DRAFT: NOT FOR QUOTATION OF CITATION WITHOUT PERMISSION

RACIAL DISCRIMINATION IN CAPITAL AND NON-CAPITAL SENTENCING WITH SPECIAL REFERENCE TO THE EVIDENCE IN MURDER AND RAPE PROSECUTIONS

David C. Baldus
College of Law
University of Iowa

September 14, 2010

It was a pleasure to review Professor Baumer’s thoughtful review of the literature on race discrimination in sentencing in the American criminal justice system. It is a timely and insightful critique of a vast and complex literature. In my remarks, I develop two themes. The first is an elaboration on Professor Baumer’s assertion that the existing “body of research … is not well organized … to detect racial discrimination by legal decision makers.” My second theme addresses what to me is a puzzle – how can we reconcile the substantial and persuasive body of evidence documenting prejudice against black males in the white population in general and within the criminal justice system with an equally persuasive body of evidence that fails to document statewide bottom line black defendant disparities in the prosecution and sentencing of murder and rape, (the crimes with which I am most familiar), and many other crimes.

In my reading of Professor Baumer’s paper, I was particularly interested in the evidence he cites of racial discrimination in non-capital sentencing. The conflicting findings and opinions he surveyed help me understand the substantial gap that I often perceive between the strong unqualified claims of racial prejudice in the system presented by its critics and the empirical evidence bearing on the issue. The reason for this gap appears to be that the literature provides support for nearly every position on the issue, or as Professor Baumer charitably puts it – the “sentencing research …is ambiguous in its conclusions.” (p.3) For example, recent commentary in the New York Times asserts that the “abominable incarceration rates among blacks are the
result of two overwhelming factors: the persistence of criminal behavior by a significant percentage of the black population, and a criminal justice system that in many respects is racially discriminatory and out of control.”¹ I am assuming that the reference here is to “purposeful” or “direct” racial discrimination on the part of prosecutors, juries, and judges. As I explain below, however, with respect to murder and rape, the empirical literature since Gregg v. Georgia (1976) contains no evidence of systematic statewide purposeful discrimination against black defendants in either capital or non-capital contexts.

These findings come as a surprise, given the substantial volume of recent empirical evidence documenting that white citizens in general and white participants in the criminal justice system in particular reveal distinct signs of anti-black prejudice² There is also compelling evidence of anti-black discrimination in the empanelling of jury venires (from which jurors are selected for jury service) and in the use of peremptory strikes by prosecutors in selecting actual juries.³

There is, of course, significant evidence of systemic and purposeful race of victim discrimination with respect to both rape and murder outcomes, but most citizens and commentators have little or no knowledge or concern about the more punitive treatment of white-victim cases. For example, Scott Turow writes off white victim discrimination, as unexceptionable socio-economic discrimination of little or no moral consequence.⁴ The real issue, it appears, is black defendant discrimination.

The story I perceive in the post-1980 empirical charging and sentencing literature, at least with respect to murder and rape, is strikingly similar to the story Gary Kleck told in his 1981 and 1985 reviews of the literature from the 1930s through the 1970s. He argued persuasively that, except for early capital rape cases in the South, there was no convincing evidence of “general or widespread overt purposeful discrimination against black offenders although there is evidence of [such] discrimination for a minority of specific jurisdictions, judges, crime types, etc.” His claims have received close scrutiny from other scholars but with only a few exceptions, they have generally held up as the methodological sophistication of the research has improved.

The pre-1980 period about which Kleck wrote has one important thing in common with the situation today. In both periods there has been a substantial body of evidence of anti-black prejudice in the white population generally and on the part of the individuals who administer the criminal justice system. Yet to the extent that this anti-black prejudice did and does exist, it has not produced statewide bottom-line black defendant sentencing disparities - an outcome I do not fully understand and address more fully below.

Professor Baumer lists four “motivations or rationales” for the research that he reviewed, one of which is the detection of “racial discrimination.” In this regard, he identifies an underlying concern with the “perceived” illegitimacy of the system. I would add to this list an explicit concern about the factual “legality” of the system under current law, i.e., to what extent is race a but for factor, consciously or unconsciously, in discretionary charging and sentencing decisions.

---

The requirements of the law in this regard are not unknown. Since the 1970s, state and federal courts have closely scrutinized the methodology of the empirical studies presented to them in support of claims of racial discrimination in a wide variety of contexts, many of which potentially apply directly to the criminal justice system. It is useful, therefore, to contrast the approaches to important methodological issues developed in the case law (in a variety of different factual situations) and in the sentencing literature that Professor Baumer reviews.

For starters, we find a common emphasis on the distinction between “disparate treatment” also known as purposeful discrimination and “disparate impact.” This distinction was only dimly understood in 1981. Today most lawyers and social scientists recognize that disparate impact claims arise from the even handed application of “facially neutral” legal standards, while disparate treatment claims arise from the discriminatory exercise of discretion by prosecutors, juries, and judges.

A good illustration of disparate impact are the consequences of the federal law, which until recently, imposed sharply different sentences for the crimes of powder cocaine possession and crack cocaine possession. As amended in the 1980s, the statute called for much more severe punishments for the crack cocaine offense. As a result, the evenhanded application of this law produced a large disparate impact in the black community because crack cocaine is much more the drug of choice among young blacks than it is among young whites. In the face of compelling evidence that the crack/powder cocaine disparate impact cannot be justified on penological or moral grounds, Congress recently modified this sentencing distinction.


In contrast to disparate impact disparities, disparate treatment disparities are the product of “purposeful” discrimination, whether overt, covert, or unconscious. The distinction between disparate impact and disparate treatment sentencing disparities is crucial, because disparate impact claims produce far less stigmatization than do comparable disparate treatment claims, which carry strong implications of overt racism.\footnote{Gerald A. Neal. Mar. 2004. Not soft on crime but strong on justice: the Kentucky Racial Justice Act, p. 9. available at http://dpa.ky.gov/library/advocate/mar04.htm (reporting strong hostile prosecutorial reactions to disparate treatment claims brought under the Kentucky racial justice act).} For example, it is very unlikely that Congress would have addressed the powder/crack cocaine distinction this summer if it had been perceived as a product of purposeful disparate treatment. Comparable evidence of disparate treatment discrimination would produce mighty resistance to any reform designed to cure or prevent it. This is exactly what happened in 1992 when Congress considered the Racial Justice Act, which would have overruled \textit{McCleskey v. Kemp}'s unattainable burden of proof for establishing purposeful discrimination in the use of the death penalty.\footnote{The measure died in the Senate with the threat of a filibuster.}

Proof of unjustifiable disparate impact, therefore, may provide a compelling basis to persuade legislatures and courts to reduce or eliminate sentencing disparities that they would be unlikely even to consider if they were perceived to be the product of purposeful disparate treatment discrimination.

But in spite of the importance of the disparate impact/disparate treatment distinction, Professor Baumer reports that the post-1990 research often fails to distinguish between “racial discrimination by legal actors, [or] disparate racial impact of non-racially discriminatory features of … the law.” p. 28. He also points out that among those studies “almost none … contain any
serious discussion of racial [purposeful] discrimination … even when the studies yield significant effects of defendant race.” P. 28

With respect to the other core methodological issues, both the case law and the literature recognize the importance of focusing on a single jurisdiction and conducting a stage by stage analysis of decision making within it.\textsuperscript{11} Professor Baumer also notes a trend in the literature, which he welcomes, of focusing principally on racial disparities that have been adjusted for non-racial case characteristics such as the defendant’s prior record, the level of violence in the case, and other factors bearing on each offender’s overall level of criminal culpability. Professor Blumstein lit up the sky on this point 30 years ago\textsuperscript{12} and the message has been generally understood and applied since then.\textsuperscript{13} However, I believe that research addressing disparate treatment issues would have substantially more credibility if the typical database included much richer evidence on the details of the offense bearing on the offender’s criminal culpability, even if the strategy resulted in substantially smaller samples of cases.

One page that I believe scholars in this field may usefully take from the case law is the distinction between (a) the different factual patterns that are accepted as sufficient to establish purposeful discrimination and (b) the different burdens of proof that are required to establish the required factual patterns. Here are some useful distinctions in the law between different forms of purposeful discrimination and their required burdens of proof:

\textsuperscript{11} However, some studies fail to include a model of the combined impact of all stages of the decision making process documenting racial disparities among all cases. In the absence of such a model it is difficult to estimate the impact of discrimination on the average offender’s risk of death, incarceration, or an extended term of years, as the case may be.


\textsuperscript{13} I also share the belief of the courts and most scholars in this field that tests of statistical significance are crucial to assessing the risk that the documented disparities may be the product of chance rather than purposeful discrimination, even when the sample embraces the universe of cases of interest.
1. First is the distinction between (a) discrimination in an individual case\textsuperscript{14} and (b) discrimination affecting a number of cases, which is known as class wide or pattern and practice discrimination.\textsuperscript{15}

2. Second is the distinction between (a) proof by a preponderance of the evidence that race “actually motivated” the adverse outcome in one or more decisions\textsuperscript{16} and (b) proof by a preponderance of the evidence of a substantial “risk” that race may have impacted such decisions.\textsuperscript{17} Proof of actual motivation standards underlie Fourteenth Amendment and state and federal civil rights claims, while substantial risk standards underlie Eighth Amendment claims of “arbitrariness.”\textsuperscript{18}

I believe that the most relevant models for proving disparate treatment are found in Title VII class action employment cases which require proof of a “pattern and practice” or “class wide evidence” of purposeful discrimination and, if the state raises the issue, proof that race motivated the adverse decisions in the plaintiff’s case. I am not suggesting that empirical researchers write

\textsuperscript{14} McDonald-Douglas v. Green, 411 U.S. 792 (1973); Raytheon Company v. Joel Hernandez, 540 U.S. 44, 49 (2003) (summarizing the elements of and burden of proof under McDonald-Douglas).


\textsuperscript{16} Raytheon Company v. Joel Hernandez, 540 U.S 44, 50 (2003) (individual claims in which an individual is “treated less favorably than others because of the race, color, religion, sex…”) International Brotherhood of Teamsters v. United States, 431 U.S. (1977) (class wide claims); the case law has not clarified the extent to which “purposeful discrimination” is limited to conscious discrimination or also embraces “unconscious” and “nonconscious” discrimination. See, e.g., Sheri Lynn Johnson. 1988. Unconscious Racism and the Criminal Law, 73 CORNELL LAW REV. 73: 1016, 1022, 1024-25 (arguing that recent cases involving both race and criminal procedure fail to take into account the effect of unconscious racism).

\textsuperscript{17} Turner v. Murray, 476 U.S. 281 (1986) (“The risk of racial prejudice infecting a capital sentencing proceeding…”).

\textsuperscript{18} United States of America v. The City of New York, 683 Fed. Supp. 2d 225, 255 (2010) (the Supreme Court has suggested and the Second Circuit Court of Appeals has ruled that the Title VII employment discrimination model of proof is the appropriate “framework to prove the existence of discriminatory purpose under the Equal Protection Clause.”
briefs or judicial opinions, only that they be more attentive to the law when they develop their research designs and analyze evidence of disparate treatment.

Before considering in more detail the contrast between the widespread evidence of anti-black bias in American society generally and the absence of bottom line evidence of black defendant discrimination in charging and sentencing studies, I will consider in a bit more detail the evidence of systemic discrimination in murder and rape prosecutions, both capital and non-capital. For both offenses, we would expect to see stronger evidence of statewide racial discrimination in capital than in non-capital contexts and in prosecutions before 1980 than after that year.

For capital murder post-Furman one can start with a comparison of the 13% black population nationwide with the current nationwide death row which is 42% black and the defendants who have been executed since 1976 who are 34% black.\textsuperscript{19} On the basis of such proof, unwarranted conclusions of purposeful black defendant discrimination have been drawn with no additional evidence.

More in keeping with traditional methods of proof, some observers have compared the 41% of blacks on death row nationwide with the 35% of blacks among those executed and the 49% of blacks among those are arrested for murder and non-negligent manslaughter.\textsuperscript{20} This evidence, they argue, is consistent with purposeful discrimination against whites rather than blacks. Moreover, a number of well controlled statewide studies of death sentencing among all death eligible cases document higher death sentencing rates for white defendants even with a


control introduced for the race of the victim.\textsuperscript{21} And no statewide murder studies, capital or non-capital, have documented significant or substantial black defendant effects.\textsuperscript{22}

The story is comparable with rape, where the working hypothesis has been that black-on-white crimes produce a substantial risk of racial prejudice, especially in capital cases. This appears to have been the case in the South before the abolition of capital punishment for rape.\textsuperscript{23} However, since then the evidence in non-capital rape cases shows little such effect. Although race effects unadjusted for legitimate case characteristics are substantial, statistical adjustments for non-racial aggravating and mitigating factors, such as risky victim behavior substantially explain away the race effects.\textsuperscript{24}

Returning to my second theme, what can explain the absence of bottom line statewide systemic black defendant disparities in the face of compelling evidence of anti-black prejudice among the nation’s white community? I see three possible explanations. The first is that anti-black prejudice is no longer a factor, systemic or otherwise, in charging and sentencing because prosecutors, jurors, and judges are without prejudice or they set it aside in their charging, judging, and sentencing capacities. Professors Blair, Judd, and Chapleau find plausibility in this


hypothesis for Florida: “over the past 20 years the state’s efforts to ensure race neutrality in sentencing … have largely been successful. Our results are also consistent with the psychological literature showing that people can effectively reduce category-based stereotyping….; it appears that judges have effectively learned to give to give sentences of the same length when Black and White offender with equivalent criminal histories come before them.”

There are, however, grounds for skepticism. The hypothesis is very much at odds with the collateral evidence of anti-black bias, in general, and on the part of important actors in the criminal justice system. However, the evidence from county based studies within states, that I describe below, suggests that anti-black discrimination in some counties may be neutralized by pro-black or no discrimination in other counties with a cancelling out of any statewide effect.

A second related possibility is that investigators are looking in the wrong places and what is needed is a narrower focus not only in terms of jurisdictions but also in terms of offender and offense characteristics. What we do know suggests that tighter geographic inquires might tell a different story. The third explanation, which I explore briefly below, is that methodological flaws confound the findings.

With respect to the second possible explanation concerning county level studies, there appears to be some truth in the claims of prosecutors that sentencing studies have credibility only when conducted at the county level. This echoes George Kendall’s elaboration on Tip O’Neill’s claim that all politics is local -- to the effect that our death penalty policy and practices are both local and political in the extreme.

25 Blair, Judd, and Chapleau supra note 22.

26 With respect to producing evidence to challenge death sentences on racial grounds, this is the opinion of an experienced capital litigator. Sheri Lynn Johnson. 2007. Litigating for Racial Fairness after McCleskey v. Kemp, 2007. COLUMBIA HUMAN RIGHTS LAW REVIEW, 39: 178-201 (“One case at a time. One juror at a time. One idiosyncratic issue at a time.”)
In this regard, it is also useful to consider “Simpson’s Paradox,” a variation on the more general omitted-variable issue. It argues that broad gauge studies in general which fail to account for differences in the policies and practices of local entities can produce significant error. A study of gender discrimination in graduate school admission policies at Berkeley makes this point well.\(^{27}\) It first documents university wide admission rates of 44% (3738/8442) for men compared to 35% (1494/4321) for equally qualified women, meaning that the expected number of men admitted exceeded the expected number of women by a statistically significant 277 men. This disparity looked compelling until the investigators accounted (a) for the very low admission rates in English and comparable departments and very high admission rates in engineering and related departments, and (b) for the fact that women applied overwhelmingly to English and related departments, while the men overwhelming applied to engineering and related departments. With the introduction of these controls, the expected number of women exceeded the expected number of men by a statistically significant 60 women.\(^{28}\)

We observed comparable but opposite effects in Texas and Pennsylvania studies of capital punishment, where statewide data reveal no black defendant effects, while closely controlled studies in Harris County, Texas and Philadelphia County, Pennsylvania reveal substantial black defendant disparities.\(^{29}\)

---


\(^{28}\) Id. at 115, 125.

George Woodworth and I encountered Simpson’s paradox in our study of the Nebraska death penalty. The statewide data revealed black defendant charging disparities, but on further investigation we found that they were strictly the product of charging decisions in Lincoln and Omaha, where 95% of the state’s death eligible black defendants were prosecuted. We further found that the introduction of controls for non-racial factors in those two counties, completely explained away the black defendant disparities at both the statewide and county levels.

With respect to the second possible explanation noted above concerning offender characteristics, several studies have focused on the interaction between the stereotypically black appearance of black defendants and their risk of being sentenced to death. The more stereotypically black the defendant’s appearance, the greater the risk of a death sentence.

Also several studies have tested the hypothesis that race disparities, including black defendant disparities, vary with the culpability level of the cases. In some studies which found no estimated statewide race effects, the data have documented the concentration of race

---

30 Baldus et al. supra note 21 at 566-79.

31 Id.

32 See e.g., Jennifer L. Eberhardt et al. 2006. Looking Death Worthy: Perceived Stereotypicality of Black Defendant Predicts Capital-Sentencing Outcomes, PSYCHOLOGICAL SCIENCE, 17: 383, 385 (One explanation for the elevated risk or racial prejudice in black-on-white murder cases is that the defendant-victim racial combination "renders race especially salient. Such crimes could be interpreted or treated as matters of intergroup conflict" whereas a black-on-black murder may "lead jurors to view the crime as a matter of interpersonal rather than intergroup conflict."); Irene V. Blair, Charles M. Judd, and Kristine M. Chapleau, 2004. The influence of afrocentric facial features in criminal sentencing. Psychological Science, 15: 674- 679 (a study based on inmate records containing their photographs); Justin D. Levinson and Danielle Young. 2010. Different shades of bias: skin tone, implicit bias, and judgments of ambiguous evidence. W. Vir. Law. Rev., 112: 307-349 (reporting the results of experiments with student subjects).
disparities (a) among the least aggravated cases, or (b) among the most aggravated cases, or (c) among the mid-range of cases in terms of offender culpability.33

The third possible explanation for the puzzle of evidence of anti-black bias but no statewide racial disparities noted above questions the validity of the statewide studies which reveal no black defendant race effects. George Woodworth and I are confident that well controlled studies based on complete and accurate data provide a good basis for inferring class-wide disparities or the absence thereof statewide and in local jurisdictions. Substantial race effects estimated in such studies give us confidence that some, although not necessarily all, of the disadvantaged racial group members were adversely affected because of their race or the race of their victim. Similarly, we have confidence that the absence of race effect in such studies presents a valid picture of the average effect statewide, which may reflect no race effects at all or a narrow concentration of effects in some locales that may be offset by no disparities in other parts of the state.

We also believe that qualitative approaches may be profitably used to establish discrimination in individual cases. For example, a comparison of a death sentenced black defendant with one or more life-sentenced white defendants with similar or greater levels of criminal culpability may support an inference of racial discrimination in an individual black defendant’s case regardless of what the class wide data reveal.34

33 Eric P. Baumer, Steven F. Messner, Richard B. Felson. 2000. The role of victim characteristics in the disposition of murder cases. Justice Quarterly, 17: 282, 285 (victim characteristics are more “influential in murder cases that are perceived to be less serious”; Baldus et al. supra note 28 at 211, fig. 5, (race of defendant effects concentrated in the mid-range of cases in terms of offender culpability); Turner v. Murray, 476 U.S. 281 (1986) (risk of racial prejudice in black defendant/white victim cases strongest in the most highly aggravated cases).

In conclusion, my review addresses only a narrow slice of the vast terrain covered by Professor Baumer. My first theme and set of recommendations focus on how the literature could provide a closer focus on proof of purposeful racial discrimination along the lines suggested by the law. My second theme and set of recommendations suggests approaches that may help explain the conflict between (a) the persuasive body of evidence purporting to document prejudice against black males in the white population in general and on the part of participants in the criminal justice system in particular, and (b) an equally persuasive body of evidence that fails to document statewide bottom line anti-black prejudice in the prosecution of murder and rape cases.