEDERAL DEATH PENALTY CASES 28–32 (2010) (reporting 4.6 times more hours worked by defense counsel and 15.7 times more spent in expert fees in federal capital cases versus death-eligible noncapital cases).

²⁹¹ *Id.* (reporting that mean capital case was \$490,000 and the mean noncapital murder case was \$76,000; the median costs for capital and noncapital cases were, respectively, \$350,000 and \$45,000).

costs.²⁹² Trial costs were five times more expensive in capital cases compared to noncapital cases (\$823,000 versus \$160,000), and appellate costs were almost four times more expensive (\$340,000 versus \$88,000).²⁹³ Another relevant study, coauthored by Senior Judge Arthur L. Alarcón of the U.S. Court of Appeals for the Ninth Circuit, suggests that federal habeas review of a state capital conviction adds between \$635,000 and \$1.58 million to each capital case.²⁹⁴ These federal costs are in addition to the \$1.2 million more California spends on each capital case at the state level.²⁹⁵

It is important to underscore that because capital cases are more expensive and time-consuming at every stage, those cases that are ultimately resolved by plea bargain may be more expensive than noncapital trials because of pretrial and pre-penalty-phase costs. According to data from the federal death penalty system, capital cases eventually disposed by plea were over four times more expensive than the median noncapital case, regardless of method of disposition.²⁹⁶ Costs for noncapital cases were not disaggregated according to trial and plea, but given the fact that capital trials were 1.3 times more expensive than the median capital case, irrespective of method of disposition (\$465,602 versus \$353,185), it may be reasonable to assume that noncapital trial costs exceed noncapital plea costs by a similar multiplier. Based on that assumption, the median capital case resolved by plea is over 3.3 times more costly than the median noncapital trial in the federal system (\$200,933 versus \$60,000).

Similar estimates have been discovered at the state level. For example, in North Carolina (2002–2006) the median cost of a capital case ultimately disposed by plea bargain was 1.6 times more expensive than the median cost of a death-eligible noncapital case disposed by trial (\$31,000 versus \$19,000).²⁹⁷ Pleas obtained in capital cases were also three times as costly as pleas obtained in death-eligible noncapital cases (\$31,000 versus \$10,000).²⁹⁸ Capital trials were nearly 4.5 times more expensive than death-eligible noncapital trials (\$82,000 versus \$19,000), and 2.6 more

²⁹² Roman et al., *supra* note 27.

²⁹³ *Id.*

²⁹⁴ See Alarcón & Mitchell, *supra* note 65, at S88.

²⁹⁵ CAL. COMM'N ON THE FAIR ADMIN. OF JUSTICE, *supra* note 94 (noting the difference between the least expensive capital trial and most expensive noncapital trial was \$1.1 million).

²⁹⁶ GOULD & GREENMAN, *supra* note 290, at 27 (reporting a total cost of \$200,993 for capital cases disposed by plea compared to \$44,809 for noncapital cases, irrespective of method of disposition).

²⁹⁷ N.C. OFFICE OF INDIGENT DEF. SERVS., FY07 CAPITAL TRIAL CASE STUDY: PAC AND EXPERT SPENDING IN POTENTIALLY CAPITAL CASES AT THE TRIAL LEVEL (2008).

²⁹⁸ *Id.*

expensive than capital cases disposed by plea (\$82,000 versus \$31,000).²⁹⁹ Even dismissals in capital cases were four times more expensive than dismissals in death-eligible noncapital cases (\$19,000 versus \$4,500).³⁰⁰ The bulk of these cost differentials can be attributed to the fact that once a case is noticed for the death penalty, most death penalty statutes (including North Carolina's) require the appointment of two attorneys for the defendant, and attorney fees comprise the bulk of the cost of capital cases.³⁰¹ It should be emphasized that these cost estimates only include the defense's attorney fees and expert witness expenses; they do not take into account expenses incurred by the government.³⁰² We would expect government expenditures to be significantly higher in capital cases, but assuming *arguendo* that the cost to the government is similar for capital and noncapital cases, the relative gap in spending would remain the same (although the multiplier would change).

Now consider a few examples to fix ideas. Using the rather conservative estimates obtained from a recent Maryland study, the government can save, at most, \$250,000 in a death-eligible noncapital case by avoiding trial and subsequent appeals.³⁰³ Of course, there are still costs associated with obtaining a plea. The Maryland study did not report estimates of costs in cases disposed by plea prior to the start of trial, but based on estimates from the federal government and North Carolina, it is reasonable to assume that significant costs are associated with pleas in both capital and noncapital cases. Discounting potential trial savings by plea costs by a conservative estimate of one-third suggests that the government can save approximately \$195,000 by avoiding trial (and subsequent appeals) in a noncapital case (plea costs = \$160,000 trial costs \times .33 = \$53,280). Also recall that a capital case in Maryland costs at least an additional \$1 million to process fully, even after considering prison costs. This suggests that the threat of the death penalty would need to deter more than five murder defendants from pursuing trial to offset the price of one capital trial. If the cost of obtaining a plea is higher—e.g., 50% as in North

²⁹⁹ *Id.*

³⁰⁰ *Id.*

³⁰¹ *Id.* North Carolina compensates capital defense attorneys at a rate of \$85 per hour, whereas the federal government paid attorneys \$125 per hour during a similar time period. *Id.* at 5; see also GOULD & GREENMAN, *supra* note 290, at 28 (explaining that attorney fees constitute a large fraction of the cost of capital cases).

³⁰² N.C. OFFICE OF INDIGENT DEF. SERVS., *supra* note 297, at 1, 10.

³⁰³ Roman et al., *supra* note 27, at 565. The estimates of costs associated with capital trials obtained from the Maryland study are likely to be biased downward. For example, the study estimated the median cost of federal habeas review at \$25,000 (std. dev. \$105,000), whereas a recent California study estimated federal habeas costs between \$635,000 and \$1.58 million.

Carolina—then the death penalty would have to deter more than six murder defendants from opting for trial. *Based on the analysis presented in this Article, the death penalty only deters between 1.5 and 2 murder defendants from opting for trial for every one capital trial in Georgia.*

Of course these calculations ignore the fact that obtaining a plea bargain in a capital case may be more expensive than the total trial and appellate costs for a noncapital death-eligible case. This would preclude the possibility of the death penalty serving any cost-saving function in light of the fact that it does not appear that there is a corresponding decrease in prison costs.³⁰⁴ More cost estimates from death penalty jurisdictions based on methodologically rigorous study designs are necessary to explore fully this counterintuitive implication. But even if the costs associated with obtaining pleas in capital cases merely constituted half of the total trial and appellate savings, then the threat of capital punishment would need to dissuade eight defendants from choosing trial. And more than twelve defendants would need to be discouraged from going to trial if plea expenses comprised two-thirds of total noncapital trial savings.

Due to the fact that most capitally prosecuted defendants are not sentenced to death, perhaps it is more illuminating to explore cost estimates that take account of cases that begin as capital cases, but do not accomplish the stated purpose of a death penalty system: executions.³⁰⁵ These “cohort costs” (i.e., per-execution costs) have been estimated at \$3.2 million in North Carolina,³⁰⁶ \$3.5 million in Texas,³⁰⁷ \$24 million in Florida,³⁰⁸ \$37.2

³⁰⁴ In fact, prison costs appear to be greater for death row inmates than non-death row inmates, although death row inmates’ total incarceration time is less. *Id.*

³⁰⁵ Nationally, only one in every three capital trials results in a death sentence (because of mercy from the judge or jury) and only one in every ten death sentences may result in an execution, so the total cost to reach that one execution may be prohibitively high. DIETER, *supra* note 49, at 14 (discussing different approaches to assessing the cost of the death penalty); *see also supra* note 127 and accompanying text (noting that in Georgia, from 1993 to 2000, only 32% of capital cases disposed by trial received the death sentence).

³⁰⁶ This figure represents the cost per execution *above* life imprisonment. The \$3.2 million price tag is the inflation-adjusted number from \$2.16 million in 1993.

³⁰⁷ This is the inflation-adjusted figure. *Costs of the Death Penalty and Related Issues: Hearing on H.B. 1094 Before the H. Comm. on the Judiciary*, 2007 Leg. (Colo. 2007) (statement of Richard C. Dieter, Executive Director, Death Penalty Information Center) (stating that Texas spends three times more on capital cases from trial through execution than for life imprisonment for forty years in a maximum security single cell). Reliable data on the cost of the death penalty are lacking, but recent reports suggest that the average nonmurder trial in Texas costs about \$3,000, whereas death penalty trials range from \$200,000 to \$1 million. Logan Carver, *Paying the Price: Death Penalty Cases More Expensive than Lifetime Imprisonment, but Local CDA Says Cost Never a Consideration*, LUBBOCK AVALANCHE-J., Dec. 13, 2009, at A1.

³⁰⁸ S.V. D ate, *The High Price of Killing Killers: Death Penalty Prosecutions Cost Taxpayers Millions Annually*, PALM BEACH POST, Jan. 4, 2000, at 1A.

million in Maryland,³⁰⁹ and \$250 million in California.³¹⁰ New York and New Jersey have spent, respectively, \$170 million and \$253 million on their capital punishment systems, but have failed to execute a single condemned inmate.³¹¹ Even so, conservative estimates based on continued annual expenditures would place cost-per-execution in the \$20–\$40 million range for these two jurisdictions.³¹² The number of plea bargains induced by the threat of the death penalty required to offset per-execution costs, obviously, would be significantly higher. Continuing with the Maryland example, and assuming no cost to obtain a plea, nearly 150 death-eligible defendants would need to be deterred from opting for trial to offset the cost of one execution. Assuming that plea costs constitute one-third of trial costs, then 190 death-eligible defendants would need to be discouraged from going to trial.

Based on these stark differences in costs between capital and noncapital cases, it is quite possible that prosecutors are fully aware that the threat of capital punishment cannot serve any cost-saving function, and they merely leverage the death penalty to impose harsher bargaining terms—most notably, life imprisonment. More than 70% of inmates serving life sentences were convicted of murder, and more than one in four of all inmates serving a life sentence have no possibility of parole.³¹³ But this use of the death penalty has important cost implications as well. The mandatory minimum for most first-degree murder (and equivalent) convictions is life with the possibility of parole,³¹⁴ and several states only allow life without parole, so it is not clear that the threat of the death penalty (and the associated expenses) is necessary to obtain a sufficiently harsh sentence (the average time served for an inmate serving life imprisonment is thirty years across all offenses). In most death penalty jurisdictions, governors and pardon and parole boards are extremely reluctant to grant convicted murderers early release, so in practice, facially

³⁰⁹ Jennifer McMenamin, *Death Penalty Costs Md. More than Life Term*, BALTIMORE SUN, Mar. 6, 2008, at 2B.

³¹⁰ Rone Tempest, *Death Row Often Means a Long Life*, L.A. TIMES, Mar. 6, 2005, at B1.

³¹¹ DIETER, *supra* note 49, at 14 (citing studies and compiling statistics).

³¹² *Id.* at 15. New York sentenced seven individuals to death from 1995 through 2004—a cost of \$24 million per death sentence. JOSEPH LENTOL ET AL., *THE DEATH PENALTY IN NEW YORK* 7 (2005). Juries in New Jersey returned sixty death sentences (\$4.2 million per death sentence), but fifty-seven were overturned on appeal and only nine inmates were on death row as of 2007 (\$28 million per death sentence). See N.J. DEATH PENALTY STUDY COMM'N, *supra* note 287, at 7.

³¹³ ASHLEY NELLIS & RYAN KING, *NO EXIT: THE EXPANDING USE OF LIFE SENTENCES IN AMERICA* (2009).

³¹⁴ See *supra* note 54 and accompanying text.

indeterminate sentences have become de facto determinate life sentences.³¹⁵

VIII. CONCLUSION

When the Supreme Court ruled in *Furman* that the death penalty was unconstitutional as applied, the majority was particularly concerned that the death penalty was not being reserved exclusively for the worst crimes and worst criminals.³¹⁶ Many commentators have remarked that the death penalty is an incredibly powerful tool at the disposal of prosecutors, especially when it is used as leverage to induce defendants to forfeit their right to trial (and accompanying rights). The empirical findings in this Article suggest that the threat of the death penalty has a substantial *causal* effect on the likelihood that a defendant accepts a plea agreement. Nevertheless, the magnitude of the effect is clearly insufficient to offset the substantial administrative and financial costs arising from the occasional capital defendant taking her chances at trial (or, in some instances, even the capital case that incurs significant pretrial or pre-penalty-phase costs prior to a plea agreement). The government's use of the death penalty to obtain convictions quickly and cheaply appears to fail on both of these dimensions—and this may be particularly true in marginal cases because the likelihood of trial, a sentence other than death, or a reversal on appeal is particularly high. Prosecutors are likely to continue to use the threat of the death penalty in this highly inefficient manner unless they are required to internalize more of the costs of making poor screening decisions on the front end. Given the tremendous human and financial costs associated with the use of the “ultimate punishment,” prosecutors must be made to think carefully about using the death penalty as a “bargaining chip” in situations when such a severe sanction may be unwarranted.³¹⁷

³¹⁵ See *supra* Part III.B.

³¹⁶ There was no controlling opinion in *Furman*. Each of the Justices comprising the 5–4 majority (Brennan, Marshall, Douglas, Stewart, and White) differed over the basis of the decision; nonetheless, three recurring themes ran through their individual opinions. The *Furman* court was primarily troubled by three glaring problems with the existing practice of capital punishment: (1) the small number of death sentences handed out relative to potentially capital crimes; (2) the lack of statutory restrictions upon the sentencing discretion of judges and jurors; and (3) sentencing disparities based on social class and race. DAVID GARLAND, PECULIAR INSTITUTION: AMERICA'S DEATH PENALTY IN AN AGE OF ABOLITION 225–30 (2010); see also Calhoun, *supra* note 268, at 632; Schornhorst, *supra* note 31, at 301.

³¹⁷ “[T]he practice of charging the death penalty solely for the purpose of obtaining plea bargains is an unethical and unconstitutional interference with a defendant's Sixth Amendment right to trial [and impermissibly] risks convicting innocent defendants who plead guilty solely to avoid the possibility of a death sentence—which has occurred on numerous occasions.” DIETER, *supra* note 49, at 17.

Appendix A
Variables

VARIABLE NAME	VARIABLE DESCRIPTION	DATA SOURCE(S)
DP Notice	Death Penalty Notice Filed (Yes=1)	CO; GCD; AJC
Plea/Verdict	Case Disposed by Trial or Plea Bargain (Trial=1)	GDC; CO; GCD; AJC
Incident Year	Date of Incident (YYYY)	GDC; CO; GCD; AJC
Offender Age	Defendant's Age at Time of Incident (in Years)	GDC; CO; GCD; AJC
Offender Sex	Defendant's sex/gender (Male=1)	GDC; CO; GCD; AJC
Offender Race	Defendant's Race (Asian, Black, Hispanic, Other, White)	GDC; CO; GCD; AJC
Offender HS Grad	Defendant Graduated from High School (Yes=1)	GDC; CO; GCD
Offender Married	Defendant's Marital Status (Married=1)	GDC
# of Children	Defendant's Number of Children	GDC
Contemp. Felony	Defendant was convicted of committing a contemporaneous felony (Yes=1)	CO; GCD
Prior Felony	Defendant had prior felony conviction (Yes=1)	CO; GCD
# of Offenders	Total Number of Codefendants	CO; GCD
Murder Conviction	Defendant Convicted of Murder (Yes=1)	GDC; CO; AJC
Statutory Aggravators	Number of statutory aggravating circumstances present in case	GDC; GCD; SHR; AJC

VARIABLE NAME	VARIABLE DESCRIPTION	DATA SOURCE(S)
# of Victims	Number of deceased victims in the case	CO; GCD; SHR; AJC
Victim Age	Victim's age at time of incident (in Years)	CO; GCD; SHR; AJC
Victim Race	Victim's race (Asian, Black, Hispanic, Other, White)	CO; GCD; SHR; AJC
Victim Sex	Victim's sex/gender (0=No; 1=Yes)	CO; GCD; SHR; AJC
Victim Stranger	Victim(s) and defendant were strangers (Yes=1)	CO; GCD; SHR; AJC
Interracial Homicide	The defendant differed in race from at least one of the victims (Yes=1)	GCD; SHR; AJC
County	County in which the trial took place	GDC; CO; GCD; AJC
Circuit	Circuit in which the trial took place	GDC; CO; GCD
DP Eligible	Defendant was eligible for the DP based on a murder conviction and the presence of <i>at least</i> one Statutory Aggravating Circumstance	GDC; GCD; SHR; AJC
<p>LEGEND: U.S. Bureau of the Census (CENSUS); Ga. Department of Corrections (GDC); Ga. Sup. Ct. Clerk's Office (CO); Ga. Bureau of Investigation (GBI); Office of the Georgia Capital Defender (GCD); Supplementary Homicide Reports (SHR); Atlanta-Journal Constitution (AJC).</p>		

Appendix B
Fixed Effects-Logit Models (Marginal Effects)

	Pr(Trial)		Pr(Notice)	
DP Notice	-0.234***	(0.038)	--	--
# of Stat Aggs	-0.042*	(0.019)	0.108***	(0.020)
# of Offenders	-0.038*	(0.017)	-0.005	(0.013)
Offender Black	-0.033	(0.048)	-0.107*	(0.044)
Offender Male	-0.149	(0.076)	-0.102	(0.065)
Offender Age	0.000	(0.002)	-0.002	(0.002)
Contemp. Felony	0.043***	(0.013)	0.029*	(0.013)
Prior Felony	0.040***	(0.010)	0.006	(0.010)
High School Grad	0.142***	(0.035)	0.032	(0.036)
Offender Employed	0.073*	(0.035)	0.072**	(0.025)
Offender Married	0.085	(0.063)	-0.029	(0.030)
# of Children	-0.045	(0.027)	0.017	(0.033)
# of Victims	0.028	(0.037)	0.104***	(0.024)
Victim White	-0.097*	(0.044)	0.115***	(0.035)
Victim Female	-0.085**	(0.033)	0.117***	(0.027)
Victim Age	-0.001	(0.001)	0.000	(0.001)
Victim Stranger	0.055	(0.036)	0.087**	(0.031)
Interracial Homicide	0.006	(0.046)	-0.012	(0.038)
<i>N</i>	852		856	
Year Dummies	Y		Y	
Pseudo <i>R</i> -Squared	0.32		0.37	

*p<.05, **p<.01, ***p<.001

Standard errors adjusted for clustering on judicial circuit.