# How Judicial Elections Increase Racial Bias: Evidence from Appellate Criminal Cases

Thomas R. Gray
University of Virginia
Department of Politics

October 4, 2016

Persistent racial disparities in incarceration rates in the United States are typically explained through a variety of structural, psychological, and institutional factors at the investigative and trial levels. In comparison, the role of appellate decision making in perpetuating racially unbalanced criminal outcomes remains under-explored. I present a behavioral theory of criminal appellate decision making in which implicit biases about the likely guilt and deservingness of protection of certain groups of defendants – specifically Black and Hispanic men – limits the ability of judges to consistently apply their attitudinal and legal preferences. I test this theory on cases from the Texas Court of Criminal Appeals between 1995 and 2010. Using a large roster of Texas prisoners, I estimate defendants' races by their names. I find that defendants with names commonly used by Black men were approximately eight percentage points less likely to receive a judge's vote than their White counterparts. Additionally, I find that this effect spikes in votes cast in a judge's election year, with Black and Hispanic men 23 and 14 percentage points less likely than Whites to receive a judge's vote, respectively. These results imply that audience pressures from punitive voters exacerbate racial biases in decision making. Finally, I show that these effects are almost exclusively found among Republican judges and *increased,* rather than decreased, between 1995 and 2010. These results present troubling evidence that appellate courts, rather than correcting the racial imbalances of the trial courts, may in fact perpetuate or exacerbate them.

I thank Jeff Jenkins, Craig Volden, David Klein, Isaac Unah, Adam Hughes, and Andrew Clarke for providing useful feedback on this paper.

The United States imprisons a larger percentage of its population than any other major country. But incarceration is not evenly shared across America's diverse population: racial minorities, men, and the poor are overrepresented in America's criminal-justice system. Research on these facts tends to focus on the immediate levels: structural determinants of criminal behavior, biases in policing and prosecutorial strategies, and racial disparities in trial-court outcomes such as conviction rates and sentences. These processes determine who enters America's "carceral state." Extant research further explores the long term political ramifications of mass incarceration (Weaver and Lerman 2010).

One important, but understudied, component is the criminal appeals process. Convicted defendants are generally entitled to an appeal by right under the law and many will receive multiple appeals before their full sentence is imposed. Yet, research on appellate judicial behavior is overwhelmingly concerned with ideological and institutional politics: the attitudes of judges and the conflicts between courts, executives, and legislators. These theoretical approaches largely assume away the possibility of racial bias at the appellate level. Defendants' identities are not integral to the attitudinal or strategic approaches to appellate decision making.

Criminal appellate law comes with internal and external pressures on judges that discourage them from consistently applying the law or their own ideological beliefs. Principally, these pressures come from judges' reluctance to help a defendant they perceive to be guilty and undeserving. Judges may believe that police should be prohibited from engaging in certain investigative practices, but exercising that belief if it results in a murder conviction being overturned raises considerable cognitive dissonance that leads to inconsistent application. This behavioral theory of criminal appellate decision making predicts that defendants from disadvantaged groups do worse in criminal appeals, because implicit biases alter judges' base

perceptions of the likelihood that the defendant is guilty and whether the defendant deserves to have their rights protected. Additionally, an elected judge's anticipation of similar prejudices in their voters can affect their decision making as they seek re-election.

I test these expectations on a set of criminal cases from the Texas Court of Criminal Appeals between 1995 and 2010. Contrary to an alternative view that the legal-theoretical framework and aloof nature of appellate process reduces the possibility of racial biases, I find evidence of systematic disadvantages for Black and, to a lesser extent, Hispanic defendants in Texas. I find that these effects occur exclusively among Republican judges and have *increased* rather than decreased over the time period analyzed. Finally, I show that these effects spike in votes taken in the year before an election, indicating that electoral pressures can transmit the racial biases of voters into appellate decision making. This spike in election years confounds alternative explanations for disparities, such as differential quality of legal counsel and social capital. The best explanation for these results is that personal prejudices and implicit biases can intrude into the supposedly objective and precedent-bound world of appellate jurisprudence.

**The Role of Criminal Appeals in a Criminal Case**

The typical backstory of a criminal case that reaches a state court of last resort is, in broad strokes, as follows: an event occurs which spurs police to investigate a suspected crime, along the way navigating a complex set of limitations and restrictions that circumscribe the ways they may investigate. A suspect is arrested and the evidence supporting that arrest is turned over to a prosecutor who determines what, if any, charges to pursue in court. After a set of hearings and procedural hurdles are cleared, and months elapse during which each side builds its case, a trial occurs. During the trial, a judge presides over an elaborate set of procedures, often with a jury empaneled to determine whether the defendant is guilty beyond a reasonable doubt. At the

completion of the trial, the jury renders a verdict on the facts as an application of the law given to them by the judge. Convicted defendants are then given a sentence from a statutory guideline.

In the preceding sequence of events, there are numerous opportunities for bias to enter into the process. Policymaking and variation in policing strategies may increase the likelihood of investigations being initiated and vigorously pursued for certain groups (Tonry 2010; Skogan 2006; Weitzer 2000). Prosecutors may be more aggressive against some populations relative to others (Smith and Levinson 2012; Songer and Unah 2006). At trial, with the defendant present and visually accessible to all of the major actors in the courtroom, various forms of explicit and implicit biases may intrude into both the judge's and the jury's decision-making (Kang et al 2012; Levinson and Young 2010; Rachlinski et al 2009; Levinson 2007). Underlying beliefs about the propensity of people to commit crimes (Levinson, Cai, and Young 2010; Dixon and Azocar 2007; Quillian and Pager 2000; Hurwitz and Peffley 1997; Peffley, Hurwitz, and Sniderman 1997) influence whether defendants are believed and the extent to which aggressive prosecution is tolerated in seeking "justice" (Eberhardt et al 2006). Even the defendant's own attorney or public defender may provide different levels of effort depending on underlying beliefs of their clients' guilt and deservingness (Richardson and Goff 2013; Eisenberg and Johnson 2004). Finally, out-group convicts may receive harsher sentences (Demuth and Steffensmeier 2004). This collection of steps, each with its own opportunity for bias, means that the trial process is fraught with difficulties for administering objective justice.

The appellate process is different in a variety of ways that imply that bias should be reduced. First, the subject matter of the case changes considerably. On appeal, the issue is typically *not* guilt or innocence, but whether the defendant's rights were protected at trial. For example, appellate courts are not ostensibly concerned with whether a defendant used a knife to

kill the victim, but instead whether the prosecution's inclusion of the knife at trial was unfairly prejudicial to the defendant, perhaps because of how the police obtained the knife during their investigation. A ruling on this question by an appellate court would resolve not only the case before them, but also all subsequent cases involving similarly obtained and introduced evidence within that court's jurisdiction. Thus, the content of the appeal is legal-theoretical, not factual. Importantly, a defendant whose rights were violated can win *even if he is guilty*.

In line with the preceding perspective, the literature on appellate judicial decision making overwhelmingly focuses on the attitudes, or policy preferences, of judges in deciding appeals (see Segal and Spaeth 2002), sometimes mitigated by strategic constraints placed on them by other actors (Canes-Wrone, Clark, and Kelly 2014; Clark 2011; Epstein, Knight, and Martin 2001). The details of the individual dispute and the impact of the ruling on the parties are thought to matter less than the final precedential policy.

Another important difference between trials and appeals is the nature of the proceedings themselves. Unlike the trial process, the appellate system is primarily contested through briefs. If there is an in-person event, it is a short hearing full of intricate and detailed arguments of a legal – rather than factual – nature. The defendant, typically not present, is not very visually accessible, an important fact given that appearance may intensify stereotyping (Dixon and Maddox 2005; Eberhardt et al 2004). Additionally, judges have law clerks who help cut through to the heart of the legal argument in the case, discarding irrelevant material.

Thus, at first glance, there are substantially fewer reasons to suspect bias to creep into criminal appeals than criminal trials. Many of the opportunities for implicit, subconscious biases are mitigated by the aloof complexity of appellate content and procedure. In fact, disadvantaged groups should do *better* than more advantaged groups at a perfectly egalitarian appellate court

exactly because appellate courts are error-correction institutions. If disadvantaged groups are more likely to have their rights infringed in the investigative and trial process, then their average appeal should be sustained *more often* in a perfectly unbiased world than similar appeals from more advantaged defendants.[1] This outcome does not rely on assuming a legalistic approach to judicial decision making; judges who rely heavily on attitudinal predispositions would also find more inconsistencies to correct for disadvantaged defendants. This provides an initial expectation: that disadvantaged groups should do at least as well or better than advantaged groups in perfectly egalitarian appellate courts. In the next section, I outline an alternative theory that predicts racial disparities in appellate jurisprudence.

**A Behavioral Theory of Criminal Appellate Decision Making**

I posit an affective version of judicial decision making in which judges attempt to consistently apply their attitudinal preferences and legal interpretations to reviews of prosecutions, but face the central tension described in the preceding section: a defendant can win on appeal even if they are guilty. On an affective level, this may be deeply dissatisfying. The "just" outcome when a defendant's rights are severely violated is to overturn the conviction, which can allow someone guilty of a horrible crime to go free (pending a retrial). However, this can feel fundamentally unjust. Though applicable to non-criminal cases as well, this tension is most pronounced in criminal law, where it is possible for a murderer or rapist to have their conviction overturned if the trial was sufficiently prejudicial.

---

[1] In theory, a perfectly egalitarian intermedia appellate court (IAC) should reduce the slant towards disadvantaged groups in an egalitarian supreme court. However, because some cases skip IACs and because IACs are unlikely to be perfectly egalitarian themselves, it remains true that egalitarian highest courts should rule for disadvantaged defendants more often than advantaged defendants. Williams (1995) provides one of the few existing studies on racial bias in criminal appeals. He analyzed IAC reviews of 183 convictions from two Florida trial courts and found that Black defendants' convictions were upheld at a higher rate than those of Whites.

I argue that a judge's willingness to apply ideological or legal preferences consistently is conditional on their belief in the defendant's actual guilt and deservingness of protection, even if that is irrelevant to the legal questions presented. The more certain a judge is that a defendant is guilty and undeserving of the protection of their rights, the less willing the judge is to consistently apply legal or ideological principles to that defendant's benefit. By finding ways to justify rejecting a valid appeal, judges avoid significant cognitive dissonance.

These inconsistencies are unlikely to be random or innocuous. If a judge is predisposed to believing that certain types of people are more likely to be guilty, or less deserving of their rights as citizens, than others, then the judge's underlying (even implicit) feelings about defendants could reduce the chance that the judge consistently applies restrictions to police, prosecutor, and trial-judge behavior. This argument neither requires nor precludes a consciously prejudiced thought process. It only requires implicit biases in the way people associate different groups with criminality and deservingness of rights and citizenship. Little research examines the role of implicit biases in appellate decision making, though Sen (2015) shows that appellate judges treat opinions written by Black judges differently than those written by White judges in ways that implicit bias may partially explain.

These implicit biases may be most pernicious with regard to race. In this paper, I argue that, because Black and Hispanic men are more strongly associated with criminality than White men, appellate judges are less likely to rule consistently in their favor, even when the law or the judge's own ideology indicate that they should. I focus on men because perceptions of crime among races are also highly gendered, creating intersectional biases. Minority men are associated with heightened criminality more than minority women are (Steffensmeier, Ulmer, and Kramer 1998). Black and Hispanic men are the most disadvantaged by stereotyping. They

make up a majority of arrestees, convicts, and prison inmates. In fact, female defendants are sufficiently infrequent that robust statistical tests are difficult without considerable multi-state data. While I focus on the impact of race among male defendants, further research on similar biases among female defendants is warranted.

There are at least two possible sources of biases in appellate outcomes: Internal Dissonance and Audience Pressure. Internal Dissonance refers to the judge's own reluctance to consistently apply legal or ideological principles in ways that may benefit those she believes are guilty and undeserving. This is the general effect I describe in the preceding paragraphs. The second, Audience Pressure, refers to the pressure placed on the judge to decide a case in a specific direction because of the beliefs the audiences to the decision have about the defendant's guilt and deservingness. Audiences can include friends and colleagues, reappointing governors and legislators, or the voters who re-elect judges back to office (Baum 2009).

If an elected judge believes that voters will find an overturned conviction unjust, she will be more hesitant to rule in favor of that defendant. If a large number of voters hold common prejudices, then it is possible for even a perfectly egalitarian judge to be incentivized through the principal-agent dynamics of the electoral system to replicate those biases. Research indicates that minorities are associated with heightened levels of criminality in America (Levinson, Cai, and Young 2010). If Black defendants are more likely to be seen as guilty than White defendants, a decision striking down a conviction or shortening a sentence for a Black defendant is more likely to be seen as an injustice committed by the court than a similar decision for a White defendant. There is even some evidence (Peffley and Hurwitz 2007) that voters may favor a system that is disproportionately punitive to out-group members.

Consider the ways that voters are most likely to interact with an appellate criminal decision: through a bombastic headline or television news report that states something like "Court Overturns Houston Man's Conviction for Violent Rape," accompanied by a mug shot and the defendant's name, perhaps also with a photo of the victim. The legal arguments will be a small or nonexistent part of this media coverage, which will instead focus overwhelmingly on the alleged (and typically irrelevant) facts of the case. In rare cases, this type of coverage can be fatal to a judge's career. Race adds an extra dimension in news coverage, often reinforcing underlying stereotypes (Dixon 2008). Newspapers are more likely to describe Black convicts in a dehumanized way (Goff et al 2008). And when the mug shot is of a Black face, and the name presented is a name frequently used by African Americans, the voting public may be more likely to believe the defendant was guilty and that injustice was done by the court (Dixon and Azocar 2007). There is also evidence that White citizens are more emotionally impacted and more likely to remember news stories about Black criminals than White criminals (Dixon and Maddox 2005). These implicit racialized messages can have a powerful impact on political behavior (Mendelberg 2001). Given that popular salience is dangerous in low-information races such as judicial elections, the potential for "lenient" decisions for minority defendants to cut through the clutter of local news and be remembered on election day threatens a judge's job security.

This theory leads to expectations that differ considerably from the starting theory that features of appellate courts limit opportunities for racial bias. If minority men are more strongly associated with criminality than White men, then the belief in the guilt and deservedness of appellants is likely correlated with race. Thus, I expect that minority men do worse than White men in criminal appeals. This is presented in the Racial Bias Hypothesis:

> **Racial Bias Hypothesis**: Judges are less likely, all else equal, to vote for Black and Hispanic defendants than White defendants in criminal appellate decisions.

Because prejudicial beliefs regarding criminality are strongest about Black men in America, I expect that the largest biases will be against Black male defendants. I expect a lesser effect against Hispanic male defendants. Existing research indicates that Hispanics fall somewhere between Whites and African Americans in their experience and perceptions of criminal injustice (Unah and Wright 2015), and I expect that my results will follow the same pattern. This is presented in the Racial Gradient Hypothesis:

> **Racial Gradient Hypothesis**: Judges are less likely, all else equal, to vote for Black defendants than Hispanic defendants in criminal appellate decisions.

I also consider the impact of Audience Pressures. While Individual Dissonance and Audience Pressures are present in each case, they vary by judge and by the timing of the case. Audience Pressure is highest when judges face re-election in the near future. A case in an election year has the highest probability of being observed by voters through media coverage and then remembered for primary and general-election votes. These cases are the ones most likely to be pivotal in a judge's re-election. Therefore, I expect that judges will be less likely to vote for minority defendants in the year before their election than in the earlier years of their judicial terms. This is stated in the Election Amplification Hypothesis:

> **Election Amplification Hypothesis**: The gaps between White and minority defendants in the likelihood, all else equal, of receiving a judge's vote is higher in election years than in non-election years.

In the next section, I explore whether these expectations are supported by evidence from one large, racially diverse state which uses frequent partisan judicial elections: Texas. I analyze more than 8,000 votes cast in cases between 1995 and 2010 in the Texas Court of Criminal Appeals. I find that Black defendants do considerably worse than White defendants in Texas, while Hispanic defendants do about as well as White defendants. These differences spike in the

year before a judge faces reelection, with both minority races doing appreciably worse than White defendants. Further, I find that the differences are split by party, with White defendants doing better among Republican judges and minority defendants doing better among Democratic judges. Finally, I show that because of these partisan correlations, racial disparities have actually increased rather than decreased as Texas has transitioned to single-party Republican governance.

**Empirical Testing**

To purely test my arguments, I would observe how a judge would vote in two different situations: both feature the same exact set of facts and legal arguments, but in one the judge perceives the defendant to be White, and in the second they perceive the defendant to be a racial minority. Given that I only observe one of these outcomes in any one case, and that I am unable to randomize racial perception experimentally (as in Bertrand and Mullainathan [2004]), I must approximate such an experiment with observational data. My research strategy is to rely on defendants' names, the strongest signal of racial identity consistently available to appellate judges. This requires picking a set of observed cases that satisfy three criteria. First, there must be sufficient quantity and quality of data to draw valid inferences. Second, there must be sufficient racial variation among defendants. And third, the state(s) used should feature electoral retention to allow for testing the Election Amplification Hypothesis. One particular state satisfies these requirements: Texas. Final criminal appeals in Texas are heard by nine judges, who are re-elected in partisan elections every six years. Texas has a substantial docket which features considerable racial diversity. Texas also provides the most thorough data. The main limitation of using Texas is that its distinct culture and politics may reduce generalizability.

Texas has two courts of last resort – the Supreme Court, which hears civil appeals, and the Court of Criminal Appeals (CCA), which hears criminal appeals. Decisions of the CCA are

only appealable to the Supreme Court of the United States, and thus the CCA is the final arbiter of justice in more than 99% of its cases. Hall and Windett (2013) provide a dataset of state court of last resort cases from across the country for the years 1995-2010, from which I take all full, published decisions coded under the topic category "Criminal Law & Procedure." This is not an exhaustive set; other criminal cases are included in other topics. This subset does represent a substantial portion of the docket, and makes for a manageable and reliable coding effort.

On top of the Hall and Windett data, I code whether defendants won their appeal. This was determined by analysis of case summaries and, where necessary, the full case decisions. A defendant is deemed to have "won" when they succeeded on at least one of their points of appeal, improving their situation relative to where they were before the CCA heard the case. Because I am interested in the behavior of judges, I transform these data into a judge-vote level dataset. The dependent variable is whether the judge cast a **Pro-Defendant Vote**, which is defined as a vote for the majority outcome (including concurrences) when the defendant wins, or votes against the majority outcome when the state wins.

The population of criminal defendants, appellants, and convicts in Texas is overwhelmingly comprised of the same three racial groups as the overall Texas population: Blacks, Whites, and Hispanics. These are the three races I analyze in this paper. Unfortunately, it is extremely difficult to link defendants with their demographic information across large ranges of time. Public records are inconsistently available and difficult to accurately match, especially for cases from the 20$^{th}$ century and defendants who ultimately won their cases. Therefore, I estimate the probability that defendants were of a particular race and sex based on the defendant's name and their presence in the Texas criminal justice system.

Estimating race probabilities for criminal-justice populations is difficult because the underlying racial distribution (the "base rate") is skewed from just about any other quantified population. Relying on national averages or local census demographics (such as recent work by Imai & Khanna [2016]) will dramatically underestimate minority probabilities. My solution to this problem is to use a demographically identified roster of more than 140,000 individuals in the Texas prison system as of May 2015. I assume that the prison population is demographically very similar to the population of defendants on appeal at the CCA. I use racial distributions within the prison roster for each given name and surname to estimate the probability that a defendant is Black, Hispanic, White, or another race. I replicate this process on given names to estimate defendants' sexes. A detailed description of this method is available in Appendix A.

This name-based probabilistic approach has several benefits.[2] First, it is consonant with my theoretical arguments. Appellate judges' most consistent interaction with defendants is through their name, making it the easiest and most widely available way to cognitively estimate a defendant's race in the appellate context. I assume that these internal estimates are correlated with the frequency of those names in the Texas criminal justice system. Second, a probabilistic approach is preferable to a forced-classification system, because this recognizes the uncertainty in name classifications.[3] A name that is overwhelmingly indicative of one race should be treated differently than a name that is only marginally indicative. The more certain a judge is of the race of the defendant in the case before them, the more pronounced the effect should be. I expect bias to increase as the likelihood that the name is uniquely associated with minority usage increases.

[2] Alternatively, one might employ Mechanical Turk workers to predict the race of a given defendant's name. This has the strength of being a direct measure of "perception," but is limited by the fact that online coders do not have the same perceptions of Texan judges, and thus do not account for the base racial distribution of the Texas criminal-defendant population.
[3] A forced classification system does produce similar results to those I present in this paper.

Thus, my primary independent variables are **Probability Black**, **Probability Hispanic**, and **Probability White**. **Probability Other**, a final variable, captures the probability that the defendant is of another race. Because there are so few defendants of other races in Texas, I make no inferences about them and they are included only so that all probabilities sum to one.

I replicate this process to estimate the sex of first names, though the distributions of these names are considerably different, with the vast majority of cases featuring names that are 95% or greater of one sex or the other. In this paper, I only analyze defendants whose names are associated with men at a rate of 95% or higher. This cutoff rule captures 93% of defendants into one of two sexes, with only 7% left over with difficult-to-classify names. In total, about 86% of defendants had names associated at least 95% of the time with men. Table 1 presents the ten typically male names with the highest estimated probability of being Black, Hispanic, and White. [Table 1 about here.]

*Case-Specific Control Variables*

In addition to defendants' demographics, I also recognize that cases are not factually and legally identical. It is difficult to translate the complexity of an alleged crime into a quantitative coding scheme, but simple facts can still help identify meaningful differences between cases.[4] Specifically, I control for the types of charges mentioned in the procedural summary attached to each case in the Hall and Windett data in a variable set named **Charges**.[5] It is possible to have multiple charges in a case, but I do not differentiate between the number of "counts" of each charge or any aggravating or mitigating factors. Even though appellate cases are ostensibly

[4] Previous research has found that the race of victims of violent crime is an important factor. There is some evidence that prosecutors seek harsher penalties on those who harm Whites and women (Songer and Unah 2006). If judges behave analogously, it may be that victim demographics are as important as those of defendants. But these data are not available.
[5] Examples include Murder, Sexual Assault, and Drug Violations.

about *legal* rather than factual questions, defendants who face very serious charges are less likely to succeed on appeal. This poses significant identification problems because minority defendants are also more likely to be charged with serious crimes such as murder in the cases I analyze. By controlling for charges, I eliminate the bias in racial outcomes owed to the racial distribution of charges, which would not be owed to the behavior of the CCA.

Given the focus on legal issues in these cases, I also attempt to control for the type of legal question that the case raises, which I take from the procedural summary text in the Hall and Windett dataset. From this text, I create a variety of dummy variables indicating that a certain issue was raised in the case, which I call **Legal Issues**.[6] To the extent that the CCA is systematically more likely to grant challenges under certain grounds, these variables control for that heterogeneity. Finally, I also include, as part of Legal Issues, a measure of whether the appeal was brought by the Defendant or the Prosecution. **Defendant Appealed** takes the value "1" when the Defendant brought the appeal and "0" otherwise. This variable provides the best indication available of whether the existing law was in favor of the defendant.

To better identify demographic effects, I also include judge and year fixed effects. Year fixed effects account for a variety of heterogeneous factors, such as the prevailing political circumstances in a given year, variations in the docket, or changes in prosecutorial strategies. While this controls out many sources of racial bias, it does so in a way that helps isolate the impact of the CCA. Judge fixed effects control for the heterogeneity in judges' predispositions towards defendants, which may include their ideological beliefs about defendants' rights or their interpretation of the law on police and prosecutor conduct.[7]

---

[6] Examples include Jury Instructions, Confrontation Clause, and Admissibility of Evidence.
[7] While judge demographics are themselves interesting, there is insufficient variation to conduct such analysis. During this time period, there was only one minority judge.

*Results*

First, I estimate a pair of logistic regressions with the previously mentioned variables for male defendants before the Texas CCA between 1995 and 2010. The dependent variable in both models is Pro-Defendant Vote. The independent variables of note are Probability Black, Probability Hispanic, and Probability White. Each ranges from 0 to 1, increasing in the probability that the defendant is of that race. Model 1 in Table 2 uses Probability White as the base category, allowing for the comparison of both minority races to White defendants. This tests the Racial Bias Hypothesis, which predicts that both groups should be less successful than White defendants, all else equal. Model 2 uses Probability Hispanic as the base category, allowing a clear comparison of Black and Hispanic defendants. This tests the Racial Gradient Hypothesis, which predicts a hierarchy of outcomes in which Black defendants do worse than Hispanic defendants, who do worse than White defendants.

[Table 2 about here.]

The initial results in Table 1 indicate partial support for both hypotheses. In comparison to White (Model 1) and Hispanic (Model 2) Probabilities, as Probability Black increased, the likelihood of a given judge voting for the defendant, all else equal, decreased. The probability that the defendant was Hispanic, relative to White, was also negatively correlated with pro-defendant votes, but not to a degree distinguishable from zero. Thus, these results do not indicate a clear gradient, but rather that Black defendants did appreciably worse with CCA judges than both Hispanic and White defendants, who were statistically indistinguishable. Figure 1 plots the effects from Model 1. In Figure 1 (and subsequent figures), the presented marginal effect is the change between a name 100% associated with being white relative to

100% associated with being the labeled race, with the probability of all other races fixed at zero.[8] Figure 1 indicates that, all else equal, a defendant with a name 100% associated with being Black would be expected to receive a judge's vote about eight percentage points less often than a defendant with a name 100% associated with being White.[9]

[Figure 1 about here.]

The eight-percentage-point effect is aggregated across all judges using the median judge's baseline propensity to rule for defendants. An alternative way to analyze the bias is in the heterogeneity of the effect across judges: the judge-level marginal effect of going from a probability zero to a probability one of being Black, for example. To do this, I re-estimate Model 1 in Table 2 with judge-specific slopes for each racial probability variable in addition to the judge-specific intercepts. From this model (not presented), I estimate judge-level effects, shown in Figures 2-A and 2-B. The effect sizes vary considerably. For the effect of frequently Black names relative to frequently White names (in Figure 2-A), judges range from a 33-percentage-point positive effect to a 26-percentage-point negative effect. The majority, and the longest serving, of the judges, have negative effects, leading to the aggregate negative effect of Probability Black presented in Table 2. For the effect of frequently Hispanic names relative to frequently White names (in Figure 2-B), the effects range from a 39-percentage-point positive effect to a 11-percentage-point negative effect. The patterns are very similar between the two graphs ($r$=0.92), with the notable difference being smaller absolute effect sizes for those who

[8] Additionally, all control variables are fixed at their medians and the median (by punitiveness) year- and judge-specific intercepts are used.

[9] In this paper, I use the phrases "Black defendants," "White defendants," and "Hispanic defendants" for ease of reading. These are most accurately understood as names which are 100% associated with that race and 0% associated with all other races. These are not unrealistic points of analysis: each race features many names with 80% or higher association, as well as names with near-zero levels of association.

ruled less often in favor of Hispanic defendants. These small negative effects by the longest serving judges in the dataset, combined with the large negative effects by judges who sat for fewer cases, combines to produce the insignificant effect for Probability Hispanic in Table 2. [Figures 2-A and 2-B about here.]

While at first glance this may appear to be a random distribution of judges, a closer look at the names reveals a troubling pattern: the dashed line at zero almost perfectly separates judges by party in each figure. The only Republican with a significant positive effect in Figure 2-A is Judge Verla Holland, who served part of a single term and did not run for re-election. Every other judge who voted for Black defendants, all else equal, at higher rates than White defendants was a Democrat. Figure 2-B is very similar. The easiest interpretation of Figures 2-A and 2-B is that judges' partisanship is correlated with how they treated Black and Hispanic defendants relative to White defendants.[10] Partisanship explains approximately 55% of the variation of effects in both Figure 2A and 2B. By comparison, common judicial ideology measures, such as SDirt scores by Hall, Harden, and Windett (2015), or CFscores by Bonica and Woodruff (2014), add little information on top of a dichotomous party indicator.[11] This means that more conservative Republicans behaved similarly to more moderate Republicans. But Republicans as a group behaved very differently than Democrats. This points to the possibility that this separation is evidence of the Audience Pressures which I discussed in conjunction with the

[10] A simple linear regression of the marginal effect estimates and party implies that Republicans on average had an effect size for the impact of having a more frequently Black name 29.3 percentage points more negative than the average Democrat. For Hispanic names, the difference was 22.1 percentage points. Both of these differences are statistically significant at $p<0.01$.

[11] Neither measure is significant in addition to the party indicator for explaining judge-specific effect sizes with regard to Probability Black. Only SDirt scores add some information on top of party to explaining effect sizes with regard to Probability Hispanic. For these analyses, I averaged SDirt scores for a judge's career to estimate a single score.

Electoral Amplification Hypothesis. If Republican voters differ from Democratic voters in how they expect criminal outcomes to be decided, then this can induce separation in the elected judges themselves. Because all nine seats on the CCA are filled via statewide elections, there is no reason to expect ideological specialization induced by region. Next, I analyze the role of elections in producing racial disparities in appellate outcomes.

*Acquiescing to Audience Pressures*

In the preceding section, I showed that Black defendants did systematically worse than White and Hispanic defendants overall between 1995 and 2010 in the Texas CCA. But I also argue that the pressure that voters put on judges means that these effects should be amplified in votes cast within a year of a judge's election. Therefore, I estimate a second set of models, including the variable **Election Year**, which takes the value "1" for votes taken within one year of the judge's re-election and "0" when the election is more than 365 days away, the judge had announced an intention to retire, or was prohibited by law from running for an additional term. Texas conducts partisan elections for CCA judges, with partisan spring primaries and runoffs followed by November general elections. Typically, three justices are up for re-election every even-numbered year. Thus, in any given election year, a third of the court is running to keep their jobs while two thirds have a reprieve for at least two more years. Judges can expect the most scrutiny of their votes during their election season, when things voters learn are more likely to be remembered and acted upon. If judges fear being associated with leniency on minority defendants and thus are more punitive of minorities, then they should have heightened biases in their election-year votes. To test this, I interact Election Year with each of the racial probability variables, and otherwise replicate Table 2. The results are presented in Table 3. Because Figures 2-A and 2-B revealed such stark differences between Republicans and Democrats, I

present three different models: one with all votes, one with votes cast by Republicans, and one with votes cast by Democrats. Because there are relatively few observations for Democrats in these data (about 15% of votes), their estimates are more uncertain.

[Table 3 about here.]

Table 3 provides several important insights. The clearest is that Republicans and Democrats voted completely differently and responded to election-year pressures differently as well. This comports with the results in Figures 2A and 2B. Because Republicans cast the overwhelming majority of votes in this time period, their results drive the overall results. Because of this, I focus on the party-specific rather than overall results. Republicans showed no statistically distinguishable bias against defendants with typically Hispanic names outside of election years, but a substantial difference emerged in election years. Republicans voted for Black defendants less often than Whites in, and out of, election years, but again, the gap increased significantly when the judge was up for re-election. For Democrats, none of these patterns hold. There are no statistically significant patterns to distinguish election years from non-election years. And, generally, Democrats voted for White defendants less often, all else equal, than Black and Hispanic defendants. Thus, I find support for the Election Amplification Hypothesis, but only among Republican judges. I present these results in Figure 3.

[Figure 3 about here.]

As Figure 3 indicates, votes cast in election years and those cast earlier in a judge's term are considerably different for Republicans. Among Republicans outside of an electoral context, the effect of a name with 100% Black association (again relative to a name with 100% white association) was -9.8 percentage points. In election years, this spiked to -36.4 percentage points. For Hispanic names, a non-election year -3.1-percentage-point (insignificant) effect became a -

23.5-percentage-point effect. A further interesting aspect of the Republican response to electoral pressures is that Whites actually did better in election-year votes with Republicans than in non-electoral contexts. In Figure 4, I present the effect of the electoral context on each Race's likelihood of receiving a favorable Republican vote.

[Figure 4 about here.]

Among Republicans, election-year votes are statistically distinguishable from non-election year votes at the $p<0.05$ level for Probability Black and Probability White, and at the $p<0.10$ level for Probability Hispanic, but the effects for Whites was the opposite of the effect for Blacks and Hispanics. White defendants were 11.9-percentage-points *more* likely to receive votes from Republican election-year judges than earlier in those judges' terms, while Black and Hispanic defendants were 8.9-percentage-points and 6.3-percentage-points less likely, respectively. Because Whites received more votes and minorities fewer votes in election years, the gaps between the races were exacerbated, as shown in Figure 3.

The results in Table 3 are striking and informative. First, they reinforce our understanding that the public can alter judicial behavior through the electoral process. Second, they show that the short memories and attention spans of voters result in considerably more influence at the end of a term than earlier in a term. Third, they show that this influence can go beyond simply communicating ideological positions about law and order, and even suggest that the public's own in-group/out-group biases shape judicial behavior. Notably, a gap between White and Hispanic defendants only emerges among Republicans in election years.

This electoral result also helps eliminate a variety of possible alternative explanations for the results in Table 2. It is possible that minority defendants are systematically different from White defendants in attributes relevant to their appellate outcomes. For example, White

defendants are likely wealthier on average and thus can afford more expensive private defense attorneys. While all states must provide public defenders, privately purchased attorneys can expend more time and effort on each case. White defendants likely also have higher levels of social capital, which might make them more likely to obtain skilled pro bono counsel or organizational assistance. And minority defendants may advance ideologically different arguments on appeal. Yet, none of these factors change dramatically from year to year. The data indicate that facing a re-election alters judicial behavior relative to the races in a way that cannot be explained by these or any other systematic differences between the races.

The only explanation left for this disparity is that Republican judges up for re-election soon do not want to be seen as having voted to overturn the conviction or stiff sentence of a minority defendant, and so vote more punitively for those defendants. That White defendants receive *more* votes in election years is surprising. Most research on election effects in criminal law implies higher levels of punitiveness (for example, Huber and Gordon 2004). Yet, my findings imply that campaigning Republican judges in Texas between 1995 and 2010 were more lenient for Whites while being more punitive for Hispanics and Blacks.

*The Changing Disparities in Racial Outcomes Over Time*

One hope in any study of racial disparities is for progress – that, over time, the biases inherent in institutions recede to the point of being undetectable. Texas presents an interesting case because the Democratic party persisted there at the state level for some time after it had disintegrated in federal races. The nine-judge CCA had at least one Democrat through 2000, but has been purely Republican since. As with much of Texas state government, the 1990s proved to be the last gasp of southern Democrats on the CCA. If racial disparities are a phenomenon of Republican judges, and Republicans have come to dominate the court more in the 21st century

than they did in the 20th, this indicates a troubling possibility: racial disparities have actually increased in Texas since the late 1990s as Democrats on the court were replaced by Republicans.

To test this, I analyze two distinct time periods separately. In Model 1, I include the years 1995 through 2000, while Model 2 includes 2001 through 2010. 2000 is a distinct year because it was the last year that a Democrat served on the court. Thus, the time period in Model 2 contains the final years of the transitional period of multi-party Texas while Model 3 contains the current single-party Republican Texas. The results are presented in Table 4.

[Table 4 about here.]

The results in Table 4 indicate that disadvantages for racial minorities in appellate outcomes are not artifacts of a slowly dissipating bias in the Texas criminal justice system, but rather that they re-emerged in the 21st century after being undetectable or even reversed in the late 1990s. Both Blacks and Hispanics were more likely to receive votes, all else equal than White defendants in the 1995-2000 period. Recall that this is possibly a "neutral" outcome, if the CCA was correcting errors in biased lower courts. However, from 2001 to 2010, Black and Hispanic defendants did considerably worse than White defendants. The results from Table 4 are presented as marginal effects in Figure 5. The effect sizes exceed 20% and 8% for Blacks and Hispanics, respectively. These large disparities defy easy alternative explanations.[12]

[Figure 5 about here.]

**Conclusion**

The criminal justice process – from investigation to incarceration and ultimately re-integration to society – is long and varied. Voluminous research has focused on the disparities

[12] A model with a linear time trend and interactions between the trend and the racial probability variables also produces significant results for the trend and each interaction. White defendants received more votes as time passed, while Black and Hispanic defendants received fewer.

that emerge between Whites and minorities in a variety of these steps, including arrests, prosecutions, trials, and incarceration. One step that has remained understudied, however, is the appellate process, in which convicted defendants appeal to appellate courts for the chance to have their conviction overturned or their sentence reduced. This stage is important because not only do these appeals help determine the outcomes for the defendants before them, but they also help shape the law that subsequent defendants will be tried under. An appeal about one investigative technique will apply to lower court interpretations of that technique going forward. Thus, it is all the more important that the racial identities of defendants impact their likelihood of winning votes on appeal. The unwillingness to let those perceived to be less deserving enjoy the full protection of their rights can erode the strength of those rights in unforeseen ways.

For judges, it may be quite difficult to overturn a conviction when the best available evidence (no matter how unlawfully obtained) indicates that the defendant committed a horrible crime. In this paper, I argue for a version of appellate decision making in criminal law in which consistent application of ideological and legal beliefs is constrained by a desire not to allow legal procedure to invalidate the convictions of guilty defendants. This approach to judicial decision making then explains how racial bias may intrude on the otherwise legal-theoretical world of appellate advocacy. If implicit racial biases alter a judge's belief about the guilt of a defendant and the extent to which they deserve the full protection of defendants' rights, this can lead to systematic biases in outcomes. Additionally, when judges anticipate that their audiences will act on similar prejudices, then those biases can be transmitted to them.

I support each of these predictions with evidence from the highest court of criminal appeals in Texas. The results indicate that Black and Hispanic defendants received substantially fewer defendants' protections than White defendants did in Texas from 1995 to 2010. And even

more troublingly, the data strongly indicate that election pressures induce even higher disparities in appellate outcomes. Judicial elections have long been the subject of concern for scholars because they place the law at the whims of voters and the donors who are necessary to pay for campaigns. Scholars have also worried that elections make judges excessively punitive in criminal law, seeking to avoid a "weak on crime" label. My findings present the most troubling aspect of judicial elections: that they may transmit the biases of American society directly into the legal system. Voters may elect judges who reward their "tribe" at the expense of others. This would be an untenable outcome for a justice system that relies on fairness and consistency.

These results also reinforce one reason why minorities persistently make up a disproportionate share of the American prison population: not only do they face an uneven playing field in law enforcement and trials, but they continue to do so on appeal as well. Instead of correcting for trial-level bias, the Texas CCA likely exacerbated racial disparities in the first decade of this century. This leads to more lives spent in prison, which significantly impacts the ways people perceive and interact with government and politics (Weaver and Lerman 2010). To the extent that racial biases in the appellate process allow for and encourage racial biases in the trial process, the lack of genuine equality at the appellate level may have far-reaching and underappreciated consequences for American political participation in minority communities.

These results defy easy solutions. One possible policy change that could reduce these outcome biases may be to make efforts to reduce the amount of defendant information that is transmitted in the appeals process. Cases should be more frequently classified by anonymous indicators prior to publication. Factual summaries and written documents should minimize the use of names, pronouns, and other indicators of demographics. This could help promote greater fairness. In theory, this should pose a minimal burden to the work of an appeals court, which is

supposed to be fundamentally about the application of legal principles that seldom vary by race. Certain types of cases – such as those that feature claims of racial bias within the case itself – may pose additional problems. But greater consistency in the application of principles will come from diminishing the amount of (irrelevant) factual information presented on appeal.

This research is an initial effort in the lengthy exploration of racial effects in criminal appeals. Though I highlight effects in Texas criminal contexts, other states in other regions can also be analyzed. Texas's selection system – partisan competitive elections – is not typical across the country. Further research could reveal that some types of selection systems help create more equality in justice outcomes. And though I research racial biases against minority male defendants, the biases against female defendants also merit analysis of their own with data that provide more volume and variation by gender.

References

Baum, Lawrence. 2009. *Judges and their audiences: A perspective on judicial behavior*. Princeton University Press.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Marker Discrimination." *The American Economic Review* 94(4): 991-1013.

Bonica, Adam, and Michael J. Woodruff. 2014. "A Common-Space Measure of State Supreme Court Ideology." *Journal of Law, Economics, and Organization* 31(3): 472-498.

Canes-Wrone, Brandice, Tom S. Clark, and Jason P. Kelly 2014. "Judicial Selection and Death Penalty Decisions," *American Political Science Review* 108(1):23-39.

Clark, Tom S. 2011. *The Limits of Judicial Independence*. New York: Cambridge University Press.

Demuth, Stephen, and Darrell Steffensmeier. 2004. "Ethnicity Effects on Sentence Outcomes in Large Urban Courts: Comparisons Among White, Black, and Hispanic Defendants." *Social Science Quarterly* 85(4): 994-1011.

Dixon, Travis L. 2008. "Crime News and Racialized Beliefs: Understanding the Relationship Between Local News Viewing and Perceptions of African Americans and Crime." *Journal of Communication* 58: 106-25.

Dixon, Travis L., and Cristina L. Azocar. 2007. "Priming Crime and Activating Blackness: Understanding the Psychological Impact of the Overrepresentation of Blacks as Lawbreakers on Television News." *Journal of Communication* 57: 229-53.

Dixon, Travis L., and Keith B. Maddox. 2005. "Skin Tone, Crime News, and Social Reality Judgments: Priming the Stereotype of the Dark and Dangerous Black Criminal." *Journal of Applied Social Psychology* 35(8): 1555-70.

Eberhardt, Jennifer L., Paul G. Davies, Valerie J. Purdie-Vaughns, and Sheri Lynn Johnson. 2006. "Looking Deathworthy: Perceived Stereotypicality of Black Defendants Predicts Capital-Sentencing Outcomes." *Psychological Science* 17(5): 383-6.

Eberhardt, Jennifer L., Phillip A. Goff, Valerie J. Purdie, and Paul G. Davies. 2004. "Seeing Black: Race, Crime, and Visual Processing." *Journal of Personality and Social Psychology* 87(6): 876-93.

Eisenberg, Theodore, and Sheri Lynn Johnson. 2004. "Implicit Racial Attitudes of Death Penalty Lawyers." *Depaul Law Review* 53: 1539-59.

Epstein, Lee, Jack Knight, and Andrew D. Martin. 2001. "The Supreme Court as a Strategic National Policymaker." *Emory Law Journal* 50: 583-611.

Goff, Phillip A., Melissa J. Williams, Jennifer L. Eberhardt, and Matthew Christian Jackson. 2008. "Not Yet Human: Implicit Knowledge, Historical Dehumanization, and Contemporary Consequences." *Journal of Personality and Social Psychology* 94(2): 292-306.

Windett, Jason H., Jeffrey J. Harden, and Matthew E. K. Hall. 2015. "Estimating Dynamic Ideal Points for State Supreme Courts." *Political Analysis* 23(3): 461-469.

Hall, Matthew, and Jason Windett. 2013. "New Data on State Supreme Court Cases." *State Politics and Policy Quarterly* 13:427-45.

Huber, Gregory, and Sanford Gordon. 2004. "Accountability and Coercion: Is Justice Blind When it runs for Office?" *American Journal of Political Science* 48(2): 247-63.

Hurwitz, Jon, and Mark Peffley. 1997. "Public Perceptions of Race and Crime: The Role of Racial Stereotypes." *American Journal of Political Science* 41(2): 375-401.

Imai, Kosuke, and Kabir Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Record." *Political Analysis* 24(2): 263-72.

Kang, Jerry, Mark Bennett, Devon Carbado, Pam Casey, Nianjana Dasgupta, David Faigman, Rachel Godsil, Anthony G. Greenwald, Justin Levinson, and Jennifer Mnookin. 2012 "Implicit bias in the courtroom." *UCLA Law Review* 59: 1124-86.

Levinson, Justin D., Huajian Cai, and Danielle Young. 2010. "Guilty by Implicit Racial Bias: The Guilty/Not Guilty Implicit Association Test." *Ohio State Journal of Criminal Law* 8: 187-208.

Levinson, Justin D., and Danielle Young. 2010. "Different Shades of Bias: Skin Tone, Implicit Racial Bias, and Judgments of Ambiguous Evidence." *West Virginia Law Review* 112: 307-50.

Mendelberg, Tali. 2001. *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton: Princeton University Press.

Peffley, Mark, Jon Hurwitz, and Paul M. Sniderman. 1997. "Racial Stereotypes and Whites' Political Views of Blacks in the Context of Welfare and Crime." *American Journal of Political Science* 41(1): 30-60.

Peffley, Mark, and Jon Hurwitz. 2007. "Persuasion and Resistance: Race and the Death Penality in America." *American Journal of Political Science* 51(4): 996-1012.

Quillian, Lincoln, and Devah Pager. 2001. "Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime." *American Journal of Sociology* 107(3): 717-67.

Rachlinksi, Jeffrey J., Sheri Lynn Johnson, Andrew J. Wistrich, and Christ Guthrie. 2009. "Does Unconscious Racial Bias Affect Trial Judges?" *Notre Dame Law Review* 84: 1195-1246.

Richardson, L. Song, and Phillip A Goff. 2013. "Implicit Racial Bias in Public Defender Triage." *Yale Law Journal* 122: 2626-49.

Segal, Jeffrey, and Harold Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge: Cambridge University Press.

Sen, Maya. 2015. "Is Justice Really Blind? Race and Reversal in US Courts." *Journal of Legal Studies* 44: 187-229.

Skogan, Wesley. 2006. "Asymmetry in the Impact of Encounters with Police." *Policing and Society* 16(2): 99-126.

Smith, Robert J., and Justin D. Levinson. 2012. "The Impact of Implicit Racial Bias on the Exercise of Prosecutorial Discretion." *Seattle University Law Review* 35: 795-826.

Songer, Michael J., and Isaac Unah. 2006. "The Effect of Race, Gender, and Location on Prosecutorial Decisions to Seek the Death Penalty in South Carolina." *South Carolina Law Review* 58: 161-211.

Steffensmeier, Darrell, Jeffery Ulmer, and John Kramer. 1998. "The Interaction of Race, Gender, and Age in Criminal Sentencing: The Punishment Cost of Being Young, Black, and Male." *Criminology* 36(4): 763-97.

Tonry, Michael. 2010. "The Social, Psychological, and Political Causes of Racial Disparities in the American Criminal Justice System." *Journal of Crime and Justice* 39: 273-312.

Unah, Isaac, and Valerie Wright. 2015. "Divided by Race: Differences in the Perception of Injustice." In *Deadly Injustice: Trayvon Martin, Race, and the Criminal Justice System*, Devon Johnson, Patricia Y. Warren, and Amy Farrell, eds. New York University Press: New York.

Weaver, Vesla, and Amy E. Lerman. 2010. "Political Consequences of the Carceral State." *American Political Science Review* 104(4): 817-33.

Weitzer, Ronald. 2000. "Racialized Policing: Residents' Perceptions in Three Neighborhoods." *Law and Society Review* 34(1): 129-55.

Williams, Jimmy J. 1995. "Race of Appellant, Sentencing Guidelines, and Decisionmaking in Criminal Appeals: A Research Note." *Journal of Criminal Justice* 23(1): 83:91.

**Table 1**. Example High-Probability Names, by Race

| Black | Hispanic | White |
|---|---|---|
| Cedric Ransom | Ramiro Ibarra | William Juvrud |
| Tyrone Leday | Santos Mireles | Jeffrey Wood |
| Kelvin Bowie | Ramiro Flores | Bradley Dyar |
| Jermaine Herron | Alejandro Mata | Randall Kellar |
| Demetrius Jackson | Esteban Arevalo | Jeffrey Marx |
| Jermaine Murphy | Raul Mata | Rex Newman |
| Maurice Bluitt | Pedro Salazar | Cody Mitten |
| Mack Modden | Marcelino Aguirre | Rick Rhoades |
| Cedric Buchanan | Francisco Marquez | Robert Blankenship |
| Dejuan Handy | Jose Barajas | Kenneth Mayer |

**Table 2.** Black, but not Hispanic, Defendants do Systematically Worse than Whites in Texas Criminal Appeals

| DV: Pro-Def. Vote | (1) White Base | (2) Hispanic Base |
|---|---|---|
| Probability Black | **-0.38 (0.17)** | **-0.28 (0.08)** |
| Probability Hispanic | -0.10 (0.11) | *Base Category* |
| Probability White | *Base Category* | 0.10 (0.11) |
| N | 8,152 | 8,152 |
| Clusters | 17 Judges | 17 Judges |

Each model also contains year and judge fixed effects, Charges, Legal Issues, and Probability Other Race. Coefficients presented are logistic regression coefficients with clustered standard errors in parentheses. Similar results are obtained using robust or unadjusted standard errors. Bolded coefficients are significant at the $p<0.05$ level, two-sided tests.

**Table 3**. Black and Hispanic Defendants do Worse than Whites in Votes by Judges Seeking Reelection

| DV: Pro-Def. Vote | (1) All Judges | (2) Republicans | (3) Democrats |
|---|---|---|---|
| Election Year | **0.35 (0.17)** | **0.54 (0.11)** | -0.61 (0.65) |
| Probability Black | -0.27 (0.17) | **-0.47 (0.17)** | **0.81 (0.30)** |
| Elec. X Prob. Black | **-0.76 (0.33)** | **-1.11 (0.21)** | 0.55 (0.99) |
| Probability Hispanic | -0.01 (0.10) | -0.15 (0.09) | **0.74 (0.23)** |
| Elec. X Prob. Hispanic | **-0.61 (0.28)** | **-0.87 (0.26)** | 0.56 (0.74) |
| N | 8,152 | 6,954 | 1,191 |
| Clusters | 17 Judges | 11 Judges | 6 Judges |

Each model also contains year and judge fixed effects, Charges, Legal Issues, and Probability Other Race. Coefficients presented are logistic regression coefficients with clustered standard errors in parentheses. Similar results are obtained with unclustered standard errors. The sample sizes in Models 2 and 3 do not sum to the sample size of Model 1 due to a small number of observations being dropped from Model 2 due to perfect prediction. Bolded coefficients are significant at the $p<0.05$ level, two-sided tests.

**Table 4.** Racial Disparities Increased Between 1995 and 2010

| DV: Pro-Def. Vote | (1) 1995 to 2000 | (2) 2001 to 2010 |
|---|---|---|
| Probability Black | **0.65 (0.18)** | **-0.97 (0.11)** |
| Probability Hispanic | **0.41 (0.20)** | **-0.38 (0.04)** |
| N | 2,974 | 5,151 |
| Clusters | 14 Judges | 10 Judges |

Each model also contains year and judge fixed effects, Charges, Legal Issues, and Probability Other Race. Coefficients presented are logistic regression coefficients with clustered standard errors in parentheses. Similar results are obtained with unclustered standard errors. Bolded coefficients are significant at the $p<0.05$ level, two-sided tests.

**Figure 1.** Black Defendants Receive Fewer Votes than White Defendants

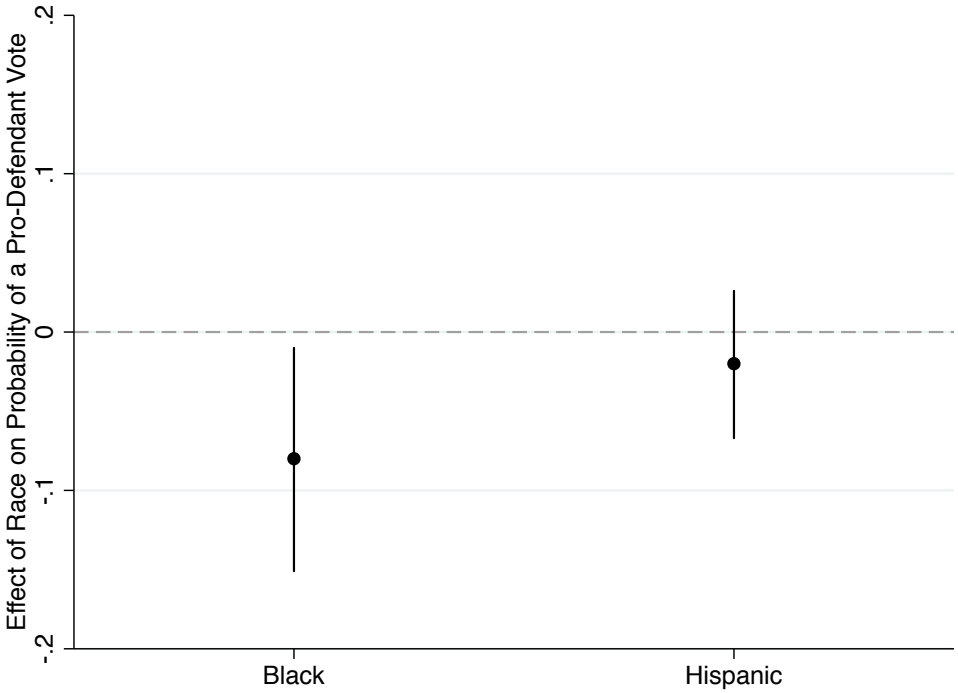


**Figure 2-A.** The Heterogeneity of Black – White Vote Bias, by Judge and Party

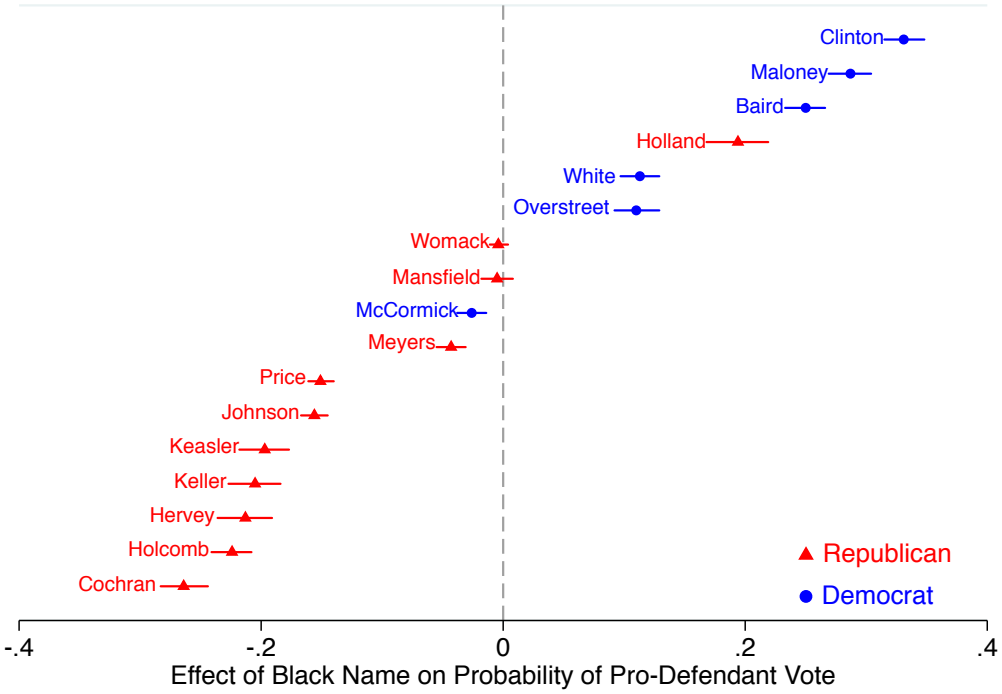



**Figure 2-B.** The Heterogeneity of Hispanic – White Vote Bias, by Judge and Party

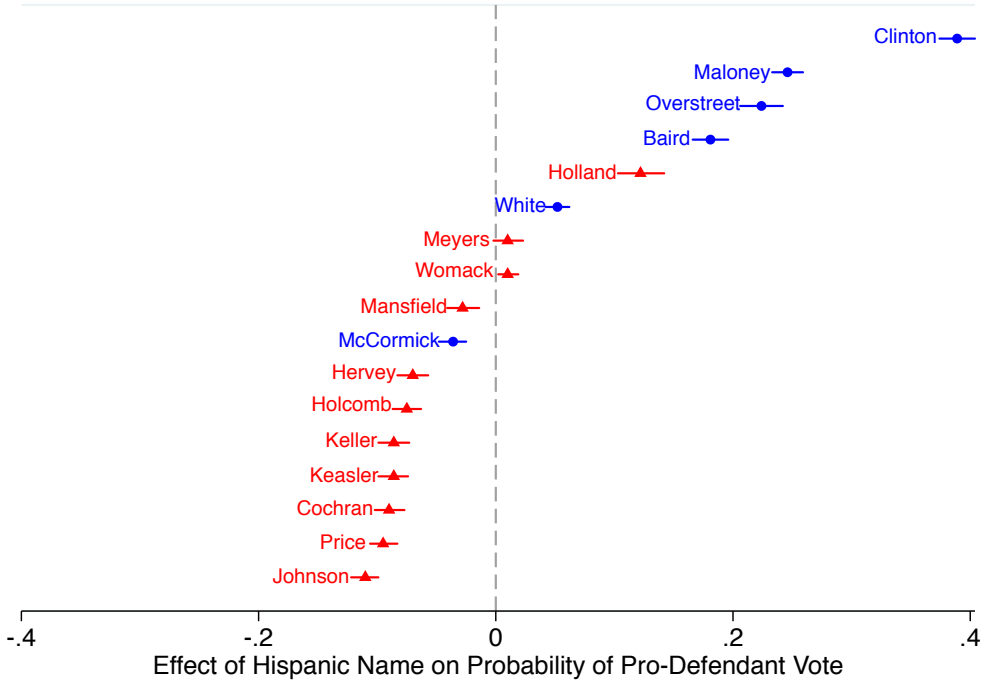


**Figure 3.** Minority Defendants do Worse with Judges in Election Years

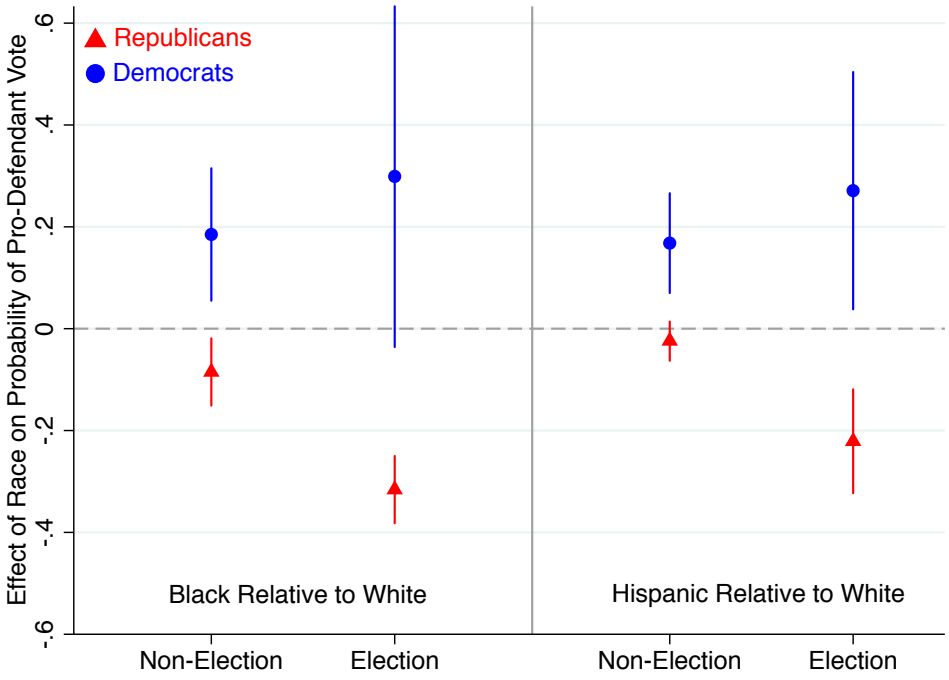



**Figure 4.** Elections Increase Success of White Defendants and Reduce Success of Minority Defendants (Republican Justices)

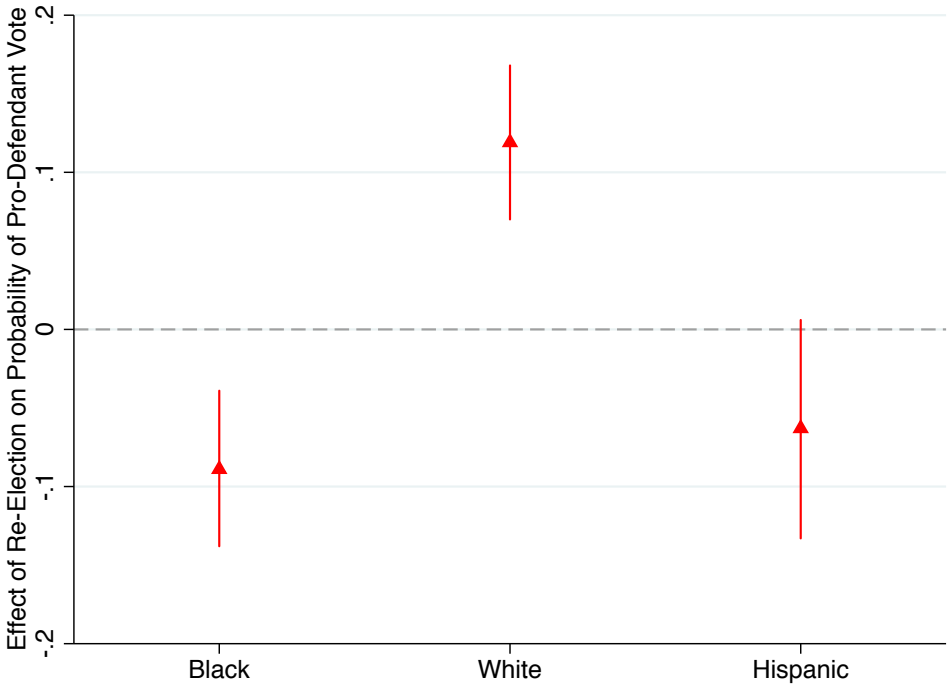
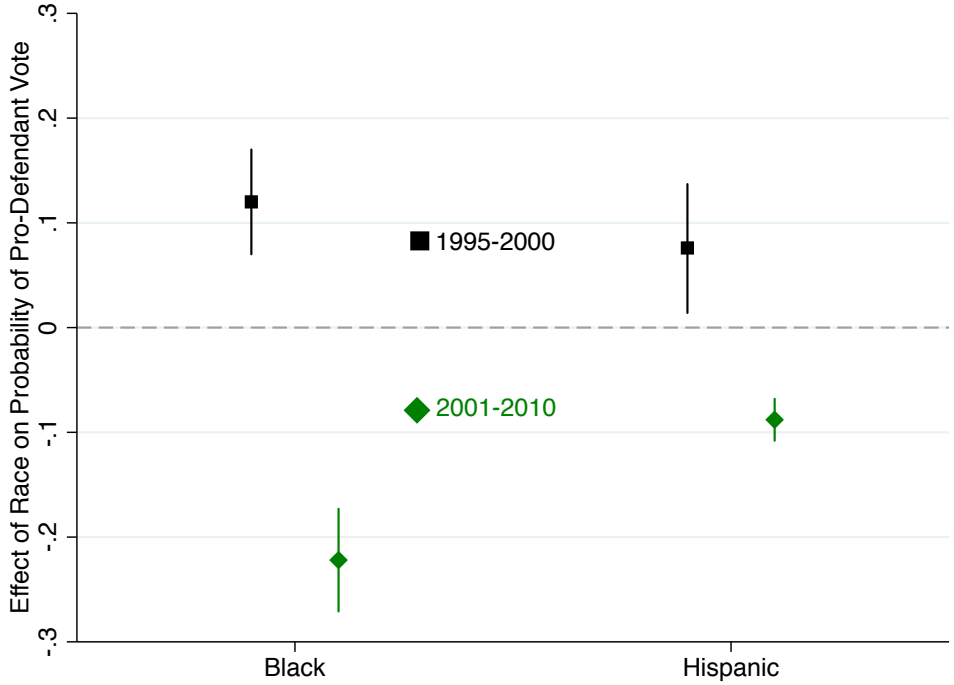


**Figure 5.** Disparities Sharply Changed as the Court Flipped from Mixed to Unanimously Republican



Appendix A: Estimating Racial Probabilities Based on Names

Names can be probabilistically associated with sex and race nationally using census data; however, this ignores the problem of base population rates in localized areas that do not match national averages (Imai & Khanna 2016). Criminal justice populations pose a distinct version of this problem because the populations of defendants in a locality neither matches national nor local population-at-large statistics. Minorities are overrepresented in the criminal justice system, thus any measure based off of even local population demographics will underestimate the probability that defendants are minorities. Fortunately, Texas provides an extensive list of names that closely matches the underlying population I am interested in: a list of all people held in Texas prisons, as of May 2016, published by the Texas Department of Criminal Justice. Using the May 2016 roster to estimate names of defendants who appeared before the CCA between 1995 to 2010 poses some measurement error risk. Naming patterns change over time and the birth-year distribution of the 2016 roster is likely later than from 1995-2010 CCA defendants. Similarly, if minority defendants receive longer sentences on average, they may be somewhat overrepresented in the prison roster than in the appellate defendant pool. Despite these limitations, the Texas prison roster remains the best available option, more closely matching the defendant pool than national or local demographics do.

While these data are insufficient to match to specific defendants in many cases, they do provide thousands of different given and surnames, each identified by sex and race, of people with significant contact with the Texas criminal-justice system. Prisons include not only convicted felons, but also those who await trial, and those with pending appeals. Thus, the population recorded in the Texas prisoner roster is the closest available match to the racial and

gender demographics of criminal appeals in Texas. From these, I estimate demographics for each name in the Hall and Windett dataset of appeals before the CCA.

Estimating based on names straddles two possible measures: perceptions of race and the actual racial identity of the defendant. Given the lack of matchable records, it is not possible to say which race or races each defendant personally identified with.[1] Perceptions of racial identity are also of interest, but difficult and expensive to validly code for a large sample of varied names. I opt for an approach that straddles these two under the assumption that perceptions of race in names are highly correlated with the frequency with which individuals with that name identify with that race. Thus, by estimating the probability that someone of a given name would be recorded in the Texas prison roster as a certain race, I construct a measure that I believe to be highly correlated with both outward perceptions of race as well as with personal identification. Because appellate judges' most frequent interaction with defendants is through their names, this measurement strategy is consonant with my theoretical arguments. This research design is in some ways analogous to studies of responses to resumes with names more commonly used by one race or another (Bertrand and Mullainathan 2004). As I cannot experimentally manipulate the names of actual defendants, I rely instead on observed variation in names.

I construct these estimates as follows: Each name has two possible components – the given name and the surname. For each defendant surname in my dataset of criminal appeals, I credit to the surname component of that defendant's probability of being any particular race the percentage of prisoners in the Texas prison roster with that name who were that race. To avoid unrealistic zero or one estimates for uncommon names, I regress each name to the prison-roster

[1] Whether public records reflect personal identification, parental identification, or the perceptions of state officials is a separate question.

mean by one observation.[2] Names that do not occur in the prison roster are credited with each race's overall share of the prison roster. Second, there are given names. I again use the observed frequency from the Texas prison roster and regress by one observation to avoid absolute estimates. Ultimately, I multiply the given name and surname probabilities for each race together for each defendant, and then rescale so that the race probabilities sum to one. Thus, the probability that any defendant is Black can be expressed as follows:

$$\frac{\Pr(Black|surname) * \Pr(Black|given\ name)}{\sum (\Pr(Race'|surname) * \Pr\ (Race'|given\ name))}$$

For each name, there are four probabilities: the probability that the defendant is White, Black, Hispanic, or another race (one percent of the Texas prison total).

Consider the name Juan Rodriguez. The surname Rodriguez occurs 1,523 times in the Texas prisoners roster. Of these, 95.7% are listed as Hispanic, 3.1% as Black, 1.1% as White, and 0.1% as another race. The given name Juan occurs 1,839 times in the Texas prisoners roster. Of these, 96.0% are listed as Hispanic, 2.0% as White, and 2.0% as Black. None are listed as another race. To create an estimate of Juan Rodriguez's probability of being each race, I multiply these probabilities[3] and divide by the total of the multiplied probabilities, as follows:

$$Pr(\text{Hispanic} \mid \text{Rodriguez, Juan}) = \frac{0.957*0.96}{(0.957*0.96)+(0.031*0.02)+(0.011*0.02)+(0.001*0.000)} = 0.999$$

[2] Thus, a surname occurring seven times in the dataset would produce an ultimate surname probability of ((7*observed prisoner rate of that race and name) + (overall rate of that race in the prison roster))/8.

[3] These probabilities include the one-observation regression to the mean, which has minimal effect on such a frequently occurring name.

Thus, I calculate that there is a 0.999 probability that a defendant named Juan Rodriguez is Hispanic. The remaining races combined make up approximately a 0.001 probability. I assume that because of this frequency, a judge who encounters the case Texas v. Juan Rodriquez will assume that the defendant is Hispanic. I also assume that there is a similarly high probability that the defendant Juan Rodriquez would himself identify as a Hispanic. I repeat this process to estimate the probability that each defendant is classified as male or female in the Texas criminal-justice system. Continuing the example of Juan Rodriguez, all 1,839 defendants named Juan in the Texas prison roster are classified as male, leading to an estimated male probability (once regressed to the mean by one observation) of 0.999. This is not unusual. 97.6% of the unique first names in the Texas prison roster featured exclusively male or exclusively female classifications.

References

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Marker Discrimination." *The American Economic Review* 94(4): 991-1013.

Imai, Kosuke, and Kabir Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Record." *Political Analysis* 24(2): 263-72.