International Review of Law & Economics 87 (2026) 106344

Contents lists available at ScienceDirect

International Review of Law & Economics

journal homepage: www.elsevier.com/locate/irle

The cognitive underpinnings of judicial bias: The role of social identity and prospect theory

Daniel L. Chen [a,*], Jimmy Graham [b], Manuel Ramos-Maqueda [c], Shashank Singh [d]

[a] Toulouse School of Economics, France
[b] NYU, United States
[c] University of Oxford, United Kingdom
[d] University of California, Berkeley, United States

ARTICLE INFO

Keywords:
In-group favoritism
Loss aversion
Behavioral economics of law

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 is consistent with a framework combining social identity and loss aversion, and we present suggestive evidence from data on the amount of damages in each case. We argue that this framework offers one plausible lens through which to interpret the cognitive underpinnings of judicial bias, and that further work is needed to establish its general applicability.

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 offer a framework that is consistent with the asymmetric pattern we document. 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 et al., 2017; Arnold et al., 2018; Knepper, 2018; Sloan, 2020; Choi et al., 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

* Corresponding author.
E-mail address: daniel.chen@iast.fr (D.L. Chen).

https://doi.org/10.1016/j.irle.2026.106344
Received 12 March 2025; Received in revised form 24 April 2026; Accepted 14 June 2026
Available online 2 July 2026

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 toward 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.

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.

We emphasize two scope conditions. First, the mechanism concerns the decision frame at the moment of adjudication: conditional on liability being contested, the ruling determines whether the court authorizes a contemporaneous transfer from defendant to plaintiff, and under a narrow decision frame the outflow from an in-group defendant may loom larger than the offsetting inflow to an in-group plaintiff. Second, we do not claim that identity is uniformly salient across all case configurations, and in particular recognize that salience may be weaker in cases where all parties share the same identity.

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 Labor 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 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.

[1] For examples see Shayo and Zussman (2011), Gazal-Ayal and Sulitzeanu-Kenan (2010), Grossman et al. (2016), Knepper (2018), Sloan (2020), Choi et al. (2021) and Ahrsjö et al. (2024).

[2] See Appendix A, Figures A2 and A3, for evidence.

[3] See http://kenyalaw.org/caselaw/.

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.[5] 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.[6]

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 data.[7] 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 Appendix A).[8]

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 3420 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.

Under Kenyan civil procedure, plaintiffs must specify their requested relief at the time of filing the suit. The Civil Procedure Rules explicitly require that "every plaint shall state specifically the relief which the plaintiff claims" (Civil Procedure Rules, Order 2 Rule 6). Where a monetary claim is involved, the plaintiff must state the precise amount claimed, except in limited situations where the amount cannot be ascertained upfront (e.g., mesne profits or accounts). In practice, this means that damages or other remedies are disclosed at the initial filing stage, before a case is registered and assigned to a judge. This institutional feature reduces concerns that damage amounts might be strategically adjusted based on the judge's identity, since the prayer for relief forms part of the plaint lodged with the registry prior to judicial allocation.

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, labor 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 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 toward 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.[9] To estimate judicial ethnic bias, we conduct the following regression:

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

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

[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] For gender, we are able to use a "majority" category because there are only two possible values (male/female), making the concept of majority meaningful and straightforward to implement. However, for ethnicity, there are many categories (e.g., 12 or 16 different ethnicities in our data), so defining a "majority" is not meaningful or feasible.

[6] 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).

[7] Of the initial 159,645 cases, 33,876 had exclusively human litigants for civil cases. An additional 4513 cases were dropped because we were unable to determine majority gender or ethnicity for the litigants in the case.

[8] The large rise in cases beginning in 2012 coincides with reforms implemented after Kenya's 2010 Constitution. The judiciary expanded its footprint by appointing new judges and creating specialized courts, while also digitizing case management and publishing a backlog of judgments through Kenya Law, leading to both greater access to courts and a surge in recorded cases (Judiciary of Kenya, 2012/13; Kenya Law, 2012).

[9] Our empirical design isolates co-ethnic effects rather than broader coalition-level dynamics. This strengthens the interpretation that the observed bias reflects shared ethnicity per se rather than alignment with a larger political coalition or the presence of inter-ethnic conflict (cf. Hjort, 2014, who shows robustness to both coalition-based and tribe-specific codings).

Table 1
Gender results.

	(1) Def. win	(2) Def. win	(3) Def. win	(4) Def. win	(5) 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. × pla. maj. fem.		0.0174	0.00695	0.00781	0.00733
		(0.0208)	(0.0172)	(0.0171)	(0.0168)
Judge maj. fem. × 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	22 787	25 602	20 383	20 383	20 383

Standard errors, in parentheses, are clustered at the judge level. For specification details, see Eq. (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.

$judge\_def\_same_{i,c,t}$ 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:

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

where $damages$ is a variable for the amount of Kenyan Shillings (in millions) requested by the plaintiff.[10] 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.

[10] 1 USD equals about 130 Kenyan Shillings.

5. Results

Table 1 presents the gender results, building up the specification across five columns. Column (1) includes only the defendant-side indicators (Judge maj. female, Def. maj. female, and their interaction), and is the most parsimonious specification of gender in-group bias toward defendants. Column (2) is the analogous specification on the plaintiff side, including only Judge maj. female, Pla. maj. female, and their interaction. Column (3) enters both the plaintiff- and defendant-side indicators and interactions simultaneously, allowing us to compare in-group bias toward the two sides within the same regression. Columns (4) and (5) progressively add fixed effects and case-level controls: column (4) augments column (3) with ethnicity fixed effects for judges, defendants, and plaintiffs, which absorb group-level differences in baseline win rates; column (5) additionally includes case-type fixed effects and controls for whether the case is an appeal and for the numbers of judges, plaintiffs, and defendants.

Several patterns are worth noting as we move across columns. First, the coefficient on Judge maj. female × Def. maj. female — our primary estimate of gender in-group bias toward defendants — is stable across specifications, ranging from 0.036 in column (1) to 0.040–0.041 in columns (3)–(4) and 0.038 in the most saturated specification in column (5). This stability indicates that the in-group estimate is not sensitive to the inclusion of plaintiff-side controls, ethnicity fixed effects, case-type fixed effects, or additional case-level covariates. Second, the plaintiff-side interaction Judge maj. female × Pla. maj. female is small (0.007–0.017) and statistically indistinguishable from zero across all specifications in which it appears, in contrast to the defendant-side interaction. Third, the level terms on Judge maj. female and Pla. maj. female are negative and significant, which is a function of the omitted reference category (all-male configurations) rather than a direct estimate of in-group bias; the quantity of interest is the interaction

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	21 827	20 966	18 949	18 949	18 949

Standard errors, in parentheses, are clustered at the judge level. For specification details, see Eq. (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
Gender analysis with judge fixed effects.

	(1) Def. win	(2) Def. win	(3) Def. win
Pla. maj. female		−0.028**	−0.049***
		(0.012)	(0.011)
Def. maj. female	−0.007		−0.001
	(0.011)		(0.011)
Judge maj. fem. × pla. maj. fem.		0.016	0.008
		(0.020)	(0.017)
Judge maj. fem. × def. maj. fem.	0.037**		0.042**
	(0.017)		(0.017)
DV mean	0.451	0.427	0.453
Judge FE	Yes	Yes	Yes
Court-year FE	Yes	Yes	Yes
Ethnicity FE	No	No	No
Case type FE	No	No	No
Observations	22,436	25,243	20,048

Notes: The dependent variable is an indicator equal to one if the defendant wins the case. All regressions include judge fixed effects and court–year fixed effects. "Maj. female" indicators equal one if the majority of the relevant party group (plaintiff, defendant, or judge panel) is female. Column (1) examines the defendant-side gender composition (Def. maj. female and its interaction with Judge maj. female); column (2) examines the plaintiff-side gender composition (Pla. maj. female and its interaction with Judge maj. female); column (3) includes both sides and both interactions simultaneously. The Judge maj. female main effect is absorbed by the judge fixed effects and is therefore not reported. Standard errors clustered at the judge level are reported in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

term. Finally, the additional controls in column (5) — appeal status, number of defendants, number of plaintiffs, and number of judges — enter with the expected signs (appeals favor defendants, larger panels tilt toward the defendant side) but do not affect the in-group estimate. For the remainder of the discussion we take column (5) as our preferred specification, while noting that the qualitative conclusions are unchanged if we use columns (3) or (4).

The significantly positive coefficients on the interaction between judge and defendant majority gender provide evidence that there is in-group gender bias from judges toward 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.

We do not find evidence of in-group bias toward plaintiffs. The interaction between judge and plaintiff majority gender is statistically indistinguishable from zero, and the sign is not consistent with a pro-plaintiff effect. Moreover, this coefficient differs significantly ($p < 0.05$) from the judge–defendant interaction term.

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 toward 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.

A natural concern is that our in-group estimates may reflect composition across judges rather than within-judge behavior — for instance, if judges who tend to rule for defendants happen to hear a disproportionate share of in-group cases. To address this, we re-estimate Eqs. (1) and (2) adding judge fixed effects, so that identification comes only from variation within a given judge's (or judge-panel's) caseload.

Results are reported in Table 3 (gender) and Table 4 (ethnicity). The gender in-group interaction is essentially unchanged: the coefficient on Judge maj. fem. × Def. maj. fem. is 0.037 (s.e. 0.017) in the specification without the plaintiff-side interaction and 0.042 (s.e. 0.017) when both interactions are included, close to the 0.036–0.040 range in Table 1. The ethnicity coefficient on Judge–defendant same ethnicity ranges from 0.049 to 0.063 across specifications with judge fixed effects (all significant at the 1 percent level), again very close to the 0.044–0.057 range in Table 2. The Judge–plaintiff same ethnicity coefficient remains statistically indistinguishable from zero. Taken together, the judge fixed-effects specifications confirm that our findings are not driven by compositional differences across judges: the asymmetric pro-defendant in-group bias operates within the decisions of individual judges.

The pooled regressions may mask important variation in who drives the aggregate effect. We therefore split the sample by judge-panel gender and re-estimate Eq. (1) separately for majority-female and majority-male judges (Tables 5 and 6). The two subsamples display a striking asymmetry. Among majority-female judges, the coefficient on Def. maj. female is positive and significant (0.028, s.e. 0.013 without plaintiff-side controls; 0.038, s.e. 0.013 in the specification that includes both litigant-side indicators), while the coefficient on Pla. maj. female is negative (−0.012 and −0.044, the latter significant at the 1 percent level). In other words, female judges are more likely to rule for female defendants, consistent with an asymmetric in-group preference operating on the defendant side. Among majority-male judges, by contrast, the Def. maj. female coefficient is effectively zero (−0.002 and 0.004, both insignificant), while Pla. maj. female is negative and significant (−0.030 and −0.052). Male judges do not exhibit a detectable in-group premium for male defendants. The gender in-group finding in Table 1 is therefore concentrated in the female-judge subsample.

This asymmetry is substantively important. It suggests that the aggregate interaction in Table 1 is not a symmetric tendency of judges to favor their own gender, but a more specific pattern in which female judges protect female defendants while male judges do not correspondingly protect male defendants. A purely information-based account, under which co-gender simply serves as an informative signal about litigant type, would not naturally predict this asymmetry, and we return to possible interpretations — including differential exposure to the stakes faced by female defendants and differences in how male and female judges weight the identities of the parties — in the discussion.

Table 4
Ethnicity analysis with judge fixed effects.

	(1) Def. win	(2) Def. win	(3) Def. win	(4) Def. win	(5) Def. win
Judge-plaintiff same ethnicity	0.005		−0.011	0.001	−0.001
	(0.013)		(0.015)	(0.016)	(0.016)
Judge-defendant same ethnicity		0.049***	0.063***	0.059***	0.059***
		(0.012)	(0.014)	(0.016)	(0.016)
Appeal					0.100***
					(0.014)
Number of defendants					0.002
					(0.003)
Number of plaintiffs					0.006**
					(0.003)
DV mean	0.449	0.453	0.452	0.452	0.452
Court-year FE	Yes	Yes	Yes	Yes	Yes
Judge FE	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	No	Yes	Yes
Case type FE	No	No	No	No	Yes
Observations	21,497	20,654	18,649	18,649	18,649

Notes: The dependent variable is an indicator equal to one if the defendant wins the case. Column (1) includes only the judge–plaintiff same-ethnicity indicator; column (2) includes only the judge–defendant same-ethnicity indicator; column (3) includes both. Columns (4) and (5) progressively add controls: column (4) adds ethnicity fixed effects for judges, defendants, and plaintiffs; column (5) adds case-type fixed effects and case-level controls (appeal indicator, number of defendants, number of plaintiffs). All regressions include court–year fixed effects and judge fixed effects. Standard errors clustered at the judge level are reported in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5
Gender analysis with majority-female judge panels only.

	(1) Def. win	(2) Def. win	(3) Def. win
Sample: Judge maj. female = 1			
Pla. maj. female	−0.012		−0.044***
	(0.016)		(0.013)
Def. maj. female		0.028**	0.038***
		(0.013)	(0.013)
DV mean	0.439	0.407	0.441
Court-year FE	Yes	Yes	Yes
Ethnicity FE	No	No	No
Case type FE	No	No	No
Observations	8155	9380	7333

Notes: The dependent variable is an indicator equal to one if the defendant wins the case. The sample is restricted to cases with a majority-female judge panel. "Maj. female" indicators are equal to one if the majority of the relevant party group (plaintiff or defendant) is female. Column (1) examines plaintiff-side gender composition only; column (2) defendant-side only; column (3) includes both. All regressions include court-year fixed effects. Standard errors clustered at the judge level are reported in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6
Gender analysis with majority-male judge panels only.

	(1) Def. win	(2) Def. win	(3) Def. win
Sample: Judge maj. female = 0			
Pla. maj. female	−0.030**		−0.052***
	(0.013)		(0.011)
Def. maj. female		−0.002	0.004
		(0.011)	(0.011)
DV mean	0.460	0.441	0.462
Court-year FE	Yes	Yes	Yes
Ethnicity FE	No	No	No
Case type FE	No	No	No
Observations	14,550	16,146	12,970

Notes: The dependent variable is an indicator equal to one if the defendant wins the case. The sample is restricted to cases with a majority-male judge panel. "Maj. female" indicators are equal to one if the majority of the relevant party group (plaintiff or defendant) is female. Column (1) examines plaintiff-side gender composition only; column (2) defendant-side only; column (3) includes both. All regressions include court-year fixed effects. Standard errors clustered at the judge level are reported in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The ethnic heterogeneity is similarly uneven. Table 7 re-estimates Eq. (2) separately by judge ethnicity. Below we focus interpretation on the six largest groups by caseload (Kikuyu, Luo, Luhya, Kamba, Kalenjin, Kisii); estimates for smaller groups are reported in the table but should be interpreted with caution given sample sizes. The pro-defendant in-group coefficient is most precisely estimated for judges from larger, politically dominant groups: among Kikuyu judges, Judge–defendant same ethnicity is 0.047 (s.e. 0.012) in the judge-defendant-only specification and 0.068 (s.e. 0.016) with both judge-plaintiff and judge-defendant indicators. Luo, Luhya, and Kamba judges show positive but less precise point estimates in the same range, while estimates for smaller groups are noisier; Kisii in particular flips sign, reflecting a small within-group sample. We interpret this pattern cautiously: the in-group effect we document is driven mainly by judges from ethnic groups that are large enough to yield both a sizeable in-group caseload and greater political salience of ethnicity as a social category. Appendix D provides additional robustness checks.

Table 8 presents the results from Eq. (3). Rather than interpreting the coefficients individually, we focus on the implied slope of the defendant's win probability with respect to damages across the four judge–defendant gender cells (holding plaintiffs at the omitted male category). Under male judges, the damages slope is approximately zero for male defendants but −0.018 per million Kenyan shillings for female defendants. Under female judges, the damages slope is approximately zero for both male and female defendants. In other words, female defendants are less likely to prevail as damages rise when their case is heard by a male judge, and this stakes-dependent gap is substantially attenuated — statistically indistinguishable from zero — when the judge is female.

This relative pattern admits two observationally equivalent interpretations, which the regression alone cannot distinguish. Under the first, which is most consistent with our broader framework, female judges extend greater protection to in-group female defendants as stakes rise, in a manner that loss aversion over in-group outflows would predict. Under the second, which is equally admissible given the identification available, male judges apply a stkes-dependent penalty to female defendants that female judges do not apply, and the resulting slope difference reflects greater gender-neutral behavior on the part of female judges rather than in-group favoritism. Because the regression identifies a relative difference in slopes across judge gender and does not anchor either

Table 7
Ethnicity analysis by ethnic group.

	(1) Def. win	(2) Def. win	(3) Def. win	(4) Def. win	(5) Def. win	(6) Def. win	(7) Def. win	(8) Def. win	(9) Def. win
Judge-plaintiff same ethnicity	0.025		0.058	−0.055		−0.072	0.003		−0.010
	(0.115)		(0.092)	(0.054)		(0.076)	(0.018)		(0.019)
Judge-defendant same ethnicity		0.043	0.055		0.085	0.107*		0.047***	0.068***
		(0.088)	(0.108)		(0.056)	(0.055)		(0.012)	(0.016)
Judge ethnicity	Kalenjin			Kamba			Kikuyu		
Specification	App	Resp	Both	App	Resp	Both	App	Resp	Both
Observations	1846	1839	1637	2700	2535	2303	5906	5752	5115
	(10) Def. win	(11) Def. win	(12) Def. win	(13) Def. win	(14) Def. win	(15) Def. win	(16) Def. win	(17) Def. win	(18) Def. win
Judge-plaintiff same ethnicity	−0.244***		−0.215***	0.013		0.013	0.008		0.016
	(0.056)		(0.074)	(0.037)		(0.046)	(0.031)		(0.035)
Judge-defendant same ethnicity		−0.036	−0.002		0.061	0.069*		0.044	0.030
		(0.046)	(0.040)		(0.038)	(0.041)		(0.039)	(0.045)
Judge ethnicity	Kisii			Luhya			Luo		
Specification	App	Resp	Both	App	Resp	Both	App	Resp	Both
Observations	1636	1576	1408	3588	3492	3205	3791	3506	3145
	(19) Def. win	(20) Def. win	(21) Def. win	(22) Def. win	(23) Def. win	(24) Def. win	(25) Def. win	(26) Def. win	(27) Def. win
Judge-plaintiff same ethnicity	0.299***		0.556***	0.074		0.110	−0.046		−0.070
	(0.013)		(0.032)	(0.094)		(0.133)	(0.269)		(0.054)
Judge-defendant same ethnicity		0.087***	−0.291***		−0.065	−0.116		0.000	0.052
		(0.000)[a]	(0.021)		(0.075)	(0.111)		(0.231)	(0.194)
Judge ethnicity	Masai			Meru			Mijikenda		
Specification	App	Resp	Both	App	Resp	Both	App	Resp	Both
Observations	117	116	104	529	492	444	264	226	204

	(28) Def. win	(29) Def. win	(30) Def. win	(31) Def. win	(32) Def. win	(33) Def. win
Judge-plaintiff same ethnicity				−0.037		−0.049
				(0.098)		(0.166)
Judge-defendant same ethnicity		−0.600***	−0.500		−0.222	−0.220
		(0.000)[a]	(·)[b]		(0.296)	(0.434)
Judge ethnicity	Pokot			Somali		
Specification	App	Resp	Both	App	Resp	Both
Observations	95	99	86	204	201	178

Notes: Dependent variable equals one if the defendant wins. Each panel restricts to cases with at least one judge from the indicated ethnic group. Within each group, "App" includes only the judge–plaintiff same-ethnicity indicator, "Resp" only the judge–defendant same-ethnicity indicator, and "Both" includes both. Empty cells indicate that the corresponding coefficient could not be estimated for that subsample due to insufficient within-group variation. All regressions include court–year fixed effects; standard errors clustered at the judge level are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

[a] Standard error rounds to 0.000 due to low variation in the subsample.

[b] Standard error is not computable (denoted "·"), typically reflecting a singular variance matrix in the subsample.

slope to an exogenously neutral benchmark, the data cannot adjudicate between these readings, and the choice between them depends on a prior about which cell represents unbiased adjudication.

We therefore treat Table 8 as suggestive of, and consistent with, the asymmetric loss-aversion mechanism we propose, rather than as standalone evidence for it. We note that even the "offsetting" reading of Table 8 requires an account of why male and female judges differ systematically in how they respond to damages in female-defendant cases, and that both the in-group and the impartiality readings imply gender-differentiated judicial behavior of the kind our framework is designed to interpret. Read alongside the defendant-side in-group interaction in Table 1 and the defendant-side ethnic in-group interaction in Table 2 — which are each consistent with the mechanism on their own terms — Table 7 adds a further piece of evidence that judicial behavior varies along gender identity lines as stakes grow, while leaving the precise cognitive channel open to alternative interpretations.

Table 9 produces the analogous three-way interaction for the ethnicity variables. Here we find no statistically significant moderating effect of damages, which may reflect the limited sample with extracted monetary damages and the additional variance introduced by a three-way interaction involving a multi-category ethnicity variable, and may also reflect substantive heterogeneity in how ethnic identification interacts with stakes across different groups. We therefore treat the ethnicity damages results as inconclusive.

An important limitation of our analysis is that several variables are imputed from text, which inevitably introduces error. However, the direction of this error is likely to attenuate our results rather than generate spurious bias. First, we exclude non-human litigants (e.g., companies, the state) using keyword filters; occasional misclassification would add noise but not systematically create in-group effects. Second, gender and ethnicity are inferred from names using survey-based mappings (Hjort, 2014) and global name databases. Misclassification here is possible but limited, and again would bias coefficients toward zero. Third, for roughly half of the cases we infer outcomes using a GloVe-based text classifier. While accuracy is high in held-out tests, any residual misclassification would reduce precision rather than inflate significance. Finally, damages data are extracted using an LLM pipeline from a smaller subsample of cases. Errors in this step are rare given the regular structure of numeric mentions, but they similarly introduce noise. Taken together, these layers of imputation may reduce power but make it unlikely that our findings are driven by artifacts of the text-processing pipeline.

Table 8
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. × Pla. maj. fem.	0.0112
	(0.0577)
Judge maj. fem. × Def. maj. fem.	−0.0156
	(0.0612)
Judge maj. fem. × Damages	−0.000464
	(0.000334)
Pla. maj. fem. × Damages	0.00205
	(0.00795)
Def. maj. fem. × Damages	−0.0181**
	(0.00721)
Judge maj. fem. × Pla. maj. fem. × Damages	−0.00109
	(0.00880)
Judge maj. fem. × Def. maj. fem. × 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 Eq. (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 9
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 × Damages	−0.000233
	(0.00178)
Judge-def. same × 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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 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 et al., 1999; Knowles et al., 2001; Plant et al., 2011; Rehavi and Starr, 2014; Knox et al., 2020).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.irle.2026.106344.

Data availability

Data will be made available on request.

References

Ahrsjö, Ulrika, Niknami, Susan, Palme, Mårten, 2024. Identity in court decision-making. Am. Econ. J.: Econ. Policy 16 (4), 142–164.
Arnold, David, Dobbie, Will, Yang, Crystal, 2018. Racial bias in bail decisions. Q. J. Econ. 133 (4), 1885–1932.
Asingo, Patrick, Biegon, Japhet, Kivuva, Joshua, Wahiu, Winluck, 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. J. Econ. Perspect. 27 (1), 173–196.
Benartzi, Shlomo, Thaler, Richard, 1995. Myopic loss aversion and the equity premium puzzle. Q. J. Econ. 110 (1), 73–92.
Berge, Lars, Bjorvatan, Kjetil, Galle, Simon, Miguel, Edward, Posner, Daniel, Tungodden, Bertil, Zhang, Kelly, 2015. How Strong are Ethnic Preferences? Working Paper.
Choi, Danny, Harris, Andy, Shen-Bayh, Fiona, 2021. Ethnic bias in judicial decision making: Evidence from criminal appeals in Kenya. Am. Political Sci. Rev. 116 (3), 1067–1080.
Depew, Briggs, Eren, Ozkan, Mocan, Naci, 2017. Judges, juveniles, and in-group bias. J. Law Econ. 60 (2), 209–239.
Gainer, Maya, 2015. Transforming the courts: Judicial sector reforms in Kenya. Princet. Univ. Innov. Success. Soc. 1 (7).

Gazal-Ayal, Oren, Sulitzeanu-Kenan, Raanan, 2010. Let my people go: Ethnic in-group bias in judicial decisions–evidence from a randomized natural experiment. J. Empir. Leg. Stud. 7 (3), 403–428.

Glynn, Adam, Sen, Maya, 2015. Identifying judicial empathy: does having daughters cause judges to rule for women's issues? Am. J. Political Sci. 59 (1), 37–54.

Grossman, Guy, Gazal-Ayal, Oren, Pimentel, Samuel, Weinstein, Jeremy, 2016. Descriptive representation and judicial outcomes in multiethnic societies. Am. J. Political Sci. 60 (1), 44–69.

Hjort, Jonas, 2014. Ethnic divisions and production in firms. Q. J. Econ. 129 (4), 1899–1946.

Kahneman, Daniel, Tversky, Amos, 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263–292.

Kastellec, Jonathan, 2013. Racial diversity and judicial influence on appellate courts. Am. J. Political Sci. 56 (1), 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. J. Labor Econ. 36 (3), 623–664.

Knowles, John, Persico, Nicola, Todd, Petra, 2001. Racial bias in motor vehicle searches: Theory and evidence. J. Political Econ. 109 (1), 203–229.

Knox, Dean, Lowe, Will, Mummolo, Jonathan, 2020. Administrative records mask racially biased policing. Am. Political Sci. Rev. 114 (3), 619–637.

Miller, Will, Kerr, Brinck, Reid, Margaret, 1999. A national study of gender-based occupational segregation in municipal bureaucracies: Persistence of glass walls? Public Adm. Rev. 218–230.

Plant, E. Ashby, Goplen, Joanna, Kunstman, Jonathan W., 2011. Selective responses to threat: The roles of race and gender in decisions to shoot. Pers. Soc. Psychol. Bull. 37 (9), 1274–1281.

Rehavi, Marit, Starr, Sonja B., 2014. Racial disparity in federal criminal sentences. J. Political Econ. 122 (6), 1320–1354.

Shayo, Moses, Zussman, Asaf, 2011. Judicial ingroup bias in the shadow of terrorism. Q. J. Econ. 126 (3), 1447–1484.

Sloan, CarlyWill, 2020. Racial Bias by Prosecutors: Evidence from Random Assignment. Working Paper.

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. J. Leg. Stud. 44 (1), 75–111.