Ethnic Bias in Judicial Decision-making: Evidence from Criminal Appeals in Kenya*

Donghyun Danny Choi,* J. Andrew Harris,** Fiona Shen-Bayh†

Forthcoming at *American Political Science Review*

Abstract

Understanding sources of judicial bias is essential for establishing due process. To date, theories of judicial decision-making are rooted in ranked societies with majority-minority group cleavages, leaving unanswered which groups are more prone to express bias and whether it is motivated by in-group favoritism or out-group hostility. We examine judicial bias in Kenya, a diverse society which features a more complex ethnic landscape. While research in comparative and African politics emphasizes instrumental motivations underpinning ethnic identity, we examine the psychological, implicit biases driving judicial outcomes. Using data from Kenyan criminal appeals and the conditional random assignment of judges to cases, we show judges grant coethnic appeals at a 3 to 5 percentage points higher rate than noncoethnic appeals. To understand mechanisms, we use word embeddings to analyze the sentiment of written judgments. Judges use more trust-related terms writing for coethnics, suggesting that in-group favoritism motivates coethnic bias in this context.

*We thank Zeinab Abbas, Laboni Bayen, Shehryar Hanif, Soohyun Hwangbo, Cindy Koyier, Fanisi Mbozi, Kaelin Miller, Khoi Ngo, Morgan Tompkins, and Cornelius Wekesa Lupao for excellent research assistance. Comments from Leonardo Arriola, Sarah Brierley, Asli Cansunar, Gidon Cohen, Justine Davis, Andrew Eggers, Aaron Erlich, Matthew Gichohi, Siri Gloppen, Robin Harding, Mai Hassan, Nahomi Ichino, Patrick Kuhn, Harrison Mbori, Susanne Mueller, Noah Nathan, George Ofosu, Cristian Pop-Eleches, Georg Picot, Pia Raffler, Lise Rakner, Pedro Rodriguez, Nelson Ruiz, Meshack Simati, Arthur Spirling, Peter Van der Windt, Neil Visalvanich, Nick Vivyan, Zach Warner, and Stephane Wolton have significantly improved this manuscript. The work also benefited from excellent comments during presentations at the University of Bergen, the London School of Economics and Political Science, the University of Oxford, Durham University, Trinity College Dublin, the Global Law and Politics Workshop, and the MIT Political Behavior Workshop.

*Choi: Assistant Professor of Political Science, University of Pittsburgh (email: dannychoi@pitt.edu). **Harris: Assistant Professor, Division of Social Science, New York University Abu Dhabi (email: andy.harris@nyu.edu). †Shen-Bayh: Assistant Professor of Government, William & Mary (email: fshenbayh@wm.edu). The authors contributed equally to this project. Their names are listed alphabetically.

††Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: https://doi.org/10.7910/DVN/JCSOB8.

Introduction

Ethnicity shapes political life in Africa, from voting (Ferree, 2006; Adida et al., 2017) and the distribution of public goods (Ejdemyr, Kramon and Robinson, 2018) to political violence (Montalvo and Reynal-Querol, 2005). But whether ethnicity plays a similar role in African courts remains under-examined. Studies of judicial decision-making in these contexts have largely overlooked the role of ethnicity in the courtroom, focusing instead on questions of socioeconomic inequality, regional disparities, and gender discrimination in judicial outcomes (e.g., Tripp, 2004; Gloppen and Kanyongolo, 2007; Ndulo, 2011). However, judges in multiethnic societies typically navigate complex political terrain (Helmke, 2002; Iaryczower, Spiller and Tommasi, 2002), particularly in new democracies where the rule of law is often perceived to be weak or under threat (O'Donnell, 2004). Given these dynamics, for Africa in particular and the Global South more broadly, exploring the link between ethnic identity and judicial outcomes may help us understand how justice is delivered in ethnically diverse societies.

To date, research on the role of ethnicity in judicial decision-making has largely focused on the American experience, highlighting how racial biases in American courtrooms undermine due process (Harris and Sen, 2019).[1] Yet, there are both theoretical and inferential benefits to studying judicial bias beyond the U.S. context. Consider that very few societies have witnessed the domination of a single ethnoracial group for as long and with such impunity as White Americans, making the U.S. a somewhat unique setting to study group-based biases. Furthermore, these studies typically center around the dominant majority-minority cleavage (White-Black), which has restricted the ability of researchers to probe the potential heterogeneity in bias across different groups, and thereby limited understanding of the mechanisms underlying group-level differences.

Against this backdrop, the questions motivating this study are twofold: does ethnic bias affect judicial decision-making in African courts? If so, what is the nature of and motivation for this bias? Building on the insights of social identity theory, we argue that ethnic favoritism in the courtroom results from the subconcious, implicit biases held among judges towards the appellants in a given

[1]Recently, Hou and Truex (2020) and Tuñón and Feierherd (2020) examine ethnic and partisan bias in the Chinese and Argentine judiciaries, respectively.

case. That is, we contend that judicial bias along ethnic lines is driven by in-group attachments and out-group antagonisms. Framing ethnic bias in these terms represents a departure from conventional theories of ethnic politics in Africa, which more often portray ethnic favoritism as being driven by political or material considerations. In contrast to these works, we treat such bias as a by-product of historical, structural, and institutional factors that shape relations among ethnic groups rather than deliberate calculations serving instrumental goals.

To test our claims, we turn to Kenya, a relatively new democracy where ethnic divisions structure partisanship and patronage across the state (Kramon and Posner, 2016), including within the legal sector (Odote and Musumba, 2016). The Kenyan judiciary has recently become a locus of democratic contestation, ruling on presidential election controversies (Kanyinga and Odote, 2019). Such cases highlight the instrumental dimensions of ethnic conflict in Kenyan superior courts, but whether identity shapes judicial decision-making in quotidian legal proceedings remains underexplored.

We focus on criminal appeals in the Kenyan High Court: cases that are not overtly political (thus unlikely to be instrumentally motivated) and reflect day-to-day Kenyan jurisprudence. We built a dataset of almost 10,000 criminal appeals at 39 Kenyan High Court stations from 2003 to 2017. Our empirical approach leverages the fact that cases filed at a court station are assigned to individual judges based on the filing date and existing caseloads, independent of other case- and court-specific characteristics like judge and appellant identity. We rely on this conditional quasi-random assignment of cases to estimate the effect of judge-defendant coethnicity on the success of criminal appeals. To better understand the motivations for bias, we use word-embeddings to measure levels of expressed trust (a marker of in-group favoritism) and disgust (a marker of out-group derogation) in written legal judgments.

Our analysis reveals significant evidence of coethnic bias in judicial decision-making in Kenya. Across a range of empirical specifications, judges are between 3–5 percentage points more likely to rule in favor of a coethnic than a non-coethnic defendant. Yet these estimates mask significant heterogeneity across groups; effects are primarily concentrated among judges who are ethnically Kikuyu, Kenya's largest and politically dominant ethnic group. We also show that judges express more trust sentiment in judgments for coethnic than non-coethnic defendants, consonant with notions that in-group favoritism and not out-group derogation motivates bias. These findings suggest that coethnic bias in Kenyan courtrooms manifests in the legal outcome and the judgment's language.

Our paper makes several contributions. To our knowledge, this study is the first to systematically examine judicial decision-making in criminal appeals in an African context. Research on African courts has predominantly focused on superior court politics, especially constitutional cases (e.g., Widner, 2001; Vondoepp and Ellett, 2011). While such cases are undeniably consequential, they are relatively rare, reflecting elite-level politics rather than how due process typically operates for everyday people. By focusing on criminal appeals, a more routine area of judicial decision-making, our study speaks to the broader challenges of ensuring free and fair justice for the citizens of new democracies in the Global South (Gibson and Caldeira, 2003; Levi, Sacks and Tyler, 2009).

Second, we study implicit bias in a real institutional setting, building on existing lab-in-the-field work (Lowes et al., 2015; Oppedal Berge et al., 2019). Whereas existing work uses games and implicit association tests (IATs) to probe coethnic bias (e.g., Blum, Hazlett and Posner, 2021), we consider identity-based bias in an important institutional setting – Kenyan appeals courts. In doing so, we show that the intensity of such bias varies across ethnic groups: consonant with social dominance theories, Kikuyu judges are the ones driving coethnic bias in appeals outcomes.[2]

Third, to differentiate explanations of in-group favoritism and out-group derogation, we consider the empirical implications of social psychology work relating emotions to in-group versus out-group biases (Brewer, 1999; Hodson et al., 2013). We evaluate the mechanisms of bias through natural language processing techniques to measure affective patterns associated with favoritism or derogation (e.g., Rice and Zorn, 2019). Our work joins recent work using text-as-data to understand the emotions, personalities, and states of mind of elites and citizens (Boussalis et al., Forthcoming; Osnabrugge, Hobolt and Rodon, 2021; Ramey et al., 2019).

[2]Results for the Kamba are not robust, as shown in the appendix.

Ethnic identity and judicial bias in comparative perspective

Beyond instrumentalism

To understand whether coethnic bias manifests in judicial decision-making, we begin by problematizing the dominant account of ethnic identity in the developing world. An influential literature in comparative politics adopts an instrumentalist framework to theorize why ascriptive group identities become salient and why agents of the state may privilege coethnics (e.g., Bates, 1974; Chandra, 2007). Focusing on contexts in which voters condition vote choice on the receipt of material inducements (Van de Walle, 2001), scholars in this tradition argue that political elites favor coethnics in the provision of public goods since doing so advances the interests of key in-group supporters (Kramon and Posner, 2013).[3]

To what extent do instrumentalist theories generalize to the judiciary? Kenyan judges are not elected but appointed by the President, usually on the advice of the Judicial Service Commission. While coethnic favoritism plausibly plays a role in judicial appointments, it is unclear why such considerations would translate into everyday judicial decision-making. This is particularly true of cases concerning individuals with limited means to exert pressure on the courts, such as low-income persons involved in petty crimes or disputes. These cases do not concern significant political players, nor do they have an overt political agenda; judges thus lack clear incentive to rule a certain way. That is to say, there is no clear strategic rationale for judges to privilege coethnic over non-coethnic defendants when adjudicating everyday disputes.

While instrumental motivations seem largely absent in quotidian cases, judges may still possess unconscious, implicit biases predisposing them to be more harsh or lenient towards certain groups (Redfield, 2017). To date, the study of implicit biases in judicial decision-making has focused on Western judiciaries (Cohen and Yang, 2019) where it has been shown that ideology is a strong predictor of judicial outcomes (and ideology is strongly correlated with race and ethnicity) (Harris and Sen, 2019). However, judicial decision-making in these contexts is not typically characterized as strate-

[3] See Letsa (2020) on expressive voting for an important exception.

gic; that is, motivated by the expectation of material reward. Such bias is seen instead as reflecting longstanding cleavages between majority and minority groups. For example, studies of U.S. courts sometimes frame the treatment of Black defendants by white judges against the broader history of Black subjugation and White supremacy (Clarke, 2018). Similar approaches have been used to understand majority-minority dynamics in other countries, such as Jewish-Arab interactions in Israeli courts (Grossman et al., 2016).

However, from a non-instrumental perspective, judges should be susceptible to group-based attachments and antagonisms just like ordinary citizens. In particular, implicit biases may lead judges to assign positive regard toward in-group members and negative regard toward out-group members (Oyserman et al., 2003; Paluck and Green, 2009). This kind of unconscious, affective bias may subsequently inform cognitive elements that shape how judges assign blame and responsibility (e.g., Fiske and Pavelchak, 1986) or perceive the moral character of the accused (e.g., Alicke, 2000; Nadler and McDonnell, 2012), which can in turn shape how judges interpret cases and render judgment.

Mechanisms of bias: in-group favoritism or out-group derogation?

The preceding discussion has highlighted how social psychological factors can shape judicial bias absent instrumental incentives. But is judicial bias a manifestation of favoritism towards members of the in-group, or hostility towards members of the out-group?

Theories of social identity and self-categorization suggest that coethnic bias is likely driven by in-group favoritism (Tajfel and Turner, 1979; Turner, Reynolds et al., 2001). Self-categorization as an in-group member requires the "assimilation of the self to the in-group category prototype and enhanced similarity to other in-group members" (Hewstone, Rubin and Willis, 2001). Research in this tradition also posits that individuals have a tendency to assign positive valence (such as trust, esteem, and positive regard) to members of their in-group without conscious reflection (Brewer, 1999; Levin and Sidanius, 1999; Otten and Wentura, 2001). Alternatively, it might be the case that coethnic bias is driven by out-group derogation. While group identification does not always lead to feelings of hostility towards outsiders, ample evidence suggests that such sentiments can be easily triggered in polarized societies that have a history of intergroup conflict (Stephan and Stephan, 2013). In such settings, members of the out-group are more likely to be perceived as a threat to the in-group, which can arouse feelings of fear, disgust, and thus antagonism towards the source of the threat (Sherif and Sherif, 1953). Whether judicial bias is driven by in-group positivity or out-group negativity generates different implications for our expectations about the affective content of written judgments, an issue we turn to in the next section.

Implicit bias, legal outcomes, and legal writing

Recent work on judicial bias has examined specific types of stereotypes, including race and gender (e.g. Rice, Rhodes and Nteta, 2019; Ash, Chen and Ornaghi, 2020). In contrast to these approaches, our aim is to uncover the affective correlates of in-group favoritism and out-group derogation. To this end, we join a growing trend in the social sciences using text-as-data to trace difficult to measure concepts like sentiment and personality (Osnabrugge, Hobolt and Rodon, 2021; Gennaro and Ash, 2021; Ramey et al., 2019).

A standard measure of judicial bias treats case outcomes as a discrete variable: whether an appeal is allowed or denied. We examine to what extent personal characteristics of the defendant (rather than legal matters of the case) affect how judges rule. However, attributing bias to either in-group or out-group attitudes is challenging if we only look at final verdicts, which lack context for understanding their motivation. That is, without a neutral control condition, we might observe that differential treatment exists, leaving unknown the motivation for that difference (Gazal-Ayal and Sulitzeanu-Kenan, 2010; Gill, Kagan and Marouf, 2017; Harris and Sen, 2019).

However, the full text of a legal decision may provide more analytical leverage on these mechanisms. Before delivering a verdict, judges summarize each case and explain their decision's logic. A written judgment can be conceived of as the final output of cognitive processes in which a judge uses evidence, legal concepts, and judicial discretion to support their decision (Simon, 1998; Maroney, 2016; Rachlinski and Wistrich, 2017).[4] Our research builds on a robust literature examining how affective

[4]See also Simon (2004) on cognitive coherence and legal decisions.

framing and cues shape legal reasoning and outcomes (Wistrich, Rachlinski and Guthrie, 2015; Black et al., 2011; Beattey Jr, Matsuura and Jeglic, 2014; Liu and Li, 2019; Wistrich, Rachlinski and Guthrie, 2015). Using this lens, we posit that if judges are indeed making decisions based on their implicit biases, such sentiments are likely reflected in their written legal judgments.

To dissect judgments for evidence of in-group versus out-group bias, we first identify what kinds of affective content we would expect to see if either mechanism were in play. Existing research finds in-group favoritism to be strongly associated with notions of trust; in the African context, such studies tend to portray trust as a basic behavioral regularity in coethnic interactions (e.g., Robinson, 2020; Arriola, Choi and Gichohi, 2021). Social psychology research similarly argues that trust underpins in-group favoritism (Allport, 1954; Brewer, 1999), whereas disgust more often accompanies feelings of out-group derogation (Hodson et al., 2013; Mackie, Devos and Smith, 2000). From these findings, it is reasonable to expect that the presence of trust-related and disgust-related language in appeals judgments would imply biases respectively suggesting in-group favoritism and out-group derogation.[5]

Are certain groups more predisposed to bias?

Our main claim is that psychological mechanisms can predispose judges to demonstrate group-based bias in judicial decision-making. However, this does not necessarily mean that judges of different groups are *equally* susceptible to such biases. Literature in social psychology and sociology suggest that prevalence of biases among groups corresponds to hierarchical status (e.g., Mullen, Brown and Smith, 1992; Hagendoorn, 1995). Two prominent theoretical strains are social identity and dominance theories and realistic group conflict theory.

Research on social identity theory and social dominance orientation predicts that processes of social comparison and social identification may lead members of higher-status groups to be more likely to discriminate between the in-group and out-group(s) (Tajfel and Turner, 1979; Sidanius et al., 2000). "Dominant" groups derive esteem from their superior status, reinforcing the value and worth they attach to their dominant position (Sachdev and Bourhis, 1987). Being at the top of the social

[5]Content-based bias in written texts is not uncommon. For instance, readily-available text sources like letters of recommendation reveal race- and gender-based differences (Heath et al., 2019; Grimm et al., 2020).

hierarchy makes such groups more predisposed to preserving the status quo as a means of sustaining their privileged access to resources and power (Sidanius and Pratto, 2001; Harkness, 2018). Evidence reveals higher levels of in-group bias among dominant group members in hierarchical societies in contexts as diverse as Israel, India, the Netherlands, Northern Ireland, and the United States (Levin and Sidanius, 1999; Levin, 2004).[6]

In the Kenyan case, this framework implies that bias should be concentrated among judges belonging to the "dominant" ethnic group: the Kikuyu. As the numerically largest ethnic group in Kenya, Kikuyus have seen their leaders occupy the presidency three times (Jomo Kenyatta, Mwai Kibaki, and Uhuru Kenyatta), comprising most of the post-independence period. Their political dominance has spread to other branches of government, including the judiciary, with multiple Kikuyu jurists serving as Chief Justice.[7]

By contrast, realistic group conflict theory in social psychology and political science argue that bias and discrimination mirror political conflicts among competing groups (Sherif and Sherif, 1953; Horowitz, 2000; Sambanis and Shayo, 2013); intergroup animosity is rooted in competition over scarce resources (Gurr, 2015). Out-group discrimination and conflict thus reflect existing grievances over the distribution of material goods (Sherif, 1988). Other work in this tradition contends that intergroup cleavages result from the subordination of certain groups, entrenching intergroup animosity and systems of marginalization and exclusion (Wimmer, Cederman and Min, 2009).

In the Kenyan context, this second perspective would lead us to expect bias to arise between specific constellations of ethnic groups. The Kikuyu and Kalenjin ethnic groups have long monopolized political power and engendered grievances among excluded groups, notably the Luo (Widner, 1992). The Kikuyu-Luo rivalry stems from the early independence struggle for national control between

[6]It is important to note that existing work has defined whether a group is "socially dominant" in political, economic, and demographic terms. Observationally, in many cases the group that holds a dominant position in one of these domains tend to also be dominant on other domains. But it is not uncommon that different forms of power reside with different groups in society. In circumstances where this is in fact the case, we would choose to apply various conceptualizations of "social dominance" (based on political versus economic versus demographic power) to the analysis, and empirically examine whether these different conceptualizations affect our inferences.

[7]Kikuyus (Mwai Kibaki, Uhuru Kenyatta) occupied the presidency throughout the entire span of our analysis. The judiciary was led by Kikuyu Chief Justice Johnson Gicheru for more than half of that time.

Kikuyu President Jomo Kenyatta and Luo Vice President Oginga Odinga.[8] Kalenjin-Luo relations have also been strained for decades, reflecting similar national-level power struggles.[9] More recently, Kikuyu-Kalenjin hostilities have intensified since the return to multiparty politics in the 1990s, culminating in the infamous post-election violence of 2007/08 that led to more than 1000 dead and more than 600,000 displaced (Lynch, 2014). Given Kenya's long history of ethnic rivalries at every level of government from the presidency downwards, these tensions might be on display in the courtroom.

Observable Implications

The preceding discussion suggests three observable implications. First, we expect that judges will exhibit identity-based implicit bias in their decisions, even absent strategic or instrumental considerations. Second, these biases may align with coethnicity and be driven by in-group favoritism; alternatively, out-group derogation may drive negative cross-ethnic biases. Patterns of trust- or disgust-related language in written judgments should thus mirror in-group or out-group motivated biases. Third, such biases might be higher among judges who belong to the dominant ethnic group; alternatively, bias may follow historical patterns of cross-ethnic rivalry.

Context: Judicial Outcomes in Kenya and Africa

In contrast to U.S. courts, with distinct federal- and state-level jurisdictions, Kenya's unitary judiciary organizes courts nationwide. We focus on criminal appeals in Kenya's Judiciary, which originate in Kenya's lowest-level magistrate courts and would be analogous to local or county-level courts in the U.S. However, while U.S. state courts abide by different laws depending on the locality, each of Kenyan high courts is governed by the same set of rules and procedures. To seek redress from lower court judgements, a defendant may appeal to the High Court, the next level in Kenya's judicial hierarchy.[10]

[8] The Kenyatta-Odinga rivalry arguably continues to structure inter-ethnic relations to this day (Branch, 2011); the two most recent presidential elections featured Kenyatta and Odinga's sons as rival candidates.

[9] Kenyatta's successor, President Daniel arap Moi, a Kalenjin, established a system of patron-client relations that was designed to perpetuate Kalenjin dominance over the Kenyan state (Throup and Hornsby, 1998).

[10] High court appeals are analogous to the U.S. state-level courts of appeals rather than the more commonly studied U.S. federal courts.

Appeals decisions are a difficult test for theories of identity-based bias due to the incentive structures confronting judges who hear appellate cases. In particular, high court decisions are made publicly available online through Kenya's National Council on Law Reporting. This transparency is partly intended to inform precedent-based judicial decision-making across the country, but publicizing such information incentivizes judges to mitigate any indication of bias in their legal reasoning. Furthermore, due to their higher status in the judiciary, high court judges generally have less anonymity than lower-level magistrates, wherein greater public recognition underscores the need to maintain at least the appearance of impartiality.

To date, little academic work focuses on judicial outcomes in lower courts in Kenya, or Africa more broadly.[11] The focus instead has been on superior court politics. For example, Widner (2001)'s study of courts in postcolonial Africa examines the development of judicial independence at the upper echelons of the judiciary, centering on the storied career of Francis Nyalali, Chief Justice of Tanzania. Other works in this field have also focused on the mindsets of High or Supreme Court justices, including Gloppen and Kanyongolo (2007)'s analysis of the High Court and the Supreme Court of Appeal in Malawi, VonDoepp (2006)'s comparative study of the High Courts of Malawi and Zambia, and Vondoepp and Ellett (2011)'s comparative analysis of executive-judicial relations in five Commonwealth African countries.

The understandable focus on superior courts among scholars of African judiciaries has helped illuminate elite-level decision-making, particularly questions of constitutional jurisprudence or judicial review, but these approaches highlight an unavoidable challenge to studying superior courts across the region: while politically salient, high-profile cases are relatively infrequent, making the relationship between legal output and judicial identity difficult to assess systematically.[12] Furthermore, the number of superior court justices is small, making inference difficult.[13] To understand the influ-

[11] Kinyanjui and Akech (2016) documents sentencing disparities in lowest-level magistrates courts. This and other work describe problems facing Kenyan courts in terms of corruption and inefficiency but do not consider how the identity of legal parties may influence judicial decision-making (e.g., JMVB, 2016; JSC, 2019).

[12] In VonDoepp (2006)'s case, there were 82 cases for Malawi and 116 for Zambia.

[13] As Gloppen and Kanyongolo (2007) observe, "when discussing the judiciary in Malawi, it is important to remember that this is just a handful of people – the higher judiciary (the High Court and the Supreme Court of Appeal) comprises only twenty-four judges in all – and, as an institution, it has but a short history."

ence of ethnic identity on more quotidian legal outcomes, more detailed data on case outcomes is required.

Data and Methods

We focus on two elements of appeals case outcomes: whether or not an appeal succeeds and the presence of trust- and disgust-related language relating to in-group favoritism and out-group derogation. For both analyses, we rely on a corpus of appeals from the Kenya Law Cases Database, an online repository of court rulings maintained by the National Council for Law Reporting (Kenya Law). We downloaded 9,545 criminal appeals rulings issued by the High Court between January 1, 2003 and December 31, 2017. Each ruling contained the full text of the judgment, including the nature of the alleged crime, the original sentence, the date of the ruling, and the county court wherein the case was heard. Using regular expressions, we created a set of case-level variables for analysis.

To classify our main dependent variable – whether the appeal was allowed or denied – we relied on regular expressions as well as a hand-coded classification scheme. Hand coding was necessary because judicial writing style varies by judge, especially with respect to their judicial logic. Wherever regular expressions could not fully capture the idiosyncracies of legal reasoning, we relied on human coders to complete our classification of appeal outcomes.

To construct the main independent variable, we collected data on the ethnicity of judges and appellants. We used appellants' names to measure ethnicity, an increasingly common approach in political science (Harris, 2015; Enos, 2016; Hassan, 2017). Our procedure leveraged information from Kenya's voter register, which identifies voter names from ethnically homogeneous areas. We created a dictionary-based ethnicity classifier to estimate the probability of ethnicity for a given last name, thereby linking each of nearly 10,000 persons' names to an ethnic group. Given the limited number of judges in the data, a member of the Kenyan legal community resolved ambiguous classifications of judge ethnicity by canvassing professional networks.[14] [15] [16]

Measuring Sentiment in Legal Judgments

We expect in-group favoritism or out-group derogation to motivate bias when judges evaluate an appeal. To test this, we use text-as-data approaches to assess the degree to which emotive reasoning appears in judicial writing. Our analysis builds on conventional dictionary methods wherein the count or proportion of key words in a given document is used to determine that document's category (Grimmer and Stewart, 2013). To this end, we generated word lists capturing our main mechanisms of interest: in-group favoritism or out-group derogation. Specifically, we identified terms related to *trust* and *disgust* to measure in-group favoritism and out-group derogation in written legal decisions.[17] We then calculated the number of *trust* and *disgust* terms as a proportion of total terms for each decision.[18]

The technical nature of judicial writing makes this approach a challenging test of our proposition. Not surprisingly, official guidelines from the Kenya Criminal Procedure Benchbook explicitly discourage judges from invoking emotive sentiments in their decisions to avoid allegations of bias: "judgment should not contain derogatory language …a dispassionate approach and clear finding of fact, are more indicative of judicial approach, and do not lay the magistrate open to a charge of possible bias. The court may express strong condemnation of the conduct of the accused, but it must be careful not to be abusive or, for example, imply that the conduct is what might be expected of those belonging to a particular race, religion, etc."[19]

While these rules are considered standard practice, the inclusion of explicit instructions suggests

[14] A key reason for ambiguity was the prevalence of female High Court judges with conflicting ethnic names due to marriage. Interethnic marriage is uncommon in the general population, but correlated with urban location and higher education (Crespin-Boucaud, 2020). This pattern is pronounced among highly educated, professional females like these judges (Bandyopadhyay and Green, 2021, Table 4). Appendix D.7 shows a robustness check accounting for uncertainty in appellant ethnicity via simulation; the results are consistent with the main results.

[15] We discuss our adherence to the American Political Science Association's 2020 Principles and Guidance for Human Subjects Research in Appendix G.

[16] See Choi, Harris and Shen-Bayh (2021) to access the data and code used to produce the figures and tables for this article.

[17] See Appendix for details regarding the generation of dictionary terms.

[18] A proportional measure is preferred to a simple count since it accounts for judgment length. This proportion is weighted by the term frequency-inverse document frequency statistic (TF-IDF), which accounts for the distribution of term usage across the corpus.

[19] See Criminal Procedure Judicial Benchbook, pg. 113.

a broader concern among legal practitioners in Kenya – that judges may indeed be discriminatory in their legal opinions and need to guard against such tendencies when rendering verdicts. Given these cautions, we expect that any expression of emotion in a written judgment will likely reflect subtle, implicit, and often unconscious biases rather than overt prejudice.

Examples from our appeals corpus corroborate these expectations, suggesting that terms of *trust* and *disgust* are subtly expressed in legal writing. In one successful appeal, the judge described the appellant's standing using terms of trust, remarking that the "magistrate erred in law and fact in disregarding the appellant's defence, which was **consistent** and **trustworthy**[.]"[20]

In another ultimately-denied judgment, the judge invoked terms of *disgust* in assessing the facts of the case: "the appellant had converted [the witness] into his wife, a **shameful** act indeed. She also physically suffered by the damage of her womanhood. The best description that this court can accord to the behavior of the appellant was that he was a **beast** to [the witness]. As rightly noted by the trial magistrate, he ought to be kept away from the society."[21] Terms of disgust written into judgments reveal a judge's personal and moral assessments of appellant character: "the offence committed was **barbaric**, **immoral** and had definitely left the complainant traumatized. I find that was a justification for passing sentence higher than the minimum provided and did not in any way offend the provisions of ... the Constitution of Kenya[.]"[22]

These excerpts illustrate that trust and disgust dictionaries capture personal evaluations of appellants, distinct from an appeal's legal merits. Given the technicality of legal writing, including the instruction to minimize perceptions of bias in the record, our approach is a hard test of the hypothesis that judicial bias is motivated by in-group favoritism or out-group derogation. If we find evidence that judges invoke terms of trust (disgust) with regard to coethnic (non-coethnic) appellants, we take it as consistent with the posited affective-cognitive mechanisms that may influence outcomes.

[20] Criminal Appeal 107 of 2017.
[21] Criminal Appeal 121 of 2014.
[22] Criminal Appeal 14 of 2016.

Research Design

To identify the effect of identity on criminal appellate judgments, we exploit features of the case assignment process in the Kenyan Courts. Cases arriving on the docket of each high court station are sorted into categories for assignment to legal divisions within the court station: family, commercial and admiralty, constitutional, land and environment, and, most relevant for our purposes, criminal. The deputy registrar, responsible for case scheduling, assigns each new incoming case to a judge. This intra-division assignment is determined by judges' calendars and existing workloads, not by case characteristics: case assignment criteria – a judge's schedule and case load – are orthogonal to case particulars like the ethnicity of defendants and judges. This provides us with quasi-random variation in the ethnic relationship between the appellant and the judge.

We use a linear model to examine the relationship between coethnicity and case outcomes:

$$Y_i = \beta_0 + \beta_1 M_i + \beta_2 X_c + \epsilon_i \tag{1}$$

where Y_i is either the binary indicator for whether the appellate judge ruled in favor of the defendant or a measure of sentiment in text, M_i takes the value of one if the case was assigned to a coethnic judge of the defendant, X_c is a vector of controls and fixed effects, including courthouse-year, judge ethnicity, individual judge, and in our most restrictive specification, courthouse-year and individual judge fixed effects. We present our baseline results with our coethnic match variable and courthouse-year fixed effects, and progressively add more restrictive sets of fixed effects and controls. Then, we conduct ethnic subgroup analyses by judge ethnicity to probe for differential bias.

Case allocation is a manual process, and it is possible that the principles of case assignments are not respected. Appendix Table B1 provides balance checks suggesting that the case assignment mechanism likely induced quasi-random variation in judge-defendant coethnicity. Conditional on the courthouse-year in which the case was heard, there is balance on most covariates. Most differences remain insignificant, except for the proportion of crimes classified as murder, manslaughter, or theft. Given this, we include regression specifications with these covariates, which do not alter our findings.

Findings

Do Kenyan appellate judges show identity-based bias in their decision-making? Table 1 summarizes the main results of the linear probability models where the outcome is equal to one if the judge ruled in the *defendant's* favor, zero otherwise. Our main focus is the coethnic match covariate, which is equal to one when the judge and the defendant share the same ethnic identity.

Table 1: Effect of Coethnic Match between Appellant and Judge.

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.042***	0.039***	0.042**	0.040**	0.036**	0.033**
	(0.015)	(0.015)	(0.018)	(0.017)	(0.014)	(0.014)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	9,545	9,545	9,545	9,545	9,545	9,545
R^2	0.001	0.009	0.079	0.085	0.105	0.110

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

The results indicate that Kenyan judges favor coethnic appellants. The basic specification without fixed effects in column (1) shows that judges are 4.2 percentage points more likely to decide in favor of a coethnic over a non-coethnic defendant. Columns (2) through (6) provide increasingly stringent empirical tests by adding fixed effects to account for factors that vary by location and time (i.e., courthouse-year) and judge, as well as case-specific controls describing the offense in question. Although the magnitude of the bias fluctuates marginally with the addition of these fixed effects, the findings remain robust.[23]

In the theory section, we generated two contrasting predictions regarding heterogeneity in coethnic bias across judges belonging to different ethnic groups. We test these predictions by running a

[23]Full results for tables in the main article with all controls are available at the APSR Dataverse.

series of subgroup analyses reported in Table 2, and Table 3. The social dominance/identity perspective, which predicts that dominant or high-status groups exhibit more in-group bias, finds strong support in Table 2; coethnic bias is observed primarily in decisions handed down by Kikuyu judges and, to a lesser extent, Kamba judges.[24] The Kikuyu ethnic group has occupied both the presidency and the position of chief justice for a significant portion of the post-independence period, including during the span of our analysis, and can thus be considered a "dominant" ethnic group in Kenya's political, economic, and legal landscape.

Given these findings, we urge caution in interpreting the Kamba estimate in Table 2 as robust (especially considering that the direction of the estimated relationship is unstable).

Table 2: Effect of Coethnic Match between Appellant and Judge, by Judge Ethnicity.

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.057^{**}	0.025	-0.018	0.024	0.079^{**}	0.096	0.057
	(0.023)	(0.022)	(0.029)	(0.061)	(0.029)	(0.130)	(0.094)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,235	1,042	2,917	1,217	760	531	843
R^2	0.169	0.146	0.132	0.091	0.087	0.223	0.203

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Our findings suggest that judges of politically dominant ethnic groups deliver more favorable outcomes to appellants of the same ethnicity. But to what extent are such judges explicitly discriminating

[24]The Kamba-specific estimates, while present in the unweighted data presented in Table 2, are unstable (e.g., switch signs) and dissipate in the inverse-probability weighting robustness checks in Appendix D.3. This contrasts with the Kikuyu-specific results, which remain stable and positive, and become larger in some of the IPW specifications. This robustness test accounts for observations that cannot plausibly be part of the quasi-experiment in question, relating coethnicity to legal outcomes, because randomization in some cases is not possible, a point we discuss below. Given these findings, we urge caution in interpreting the Kamba estimate in Table 2 as robust (especially considering that the direction of the estimated relationship is unstable). The sensitivity of the result related to Kamba judges suggests that the results may be driven by situations in which the distribution of judges at a given court station did not afford (or always afforded) the possibility of an ethnic match.

against members of the political outgroup? Recall that if the political rivalry hypothesis were true, we would expect to see appeals outcomes vary depending on whether judges and appellants come from groups with a history of inter-ethnic conflict; in particular, judges should rule more harshly against appellants from politically rivalrous groups. In Table 3, we examine the decision-making patterns of judges belonging to three large and politically significant ethnic groups in Kenya— the Kikuyu, Kalenjin, and Luo—that have recently experienced divisive contests for control of the state, including violent conflicts over the outcomes of national elections. The results provide limited support for the idea that bias against outgroups follows the logic of political rivalry and conflict. While we see a marginally significant effect on Luo appellants for Kalenjin judges, this finding dissipates when indicators for Kikuyu and Kalenjin appellants are included in the regression. Our findings thus suggest that the Kikuyu, Kalenjin, and Luo judges do not systematically penalize appellants from rival ethnic groups.

Table 3: Dyad-specific effects between Appellant and Judge

	Dependent variable:											
	Outcome: Judgement for the Defendant											
	Kikuyu Judges				Luo Judges				Kalenjin Judges			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Kikuyu Appel.	0.06**			0.05**	−0.02			−0.01	0.05			0.04
	(0.02)			(0.02)	(0.05)			(0.04)	(0.04)			(0.05)
Kalenjin Appel.		−0.06		−0.05		0.05		0.05		0.02		0.03
		(0.06)		(0.06)		(0.09)		(0.08)		(0.02)		(0.03)
Luo Appel.			−0.002	0.01			0.02	0.02			−0.07*	−0.06
			(0.08)	(0.07)			(0.06)	(0.06)			(0.04)	(0.04)
Observations	2,235	2,235	2,235	2,235	1,217	1,217	1,217	1,217	1,042	1,042	1,042	1,042
R^2	0.17	0.17	0.17	0.17	0.09	0.09	0.09	0.09	0.15	0.15	0.15	0.15

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. All models contain Courthouse-Year FE, Judge FE, and case-specific controls.

In the Online Appendix, we present robustness tests related to Tables 1 and 2. Appendix D.1 and D.2 replicates the main results using inverse probability weighting to account for unit-level differences in treatment probability, and to exclude observations that had no chance at randomization (e.g., Kamba appellants in court stations staffed exclusively by Kamba judges or, alternatively, Kamba appellants who had no chance of being assigned to a Kamba judge). Despite this loss of power due to a smaller sample size, the models recover similar point estimates and, for the more reasonable weighting specifications, retain statistical significance. Appendix D.3 applies the same IPW approach to the findings reported in Table 2, to similarly recover point estimates for Kikuyu judges that are statistically significant and similar in magnitude for the more reasonable weighting specifications. Kamba judges also show positive and significant results in Table 2. However, the IPW estimates for Kamba judges in Appendix D.3 lose significance, switch signs, and attenuate in the more reasonable weighting specifications. Given this, we do not consider these estimates to be robust, and put little weight on the Kamba judge estimates. Appendix D.7 incorporates uncertainty over appellants' ethnicity classification, given our probabilistic name-based approach, showing the full distribution of estimates over 10,000 iterations, for column 6 of Table 1 and each ethnic group in Table 2, both with and without inverse probability weights. Again, the results comport with the main results. Appendix D.6 reruns the results above, aggregating ethnic groups commonly associated with one another, finding little change in the results.[25]

Identifying Mechanisms of Coethnic Bias: Favoritism or Derogation?

To elucidate whether in-group favoritism or out-group derogation drives coethnic bias in judicial decision-making, we examine the texts of appeals judgments. We calculate the proportion of trust, disgust, positive, and negative sentiments for each written judgment using word embeddings. We estimate the effect of coethnicity using OLS as described in equation 1 above.

Table 4 shows that only trust has a positive and significant relationship with the coethnic match variable. Judgments for coethnics contain approximately 0.12 standard deviations more trust-related words than judgments written for non-coethnics. A judge invokes terms of confidence, credibility, and honesty more often when writing a judgment for a coethnic. The insignificance for the disgust

[25] In these analyses, we pool ethnic groups that might be considered functionally coethnic: Samburu and Masai; Embu and Kikuyu; Taita and Mijikenda; Pokot and Kalenjin. We also run robustness checks to account for cases that were administratively purged from the court docket after years of inactivity, presumably because the appellant lost interest in the appeal. Appendix D.4 shows that dropping "lost interest" cases strengthens our main results. We also conduct robustness test to account for the possibility that findings of ethnic bias are driven by judges who have been deemed ineligible for higher posts by the Judicial Service Commission due to reasons of corruption or inappropriate behavior. Results are reported in Appendix D.5. Our main findings are unchanged when we either control for cases adjudicated by these judges (Table D13) or omit the cases adjudicated from them from the analysis (Table D14).

Table 4: Coethnic Bias in Written Judgments: Corpus Seeds, GloVe Vectors

	Dependent variable:			
	Sentiment			
	Trust	Disgust	Positive	Negative
	(1)	(2)	(3)	(4)
Coethnic Match	0.115***	0.015	0.046	0.040
	(0.038)	(0.040)	(0.037)	(0.038)
Individual Judge FE	Yes	Yes	Yes	Yes
Courthouse-Year FE	Yes	Yes	Yes	Yes
Observations	9,545	9,545	9,545	9,545
R^2	0.228	0.221	0.211	0.227

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

score suggests that judges use similar amounts of disgust-related language when writing for coethnics and non-coethnics. We interpret this as evidence that judicial bias is a manifestation of in-group favoritism rather than out-group derogation.

We draw two main takeaways from this dictionary analysis, both of which should be interpreted with caution. First, despite the fact that judges are explicitly instructed to refrain from issuing decisions based on non-legal considerations, they are still likely to invoke emotive sentiments (trust) when regarding coethnic appellants. This has important implications for legal reform efforts in Kenya, the bulk of which have either implicitly or explicitly assumed that judicial bias falls along gender, income, or other socioeconomic dimensions. Our findings suggest that more consideration should be directed instead toward the conscious or subconscious favoritism judges may harbor for coethnics. Second, our analysis reveals the utility of text-as-data approaches on judicial writing, especially the applicability of minimally supervised dictionary methods on niche corpora. Our main findings suggest that conventional dictionary methods combined with word embedding models elucidate different sources of implicit, affective bias in legal language and help assess the weight of one type of bias against another.[26] While these findings are more suggestive than definitive, they provide a framework

[26] In the Appendix, we illustrate that our main results are consistent using different dictionaries and embedding models.

for understanding some nuances of judicial decision-making.

Conclusion

To our knowledge, this paper is the first detailed analysis of the relationship between ethnic identity and legal outcomes in an African context. We provide evidence of a coethnic bonus in appeals decisions using a new dataset of almost 10,000 criminal cases from the Kenyan High Court. We find that, when faced with a coethnic appellant, a judge is about 3 to 5 percentage points more likely to grant the appeal. These effects are concentrated among ethnic Kikuyu – the "dominant" ethnic group in both the political and judicial arena – where the coethnic bonus is around 6 to 10 percentage points. Judgments written for coethnics also show higher concentrations of trust-related words – an indicator of in-group favoritism. We find little evidence in written judgments of out-group derogation (i.e. words related to disgust). These findings echo the observation of U.S.-based legal practitioners that "ingroup and outgroup are differentiated more as targets of positive than of negative feelings" (Greenwald and Pettigrew, 2017, p. 161).

How does the magnitude of our results relate to similar work on judiciaries in other contexts? While the magnitude of the effects we discover fall within the range of ethnic bias observed in other studies, it is important to note that it is moderate in comparison to work on criminal appeals outcomes in Israel and the U.S., suggesting that the guidelines set by the Kenyan Criminal Procedure Judicial Benchbook judiciary to guard against judicial bias may be effective to some extent.[27,28]

In addition, Appendix F presents the word-embeddings validation exercise described in Rodriguez and Spirling (N.d.). This shows that human coders from mTurk are unable to distinguish between word relations generated by the word embeddings we employ and human-generated word relationships. Our sample of mTurkers were drawn from a variety of English-speaking contexts and instructed to carry out the task, focusing on word use in legal and courtroom settings. While we see little reason *ex ante* that coders from Kenya specifically would employ different intuitive word mappings from coders of other nationalities, future research in text analysis might consider how validation exercises like the one in Rodriguez and Spirling (N.d.) might vary from context to context to better understand asymmetries between the context of the written language and the nationalities of coders.

[27]Several studies allow us to contextualize our estimates in broader perspective. In a study of Israeli courts, Shayo and Zussman (2011) finds that Jewish judges are around 14% points more likely to rule in favor of Jewish plaintiffs than Arab plaintiffs. In the same setting, Grossman et al. (2016) shows that the presence of an Arab judge on a criminal panel decreases incarceration for Arab defendants by 14–20 percentage points. Meanwhile in the U.S., Alesina and La Ferrara (2014) finds bias of between 3–9 percentage points against minority defendants who killed white victims in a study of capital sentence reversal patterns.

[28]Our estimates align well with coethnic bias documented outside the legal domain. Also in Kenya, Kramon and Pos-

These results have important implications for the internal governance of the judicial branch. In particular, we provide evidence that in-group favoritism may privilege some defendants' appeals over others. In such circumstances, the composition of the local judiciary and the geographic reach of the judicial services may determine how frequently this type of bias might occur. While shuffling judicial appointments to different court stations can help encourage impartiality to a certain extent, our findings suggest that more fundamental factors such as identity can still affect legal outcomes. The results presented here thus suggest that future Kenyan reform efforts with regards to judicial vetting and legal training processes might pay greater attention to the potential for bias in judicial decision-making.

Our findings also complement existing research on the political drivers of judicial outcomes. Much of this work assumes that judges are strategic actors who often must navigate uncertain terrain in order to ensure their own political survival (Helmke, 2002). Such strategies can be especially important in autocratic or weakly democratic regimes where leaders have greater wherewithal to manipulate courts to further their political agenda or hold onto power (Shen-Bayh, 2018). Much like the broader literature on African courts, these studies tend to focus on high-profile cases relating to constitutional or national security questions (Ginsburg, 2003). Our study expands this literature by shifting attention to lower-level courts: institutions that not only hear lower-visibility cases much more frequently, but are also often the *only* judicial institutions directly interacting with regular citizens.[29]

Perhaps more crucially, our analysis suggests that judicial bias in a multiethnic society such as Kenya is primarily concentrated among judges from the "socially dominant" group. This finding has theoretical and practical implications for how we understand implicit bias in both the legal domain and beyond. Theoretically, it urges us to revise our models of implicit bias and pay greater attention to how contextual variation in group status and hierarchies across societies can condition bias. That is, rather than assume that implicit bias is an inherent feature of group identity (and one that manifests

ner (2016) shows that coethnics of the president are 3–5 percentage points more likely to receive or complete primary education.

[29]Recent work on African courts has begun probing related questions on citizen evaluations of judicial institutions (Bartels and Kramon, 2020; Kerr and Wahman, N.d.).

evenly across all groups), such bias is potentially heterogeneous and may vary depending on which groups have historically exercised sociopolitical and economic dominance over others. This is not to say that socially non-dominant groups do not hold psychological predispositions towards ingroup or outgroup members or that they are not susceptible to bias *per se*. But the fact that their biases are not directly observed in the legal decision-making process should prompt future investigations as to why this is the case.

Practically, rethinking implicit bias along these terms means also revisiting the design of policy interventions intending to bring fairness and impartiality to judicial decision-making. Prior scholarship has shown that diversifying the bench can help reduce judicial discrimination against members of minority groups.[30] These works show that in societies dominated by a single ethnic group, improving diversity on the bench increases the frequency and intensity of "contact" between members of the dominant group and outgroup, which may help reduce implicit biases among dominant group judges in the medium to long term. However, it remains unclear whether diversifying the bench in more ethnically diverse societies (where multiple groups have either dominated government or shared political control) will have the same mitigating effect on implicit bias in judicial decision-making. In particular, policies that are meant to increase representation of non-dominant ethnic groups on the bench may not in and of themselves be sufficient in reducing judicial bias unless there is also change in the overarching power relations among groups.

Our findings also suggest that policy interventions that are designed to reduce implicit bias in the courts should be attuned to heterogeneity that manifests as a result of status or power structures and hierarchies. That is, if judges from socially dominant groups are more susceptible to biases, interventions that aim to reduce the prejudices that dominant group judges hold towards minority defendants and appellants may be one of the most effective means through which discrimination can be curbed in the judicial realm (e.g., Redfield, 2017, chapters 11 and 12).

There are important scope considerations to our findings. In particular, our analysis focuses on

[30] For example, Grossman et al. (2016) shows that in the context of Israel, serving alongside non-coethnic peers on the bench moderated anti-Muslim bias among Israeli judges. Harris (2021) similarly finds that in the U.S., white-black mixed benches reduce disparity in criminal sentencing of white and black defendants.

bias at the appeals stage rather than courts of first instance. This means that we do not account for whether the identity of *lower* court magistrates affects judicial decision-making during the initial trial, or whether magistrate ethnicity shapes the opinions of higher court justices upon appeal. It is possible that co-ethnicity plays a similar biasing role at the magistrates level, which may in turn affect the likelihood of appeals success at higher courts (e.g. Alesina and La Ferrara, 2014; Sen, 2015).[31] We leave it to future studies to examine these dynamics in greater detail.

Given the relative infancy of studies of bias in legal decision-making in African contexts, there is ample opportunity to build upon our findings and contribute to existing debates on these themes in other geographic contexts. We see three particularly promising areas for future research. First, what are the building blocks of bias? Building on the findings of scholars such as Liu and Li (2019) and Wistrich, Rachlinski and Guthrie (2015), future studies may choose to experimentally manipulate characteristics of hypothetical cases in order to better understand how identity-based, moral, and emotional factors shape judicial decision-making in the Global South. Second, another promising line of research might explore techniques for bias reduction such as simple informational interventions that create awareness about bias and may aid in its reduction (Liu, 2018; Redfield, 2017). Third, do resource constraints exacerbate bias? Due to lack of infrastructure and operating funds, judicial officers in lower income countries are often overworked and under-resourced.[32] Given the implicit, heuristic nature of in-group biases, it seems reasonable to conclude that excessive workloads may exacerbate bias by forcing judges to produce judgments quickly, as opposed to spending the time necessary to deliver circumspect, carefully-reasoned judgements. To the extent that such constraints can be addressed, examining the logistical burdens of judicial decision-making may help elucidate whether and how institutional reforms can enhance due process.

[31]For example, it might be the case that appeals judges base reversals on the identity of lower-court judges whereby appeals judges seek to undermine the opinions of non-coethnic judges.

[32]For instance, see Dimitrova-Grajzl et al. (2016) and Grajzl and Silwal (2020) on quality-quantity trade-offs in resource constrained judiciaries.

References

Adida, Claire, Jessica Gottlieb, Eric Kramon and Gwyneth McClendon. 2017. "Reducing or Reinforcing In-Group Preferences? An Experiment on Information and Ethnic Voting." *Quarterly Journal of Political Science* 12(4):437–477.

Alesina, Alberto and Eliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing." *The American Economic Review* 104(11):3397–3433.

Alicke, Mark D. 2000. "Culpable Control and the Psychology of Blame The motivation to maintain favorable identities View project." *Psychological Bulletin* 126(4):556–574.

Allport, G.W. 1954. *The nature of prejudice*. Addison-Wesley.

Antoniak, Maria and David Mimno. 2018. "Evaluating the Stability of Embedding-based Word Similarities." *Transactions of the Association for Computational Linguistics* 6:107–119.

Arriola, Leonardo R, Donghyun Danny Choi and Matthew K Gichohi. 2021. "Increasing Intergroup Trust: Endorsements and Voting in Multiethnic Societies." *Working Paper* .

Ash, Elliott, Daniel L. Chen and Arianna Ornaghi. 2020. ""Gender Attitudes in the Judiciary: Evidence from U.S. Circuit Courts"." *Working Paper* .

Bandyopadhyay, Sanghamitra and Elliott Green. 2021. "Explaining inter-ethnic marriage in Sub-Saharan Africa." *Journal of International Development* 33(4):627–643.

Bartels, Brandon L and Eric Kramon. 2020. "Does Public Support for Judicial Power Depend on Who is in Political Power? Testing a Theory of Partisan Alignment in Africa." *American Political Science Review* 114(1):144–163.

Bates, Robert H. 1974. "Ethnic competition and modernization in contemporary Africa." *Comparative political studies* 6(4):457–484.

Beattey Jr, Robert A, Taiki Matsuura and Elizabeth L Jeglic. 2014. "Judicial bond-setting behavior: The perceived nature of the crime may matter more than how serious it is." *Psychology, Public Policy, and Law* 20(4):411.

Black, Ryan C, Sarah A Treul, Timothy R Johnson and Jerry Goldman. 2011. "Emotions, Oral Arguments, and Supreme Court Decision Making." *The Journal of Politics* 73(2):572–581.

Blum, Ashley, Chad Hazlett and Daniel N. Posner. 2021. "Measuring Ethnic Bias: Can Misattribution-Based Tools from Social Psychology Reveal Group Biases that Economics Games Cannot?" *Political Analysis* 29(3):385–404.

Boussalis, Constantine, Travis Coan, Mirya Holman and Stefan Müller. Forthcoming. "Gender, Candidate Emotional Expression, and Voter Reactions During Televised Debates." *American Political Science Review* pp. 1–15.

Branch, Daniel. 2011. *Kenya: between hope and despair, 1963-2011.* Yale University Press.

Brewer, Marilynn B. 1999. "The psychology of prejudice: Ingroup love or outgroup hate?" *Journal of Social Issues* 55(3):429–444.

Chandra, Kanchan. 2007. *Why ethnic parties succeed: Patronage and ethnic head counts in India*. Cambridge University Press.

Choi, Donghyun Danny, J. Andrew Harris and Fiona Shen-Bayh. 2021. "Replication Data for: Ethnic Bias in Judicial Decision-making: Evidence from Criminal Appeals in Kenya.".
URL: *https://doi.org/10.7910/DVN/JCSOB8*

Clarke, Jessica A. 2018. "Explicit Bias." *Northwestern University Law Review* 113(3):505–586.

Cohen, Alma and Crystal S Yang. 2019. "Judicial politics and sentencing decisions." *American Economic Journal: Economic Policy* 11(1):160–91.

Crespin-Boucaud, Juliette. 2020. "Interethnic and interfaith marriages in sub-Saharan Africa." *World Development* 125:104668.

Dimitrova-Grajzl, Valentina, Peter Grajzl, Atanas Slavov and Katarina Zajc. 2016. "Courts in a transition economy: Case disposition and the quantity–quality tradeoff in Bulgaria." *Economic Systems* 40(1):18–38.

Ejdemyr, Simon, Eric Kramon and Amanda Lea Robinson. 2018. "Segregation, Ethnic Favoritism, and the Strategic Targeting of Local Public Goods." *Comparative Political Studies* 51(9):1111–1143.

Enos, Ryan D. 2016. "What the Demolition of Public Housing Teaches Us about the Impact of Racial Threat on Political Behavior." *American Journal of Political Science* 60(1):123–142.

Ferree, Karen E. 2006. "Explaining South Africa's Racial Census." *Journal of Politics* 68(4):803–815.

Fiske, Susan T and Mark A Pavelchak. 1986. "Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect.".

Gazal-Ayal, Oren and Raanan Sulitzeanu-Kenan. 2010. "Let My People Go: Ethnic In-Group Bias in Judicial Decisions—Evidence from a Randomized Natural Experiment." *Journal of Empirical Legal Studies* 7(3):403—-428.

Gennaro, Gloria and Elliott Ash. 2021. "Emotion and Reason in Political Language." *Center for Law & Economics Working Paper Series* 2021(02).

Gibson, James L. and Gregory A. Caldeira. 2003. "Defenders of democracy? Legitimacy, popular acceptance, and the South African Constitutional Court." *Journal of Politics* 65(1):1–30.

Gill, Rebecca D, Michael Kagan and Fatma Marouf. 2017. "The impact of maleness on judicial decision making: masculinity, chivalry, and immigration appeals." *Politics, Groups, and Identities* 7(3):509–528.

Ginsburg, Tom. 2003. *Frontmatter*. Cambridge University Press.

Gloppen, Siri and Fidelis Edge Kanyongolo. 2007. "Courts and the poor in Malawi: Economic marginalization, vulnerability, and the law." *International Journal of Constitutional Law* 5(2):258–293.

Grajzl, Peter and Shikha Silwal. 2020. "The functioning of courts in a developing economy: evidence from Nepal." *European Journal of Law and Economics* 49(1):101–129.

Greenwald, Anthony and Thomas Pettigrew. 2017. *With Malice Toward None and Charity for Some: Ingroup Favoritism Enables Discrimination.* Chicago: American Bar Association pp. 87 – 130.

Grimm, Lars J., Rebecca A. Redmond, James C. Campbell and Ashleigh S. Rosette. 2020. "Gender and Racial Bias in Radiology Residency Letters of Recommendation." *Journal of the American College of Radiology* 17(1):64–71.

Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–297.

Grossman, Guy, Jeremy M Weinstein, Oren Gazal-Ayal and Samuel D Pimentel. 2016. "Descriptive Representation and Judicial Outcomes in Multiethnic Societies." *American Journal of Political Science* 60(1):44–69.

Gurr, Ted Robert. 2015. *Why men rebel.* Routledge.

Hagendoorn, Louk. 1995. "Intergroup biases in multiple group systems: The perception of ethnic hierarchies." *European review of social psychology* 6(1):199–228.

Harkness, Kristen A. 2018. *When soldiers rebel: ethnic armies and political instability in Africa.* Cambridge University Press.

Harris, Allison P. 2021. "Can Racial Diversity Among Judges Affect Sentencing Outcomes." *Working Paper* .

Harris, Allison P and Maya Sen. 2019. "Bias and Judging." *Annual Review of Political Science* 22:241–259.

Harris, Andrew. 2015. "What's in a Name? A Method for Extracting Information about Ethnicity from Names." *Political Analysis* 23(2):212–224.

Hassan, Mai. 2017. "The strategic shuffle: Ethnic geography, the internal security apparatus, and elections in Kenya." *American Journal of Political Science* 61(2):382–395.

Heath, Janae K., Gary E. Weissman, Caitlin B. Clancy, Haochang Shou, John T. Farrar and C. Jessica Dine. 2019. "Assessment of Gender-Based Linguistic Differences in Physician Trainee Evaluations of Medical Faculty Using Automated Text Mining." *JAMA Network Open* 2(5):e193520.

Helmke, Gretchen. 2002. "The Logic of Strategic Defection: Court–Executive Relations in Argentina Under Dictatorship and Democracy." *American Political Science Review* 96(2):291–303.

Hernan, MA and JM Robins. 2020. *Causal Inference: What If.* Chapman and Hall/CRC.

Hewstone, Miles, Mark Rubin and Hazel Willis. 2001. INTERGROUP BIAS. Technical report.

Hodson, Gordon, Becky L. Choma, Jacqueline Boisvert, Carolyn L. Hafer, Cara C. MacInnis and Kimberly Costello. 2013. "The role of intergroup disgust in predicting negative outgroup evaluations." *Journal of Experimental Social Psychology* 49(2):195–205.

Horowitz, Donald L. 2000. *Ethnic groups in conflict, updated edition with a new preface*. Univ of California Press.

Hou, Yue and Rory Truex. 2020. "Ethnic Discrimination and Authoritarian Rule: An Analysis of Criminal Sentencing in China." *Available at SSRN 3481448* .

Iaryczower, Matías, Pablo T. Spiller and Mariano Tommasi. 2002. "Judicial Independence in Unstable Environments, Argentina 1935-1998." *American Journal of Political Science* 46(4):699–716.

JMVB. 2016. Restoring Confidence in the Judiciary: Vetting of Judges and Magistrates in Kenya. Technical report The Judiciary of Kenya Nairobi: .

JSC. 2019. Judicial Service Commission Press Statement on Petitions Against Judges of Superior Courts and Complaints Against Magistrates. Technical report Judicial Service Commission Nairobi: .

Kanyinga, Karuti and Collins Odote. 2019. "Judicialisation of Politics and Kenya's 2017 Elections." *Journal of Eastern African Studies* 13(2):235—-252.

Kerr, Nicholas and Michael Wahman. N.d. "Electoral Rulings and Public Trust in African Courts and Elections." *Comparative Politics*. Forthcoming.

Kinyanjui, Sarah Muringa and Migai Akech. 2016. "Towards Structured Sentencing in Kenya: A Case for Reform." *African Journal of Criminology and Justice Studies* 9(1):266—-278.

Kramon, Eric and Daniel N. Posner. 2013. "Who Benefits from Distributive Politics? How the Outcome One Studies Affects the Answer One Gets." *Perspectives on Politics* 11(2):461–474.

Kramon, Eric and Daniel N. Posner. 2016. "Ethnic favoritism in education in Kenya." *Quarterly Journal of Political Science* 11(1):1–58.

Letsa, Natalie Wenzell. 2020. "Expressive Voting in Autocracies: A Theory of Non-Economic Participation with Evidence from Cameroon." *Perspectives on Politics* 18(2):439–453.

Levi, Margaret, Audrey Sacks and Tom Tyler. 2009. "Conceptualizing Legitimacy, Measuring Legitimating Beliefs." *American Behavioral Scientist* 53(3):354–375.

Levin, Shana. 2004. "Perceived group status differences and the effects of gender, ethnicity, and religion on social dominance orientation." *Political Psychology* 25(1):31–48.

Levin, Shana and Jim Sidanius. 1999. "Social dominance and social identity in the United States and Israel: Ingroup favoritism or outgroup derogation?" *Political Psychology* 20(1):99–126.

Liu, John Zhuang and Xueyao Li. 2019. "Legal Techniques for Rationalizing Biased Judicial Decisions: Evidence from Experiments with Real Judges." *Journal of Empirical Legal Studies* 16(3):630–670.

Liu, Zhuang. 2018. "Does Reason Writing Reduce Decision Bias? Experimental Evidence from Judges in China." *The Journal of Legal Studies* 47(1):83–118.

Loughran, Tim and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66(1):35–65.

Lowes, Sara, Nathan Nunn, James A Robinson and Jonathan Weigel. 2015. "Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT)." *American Economic Review: Papers & Proceedings* 105(5):340–345.

Lynch, Gabrielle. 2014. "Electing the 'alliance of the accused': the success of the Jubilee Alliance in Kenya's Rift Valley." *Journal of Eastern African Studies* 8(1):93–114.

Mackie, Diane M., Thierry Devos and Eliot R. Smith. 2000. "Intergroup emotions: explaining offensive action tendencies in an intergroup context." *Journal of Personality and Social Psychology* 79(4):602–616.

Maroney, Terry A. 2016. "A Field Evolves: Introduction to the Special Section on Law and Emotion." *Emotion Review* 8(1):3 – 7.

Montalvo, José G and Marta Reynal-Querol. 2005. "Ethnic polarization, potential conflict, and civil wars." *American economic review* 95(3):796–816.

Mullen, Brian, Rupert Brown and Colleen Smith. 1992. "Ingroup bias as a function of salience, relevance, and status: An integration." *European journal of social psychology* 22(2):103–122.

Nadler, Janice and Mary Hunter McDonnell. 2012. "Moral character, motive, and the psychology of blame." *Cornell Law Review* 97(2):255–304.

Ndulo, Muna. 2011. "African Customary Law, Customs, and Women's Rights." *Indiana Journal of Global Legal Studies* 18(1):87–120.

O'Donnell, Guillermo. 2004. "Why the Rule of Law Matters." *Journal of Democracy* 15:32–46.

Odote, Collins and Linda Musumba. 2016. *Balancing the Scales of Electoral Justice: 2013 Kenyan Election Disputes Resolution and Emerging Jurisprudence*. Nairobi: .

Oppedal Berge, Ivar Lars, Kjetil Bjorvatn, Simon Galle, Edward Miguel, Daniel N Posner, Bertil Tungodden and Kelly Zhang. 2019. "Ethnically Biased? Experimental Evidence from Kenya." *Journal of the European Economic Association* 17(0):1–31.

Osnabrugge, Moritz, Sara B. Hobolt and Toni Rodon. 2021. "Playing to the Gallery: Emotive Rhetoric in Parliaments." *American Political Science Review* p. 1–15.

Otten, Sabine and Dirk Wentura. 2001. "Self-anchoring and in-group favoritism: An individual profiles analysis." *Journal of Experimental Social Psychology* 37(6):525–532.

Oyserman, Daphna, Markus Kemmelmeier, Stephanie Fryberg, Hezi Brosh and Tamera Hart-Johnson. 2003. "Racial-ethnic self-schemas." *Social Psychology Quarterly* pp. 333–347.

Paluck, Elizabeth Levy and Donald P Green. 2009. "Prejudice reduction: What works? A review and assessment of research and practice." *Annual review of psychology* 60:339–367.

Pennington, Jeffrey, Richard Socher and Christopher D. Manning. 2014. "GloVe: Global Vectors for Word Representation Jeffrey." *Proceedings ofthe 2014 Conference on Empirical Methods in Natural Language Processing* pp. 1532—-1543.

Rachlinski, Jeffrey J. and Andrew J. Wistrich. 2017. "Judging the Judiciary by the Numbers: Empirical Research on Judges." *Annual Review of Law and Social Science* 13(1):203–229.

Ramey, Adam, Gary E Hollibaugh, Jonathan D Klingler and Matthew Hall. 2019. "Attributes Beyond Attitudes: Personality Traits on the US Supreme Court: Personality traits on the US Supreme Court." *Journal of Law and Courts* .

Redfield, Sarah E., ed. 2017. *Enhancing Justice: Reducing Bias*. Chicago: American Bar Association Judicial Division.

Rice, Douglas, Jesse Rhodes and Tatishe Nteta. 2019. "Racial Bias in Language." *Research and Politics* pp. 1–7.

Rice, Douglas R. and Christopher Zorn. 2019. "Corpus-based dictionaries for sentiment analysis of specialized vocabularies." *Political Science Research and Methods* pp. 1–16.

Robinson, Amanda Lea. 2020. "Ethnic Diversity, Segregation and Ethnocentric Trust in Africa." *British Journal of Political Science* 50(1):217–239.

Rodriguez, Pedro L and Arthur Spirling. N.d. "Word Embeddings What works, what doesn't, and how to tell the difference for applied research." *Journal of Politics*. Forthcoming.

Sachdev, Itesh and Richard Y Bourhis. 1987. "Status differenttals and intergroup behaviour." *European journal of social psychology* 17(3):277–293.

Sambanis, Nicholas and Moses Shayo. 2013. "Social identification and ethnic conflict." *American Political Science Review* pp. 294–325.

Sen, Maya. 2015. "Is justice really blind? race and reversal in US courts." *Journal of Legal Studies* 44(S1):S187–S229.

Shayo, Moses and Asaf Zussman. 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." *Quarterly Journal of Economics* 126:1447–1484.

Shen-Bayh, Fiona. 2018. "Strategies of Repression: Judicial and Extrajudicial Methods of Autocratic Survival." *World Politics* 70(3):321–357.

Sherif, Muzafer. 1988. *The robbers cave experiment: Intergroup conflict and cooperation.[Orig. pub. as Intergroup conflict and group relations]*. Wesleyan University Press.

Sherif, Muzafer and Carolyn W Sherif. 1953. "Groups in harmony and tension; an integration of studies of intergroup relations.".

Sidanius, Jim and Felicia Pratto. 2001. *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Sidanius, Jim, Shana Levin, James Liu and Felicia Pratto. 2000. "Social dominance orientation, anti-egalitarianism and the political psychology of gender: An extension and cross-cultural replication." *European journal of social psychology* 30(1):41–67.

Simon, Dan. 1998. "A Psychological Model of Judicial Decision Making." *Rutgers Law Journal* 30(1):1—-142.

Simon, Dan. 2004. "A Third View of the Black Box: Cognitive Coherence in Legal Decision Making." *The University of Chicago Law Review* 2(71):511–586.

Stephan, Walter S and Cookie White Stephan. 2013. An integrated threat theory of prejudice. In *Reducing prejudice and discrimination*. Psychology Press pp. 33–56.

Tajfel, Henri and John Turner. 1979. An Integrative Theory of Intergroup Conflict. In *The Social psychology of Intergroup Relations*, ed. Stephen Worchel and William Austin. Monterey, CA: Cole pp. 33–47.

Throup, David and Charles Hornsby. 1998. *Multi-party politics in Kenya: The Kenyatta & Moi States & the triumph of the system in the 1992 election*. Ohio University Press.

Tripp, Aili Marie. 2004. Women's Movements, Customary Law, and Land Rights in Africa: The Case of Uganda.

Turing, A. M. 1950. "Computing Machinery and Intelligence." *Mind* LIX(236):433–460.

Turner, John C, Katherine J Reynolds et al. 2001. "The social identity perspective in intergroup relations: Theories, themes, and controversies." *Blackwell handbook of social psychology: Intergroup processes* 4:133–152.

Tuñón, Guadalupe and German Feierherd. 2020. "The Selective Promotion of Corruption Probes: Lessons and Evidence from the Argentine Courts." *Working Paper* .

Van de Walle, Nicolas. 2001. *African economies and the politics of permanent crisis, 1979-1999*. Cambridge University Press.

VonDoepp, Peter. 2006. "Politics and Judicial Assertiveness in Emerging Democracies: High Court Behavior in Malawi and Zambia." *Political Research Quarterly* September:389—-399.

Vondoepp, Peter and Rachel Ellett. 2011. "Reworking Strategic Models of Executive-Judicial Relations: Insights from New African Democracies." *Comparative Politics* 43(2):147–165.

Widner, J.A. 2001. *Building the Rule of Law*. W.W. Norton.

Widner, Jennifer A. 1992. *The rise of a party-state in Kenya*. University of California Press.

Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. "Ethnic politics and armed conflict: A configurational analysis of a new global data set." *American Sociological Review* 74(2):316–337.

Wistrich, Andrew J., Jeffrey J. Rachlinski and Chris Guthrie. 2015. *Heart versus head: Do judges follow the law or follow their feelings?* Vol. 93.

Supplementary Appendix

A Data

We assembled data on appeals from the Kenya Law Cases Database, an online repository of court rulings maintained by the National Council for Law Reporting (Kenya Law).[33] The full texts of court rulings are available in XML format for individual download. Using a web driver, we crawled and downloaded the XML files of 9,545 criminal appeals rulings issued by the High Court between January 1, 2003 and December 31, 2017. We then used regular expression text extractions to identify critical case features.

For each criminal appeal, we used XML header tags to extract the date of ruling, case number, and county court where the case was heard. We also identified the names of the appellant and respondent from the case citation where the appellant is always listed before the respondent, e.g. Mithungi [appellant] v. Republic [respondent]. To extract the names of individual judges, we utilized the fact that judge names are typically located near the end of a ruling, often following the word "Judge." Furthermore, when multiple judges vote together in a case, their names are located next to one another in the text. We leveraged these spatial patterns to extract 133 unique judges.

We then used the names of appellants to estimate their ethnic identity. Given the relatively limited number of judges in the data, a member of the Kenyan legal community coded ambiguous judges ethnicities by canvassing their professional networks to learn about the judges' ethnicity. In contrast, names for the appellants and respondents were too numerous to code by hand. To solve this problem, we build upon data and methods in Harris (2015) to estimate the ethnicity of participants in legal proceedings. Our approach leverages information from Kenya's voter register, which identifies voter names from ethnically homogeneous areas. We use this data to create a dictionary-based ethnicity classifier that estimates the probability of ethnicity for a given name. Then, we use these probabilities to link each person's name to an ethnic group.[34]

Extracting the final outcome of each appeal presented complications. The language of appeal decisions does not always follow a consistent pattern. While some phrases remain relatively constant (e.g. "This court hereby finds..."), patterns of judicial speech sometimes vary considerably by judge and year. Furthermore, some texts feature summaries of a ruling made in a previous case, using language that might be captured by our regular expressions extractor and thus incorrectly classified as a final appeal outcome. To address these concerns, we compiled two mutually exclusive lists of expressions designed to capture appeals that were either allowed or denied.[35] We also utilized the fact

[33] Originally established by The National Council for Law Reporting Act (1994), Kenya Law Cases is the most comprehensive legal database in Kenya. It contains the full text decisions of civil and criminal cases delivered by magistrate courts, High Courts, the Appeals Court, the Supreme Court, and other special tribunals.

[34] For simplicity, we include 12 (of ~ 42) ethnic groups in Kenya: Kalenjin, Kamba, Kikuyu, Kisii, Luhya, Luo, Masai, Meru, Mijikenda, Pokot, Somali, and Turkana. We make this simplification for two reasons. These 12 groups make up over 90% of the population. And for many of the smaller ethnic groups, a lack of group-specific naming conventions or simply a small numbers of individuals means that they are often indistinguishable from other nearby groups, or intermixed in a way that makes a name-based approach to identity futile.

[35] Successful appeals tend to feature words such as "allowed", "succeeds", "finds merit", and "conviction overturned", while rejected appeals tend to feature phrases such as "is denied", "is dismissed", "is rejected", "finds no merit", and "conviction upheld". These phrases represent a small subset of the full list used (see appendix).

that previous rulings tended to be summarized in the opening paragraph of each decision, whereas final outcomes tended to appear in the closing paragraph. By restricting our text extractor to the last few paragraphs of each decision, we thus limited misclassification of previous rulings as final outcomes.[36] Taking these steps enabled us to code the outcomes for approximately 80% the appeals in our sample.[37]

Finally, we used regular expressions to extract details related to whether the defendant was convicted of crimes relating to persons, property, health, or morality,[38] as well as whether they were sentenced to death, imprisonment, or corporal punishment. Details about the original case were often featured in the opening lines of the decision, as described above. We converted these crime and sentencing features into fixed effects in the analyses that follow.

B Balance Tests

Table B1: Balance Tests

Term	Estimate	Std.Error	T Statistic	P-Value
Case Type				
Murder	-0.025	0.015	-1.739	0.083
Manslaughter	0.046	0.020	2.233	0.026
Violence	0.007	0.007	0.893	0.373
Vehicle	0.020	0.013	1.480	0.140
Arson	-0.004	0.011	-0.367	0.714
Drug	-0.002	0.015	-0.151	0.880
Theft	-0.017	0.008	-2.110	0.036
Public order	0.033	0.031	1.073	0.284
Prior Sentence				
Death	-0.027	0.018	-1.451	0.148
Prison	0.002	0.009	0.251	0.802
Stroke	0.006	0.026	0.215	0.830

[36]There was an important trade-off between the size of the text window and classification accuracy. Shrinking the text window reduced the number of false classifications, but at the risk of truncating relevant information about case outcomes and producing null results. Expanding the text window raised the number of positive classifications, but at the risk of capturing information about a previous case rather than a final outcome. To account for this trade-off, the window size was manually adjusted for each year in the sample (2003-2017) in order to minimize the number of null results and false positives.

[37]The remaining 20% either had outcomes that were contained earlier in the text (which was removed by our length restrictions) or used idiosyncratic or misspelled language that were not captured by regular expressions.

[38]These classifications are based on the Kenyan Penal Code.

C Dictionaries

A key consideration of any dictionary method is which words best represent our sentiment of interest. Yet, word selection is often an arbitrary process and many studies do not offer guidance on optimal selection criteria (Grimmer and Stewart, 2013). Recent advances in natural language processing reveal how word embeddings can be used to populate dictionaries with minimal supervision.[39] The intuition of this approach is that every word in a corpus can be represented as a vector, the mapping of which can be interpreted as a spatial representation of the sentiment of that word; words that are closer together in the same vector space can thus be thought of as syntactically similar (Pennington, Socher and Manning, 2014). Furthermore, by vectorizing the vocabulary, the semantic similarity between any two words can be quantified using vector operations, specifically the cosine of the angle between two word vectors. Leveraging these properties, we build a sentiment dictionary with a small set of seed words and using cosine similarity scores to populate each dictionary with a list of most-similar terms.

Given the trade-offs of different dictionary approaches,[40] we evaluate sentiment using two sets of dictionaries: one derived from the corpus itself and another derived from an off-the-shelf vocabulary.[41] The findings from the corpus-derived dictionaries our featured in our main results section; findings from the off-the-shelf dictionaries, which are consistent with our main results, are saved for the Appendix (see Robustness: Text Analysis).

Our main text analysis uses a word embeddings model to build a minimally supervised dictionary for each sentiment of interest. For sentiment category s, we randomly sampled three seed words from our lists of trust- and disgust-related terms. We then used these seed words to retrieve the 2000 most-similar words from the corpus, where similarity was calculated using the cosine of the angle between two word vectors. Our word vectors were derived using the Global Word Vectors (GloVe) model. Because word embeddings require a vast amount of training data in order to produce stable vector representations (Antoniak and Mimno, 2018), we used the Common Crawl GLoVe model that was trained using a 1.9 million word vocabulary. Rodriguez and Spirling (N.d.) further find that pretrained word vectors perform well against both locally trained vectors and human coders.

Using the embeddings procedure described above, we derived sentiment dictionaries as shown in Table C1 where each column shows the most-similar terms derived from a set of sentiment seed words. The top three rows are the three randomly selected seed words from each category; the bottom rows show the top ten most similar words, where similarity is calculated using the cosine similarity scores derived from the GloVe model. We show the embeddings-derived dictionaries for the off-the-shelf seed terms below under Robustness: Text Analysis.

[39] Our approach builds on the work of Rice and Zorn (2019), who use such techniques to develop polarity scores measuring the proportion difference in opposing sentiments (e.g., positive versus negative).

[40] Corpus-specific dictionaries are better attuned to how keywords are used in context, but can be more difficult to validate (Loughran and McDonald, 2011); off-the-shelf dictionaries are more comprehensive, but more general vocabularies can be misapplied to niche corpora (Rice and Zorn, 2019; Grimmer and Stewart, 2013).

[41] The NRC Emotion Lexicon. See https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm.

Table C1: Seeds and Retrieved Words for Sentiment Categories:
Corpus-Derived Words

	Trust	Disgust	Positive	Negative
Seed Words	trustworthiness	cruel	positive	negative
	honest	heinous	good	bad
	confident	immoral	correct	incorrect
Retrieved (Top 10)	truthful	brutal	better	wrong
	confidence	inhuman	should	worse
	respectful	shameful	possible	unfortunately
	legitimate	evil	kind	mistaken
	genuinely	stupid	right	problem
	caring	vicious	proper	poor
	reliable	ridiculous	fine	lack
	assured	absurd	truth	worst
	competent	dishonest	consistently	misleading
	integrity	irresponsible	correctly	avoid

D Robustness: Judgment for Defendant

In this appendix, we consider two extensions of our analyses in order to assess the robustness of our results. First, we use the ethnic profiles of judges at court-stations in given years to estimate the probability of treatment – being matched to a coethnic judge – for each appellant. Second, we consider the robustness of our results taking into account estimation uncertainty related to the appellant ethnicity categorization.

D.1 Inverse Probability Weighting

In this appendix subsection, we replicate the main results using inverse probability weighting (IPW), following the discussion in Hernan and Robins (2020, p. 23). The logic behind IPW lies in adjusting our sample to account for varying probabilities of treatment which might bias our estimates of treatment, effectively reweighting the sample. For a given probability of treatment p_i for unit i, the inverse probability of treatment weight for a treated unit is $\frac{1}{p_i}$ and for a control unit is $\frac{1}{1-p_i}$.

In many applications, these probabilities are estimated with covariates using standard propensity score techniques. In our case, we can approximate the *actual* probability of treatment for each appellant by calculating the distribution of judge ethnicities at the court station where the appeal was filed. This is important because, conditional on the appellant, the distribution of ethnicities of available judges defines the probability of treatment for that appellant at a given station. For instance, a Kikuyu appellant faced with a pool of judges – a Kikuyu, two Luos, and a Mijikenda – has a probability of treatment equal to 0.25.

Inverse probability weighting has an appealing property, in that it allows us to identify and drop observations that violate the positivity assumption (e.g., observations that have probability of treatment equal to zero or one.) Such observations are not properly randomized, since they will always (or never) receive treatment. In our application, this is particularly pertinent, since the ethnic geography of Kenya means that certain court stations will likely hear cases from appellants from the locally predominant ethnic group. If the judges hearing cases all come from that same locally predominant ethnic group, then many appellants will face a probability of treatment equal to one. The cost is sample size: units that never or always receive treatment are dropped from the analysis.

Given that we have no information about when an appeal was filed and assigned to a judge, we cannot know for certain the *precise* distribution of judge ethnicities at a given court station. We start by approximating the probability of treatment for each appellant by calculating the distribution of judge ethnicities at each court station in the year in which the appellant's judgment is delivered. Given case backlogs, the length of time it takes to obtain a judgment, and rotation of judges in and out of a court station over time, it is likely that the actual distribution of judge ethnicities when the case was filed/assigned would be different from that distribution when the judgment was delivered. Thus, in addition to calculating this distribution using the set of judges at station j in year t, we also calculate that distribution using judges from years $t-1$ and t (table xx), $t-2$ and t (table xx), $t-3$ and t (table xx), and $t-4$ and t (table xx) below. In short, we examine how the results change across a range of plausible IPW's derived from reasonable approximations of appellants' individual probabilities of treatment. We find positive point estimates that largely comport with the main results in the text. Moreover, as the timespan of the judge-ethnicity distribution increases, we discard fewer units due to the positivity issue discussed above, and statistical significance approaches that of the primary results presented in main text.

D.2 Replication of Table 1 with IPW

Table D1: Coethnic Bias in Criminal Appeal Decisions – 1-year IPW

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.046***	0.048***	0.040*	0.041*	0.037	0.030
	(0.018)	(0.018)	(0.022)	(0.021)	(0.028)	(0.029)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	3,008	3,008	3,008	3,008	3,008	3,008
R^2	0.002	0.015	0.108	0.119	0.153	0.161

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with one-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D2: Coethnic Bias in Criminal Appeal Decisions – 2-year IPW

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.046***	0.045***	0.043*	0.044**	0.036	0.031
	(0.016)	(0.016)	(0.022)	(0.021)	(0.026)	(0.026)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	3,863	3,863	3,863	3,863	3,863	3,863
R^2	0.002	0.014	0.104	0.113	0.143	0.150

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with two-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D3: Coethnic Bias in Criminal Appeal Decisions – 3-year IPW

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.052^{***}	0.049^{***}	0.058^{***}	0.059^{***}	0.054^{**}	0.048^{*}
	(0.015)	(0.015)	(0.022)	(0.021)	(0.027)	(0.026)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	4,436	4,436	4,436	4,436	4,436	4,436
R^2	0.003	0.015	0.105	0.114	0.141	0.149

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Coefficients estimated using OLS with three-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D4: Coethnic Bias in Criminal Appeal Decisions – 4-year IPW

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.052^{***}	0.048^{***}	0.057^{**}	0.057^{**}	0.052^{**}	0.047^{*}
	(0.014)	(0.014)	(0.023)	(0.022)	(0.025)	(0.025)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	4,734	4,734	4,734	4,734	4,734	4,734
R^2	0.003	0.016	0.105	0.115	0.141	0.150

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Coefficients estimated using OLS with four-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D5: Coethnic Bias in Criminal Appeal Decisions – 5-year IPW

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.048^{***}	0.043^{***}	0.057^{**}	0.056^{**}	0.050^{**}	0.046^{*}
	(0.014)	(0.014)	(0.023)	(0.023)	(0.024)	(0.024)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	4,960	4,960	4,960	4,960	4,960	4,960
R^2	0.002	0.016	0.104	0.115	0.140	0.150

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with five-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

D.3 Replication of Table 2 with IPW

Table D6: Effect of Coethnic Match between Appellant and Judge, by Judge Ethnicity and One-year IPW.

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.047	−0.013	0.001	0.106	−0.055	0.028	0.065
	(0.051)	(0.045)	(0.040)	(0.081)	(0.128)	(0.156)	(0.245)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	691	335	984	406	190	131	271
R^2	0.209	0.263	0.208	0.175	0.196	0.409	0.264

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with one-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D7: Effect of Coethnic Match between Appellant and Judge, by Judge Ethnicity with Two-year IPW.

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.056	0.023	−0.014	0.072	−0.020	0.001	−0.022
	(0.038)	(0.049)	(0.044)	(0.064)	(0.077)	(0.145)	(0.263)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	822	420	1,314	566	250	151	340
R^2	0.207	0.282	0.177	0.147	0.166	0.397	0.259

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with two-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D8: Effect of Coethnic Match between Appellant and Judge, by Judge Ethnicity with Three-Year IPW.

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.095^{**}	0.063	0.022	0.064	−0.040	−0.006	0.030
	(0.042)	(0.037)	(0.044)	(0.067)	(0.070)	(0.155)	(0.233)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	935	504	1,511	653	291	176	366
R^2	0.210	0.277	0.169	0.143	0.148	0.396	0.277

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with three-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D9: Effect of Coethnic Match between Appellant and Judge, by Judge Ethnicity with Four-Year IPW.

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.107^{***}	−0.034	0.022	0.054	−0.006	−0.002	0.032
	(0.036)	(0.042)	(0.045)	(0.066)	(0.056)	(0.140)	(0.229)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,003	588	1,552	700	308	185	398
R^2	0.221	0.255	0.168	0.131	0.148	0.405	0.285

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with four-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

Table D10: Effect of Coethnic Match between Appellant and Judge, by Judge Ethnicity with Five-Year IPW.

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.093**	−0.011	0.027	0.054	−0.001	−0.040	0.050
	(0.034)	(0.030)	(0.045)	(0.063)	(0.052)	(0.146)	(0.232)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,038	645	1,568	755	332	191	431
R^2	0.220	0.242	0.169	0.143	0.145	0.407	0.294

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS with five-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise.

D.4 Dropping "Lost Interest" Cases

Tables D11 and D12 drop observations that represent administrative clearance of appeals that had not seen regular activity in several years. These 45 cases represent situations where the judge asserts that, given the lack of activity from the appellant's side, the appellant has lost interest in pursuing the appeal. When these cases are dropped, the results marginally strengthen.

Table D11: Effect of Coethnic Match, Dropping "Lost Interest" Cases

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.044***	0.041***	0.043**	0.042**	0.039***	0.036**
	(0.015)	(0.015)	(0.018)	(0.017)	(0.014)	(0.014)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	9,500	9,500	9,500	9,500	9,500	9,500
R^2	0.001	0.008	0.078	0.084	0.104	0.108

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise. Excludes cases dropped due to loss of interest from appellant.

Table D12: Effect of Coethnic Match, Dropping "Lost Interest" Cases, by Judge Ethnicity

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.060**	0.025	−0.018	0.024	0.079**	0.096	0.057
	(0.023)	(0.022)	(0.029)	(0.061)	(0.029)	(0.130)	(0.094)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,192	1,042	2,916	1,216	760	531	843
R^2	0.164	0.146	0.132	0.091	0.087	0.223	0.203

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Coefficients estimated using OLS with five-year inverse probability weights. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise. Excludes cases dropped due to loss of interest from appellant.

D.5 Accounting for judges who were deemed unsuitable for higher positions in the judiciary due to corruption

Table D13: Effect of Coethnic Match with Controls for Judges Ineligible for Promotion

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.042***	0.039***	0.043**	0.041**	0.036**	0.033**
	(0.015)	(0.015)	(0.018)	(0.018)	(0.014)	(0.014)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	9,545	9,545	9,545	9,545	9,545	9,545
R^2	0.001	0.009	0.079	0.085	0.105	0.110

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise. Specifications include dichotomous variables for judges declared ineligible for higher judicial posts by the Judicial Service Commission review.

Table D14: Effect of Coethnic Match Omitting Judges Ineligible for Promotion

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.048***	0.044***	0.048**	0.046**	0.037**	0.034**
	(0.015)	(0.015)	(0.019)	(0.018)	(0.015)	(0.015)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	9,267	9,267	9,267	9,267	9,267	9,267
R^2	0.001	0.010	0.080	0.086	0.104	0.109

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise. Specifications exclude judges declared ineligible for higher judicial posts by the Judicial Service Commission review.

D.6 Alternative Aggregation of Ethnic Groups

Table D15: Effect of Coethnic Match between Appellant and Judge.

	Outcome: Judgement for the Defendant					
	(1)	(2)	(3)	(4)	(5)	(6)
Coethnic Match	0.041***	0.039***	0.043**	0.042**	0.037**	0.035**
	(0.015)	(0.015)	(0.018)	(0.017)	(0.014)	(0.014)
Courthouse-Year FE	No	No	Yes	Yes	Yes	Yes
Individual Judge FE	No	No	No	No	Yes	Yes
Case-specific Controls	No	Yes	No	Yes	No	Yes
Observations	9,545	9,545	9,545	9,545	9,545	9,545
R^2	0.001	0.009	0.079	0.085	0.105	0.110

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise. Uses more aggregated version of ethnic match variable.

Table D16: Effect of Coethnic Match between Appellant and Judge, by Judge Ethnicity.

	Dependent variable:						
	Outcome: Judgement for the Defendant						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coethnic Match	0.057**	0.054	−0.018	0.024	0.079**	0.096	0.042
	(0.023)	(0.032)	(0.029)	(0.061)	(0.029)	(0.130)	(0.083)
Sample	Kikuyu	Kalenjin	Luhya	Luo	Kamba	Kisii	Other
Courthouse-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Judge FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Case-specific Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,235	1,042	2,917	1,217	760	531	843
R^2	0.169	0.147	0.132	0.091	0.087	0.223	0.203

Note: *p<0.1; **p<0.05; ***p<0.01. Coefficients estimated using OLS. "Coethnic Match" is a binary variable equal to one if the judge and appellant share the same ethnic group, zero otherwise. Uses more aggregated version of ethnic match variable.

D.7 Uncertainty over appellant ethnicity

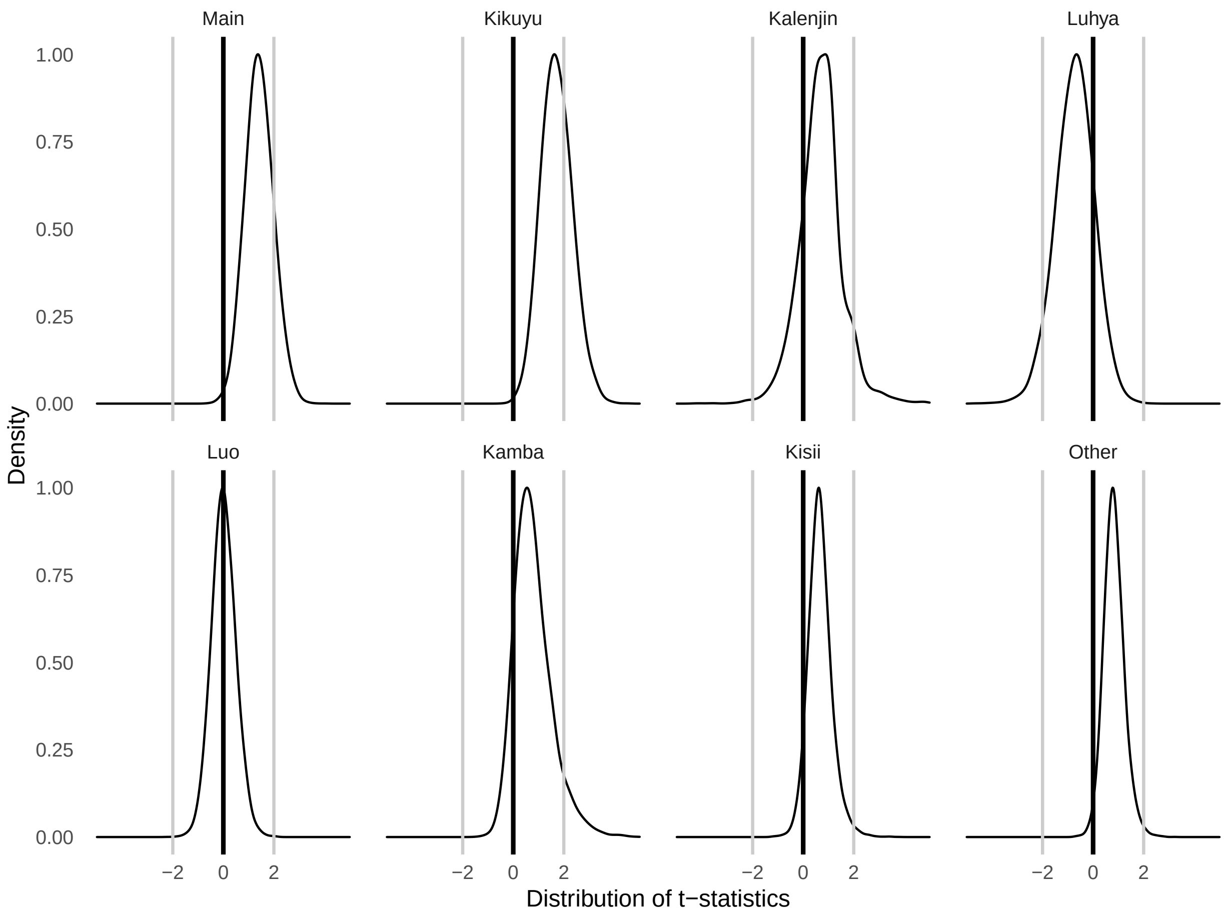
In the main text, we take the estimated highest probability ethnic group associated with the appellant, and classify that group as the appellant's group. This approach underestimates the measurement uncertainty, since there is some chance that the appellant could identify with another group. In this subsection, we simulate this uncertainty to see how it affects the primary results. Our expectation is that the results will retain the observed sign but, given the additional noise, the results will attenuate.

To simulate, we proceed as follows for 10000 iterations.

1. For appellant i, calculate the probability of membership to each ethnic group g.
2. Draw a random variable from the distribution defined by the calculated categorical probabilities. This draw represents one possible ethnic assignment for the appellant.
3. Create the ethnic match variable based on this random draw for the appellant.
4. Estimate regression coefficients based on this new ethnic match variable.
5. Store the coefficients and standard errors and repeat.

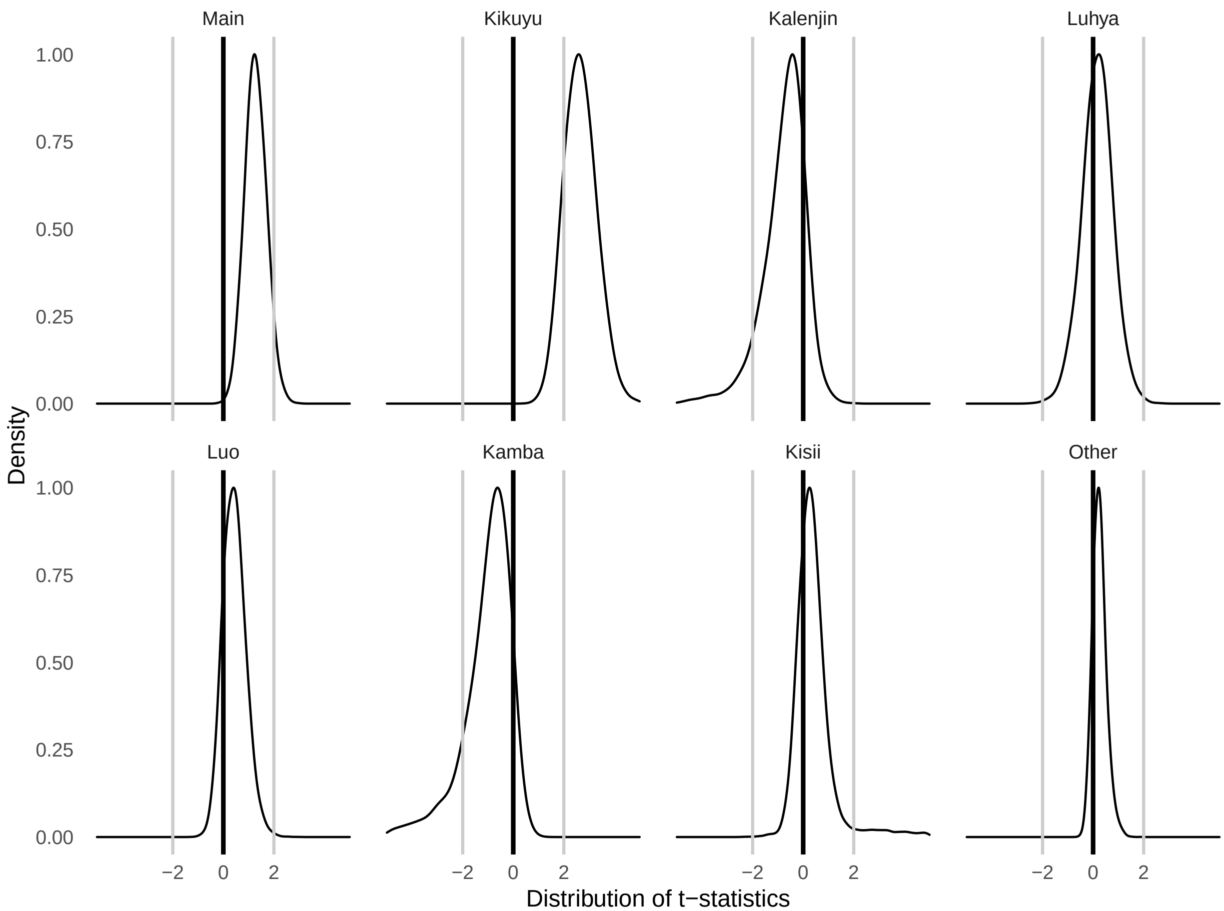
Figure D1 presents the distribution of t-statistics from these 80000 regressions (10000 x 8 table columns) for the rightmost column in table 1 (containing the most stringent fixed effect specification) and all columns of table 3 (by judge-ethnic group). The figure largely supports the main results. The modes for both the main result in table 1 and the result from table 3 on Kikuyu judges remain near the deterministic results, while the other analyses remain mostly below standard levels of statistical significance. Figure D2 replicates Figure D1 using the four-year inverse-probability weights.

Figure D1: Appeal Outcome: Robustness Check



Notes: The figure shows the distribution of t-statistics for each regression model, where the match variable incorporates uncertainty about appellant ethnicity by randomly selecting appellant ethnicity from the estimated distribution across ethnic groups.

Figure D2: Appeal Outcome: Robustness Check, Four-Year IPW.



Notes: The figure shows the distribution of t-statistics for each regression model, where the match variable incorporates uncertainty about appellant ethnicity by randomly selecting appellant ethnicity from the estimated distribution across ethnic groups.

E Robustness: Text Analysis

We validated our main text analysis two ways: first, we ran a GloVe model on our corpus to derive a corpus-specific set of word embeddings; second, we created a second set of dictionaries using the NRC Emotion Lexicon. We then evaluated our analysis using every possible combination of vectors and dictionaries – pretrained vectors, corpus-trained vectors, corpus seeds, NRC seeds. Table E1 reveals that our main results are unchanged using NRC seed words, wherein judges use approximately 7% more trustworthy terms when hearing the case of a coethnic appellant than a non-coethnic; this estimate is statistically significant at the 5% confidence level. This relationship appears even stronger when using vectors modeled directly from the corpus, as shown in Table E2 – the coethnic match variable is now approximately 12% at the 1% confidence level. Our results are consistent when we combine corpus-trained vectors with corpus-derived seed words, as shown in Table E3, where the coethnic match variable is approximately 11% and still statistically significant at the 1% confidence level. The variability of our results with the corpus-trained vectors is in line with broader research on the challenges of running such models on smaller bodies of text. As Antoniak and Mimno (2018) observe, cosine similarity scores are more volatile when trained on relatively small, niche corpora. It thus makes sense that our findings from the pretrained GloVe vectors – estimated using nearly 2 billion words – are considerably more stable.

Table E1: Coethnic Bias in Written Judgments: NRC Seeds, GloVe Vectors

	Dependent variable:			
	Sentiment			
	Trust	Disgust	Positive	Negative
	(1)	(2)	(3)	(4)
Coethnic Match	0.074**	0.022	0.082**	0.011
	(0.031)	(0.038)	(0.037)	(0.033)
Individual Judge FE	Yes	Yes	Yes	Yes
Courthouse-Year FE	Yes	Yes	Yes	Yes
Observations	9,545	9,545	9,545	9,545
R^2	0.235	0.214	0.221	0.238
Note:				*p<0.1; **p<0.05; ***p<0.01

Table E2: Coethnic Bias in Written Judgments: NRC Seeds, Corpus-Derived Vector

	Dependent variable:			
	Sentiment			
	Trust	Disgust	Positive	Negative
	(1)	(2)	(3)	(4)
Coethnic Match	0.122***	0.052*	0.022	0.022
	(0.040)	(0.027)	(0.035)	(0.036)
Individual Judge FE	Yes	Yes	Yes	Yes
Courthouse-Year FE	Yes	Yes	Yes	Yes
Observations	9,545	9,545	9,545	9,545
R^2	0.194	0.233	0.168	0.227
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table E3: Coethnic Bias in Written Judgments: Corpus Seeds, Corpus-Derived Vectors

	Dependent variable:			
	Sentiment			
	Trust	Disgust	Positive	Negative
	(1)	(2)	(3)	(4)
Coethnic Match	0.111***	0.045	0.027	0.024
	(0.035)	(0.036)	(0.036)	(0.029)
Individual Judge FE	Yes	Yes	Yes	Yes
Courthouse-Year FE	Yes	Yes	Yes	Yes
Observations	9,545	9,545	9,545	9,545
R^2	0.154	0.165	0.241	0.178
Note:	*p<0.1; **p<0.05; ***p<0.01			

F Robustness: Validation of Word Embeddings

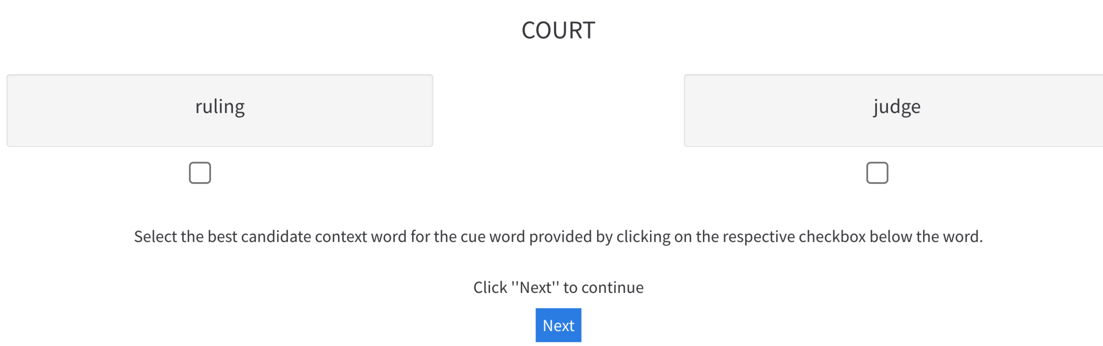
In this appendix, we report on a validation exercise designed to evaluate the degree to which the word embeddings we employ[42] approximate human usage of words.

Word embeddings systematically encode how words relate to one another. The degree to which word embeddings approximate human use of words is an open question within any given application. We use the Turing test described in Rodriguez and Spirling (N.d.) to investigate this question. In general, a Turing test is an "imitation game" (Turing, 1950). A human is confronted with two signals – sentences, for instance – one generated by a human and the other generated by a computer. If the human is unable to distinguish which signal was generated by which source, then the computer has passed the Turing test.

In our case, word embeddings take the role of the computer. In the first step, the word embeddings are used to retrieve ten context words most similar to various legally-relevant cue words.[43] Similarly, we task a set of mTurk workers with listing their top ten context words for the same cue words, and taking the top ten words most frequently mentioned by these workers. These represent the "human" signal.[44]

In the second step, we presented approximately 200 mTurkers with a set of cue words. For each cue word, we also presented two context words, one human- and one embedding-generated. These words were unlabelled, so that the worker had no information about the source of the context word. Then, we asked the mTurker to indicate the context word that better corresponded with the cue word. Figure F1 provides an example of this "triad task."

Figure F1: Word Embeddings Validation Turing Test: In this triad task, the human respondent observes the cue word "court," and must decide whether "ruling" or "judge" is a more suitable context word.



For each cue word, we calculated the expected probability that the embedding-generated context word was preferred by mTurkers over the human-generated word. Again, following Rodriguez and Spirling (N.d.), we divide this probability by 0.5 to create a metric ranging from 0 to 2. On this scale, a "1" represents human-rater indifference between embedding- and human-generated context words.

[42]We use the 300d, 1.9million dimension vectors available at http://nlp.stanford.edu/data/glove.42B.300d.zip.

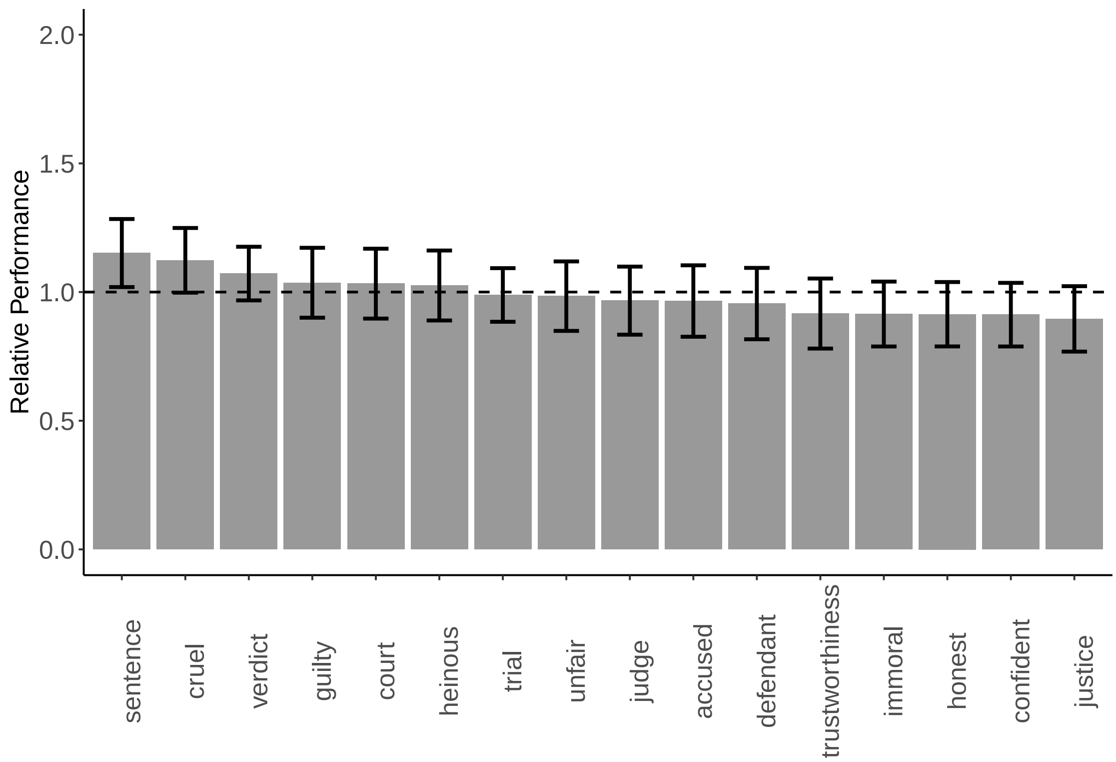
[43]Following the literature, we use the top ten as defined by cosine similarity for the 300d vector.

[44]We thank Pedro Rodriguez for making his Rshiny code available via github for both the generation of human context words and for the Turing test.

Values above one suggest humans prefer the embeddings-generated word; values below one suggest humans prefer the human-generated word.

Figure F2 presents results from the Turing test. Across sixteen words relevant to the legal context in question, we find evidence that the word embeddings compare favorably to human-generated text. For all but one of the 95% bootstrapped confidence intervals contain one, suggesting that mTurkers do not systematically prefer human-generated context words over machine-generated ones.

Figure F2: Word Embeddings Validation Turing Test



Notes: On the y-axis, values equal to one suggest that mTurkers are indifferent between embeddings context words and the human-generated context words on average. Values above one suggest humans prefer the embeddings-generated word; values below one suggest humans prefer the human-generated word. Error bars represent 95% bootstrap confidence intervals.

G Research Ethics

This appendix discusses ethical considerations related to this research. Below, we discuss issues related to the mTurk Turing Test (subsection G.1 and the observational data subsectionG.2), focusing in each subsection on the specific principles considered during the research process. The authors affirm adherence to American Political Science Association (APSA)'s 2020 Principles and Guidance for Human Subjects Research, and have no deviations to report.

G.1 mTurk Turing Test

Principle 11 (Ethical Review) Prior to conducting the mTurk data collection, one of the authors corresponded with the institutional review board regarding whether or not this data collection should undergo full human subjects review. The IRB team advised that as long as mTurk workers are engaged "to provide human validation, or perform a human task that will help validate a statistical model, without 'studying' them [i.e., the mTurkers] in any way (i.e. not collecting any data about them personally, not asking for their perspectives, opinions, beliefs, etc.)," then human subjects review would not be required, though including consent language was recommended. Accordingly, we did not collect any personal data or information on their perspectives, opinions, or beliefs, and we did include consent language.

Principle 5 (Consent) To obtain consent for the human generation of words for the Turing test, we included the following text on the landing page of the data collection website: "This is an academic research project to understand words and their contexts. If you consent to participate in this study, please enter your MTurk ID and press 'Start'."

To obtain consent for the Turing test itself, we included the following text on the landing page of the data collection website: "This is an academic research project to understand how words relate to each other. If you consent to participate in this study, please enter your MTurk ID and press 'Start'."

Additionally, both landing pages contained the following note on confidentiality: "Confidentiality: responses are anonymous, we have no way of linking the data to individual identities."

Principle 6 (Deception) No deception was used during the mTurk Turing Test.

Principle 7 & 8 (Harm and Trauma), Principle 9 (Confidentiality) The Turing Test did not involve any tasks with any potential to inflict harm or trauma on respondents. Furthermore, we did not collect any identifying information for participants; there should be no potential for a breach in confidentiality.

Compensation: Our aim in compensation was to provide a payment equivalent to at least a $10 an hour wage, which would be about 38% higher than U.S. minimum wage. To estimate the time for task completion, one co-author and two undergraduate RAs completed the task.
Word Generation Task: We estimated ex ante that the word generation task would take between 15 and 20 minutes. We paid mTurk workers $3.50 to complete this task, which translated to an hourly wage of between $10.50 and $14. On average, post-completion data show that mTurk respondents took 20 minutes to complete the task, resulting in an average ex post hourly wage equivalent of $10.50.
Turing Test Task: We estimated ex ante that the Turing Test task would take approximately 5 minutes to complete. We paid mTurk workers $1.25 to complete this task, which translated to an hourly wage of between $15 and $18.75 per hour. In practice, this task took approximately 7.77 minutes to complete, resulting in an ex post hourly wage equivalent of $9.65.

G.2 Observational Data

Principle 7 & 8 (Harm and Trauma), Principle 9 (Confidentiality) Although our observational data analysis does not fall under traditional definitions of human subjects research, we nonetheless briefly discuss the ethical implications of using data on criminal appeals, especially as it pertains to issues of harm, trauma, and confidentiality. Our legal judgements data were accessed on the free, publicly available database at http://kenyalaw.org/caselaw/. This repository is accessible by anyone around the world and was explicitly designed in cooperation with the Kenyan Law Society to make legal decisions in Kenya more transparent. In fact, users can search for cases based on citation, judge name, litigant name, date, and other keywords. For our analysis, we compiled raw text judgements as a spreadsheet, which included the already-public judge and defendant names, along with the requisite indicator variables for case types and court stations. Because our data are derived from a more detailed, structured, and easily accessible public data source, our dataset does not increase existing risks to participants in these legal proceedings. More specifically, the data we generate from these publicly-available records will be less accessible than the easily searchable Kenya Law website (in fact, entering citation information directly into google search redirects you to the Kenya Law website), since the file itself will be contained within a (relatively obscure) political science data archive. Simply put, our data is less accessible, less Google-searchable, and less digestible than the documents that can be found on the public Kenya Law case search website from which our data are derived. We therefore believe that the release of the replication data should not pose any additional harm or trauma (Principles 7 and 8) or breach of confidentiality (Principle 9) beyond the risk imposed by the release of the source data on http://kenyalaw.org/caselaw/.