The Cognitive Underpinnings of Judicial Bias: The Role of Social Identity and Prospect Theory

Daniel L. Chen, Jimmy Graham, Manuel Ramos-Maqueda, Shashank Singh*

March 6, 2025

Abstract

Under what conditions do judges favor their own group? Collecting the available universe of Superior Court decisions in Kenya, we leverage the random assignment of cases to judges to evaluate the extent of judicial in-group bias along gender and ethnic lines. We find that defendants are 4 or 6 percentage points more likely to win if they share the judge's gender or ethnicity, respectively, and that judges are significantly more biased in favor of defendants than plaintiffs. Our findings highlight the need to re-examine the emerging consensus that judges uniformly favor their own group and continue investigating the mechanisms driving judicial bias. We propose that the uneven application of bias can be explained by a framework of social identity and loss aversion, and we support this claim with data on the amount of damages in each case, which we extract using a large language model. We argue that our framework could serve as a more complete lens through which to decipher the cognitive underpinnings of judicial biases.

*Daniel L. Chen (Toulouse School of Economics, daniel.chen@iast.fr); Jimmy Graham (NYU), Manuel Ramos-Maqueda (Oxford), Shashank Singh (University of Chicago)

Significance Statement:

This study analyzes judicial bias in Kenya's Superior Courts, using 29,363 civil cases from 1976-2020. We reveal judges' significant in-group bias toward defendants based on gender and ethnicity through quasi-random case assignments. Our findings challenge the belief that judicial bias is uniformly applied, suggesting instead that selective biases exist. By combining social identity theory with prospect theory, we propose that loss aversion underlies judicial behavior. These insights deepen understanding of cognitive influences on judicial decisions and underscore the need for reforms to enhance fairness and equity in judicial systems.

1 Introduction

Judges often exhibit bias toward specific groups (Shayo and Zussman 2011; Gazal-Ayal and Sulitzeanu-Kenan 2010; Kastellec 2013; Glynn and Sen 2015; Grossman et al. 2016; Yang 2015; Depew, Eren, and Mocan 2017; Arnold, Dobbie, and Yang 2018; Knepper 2018; Sloan 2020; Choi, Harris, and Shen-Bayh 2021). This bias can have serious consequences. It can further disadvantage marginalized groups and undermine the effectiveness and inclusivity of courts, which are important for a well-functioning economy and a state's ability to maintain legitimacy and trust among citizens (World Bank 2017). However, there are still many unknowns regarding the scope and drivers of judicial bias. Without a better understanding of where judicial bias comes from and who is targeted by it, it is difficult to develop solutions to mitigating bias.

In this paper, we seek to explore the extent to which judges favor their own group in terms of gender and ethnicity—and the conditions under which they do so. Our study takes place within the context of the Kenyan judiciary. We construct our main data source by scraping the Kenyan Judiciary's publicly available database for Superior Court civil cases over the period 1976-2020 and using machine learning techniques to extract case outcome information from written judicial decisions. Our final sample includes 29,363 cases. To determine the causal effect of an in-group relationship between judges and litigants, we rely on the quasi-random assignment of cases to Kenyan judges. In Kenya, cases filed to a court are assigned to judges based on their existing caseload and the date of filing, which is orthogonal to any other characteristics of the case. To confirm random assignment, we show that male and female defendants and plaintiffs are balanced across judges. We also show that defendants and plaintiffs across all ethnicities are not more likely to be assigned judges from their ethnic group.

We find that judges in Kenya display both gender and ethnic in-group bias toward defendants. Our results suggest that defendants are about 4 percentage points more likely to win if they share the judge's gender and about 6 percentage points more likely to win if they share the judge's ethnicity relative to a baseline win rate of about 44 percent. However, we find no evidence of in-group bias toward plaintiffs. This gap in biases is driven by cases in which the defendant and plaintiff have the same identity. For example, in cases where both the defendant and plaintiff are female, the defendant is more likely to win if the judge is female. Although we are not able to disentangle whether these results are due to a lack of bias toward plaintiffs or a stronger bias toward defendants relative to plaintiffs, the results clearly suggest that judicial in-group bias is not uniformly applied across actors in court cases.

The idea that judges do not uniformly apply bias across members of their group has important implications. For one, it adds important nuance to an emerging consensus that judges and jurors display bias toward their in-group.[1] While generally consistent with this consensus, our results suggest that judges may not be biased toward all in-group members and, at the very least, do not apply bias evenly across all actors. Indeed, our research raises the possibility that judicial in-group bias is applied differently across other actors in court cases, including victims and attorneys. Therefore, the simple claim that judges tend to be biased toward their in-group is incomplete, as a judge's application of bias towards their group clearly depends on the context. This has important implications for judicial outcomes, as different actors in court cases may receive greater advantages than others. Recognizing variation in the application of bias throughout judicial systems is an important first step to devising means to reduce bias and create more just systems.

To our knowledge, no other study of judicial in-group bias has examined cases that include human plaintiffs and defendants of different identities. Most studies focus on criminal cases, in which the plaintiff is usually the state. Other studies have looked at civil cases, but only those where the defendant is an organization or the litigants differ in terms of the relevant in-group identity (Shayo and Zussman 2011; Knepper 2018). As such, our study is unique in its ability to uncover judges' unequal application of bias.

[1] For examples see Shayo and Zussman 2011; Gazal-Ayal and Sulitzeanu-Kenan 2010; Grossman et al. 2016; Knepper 2018; Sloan 2020; Choi, Harris, and Shen-Bayh 2021; Ahrsjo, Niknami, and Palme forthcoming

We propose that this variation in bias across in-group litigants can be explained by social identity theory and loss aversion. Previous research has shown that people are typically more sensitive to losses than gains (Kahneman and Tversky 1979; Benartzi and Thaler 1995). This phenomenon is called loss aversion. To apply this concept to the judiciary, note that any loss for the defendant in a civil case constitutes an equal gain for the plaintiff. It follows that if the judge shares a social identification with both the plaintiff and the defendant, any transfer from the defendant to the plaintiff should constitute a net loss for the group, as the loss felt from one group member (the defendant) will be greater than the gain felt by another group member (the plaintiff). In other words, a judge may feel that the harm of assigning guilt to an in-group member may outweigh the benefit of validating the claims of another in-group member. Therefore, if the judge's goal is to maximize the standing of their group, they may choose to rule in favor of the defendant, as doing so would avoid the sensation that their group is experiencing loss.

To provide evidence of this mechanism, we leverage original data on damages requested from each case, which we extract using a large language model. Loss aversion theory (also known an prospect theory) suggests that, as the size of losses grows, so too does the loss in utility. As the size of gains grows, the gain in utility also grows. However, the rate at which gain translates into utility is slower than the rate at which loss translates into lost utility (Barberis 2013). Therefore, if loss aversion is driving the discrepancy in bias between defendants and plaintiffs, we should see the discrepancy widen for larger losses. In other words, for cases invoving larger amounts of damages, judges should be even more biased toward defendants relative to plaintiffs. We present evidence consistent with this idea.

2 The Kenyan context

Kenya provides an ideal setting for studying judicial bias. First, economic inequalities and political allegiance are distributed across regions and ethnicity, with political parties and coalitions created along clear ethnic lines. So, ethnicity is a highly salient topic that could influence decision-making (Asingo et al. 2018). Second, according to our data, women and certain ethnic groups are underrepresented in the judiciary.[2] If in-group bias is widespread, it may disproportionately harm these underrepresented groups. This means the stakes of judicial bias are high. Third, there is an ongoing debate regarding the extent to which co-ethnic bias affects decision-making in the context of Africa generally and Kenya specifically (Berge et al. 2015).

The Kenyan judiciary is divided into two main court types: Superior and Subordinate Courts. The vast majority of our data covers the Superior Courts, which include High Courts, which hear both criminal and civil cases and appeals from Subordinate Courts; Environment and Land Courts; Employment and Labour Relations Courts; the Court of Appeal, which hears appeals from other Superior Courts; and the Supreme Court, which hears appeals from the Court of Appeal and other high-level cases (Kenyan Judiciary 2021).

The Kenyan judiciary does not employ a jury system. This means that judges alone are able to decide the outcomes of cases, which implies that bias among judges can have especially serious consequences. For most cases in most courts, there is only one judge. An exception is in Courts of Appeal, where the majority of cases are composed of multi-judge panels.

3 Data

The main data source used in our analysis is the Kenyan Judiciary's publicly available database for court cases.[3] The database includes 159,645 cases, almost exclusively from the Superior Courts, over the period of 1976 to 2020. Kenya Law, an organization within the Kenya Judiciary, began uploading case information in 2006. They upload all cases that are sent to them from the individual courts. Judicial officers in Superior Courts have a mandate

[2]See appendix A for evidence.
[3]See http://kenyalaw.org/caselaw/.

to send cases to Kenya Law. For cases prior to 2006, Kenya Law has made (and continues to make) efforts to gather and upload case information.

In order to build our dataset for analysis from this database, we scraped the metadata and full text decision associated with each case. These data allowed us to directly extract the following for most cases: the names of plaintiffs, defendants, and judges; the type of case; the court in which the case was heard; and the year the judgment was delivered.

Cases with non-human plaintiffs or defendants (i.e., companies, organizations, or the state) were dropped from the sample. By focusing on cases with human litigants, we are able to examine how judicial in-group bias varies across actors involved in a case. This approach leaves us with only civil cases and no criminal cases in the sample.

To determine gender and ethnicity and remove non-human cases, we used the name information scraped from the database. Cases without gender or ethnicity information for judges and either plaintiffs or defendants were dropped.[4] In some cases, multiple defendants (or plaintiffs) share the same side; in these situations, we classify the side's 'majority' gender or 'plurality' ethnicity. While grouping litigants this way is not entirely conventional, it captures the predominant demographic identity on each side and fits our main question of whether judges favor in-group litigants overall.[5]

To determine the winner of each case, we first scraped the case outcome information from the metadata. However, for 58,622 cases, the outcome was not stated. For these cases, we used a Binary Classification Machine Learning Model (described in appendix B) to analyze the text decisions of each case and determine the outcome. In the test set, the model was about 93 percent accurate.

Our final sample includes 29,363 cases with human litigants with gender or ethnicity

[4]The process for removing non-humans and determining gender and ethnicity, as well as the reasons for missing information, are discussed in appendix B.

[5]If no majority could be determined for gender, the majority gender was coded as missing. If no plurality could be determined for ethnicity, the plurality ethnicity was coded as "no plurality" and kept in the sample. This difference in coding was necessary because the main specification for gender in-group analysis requires binary coding, while the main specification for ethnicity in-group analysis does not (see below).

data.[6] This data covers 95 courts and 392 judges. This sample cover the years 1976 to 2020, with an increase in cases over time (see figure A1 in the appendix).

To unpack the mechanisms driving bias, we also extract data on the monetary value of damages claimed for each case. The data extraction process utilized an advanced large language model (LLM), specifically LLama-Index, powered by the OpenAI API. The LLM was tasked with identifying and extracting phrases indicating amounts requested, followed by regular expressions to capture the precise figures in Kenyan shillings. We successfully extracted monetary damages for 3,420 cases, about 12 percent of the full sample. The low extraction rate is driven by a lack of mention of damage amounts in many cases and the fact that not all cases involve monetary damages. For example, many cases involve property disputes.

4 Empirical strategy

To evaluate the existence of in-group bias, we exploit the quasi-random assignment of cases to judges. In Kenyan Superior Courts, cases filed in a court are categorized by court type and sent to the deputy registrar of the relevant court division (family, commercial and admiralty, labour and employment, constitutional, land and environment, or criminal). The deputy registrar then assigns the case to a judge based on the judge's caseload and calendar, without considering case characteristics. This exogenous assignment is orthogonal to case characteristics such as the gender or ethnicity of the parties. Thus, this system produces as-good-as-random assignment of plaintiffs and defendants to judges, conditional on court division. In appendix C, we present balance tests to support our claim of quasi-randomness. It must be noted that exogenous case assignment may be a recent phenomenon, as introducing randomization was one of the goals of the reform team following the implementation of the 2010 Kenyan Constitution (Gainer 2015). Therefore, in addition to conducting the tests

[6] Of the initial 159,645 cases, 33,876 had exclusively human litigants for civil cases. An additional 4,513 cases were dropped because we were unable to determine majority gender or ethnicity for the litigants in the case.

on the full sample, we split the samples into before 2011 and after 2010. As a robustness check in appendix D, we conduct the main analysis for years after 2010.

To estimate judicial gender bias, we conduct the following regression:

$$\begin{aligned} Def\_win_{i,c,t} = \alpha + \beta_1 judge\_maj\_female_{i,c,t} + \beta_2 def\_maj\_female_{i,c,t} + \\ \beta_3 judge\_maj\_female_{i,c,t} * def\_maj\_female_{i,c,t} + \Phi_{c,t} + \lambda' X_{i,c,t,} + \epsilon_{i,c,t} \end{aligned} \tag{1}$$

where $def\_win_{i,c,t}$ is an indicator that equals 1 if the defendant won the case, for case i filed in court c at time t. $judge\_maj\_female$ and def_maj_female are binary variables indicating whether judge panels and defendant groups, respectively, are majority female. The main outcome of interest is the interaction term, which indicates in-group bias. The specification used to test for in-group bias towards plaintiffs is identical to (1), except a binary variable for plaintiff majority gender, pla_maj_female substitutes def_maj_female. An alternate specification includes both variables and their interactions. Φ captures court-year fixed effects and X is a vector of covariates, which includes binary variables for judge, defendant, and plaintiff plurality ethnicity; variables for the numbers of judges, plaintiffs, and defendants; a binary variable indicating whether the case is an appeal; and binary variables indicating the case type. Court-year fixed effects are used to ensure that we are comparing defendants and plaintiffs that are in the same court at the same time. For this and all other models, we cluster standard errors at the judge level.

For the ethnicity in-group bias analysis, we use a slightly different econometric specification in order to account for the fact that there are many more categories of ethnicity. To estimate judicial ethnic bias, we conduct the following regression:

$$Def\_win_{i,c,t} = \alpha + \beta_1 judge\_pla\_same_{i,c,t} + \beta_2 judge\_def\_same_{i,c,t} + \Phi_{c,t} + \lambda' X + \epsilon_{i,c,t} \tag{2}$$

where $judge\_pla\_same_{i,c,t}$ is a binary variable indicating whether the judge ethnic plurality is the same as the plaintiff ethnic plurality, and $judge\_def\_same_{i,c,t}$ is a binary variable indicating whether the judge ethnic plurality is the same as the defendant ethnic plurality.

To explore mechanisms, we incorporate information on damages requested in each case into the analysis. As described above, prospect theory suggests that, as damage amounts grow, judges should exhibit increasingly stronger bias toward defendants relative to plaintiffs. To test this idea, we conduct the following regression for gender:

$$\begin{gathered} Def\_win_{i,c,t} = \alpha + \beta_1 judge\_maj\_female_{i,c,t} + \beta_2 def\_maj\_female_{i,c,t} + \\ \beta_3 pla\_maj\_female_{i,c,t} + \beta_4 damages_{i,c,t} + \\ \beta_5 judge\_maj\_female_{i,c,t} * damages_{i,c,t} + \\ \beta_6 judge\_maj\_female_{i,c,t} * def\_maj\_female_{i,c,t} + \\ \beta_7 damages_{i,c,t} * def\_maj\_female_{i,c,t} + \\ \beta_8 judge\_maj\_female_{i,c,t} * pla\_maj\_female_{i,c,t} + \\ \beta_9 damages_{i,c,t} * pla\_maj\_female_{i,c,t} + \\ \beta_{10} judge\_maj\_female_{i,c,t} * def\_maj\_female_{i,c,t} * damages_{i,c,t} + \\ \beta_{11} judge\_maj\_female_{i,c,t} * pla\_maj\_female_{i,c,t} * damages_{i,c,t} + \Phi_{c,t} + \epsilon_{i,c,t} \end{gathered} \tag{3}$$

where *damages* is a variable for the amount of Kenyan Shillings (in millions) requested by the plaintiff.[7] Loss aversion predicts that β_{10} should be positive and β_{11} should be null. Using the variables from equation two, we also conduct an analogous three-way interaction for ethnicity.

[7] 1 USD equals about 130 Kenyan Shillings.

5 Results

The gender regression results are presented in table 1. The significantly positive coefficients on the interaction between judge and defendant majority gender provide evidence that there is in-group gender bias from judges towards defendants. The significant results suggest that, all else equal, defendants are about 4 percentage points more likely to win if they have the same majority gender as the judges.

The results do not provide evidence for in-group bias toward plaintiffs. The coefficient on the interaction between judge and plaintiff majority gender is null. The sign is also not in a direction that is indicative of bias; flipping the sign, we find that the judge-plaintiff interaction is significantly different ($p<0.05$) than the judge-defendant interaction. To make sense of these results, it is important to note that if we were to restrict the sample to cases where the defendant and plaintiff are different genders, the coefficient on a judge-plaintiff interaction from a regression based on equation 3 would be identical (of an opposite sign) to the coefficient on a judge-defendant interaction from a separate regression based on equation 3. Therefore, this sample of cases cannot be driving the difference in results for plaintiffs and defendants. Now note that if we were to restrict the sample to cases where the defendant and plaintiff are the same gender, the coefficient on a judge-plaintiff interaction from a regression based on equation 3 would be identical (of the same sign) to the coefficient on a judge-defendant interaction from a separate regression based on equation 3. For this sample of cases, a positive coefficient indicates that when judges are the same gender as plaintiffs and defendants, they are more likely to rule in favor of defendants *compared to cases where judges are a different gender than both the plaintiffs and defendants*. Thus, the greater bias toward defendants relative to plaintiffs that we observe is driven by cases where defendants and plaintiffs are the same gender. Importantly, in cases where defendants and plaintiffs are different genders, a significant positive coefficient still indicates in-group bias, but we cannot determine whether it is driven by in-group bias in favor of the defendants, plaintiffs, or both. We therefore cannot claim that judges do not exhibit gender in-group bias toward plaintiffs, only that they exhibit more gender bias toward defendants in cases where the defendants and plaintiffs are the same gender.

The ethnicity results are presented in table 2. They show that defendants are about 6 percentage points more likely to win if they share an ethnicity with the judge. This is evidence of in-group bias among judges towards defendants. The finding is robust to all of the specifications presented. As with gender, we find no evidence of in-group bias toward plaintiffs, and the coefficients for plaintiff and defendant in-group bias are significantly different ($p<0.01$). Once again, we cannot claim that judges do not exhibit in-group ethnic bias toward plaintiffs, only that they exhibit more bias toward defendants. Appendix D provides various robustness checks.

Table 3 presents the results from equation 3. As predicted by loss aversion theory, it shows that the amount of damages requested do indeed moderate the effect of in-group bias for defendants, but not plaintiffs. Table 4 produces results for the interaction between damages and the ethnicity variables. Here we find no moderating effect of damages. The null results may be due in part to the limited sample size for a three-way interaction. They may also suggest the need for future research to explore other explanations for the unevenly applied bias. Although the results in aggregate support prospect theory, there may be additional factors at play.

6 Discussion

In this paper, we have examined the presence of judicial in-group bias in Kenya along ethnic and gender lines, toward plaintiffs and defendants. Our data cover Kenyan Superior Court cases spanning 1976-2020, and our identification strategy relies on the random assignment of judges to cases. We have shown that judges in Kenya's Superior Courts exhibit in-group bias in terms of both gender and ethnicity. Specifically, defendants are about 4 percentage points more likely to win if they share the judge's gender and about 6 percentage points more likely

Table 1: Gender results

	(1)	(2)	(3)	(4)	(5)
	Def. win	Def. win	Def. win	Def. win	Def. win
Judge maj. female	-0.0381***	-0.0397**	-0.0469***	-0.0472***	-0.0412***
	(0.0122)	(0.0197)	(0.0133)	(0.0134)	(0.0133)
Pla. maj. female		-0.0297**	-0.0530***	-0.0536***	-0.0434***
		(0.0128)	(0.0109)	(0.0110)	(0.0108)
Def. maj. female	-0.00502		0.00123	0.000774	0.00650
	(0.0105)		(0.0108)	(0.0108)	(0.0107)
Judge maj. fem. X pla. maj. fem.		0.0174	0.00695	0.00781	0.00733
		(0.0208)	(0.0172)	(0.0171)	(0.0168)
Judge maj. fem. X def. maj. fem.	0.0359**		0.0404**	0.0405**	0.0383**
	(0.0166)		(0.0168)	(0.0168)	(0.0167)
Appeal					0.0882***
					(0.0138)
Number of defendants					0.00678**
					(0.00280)
Number of plaintiffs					0.00311
					(0.00316)
Number of judges					0.0665***
					(0.0145)
DV mean	0.452	0.429	0.454	0.454	0.454
Court-year FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	22787	25602	20383	20383	20383

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 1. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge maj. female is an indicator that equals 1 if the majority of the judge panel is female. Pla. maj. female is an indicator that equals 1 if the majority of the plaintiffs are female. Def. maj. female is an indicator that equals 1 if the majority of the defendants are female. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls or FE (such as case type) include a dummy that denotes if data is missing/unknown. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 2: Ethnicity results

	(1) Def. win	(2) Def. win	(3) Def. win	(4) Def. win	(5) Def. win
Judge-pla. same	0.00655		-0.0136	-0.00543	-0.00572
	(0.0127)		(0.0139)	(0.0153)	(0.0154)
Judge-def. same		0.0437***	0.0571***	0.0532***	0.0550***
		(0.0122)	(0.0144)	(0.0159)	(0.0158)
Appeal					0.0965***
					(0.0136)
Number of defendants					0.00208
					(0.00287)
Number of plaintiffs					0.00717**
					(0.00312)
Number of judges					0.0291*
					(0.0166)
DV mean	0.450	0.453	0.453	0.453	0.453
Court-year FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	21827	20966	18949	18949	18949

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 2. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge-pla. same and Judge-def. same are indicators that equal 1 if the judge and plaintiff of defendant, respectively, are the same plurality ethnicity. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for both defendants and plaintifs. To prevent a loss of observations, all categorical controls (such as case type)include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * p<0.1, ** p<0.05, *** p<0.01

Table 3: Moderating effect of damages (gender)

	(1)
	Def. win
Judge maj. female	-0.0121
	(0.0377)
Pla. maj. female	-0.0876**
	(0.0371)
Def. maj. female	0.0275
	(0.0373)
Damages	-0.0000592
	(0.000271)
Judge maj. fem. X Pla. maj. fem.	0.0112
	(0.0577)
Judge maj. fem. X Def. maj. fem.	-0.0156
	(0.0612)
Judge maj. fem. X Damages	-0.000464
	(0.000334)
Pla. maj. fem. X Damages	0.00205
	(0.00795)
Def. maj. fem. X Damages	-0.0181**
	(0.00721)
Judge maj. fem. X Pla. maj. fem. X Damages	-0.00109
	(0.00880)
Judge maj. fem. X Def. maj. fem. X Damages	0.0180**
	(0.00720)
DV mean	0.341
Court-year FE	Yes
Observations	2365

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 3. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge maj. female is an indicator that equals 1 if the majority of the judge panel is female. Pla. maj. female is an indicator that equals 1 if the majority of the plaintiffs are female. Def. maj. female is an indicator that equals 1 if the majority of the defendants are female. Damages is a variable for the amount of Kenyan Shillings (in millions) requested by the plaintiff. * p<0.1, ** p<0.05, *** p<0.01

Table 4: Moderating effect of damages (ethnicity)

	(1)
	Def. win
Judge-pla. same	-0.0261
	(0.0486)
Judge-def. same	0.129***
	(0.0451)
Damages	-0.000219
	(0.000399)
Judge-pla. same X Damages	-0.000233
	(0.00178)
Judge-def. same X Damages	-0.000505
	(0.000403)
DV mean	0.344
Court-year FE	Yes
Observations	2011

Standard errors, in parentheses, are clustered at the judge level. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge-pla. same and Judge-def. same are indicators that equal 1 if the judge and plaintiff of defendant, respectively, are the same plurality ethnicity. Damages is a variable for the amount of Kenyan Shillings (in millions) requested by the plaintiff. * p<0.1, ** p<0.05, *** p<0.01

to win if they share the judge's ethnicity. But there is no evidence of bias toward in-group plaintiffs, suggesting at least that judges are more biased in favor of in-group defendants, and potentially that they are not biased in favor of in-group plaintiffs at all. We propose that social identity theory, in combination with the concept of loss aversion, can be used to explain the uneven application of bias across judges' in-group members, and we provide evidence consistent with this proposition. However, the results also suggest that additional factors may be driving the uneven application of bias; understanding these additional factors is an important area for future work.

By integrating the paradigms of social identity and loss aversion, we aim to provide a more complete lens through which to decipher the cognitive underpinnings of judicial biases. This interdisciplinary approach—drawing on political science, behavioral economics, and social psychology—not only offers a synthesis of divergent research trajectories but also foregrounds a more nuanced blueprint for future interventions designed to mitigate inherent biases within the judicial system. As far as we know, this paper is the first to apply a framework that integrates social identity and loss aversion. This framework can be applied not only to the judiciary, but also to other contexts where individuals have the opportunity to favor one group member over another as a means to achieve greater overall group standing. As such, the paper contributes to the broader literature on the drivers of bias among civil servants (Miller, Kerr, and Reid 1999; Knowles, Persico, and Todd 2001; Plant, Goplen, and Kunstman 2011; Rehavi and Starr 2014; Knox, Lowe, and Mummolo 2020).

References

Ahrsjo, Ulrika, Susan Niknami, and Marten Palme (forthcoming). "Identity in Court Decision-Making". In: *American Economic Journal: Economic Policy.*

Arnold, David, Will Dobbie, and Crystal Yang (2018). "Racial Bias in Bail Decisions". In: *The Quarterly Journal of Economics* 133.4, pp. 1885–1932.

Asingo, Patrick et al. (2018). *Ethnicity and Politicization in Kenya.* Kenya Human Rights Commission.

Barberis, Nicholas C. (2013). "Thirty years of prospect theory in economics: A review and assessment". In: *Journal of Economic Perspectives* 27.1, pp. 173–196.

Benartzi, Shlomo and Richard Thaler (1995). "Myopic loss aversion and the equity premium puzzle". In: *The quarterly journal of Economics* 110.1, pp. 73–92.

Berge, Lars et al. (2015). "How strong are ethnic preferences?" In: *Working paper.*

Choi, Danny, Andy Harris, and Fiona Shen-Bayh (2021). "Ethnic Bias in Judicial Decision Making: Evidence from Criminal Appeals in Kenya". In: *American Political Science Review* 116.3, pp. 1067–1080.

Depew, Briggs, Ozkan Eren, and Naci Mocan (2017). "Judges, juveniles, and in-group bias". In: *The Journal of Law and Economics* 60.2, pp. 209–239.

Gainer, Maya (2015). "Transforming the Courts: Judicial Sector Reforms in Kenya". In: *Princeton University Innovations for Successful Societies* 1.7.

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". In: *Journal of Empirical Legal Studies* 7.3, pp. 403–428.

Glynn, Adam and Maya Sen (2015). "Identifying judicial empathy: does having daughters cause judges to rule for women's issues?" In: *American Journal of Political Science* 59.1, pp. 37–54.

Grossman, Guy et al. (2016). "Descriptive Representation and Judicial Outcomes in Multi-ethnic Societies". In: *American Journal of Political Science* 60.1, pp. 44–69.

Hjort, Jonas (2014). “Ethnic Divisions and Production in Firms”. In: *The Quarterly Journal of Economics* 129.4, pp. 1899–1946.

Hochreiter, Sepp and Jurgen Schmidhuber (1997). “Long short-term memory”. In: *Neural computation* 9.8, pp. 1735–1780.

Kahneman, Daniel and Amos Tversky (1979). “Prospect Theory: An Analysis of Decision under Risk”. In: *Econometrica* 47.2, pp. 263–292.

Kastellec, Jonathan (2013). “Racial diversity and judicial influence on appellate courts”. In: *American Journal of Political Science* 56.1, pp. 167–183.

Kenyan Judiciary (2021). *Courts: Overview.* URL: `https://www.judiciary.go.ke/courts/`.

Knepper, Matthew (2018). “When the shadow is the substance: Judge gender and the outcomes of workplace sex discrimination cases”. In: *Journal of Labor Economics* 36.3, pp. 623–664.

Knowles, John, Nicola Persico, and Petra Todd (2001). “Racial bias in motor vehicle searches: Theory and evidence”. In: *Journal of Political Economy* 109.1, pp. 203–229.

Knox, Dean, Will Lowe, and Jonathan Mummolo (2020). “Administrative records mask racially biased policing”. In: *American Political Science Review* 114.3, pp. 619–637.

Miller, Will, Brinck Kerr, and Margaret Reid (1999). “A national study of gender-based occupational segregation in municipal bureaucracies: Persistence of glass walls?” In: *Public Administration Review*, pp. 218–230.

Pennington, Jeffrey, Richard Socher, and Christopher Manning (2014). “Glove: Global vectors for word representation”. In: *Proceedings of the 2014 conference on empirical methods in natural language processing.*

Plant, E. Ashby, Joanna Goplen, and Jonathan W. Kunstman (2011). “Selective responses to threat: The roles of race and gender in decisions to shoot”. In: *Personality and Social Psychology Bulletin* 37.9, pp. 1274–1281.

Rehavi, Marit and Sonja B. Starr (2014). "Racial disparity in federal criminal sentences". In: *Journal of Political Economy* 122.6, pp. 1320–1354.

Shayo, Moses and Asaf Zussman (2011). "Judicial ingroup bias in the shadow of terrorism". In: *The Quarterly Journal of Economics* 126.3, pp. 1447–1484.

Sloan, CarlyWill (2020). "Racial bias by prosecutors: Evidence from random assignment". In: *Working paper*.

Spirling, Arthur and Pedro Rodriguez (2019). "Word embeddings: What works, what doesn't, and how to tell the difference for applied research". In: *Journal of Politics*.

World Bank (2017). *World Development Report: Governance and the Law.* The World Bank Group.

Yang, Crystal (2015). "Free at Last? Judicial Discretion and Racial Disparities in Federal Sentencing". In: *The Journal of Legal Studies* 44.1, pp. 75–111.

Competing interests

Competing interests: The authors declare none

Appendix

Appendix A: Descriptive statistics

Figure A1 indicates the number of cases in our dataset from each year over time. Figures A2 and A3 indicate the women and certain ethnicities are underrepresented in the judiciary, which has implications for the impact of in-group bias.

Figure A1: Frequency of cases in the dataset over time

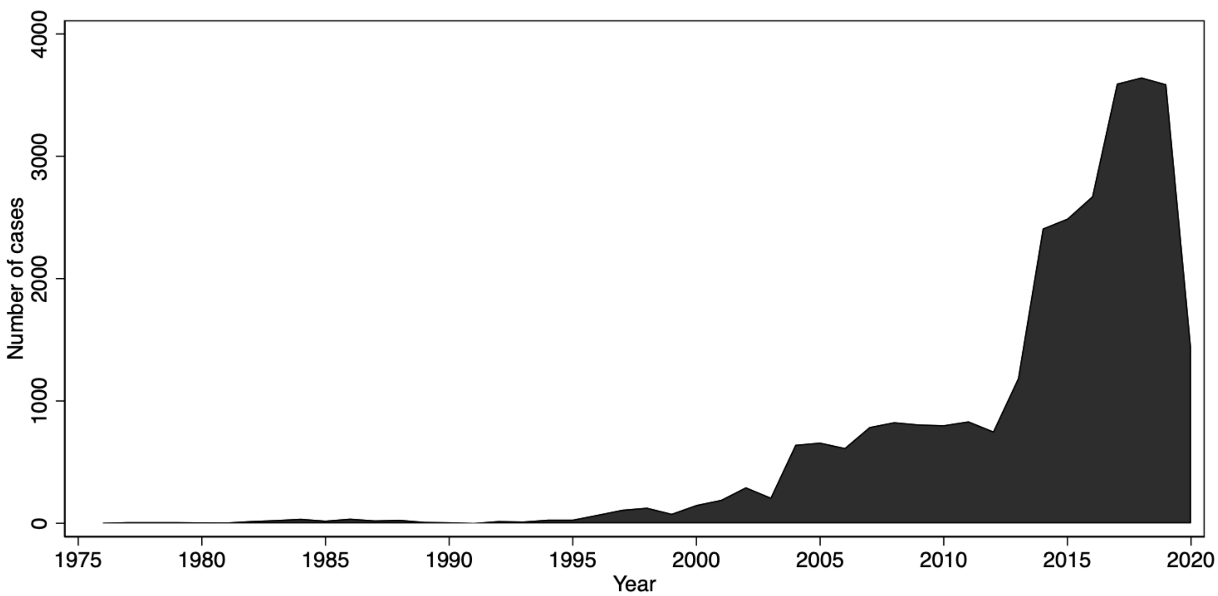


Figure A2: Total number of cases, by majority gender and role in the case

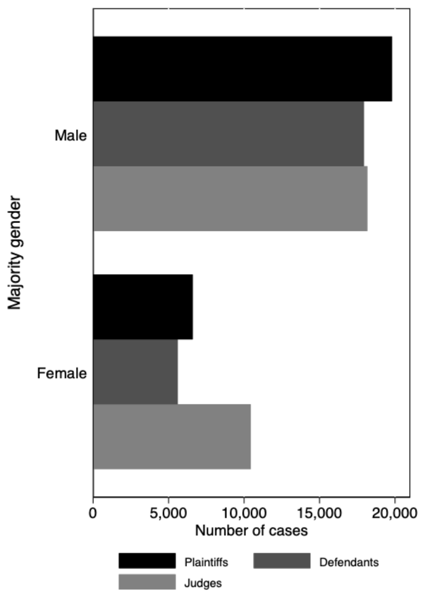
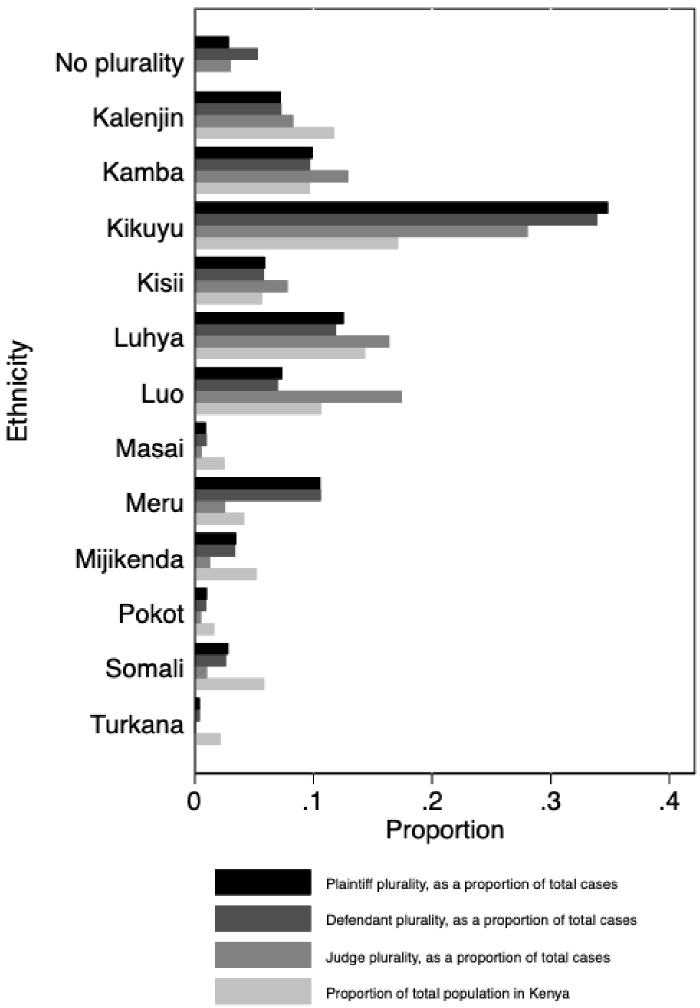


Figure A3: Ethnicities as a proportion of total cases (by role in the case) and the total population in Kenya



Proportions of the total population are derived from the 2019 census. Kalenjin refers to non-Pokot Kalenjin.

Appendix B: Variable construction

Constructing variables with judge, defendant, and plaintiff information

The names of judges, defendants, and plaintiffs were used to remove non-humans and to extract additional information for each case, including gender, ethnicity, and the number of judges and litigants. Cases were identified as non-human and removed if either the plaintiff or defendant name included any of a long list of key words, such as "republic," "company," or "medical." A full list of the keywords can be found in the cleaning scripts in the replication materials posted online.

Afterwards, we could determine the gender of each individual using their first name and the ethnicity of each individual using their last name. To assign gender based on first names we used the genderize.io API and Gender API, both of which use global databases of names and genders to probabilistically assign gender to names.[8] One exception was for the judges, for whom gender was assigned manually.

To assign ethnicity based on last names, we used data available on Harvard Dataverse that links names to ethnicities (Hjort 2014). This data could be used to identify 12 ethnic groups (Meru, Kisii, Kalenjin, Kamba, Luo, Turkana, Mijikenda, Luhya, Kikuyu, Somali, Masai, and Pokot). This includes one ethnic sub-group, the Pokot, which is a sub-group of the Kalenjin. Throughout our analysis, Kalenjin refers to non-Pokot Kalenjin. Together, these groups account for about 91 percent of the population of Kenya. Of the other 29 major ethnic groups (i.e., non-subgroups) identified in the 2019 census, the largest group accounts for only about 0.9 percent of the population.

Gender and ethnicity could not be determined for all individuals in all cases. Gender could not be determined if the first name was either abbreviated (i.e., if only initials were given), it did not clearly match to a single gender, or it was not included in the API datasets. Ethnicity could not be determined if the last name was not included in the ethnicity dataset. For some of the cases included in analysis, information could be extracted for plaintiffs but

[8] See the following websites: https://genderize.io/; https://gender-api.com/.

not defendants (and vice versa) and for gender but not ethnicity (and vice versa).

It is important to note that there is the possibility of a small amount of error resulting from the automated process of removing non-humans and determining gender and ethnicity. For example, although the list of key words for non-humans is long and we have manually scanned the data for non-humans, it is still possible that some non-humans remain. It is also possible that gender and/or ethnicity has been assigned to non-humans with certain key words included in the organization name. Similarly, if names were separated in an unusual way, it is possible that the number of defendants or plaintiffs was incorrectly counted, possibly resulting in an incorrect assignment of majority/plurality gender/ethnicity. However, having thoroughly scanned the data, we are confident that the number of such errors is insignificant.

Using the Binary Classification Machine Learning Model to construct the defendant_win outcome variable

To determine the winner of each case, we created a Binary Classification Machine Learning Model using the Global Vectors for Word Representation (GloVe) algorithm (Pennington, Socher, and Manning 2014). The objective function of GloVe can be written as follows:

$$J(w) = \sum f(X_{ij})(w_i^t w_j - log X_{ij}))^2 \tag{4}$$

where X_{ij} denotes the co-occurrence count between words i and j, and $f(\cdot)$ is a weighting function that serves to down-weight particularly frequent words. The objective function $J(\cdot)$ trains the word vectors to minimize the squared difference between the dot product of the vectors representing two words and their empirical co-occurrence in the corpus. The algorithm requires two hyperparameters, *dimensionality* of the vectors and the *window* size for computing co-occurrence statistics. Prior research has found 300 to be the optimum size in many a cases and that increasing dimensionality beyond 300 has negligible improvements for downstream tasks (Pennington, Socher, and Manning 2014; Spirling and Rodriguez 2019).

Table B1: Model outcomes

Training set accuracy	92.44%
Validation set accuracy	91.92%
Test set accuracy (on previously unseen data)	92.83%
Accuracy	0.928388
Precision	0.896705
Recall	0.959647
F1 score	0.927109

Following that literature, we train 300 dimensional vectors. We used a standard 10-word window size, in between a shorter window size (which tends to capture syntactic/functional relations between words) and a longer window size (which tends to capture topical relations between words). To improve accuracy, the classification model was also comprised of a Long Short-Term Memory layer in addition to the fully connected neural network layers and the initial embedding layer (Hochreiter and Schmidhuber 1997).

Applying this model to our data, we used the bottom 500 words of the case judgments, since the outcomes were found to be present towards the bottom of the judgments. As a training dataset, we applied the model to cases for which we could determine the outcome (in favor or against the defendant) directly from the case outcome variable of the meta-data. There were 49,706, 6,214, and 6,213 cases in the training, testing, and validation sets, respectively. The results of the model are presented in table B1.

Appendix C: Balance tests

We use the following balance test for the analysis sample, for case i filed in court c at time t:

$$
\begin{gathered}
judge\_maj\_female_{i,c,t} = \beta_1 def\_maj\_female_{i,c,t} + \\
\beta_2 pla\_maj\_female_{i,c,t} + \Phi_{c,t} + \lambda' X_{i,c,t,} + \epsilon_{i,c,t}
\end{gathered}
\tag{5}
$$

The results below indicate that male- and female-majority defendant groups are equally likely to be assigned male- and female-majority judge panels. Likewise, male- and female-majority plaintiff groups are equally likely to be assigned male- and female-majority judge panels. Tables C2 and C3 present balance tests for pre-2011 and since 2011, respectively. The results are consistent with Table C1.

To confirm that judge assignment to cases is random in terms of ethnic majority, we use variations of the following balance test:

$$
\begin{gathered}
judge\_plur\_kikuyu_{i,c,t} = \beta_1 def\_plur\_kikuyu_{i,c,t} + \\
\beta_2 pla\_plur\_kikuyu_{i,c,t} + \Phi_{c,t} + \lambda' X_{i,c,t,} + \epsilon_{i,c,t}
\end{gathered}
\tag{6}
$$

where$judge_plur_kikuyu_{i,c,t}$ is a binary variable indicating whether the judge plurality is the Kikuyu ethnic group, $def\_plur\_kikuyu_{i,c,t}$ is a binary variable indicating whether the defendant plurality is the Kikuyu ethnic group, and $pla\_plur\_kikuyu_{i,c,t}$ is a binary variable indicating whether the plaintiff plurality is the Kikuyu ethnic group. We run a series of 12 tests, with each test using binary variables for different ethnicities.

Tables C4 through C7 below report the results of the tests. They show that defendants and plaintiffs across all ethnicities are not more likely to be assigned judges from their ethnic group. One exception is Luhya defendants, as table C5 shows. Balance tests for both pre-2011 and since 2011 are also presented in Tables C8 through C15. They show that there are significant coefficients for Luhya defendants in the 2011-2020 period and Kamba for the 1976-2010 period. We conduct a robustness check of the main analysis that drops all Luhya and Kamba individuals. Appendix D presents these results. A comparison between these results and the main results below show that the in-group bias we observe is not driven by any possible bias in Luhya or Kamba case assignment.

Table C1: Gender randomization checks

	(1)	(2)	(3)
	Judge maj. female	Judge maj. female	Judge maj. female
Pla. maj. female	0.0115	0.0117	0.00747
	(0.00861)	(0.00859)	(0.00678)
Def. maj. female	0.00350	0.00335	-0.000434
	(0.00763)	(0.00755)	(0.00650)
Appeal			-0.00439
			(0.0158)
Number of defendants			0.000381
			(0.00158)
Number of plaintiffs			-0.00268
			(0.00201)
Number of judges			-0.0301
			(0.0281)
DV mean	0.363	0.363	0.363
Court-year FE	Yes	Yes	Yes
Ethnicity FE	No	Yes	Yes
Case type FE	No	No	Yes
Observations	20383	20383	20383

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 5. Judge maj. female is an indicator that equals 1 if the majority of the judge panel is female. Pla. maj. female is an indicator that equals 1 if the majority of the plaintiffs are female. Def. maj. female is an indicator that equals 1 if the majority of the defendants are female. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls or FE (such as case type) include a dummy that denotes if data is missing/unknown. * p<0.1, ** p<0.05, *** p<0.01

Table C2: Gender randomization checks, before 2011

	(1) Judge maj. female	(2) Judge maj. female	(3) Judge maj. female
Pla. maj. female	0.0224	0.0213	0.0118
	(0.0189)	(0.0188)	(0.0153)
Def. maj. female	-0.00693	-0.00616	-0.0114
	(0.0137)	(0.0133)	(0.0111)
Appeal			-0.00852
			(0.0379)
Number of defendants			0.00363
			(0.00425)
Number of plaintiffs			0.00278
			(0.00665)
Number of judges			-0.0305*
			(0.0184)
DV mean	0.299	0.299	0.299
Court-year FE	Yes	Yes	Yes
Ethnicity FE	No	Yes	Yes
Case type FE	No	No	Yes
Observations	4706	4706	4706

Sample is restricted to the years 1976-2010. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 5. Judge maj. female is an indicator that equals 1 if the majority of the judge panel is female. Pla. maj. female is an indicator that equals 1 if the majority of the plaintiffs are female. Def. maj. female is an indicator that equals 1 if the majority of the defendants are female. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls or FE (such as case type) include a dummy that denotes if data is missing/unknown. * p<0.1, ** p<0.05, *** p<0.01

Table C3: Gender randomization checks, 2011 and after

	(1) Judge maj. female	(2) Judge maj. female	(3) Judge maj. female
Pla. maj. female	0.00827	0.00850	0.00514
	(0.00968)	(0.00951)	(0.00735)
Def. maj. female	0.00674	0.00617	0.00325
	(0.00845)	(0.00832)	(0.00719)
Appeal			-0.00317
			(0.0190)
Number of defendants			0.0000425
			(0.00167)
Number of plaintiffs			-0.00360*
			(0.00216)
Number of judges			-0.0320
			(0.0511)
DV mean	0.382	0.382	0.382
Court-year FE	Yes	Yes	Yes
Ethnicity FE	No	Yes	Yes
Case type FE	No	No	Yes
Observations	15677	15677	15677

Sample is restricted to the years 2011-2020. Sample is restricted to the years 1976-2010. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 5. Judge maj. female is an indicator that equals 1 if the majority of the judge panel is female. Pla. maj. female is an indicator that equals 1 if the majority of the plaintiffs are female. Def. maj. female is an indicator that equals 1 if the majority of the defendants are female. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls or FE (such as case type) include a dummy that denotes if data is missing/unknown. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table C4: Ethnicity randomization checks 1

	(1) Judge plur. Kalenjin	(2) Judge plur. Kamba	(3) Judge plur. Kikuyu
Pla. plur. Kalenjin	0.00699		
	(0.00894)		
Def. plur. Kalenjin	-0.00933		
	(0.0103)		
Pla. plur. Kamba		-0.00785	
		(0.00651)	
Def. plur. Kamba		0.00337	
		(0.00771)	
Pla. plur. Kikuyu			0.00641
			(0.00760)
Def. plur. Kikuyu			0.00166
			(0.00759)
DV mean	0.0895	0.125	0.275
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	14592	14592	14942

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table C5: Ethnicity randomization checks 2

	(1) Judge plur. Kisii	(2) Judge plur. Luhya	(3) Judge plur. Luo
Pla. plur. Kisii	-0.0111		
	(0.00753)		
Def. plur. Kisii	0.00201		
	(0.00795)		
Pla. plur. Luhya		0.00489	
		(0.00674)	
Def. plur. Luhya		0.0112*	
		(0.00589)	
Pla. plur. Luo			0.000716
			(0.00907)
Def. plur. Luo			0.00195
			(0.0104)
DV mean	0.0765	0.174	0.173
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	14942	14942	14942

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table C6: Ethnicity randomization checks 3

	(1) Judge plur. Masai	(2) Judge plur. Meru	(3) Judge plur. Mijikenda
Pla. plur. Masai	-0.000333		
	(0.00218)		
Def. plur. Masai	0.000336		
	(0.000782)		
Pla. plur. Meru		0.00203	
		(0.00313)	
Def. plur. Meru		-0.000890	
		(0.00268)	
Pla. plur. Mijikenda			-0.000673
			(0.00299)
Def. maj. Mijikenda			0.00369
			(0.00408)
DV mean	0.00589	0.0260	0.0110
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	14942	14942	14942

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant.
* p<0.1, ** p<0.05, *** p<0.01

Table C7: Ethnicity randomization checks 4

	(1) Judge plur. Pokot	(2) Judge plur. Somali	(3) Judge plur. Turkana
Pla. plur. Pokot	-0.00444		
	(0.00458)		
Def. plur. Pokot	0.00327		
	(0.00832)		
Pla. plur. Somali		0.00474	
		(0.00516)	
Def. plur. Somali		-0.00715	
		(0.00518)	
Pla. plur. Turkana			0.0000973
			(0.000118)
Def. plur. Turkana			-0.000940
			(0.00102)
DV mean	0.00482	0.00964	0.000468
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	14942	14942	14942

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant.
* p<0.1, ** p<0.05, *** p<0.01

Table C8: Ethnicity randomization checks, before 2011, 1

	(1) Judge plur. Kalenjin	(2) Judge plur. Kamba	(3) Judge plur. Kikuyu
Pla. plur. Kalenjin	0.000864		
	(0.00136)		
Def. plur. Kalenjin	-0.00108		
	(0.00133)		
Pla. plur. Kamba		-0.0483***	
		(0.0181)	
Def. plur. Kamba		0.0687***	
		(0.0233)	
Pla. plur. Kikuyu			0.00687
			(0.0109)
Def. plur. Kikuyu			-0.00966
			(0.0101)
DV mean	0.0438	0.128	0.169
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	3104	3104	3104

Sample is restrictd to the years 1976-2010. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * p<0.1, ** p<0.05, *** p<0.01

Table C9: Ethnicity randomization checks, before 2011, 2

	(1) Judge plur. Kisii	(2) Judge plur. Luhya	(3) Judge plur. Luo
Pla. plur. Kisii	-0.0345**		
	(0.0164)		
Def. plur. Kisii	-0.0140		
	(0.0159)		
Pla. plur. Luhya		0.0141	
		(0.0285)	
Def. plur. Luhya		-0.00398	
		(0.0188)	
Pla. plur. Luo			0.0197
			(0.0268)
Def. plur. Luo			-0.00101
			(0.0268)
DV mean	0.0802	0.216	0.221
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	3104	3104	3104

Sample is restrictd to the years 1976-2010. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table C10: Ethnicity randomization checks, before 2011, 3

	(1) Judge plur. Masai	(2) Judge plur. Meru	(3) Judge plur. Mijikenda
Pla. plur. Masai	0.000169		
	(0.000474)		
Def. plur. Masai	0.000227		
	(0.000414)		
Pla. plur. Meru		0.00370	
		(0.00942)	
Def. plur. Meru		-0.0123	
		(0.00907)	
Pla. plur. Mijikenda			-0.0129
			(0.0129)
Def. maj. Mijikenda			0.000791
			(0.00325)
DV mean	0.000966	0.0599	0.0168
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	3104	3104	3104

Sample is restrictd to the years 1976-2010. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table C11: Ethnicity randomization checks, before 2011, 4

	(1) Judge plur. Pokot	(2) Judge plur. Somali	(3) Judge plur. Turkana
Pla. plur. Pokot	-0.0102		
	(0.00724)		
Def. plur. Pokot	0.0236		
	(0.0701)		
Pla. plur. Somali		0.0142	
		(0.0211)	
Def. plur. Somali		-0.0308	
		(0.0226)	
Pla. plur. Turkana			0
			(.)
Def. plur. Turkana			0
			(.)
DV mean	0.0110	0.0174	0
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	3104	3104	3104

Sample is restrictd to the years 1976-2010. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table C12: Ethnicity randomization checks, 2011 and after, 1

	(1)	(2)	(3)
	Judge plur. Kalenjin	Judge plur. Kamba	Judge plur. Kikuyu
Pla. plur. Kalenjin	0.00864		
	(0.0108)		
Def. plur. Kalenjin	-0.0113		
	(0.0123)		
Pla. plur. Kamba		-0.000521	
		(0.00806)	
Def. plur. Kamba		-0.00935	
		(0.00840)	
Pla. plur. Kikuyu			0.00681
			(0.00987)
Def. plur. Kikuyu			0.00469
			(0.00970)
DV mean	0.100	0.125	0.303
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	11838	11838	11838

Sample is restrictd to the years 2011-2020. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * p<0.1, ** p<0.05, *** p<0.01

Table C13: Ethnicity randomization checks, 2011 and after, 2

	(1) Judge plur. Kisii	(2) Judge plur. Luhya	(3) Judge plur. Luo
Pla. plur. Kisii	-0.00434		
	(0.00743)		
Def. plur. Kisii	0.00863		
	(0.00983)		
Pla. plur. Luhya		0.00416	
		(0.00644)	
Def. plur. Luhya		0.0143**	
		(0.00607)	
Pla. plur. Luo			-0.00583
			(0.00963)
Def. plur. Luo			0.00269
			(0.0106)
DV mean	0.0755	0.163	0.160
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	11838	11838	11838

Sample is restrictd to the years 2011-2020. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * p<0.1, ** p<0.05, *** p<0.01

Table C14: Ethnicity randomization checks, 2011 and after, 3

	(1)	(2)	(3)
	Judge plur. Masai	Judge plur. Meru	Judge plur. Mijikenda
Pla. plur. Masai	-0.000695		
	(0.00261)		
Def. plur. Masai	0.000584		
	(0.000898)		
Pla. plur. Meru		0.00183	
		(0.00265)	
Def. plur. Meru		0.00195	
		(0.00287)	
Pla. plur. Mijikenda			0.00181
			(0.00255)
Def. maj. Mijikenda			0.00498
			(0.00480)
DV mean	0.00718	0.0171	0.00946
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	11838	11838	11838

Sample is restrictd to the years 2011-2020. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * p<0.1, ** p<0.05, *** p<0.01

Table C15: Ethnicity randomization checks, 2011 and after, 4

	(1) Judge plur. Pokot	(2) Judge plur. Somali	(3) Judge plur. Turkana
Pla. plur. Pokot	-0.00227		
	(0.00281)		
Def. plur. Pokot	-0.000560		
	(0.00171)		
Pla. plur. Somali		0.00220	
		(0.00364)	
Def. plur. Somali		-0.00161	
		(0.00244)	
Pla. plur. Turkana			0.000157
			(0.000191)
Def. plur. Turkana			-0.00111
			(0.00120)
DV mean	0.00321	0.00760	0.000591
Court-year FE	Yes	Yes	Yes
Other controls and FE	Yes	Yes	Yes
Observations	11838	11838	11838

Sample is restrictd to the years 2011-2020. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 6. Other controls and FE include case type dummies, a dummy for an appeal case, and variables for the numbers of defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls (such as case type) include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. * p<0.1, ** p<0.05, *** p<0.01

Appendix D: Main regression outputs and robustness checks

Table D1: Gender results, 2010 and after

	(1)	(2)	(3)	(4)	(5)
	Def. win	Def. win	Def. win	Def. win	Def. win
Judge maj. female	-0.0381***	-0.0397**	-0.0469***	-0.0472***	-0.0412***
	(0.0122)	(0.0197)	(0.0133)	(0.0134)	(0.0133)
Pla. maj. female		-0.0297**	-0.0530***	-0.0536***	-0.0434***
		(0.0128)	(0.0109)	(0.0110)	(0.0108)
Def. maj. female	-0.00502		0.00123	0.000774	0.00650
	(0.0105)		(0.0108)	(0.0108)	(0.0107)
Judge maj. fem. X pla. maj. fem.		0.0174	0.00695	0.00781	0.00733
		(0.0208)	(0.0172)	(0.0171)	(0.0168)
Judge maj. fem. X def. maj. fem.	0.0359**		0.0404**	0.0405**	0.0383**
	(0.0166)		(0.0168)	(0.0168)	(0.0167)
Appeal					0.0882***
					(0.0138)
Number of defendants					0.00678**
					(0.00280)
Number of plaintiffs					0.00311
					(0.00316)
Number of judges					0.0665***
					(0.0145)
DV mean	0.452	0.429	0.454	0.454	0.454
Court-year FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	22787	25602	20383	20383	20383

Years before 2010 are dropped. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 1. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge maj. female is an indicator that equals 1 if the majority of the judge panel is female. Pla. maj. female is an indicator that equals 1 if the majority of the plaintiffs are female. Def. maj. female is an indicator that equals 1 if the majority of the defendants are female. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls or FE (such as case type) include a dummy that denotes if data is missing/unknown. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table D2: Gender results, robust standard errors

	(1)	(2)	(3)	(4)	(5)
	Def. win	Def. win	Def. win	Def. win	Def. win
Judge maj. female	-0.0381***	-0.0397***	-0.0469***	-0.0472***	-0.0412***
	(0.0106)	(0.00956)	(0.0122)	(0.0123)	(0.0122)
Pla. maj. female		-0.0297***	-0.0530***	-0.0536***	-0.0434***
		(0.00918)	(0.0102)	(0.0102)	(0.0102)
Def. maj. female	-0.00502		0.00123	0.000774	0.00650
	(0.00986)		(0.0105)	(0.0105)	(0.0105)
Judge maj. fem. X pla. maj. fem.		0.0174	0.00695	0.00781	0.00733
		(0.0148)	(0.0164)	(0.0164)	(0.0164)
Judge maj. fem. X def. maj. fem.	0.0359**		0.0404**	0.0405**	0.0383**
	(0.0163)		(0.0172)	(0.0172)	(0.0171)
Appeal					0.0882***
					(0.0106)
Number of defendants					0.00678***
					(0.00250)
Number of plaintiffs					0.00311
					(0.00325)
Number of judges					0.0665***
					(0.0145)
DV mean	0.452	0.429	0.454	0.454	0.454
Court-year FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	22787	25602	20383	20383	20383

Includes robust standard errors. For specification details, see equation 1. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge maj. female is an indicator that equals 1 if the majority of the judge panel is female. Pla. maj. female is an indicator that equals 1 if the majority of the plaintiffs are female. Def. maj. female is an indicator that equals 1 if the majority of the defendants are female. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for defendants, plaintiffs, and judges. To prevent a loss of observations, all categorical controls or FE (such as case type) include a dummy that denotes if data is missing/unknown. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table D3: Ethnicity results, no Kamba or Luhya

	(1)	(2)	(3)	(4)	(5)
	Def. win	Def. win	Def. win	Def. win	Def. win
Judge-pla. same	0.00982		-0.00482	0.00180	0.00240
	(0.0154)		(0.0170)	(0.0207)	(0.0207)
Judge-def. same		0.0443***	0.0526***	0.0479**	0.0500**
		(0.0145)	(0.0179)	(0.0235)	(0.0235)
Appeal					0.108***
					(0.0171)
Number of defendants					0.00186
					(0.00353)
Number of plaintiffs					0.00993**
					(0.00391)
Number of judges					0.0422**
					(0.0203)
DV mean	0.450	0.451	0.452	0.452	0.452
Court-year FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	11360	10983	9739	9739	9739

Kamba and Luhya are dropped due to some evidence of lack of balance for these ethnicities in the balance tests. Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 2. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge-pla. same and Judge-def. same are indicators that equal 1 if the judge and plaintiff of defendant, respectively, are the same plurality ethnicity. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for both defendants and plaintifs. To prevent a loss of observations, all categorical controls (such as case type)include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. Cases where the judge or either litigant is majority Luhya or Kamba are dropped. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table D4: Ethnicity results, 2010 and after

	(1)	(2)	(3)	(4)	(5)
	Def. win	Def. win	Def. win	Def. win	Def. win
Judge-pla. same	0.00585		-0.0143	-0.00466	-0.00424
	(0.0142)		(0.0153)	(0.0169)	(0.0170)
Judge-def. same		0.0456***	0.0589***	0.0533***	0.0561***
		(0.0137)	(0.0157)	(0.0178)	(0.0176)
Appeal					0.102***
					(0.0150)
Number of defendants					0.000901
					(0.00280)
Number of plaintiffs					0.00606*
					(0.00325)
Number of judges					0.0461**
					(0.0214)
DV mean	0.447	0.452	0.451	0.451	0.451
Court-year FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	17877	17194	15555	15555	15555

Standard errors, in parentheses, are clustered at the judge level. For specification details, see equation 2. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge-pla. same and Judge-def. same are indicators that equal 1 if the judge and plaintiff of defendant, respectively, are the same plurality ethnicity. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for both defendants and plaintifs. To prevent a loss of observations, all categorical controls (such as case type)include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. Years before 2010 are dropped. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table D5: Ethnicity results, robust standard errors

	(1) Def. win	(2) Def. win	(3) Def. win	(4) Def. win	(5) Def. win
Judge-pla. same	0.00655		-0.0136	-0.00543	-0.00572
	(0.0115)		(0.0142)	(0.0151)	(0.0151)
Judge-def. same		0.0437***	0.0571***	0.0532***	0.0550***
		(0.0120)	(0.0145)	(0.0155)	(0.0155)
Appeal					0.0965***
					(0.0111)
Number of defendants					0.00208
					(0.00261)
Number of plaintiffs					0.00717**
					(0.00320)
Number of judges					0.0291*
					(0.0166)
DV mean	0.450	0.453	0.453	0.453	0.453
Court-year FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	21827	20966	18949	18949	18949

Includes robust standard errors. For specification details, see equation 2. The dependent variable is an indicator that equals 1 if the defendant won the case. Judge-pla. same and Judge-def. same are indicators that equal 1 if the judge and plaintiff of defendant, respectively, are the same plurality ethnicity. Ethnicity fixed effects (FE) include binary variables indicating whether a given ethnicity is the plurality, one for each ethnicity, for both defendants and plaintifs. To prevent a loss of observations, all categorical controls (such as case type)include a dummy that denotes if data is missing/unknown. Pla. = plaintiff, def. = defendant. Cases where the judge or either litigant is majority Luhya or Kamba are dropped. * p<0.1, ** p<0.05, *** p<0.01.