Title
Disentangling Disparity: Exploring Racially Disparate Effect and Treatment in Capital Charging

Journal
American Journal of Criminal Law, 25(1)

Author
Thaxton, Sherod

Publication Date
2018

UCLA School of Law

PUBLIC LAW & LEGAL THEORY RESEARCH PAPER NO. 18-46

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

BY

SHEROD THAXTON
PROFESSOR OF LAW

In 45 AMERICAN JOURNAL OF CRIMINAL LAW 95 (2018)

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

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 Rehavi & 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 capitally 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 OBESITY 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 serves[55] 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 additional 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 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 sensitively 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:

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

where $P(N)$ is the probability that a death notice is filed, X_k is a vector of k explanatory variables, CV is an 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, $log_e\,(\frac{P(N)}{1-P(N)})$ (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 disparity[181] *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 qualitative[186] 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.

quantitative[187] 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":

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

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 field 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 the by 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, $\overline{N}$, between W and B is: $G = \overline{N}_W - \overline{N}_B = P(N_W) - P(N_B)$, 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:

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

[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, $\mathcal{F}(\cdot) = \frac{e^{(\cdot)}}{1+e^{(\cdot)}}$, 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:

$$G = \underbrace{\mathcal{F}\overline{(X_W \times \beta_W)} - \mathcal{F}\overline{(X_B \times \beta_W)}}_{E} + \underbrace{\mathcal{F}\overline{(X_W \times \beta_W)} - \mathcal{F}\overline{(X_W \times \beta_B)}}_{C}, \tag{4.1}$$

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:

$$G = \underbrace{\mathcal{F}\overline{(X_W - X_B)\beta_W}}_{E} + \underbrace{\mathcal{F}\overline{\{X_W \times (\beta_W - \beta_B)\}}}_{C}. \tag{4.2}$$

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:

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

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:

$$G = \underbrace{\mathcal{F}\overline{(X_W \times \beta_P)} - \overline{\mathcal{F}(X_B \times \beta_P)}}_{E} + \underbrace{\mathcal{F}\{\overline{[X_W \times (\beta_W - \beta_P)]} - \overline{[X_B \times (\beta_P - \beta_B)]}\}}_{C}. \tag{6.1}$$

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:

$$G = \mathcal{F}\,\underbrace{\overline{(X_W - X_B)}\beta_P}_{E} + \underbrace{\mathcal{F}\{\overline{[X_W \times (\beta_W - \beta_P)]} - \overline{[X_B \times (\beta_P - \beta_B)]}\}}_{C}. \tag{6.2}$$

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 who 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 inappropriate 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 Causcasian-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:

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

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 & 217and 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 ÷ 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 ÷ 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 ÷ 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 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 defendants accused of murdering African Americans (or an 230% increase 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 effect[269] 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 Appeals[292] and the Supreme Court[293] 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 treatment[317] 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: <u>Top Panel</u>: $P(N_W)$ 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
First Appearance	Accused presented before a magistrate judge within 48 (warrant) or 72 (without warrant) hours. Unif. Super. Ct. R. 26.1 (2007).
Grand Jury Indictment	Grand jury returns an indictment charging a capital offense within 180 days. Ga. Code Ann. § 17-7-50.
Appointment of Counsel	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).
Pretrial Conference	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).
Arraignment	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; *nolo contendere* pleas are disallowed. Ga. Code Ann. § 17-7-95.
Capital Voir Dire	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.
Capital Trial	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 *may* 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.
Post-Sentencing and Direct Appellate Proceedings	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.
State Post-Conviction Proceedings (Habeas Corpus)	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.
Federal Post-Conviction Proceedings (Habeas Corpus)	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.
Clemency	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).
Execution	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 if found to be mentally incompetent. Ga. Code Ann. § 17-10-40, 17-10-61.