IFS

Institute for Fiscal Studies

Omry Yoresh
Weijian Zou

Working paper

26/12

The lifecycle of judicial bias

Economic and Social Research Council

The Lifecycle of Judicial Bias

Omry Yoresh
LSE and CEP

Weijian Zou
LSE, IFS, and STICERD

February 2026

Abstract

How does judicial bias arise and persist in the legal system? We trace the lifecycle of judicial bias in India's criminal courts, leveraging the quasi-random assignment of judges to courts and cases. We first document substantial variation in bias: within the same court, assigning a same-religion defendant from a judge at the 25th to 75th percentile of bias increases acquittal probability by 7.5 percentage points, 45% of the mean acquittal rate. We then examine how different triggers across judges' professional careers affect their bias. Exposure to Hindu-Muslim communal riots during a judge's first five years of service leads them to increase same-religion acquittal rates by 17.5%. Yet, judges who work alongside colleagues from different-religions during this critical period show no such effect. Next, social interactions between judges reinforce judicial bias. On-the-job exposure to biased colleagues leads judges to reduce acquittal rates for defendants of different-religions by 3.2 percentage points, with effects that persist for over a year. These findings suggest that judicial bias is neither innate nor inevitable, but rather shaped by experiences and relationships, and that thoughtful institutional design can prevent and mitigate discriminatory practices across various professions.

Keywords: Judicial Bias, Discrimination, Peer Effects, Contact Theory
JEL Codes: K41, J15, D91

Yoresh, Department of Economics, London School of Economics and Political Science, o.yoresh@lse.ac.uk. Zou, Department of Economics, London School of Economics and Political Science, w.zou2@lse.ac.uk. Tables and figures are presented in the order they are referenced in the paper. Online Appendix can be found here.

1. Introduction

Fighting structural inequalities based on individual identities and characteristics has become a top policy priority worldwide. A key part of reducing structural inequalities lies in creating a fair and well-functioning judicial system that enforces laws and policies aimed at reducing discrimination and ensuring equal representation (Sandefur and Siddiqi 2013). Beyond this, the right of any defendant to a fair trial in a judicial system is often enshrined in the constitutions of various countries, such as India, the country this study focuses on.[1]

Growing evidence points to the existence of judicial bias along gender, religious, and/or ethnic lines in some countries. Nonetheless, little is known about how such biases are formed, transmitted, or influenced by policies. Existing work suggests that exposure to ethnic or religious conflicts on the job affects judges' rulings towards their in-groups (Shayo and Zussman 2011, 2017). A nascent literature identifies the role of gender, ethnic, and ideological composition in influencing panels' judicial decisions in various settings, including US Supreme Court judgments, jury decisions, and scientific committees (Anwar et al. 2012; Bagues et al. 2017; Anwar et al. 2019; Holden et al. 2021).[2] Another strand of criminology research examines how institutional features, like judicial rotation and court culture, create uniformity in sentencing practices across courts (e.g., Hester (2017)).

In this paper, we begin to unpack the black box of judicial bias formation by providing a novel, comprehensive account of how judicial bias emerges through experiences and spreads through social interactions. As tracing the origins of discriminatory practices is a challenging task, India's judicial system provides an excellent setting to study the question at hand. Judges enter the system at a young age, creating an overlap between their professional first years and their early formative years. Moreover, judges are quasi-randomly assigned to courts throughout their careers, with a mandatory rotation roughly every two to three years. This institutional structure generates exogenous variation in two key experiences: exposure to local inter-group conflict and the composition of their colleagues.

We start by quantifying the level of heterogeneity in judicial bias, measured as judge fixed effects on the differential probability to acquit in-group defendants. We find that within the same court, between-judge variability in in-group (religious) bias is larger than across-district variability, the basic spatial unit in which courts are organized. This implies that, even conditional on being assigned to the same court, defendants may experience vastly different outcomes depending on the judge assigned to their case. As one striking example,

[1] See Article 21 of the Indian Constitution, which protects defendants to the right to a fair trial.

[2] There is a separate literature that studies the role of the workplace in shaping worker outcomes, focusing on the role of social interactions between workers and managers (e.g: Haegele 2022; Cullen and Perez-Truglia 2023).

being assigned a judge at the 75th percentile (relative to the 25th percentile) of the bias distribution increases the differential probability that a same-religion defendant is acquitted by 7.5 percentage points, a 45% increase from the mean acquittal rate. This large within-court heterogeneity motivates our investigation into how such disparate biases form and spread.

We document the role of different professional triggers on judicial bias. First, bias forms during a critical professional window, the first five years of service, when local exposure to Hindu-Muslim communal riots creates lasting in-group bias (17.5% increase). In contrast, exposure to similar conflicts after the initial five-year period shows no significant impact on judicial decisions, highlighting the importance of early career periods in shaping judge behaviours.

Second, we find that early-career contact with out-group peers during these formative years can counteract the effects of riot exposure. Hindu judges working alongside Muslim colleagues during these formative years show much less riot-induced bias in their treatment of defendants. The effectiveness of the contact seems to depend on the judges' social proximity, such as same-gender or close-in-age relationships. When judges experience these demographic similarities with a (religious) out-group colleague, it eliminates the development of in-group bias. This suggests that out-group exposure operates through meaningful interactions to counteract the polarizing effects of conflict rather than through mere presence.

Third, a reinforcement effect further stresses the importance of early experiences and highlights a mechanism at play. While childhood exposure to communal conflict shows no direct association with discriminatory practices, it creates vulnerabilities that later events might trigger. Judges who experienced conflict in their early childhood show a much stronger riot-induced bias when exposed again to conflict in their early careers (75% from the baseline mean). This result suggests that early-childhood exposure primes judges to respond more strongly to later conflict exposure, indicating that past experiences compound bias through repeated exposure.

Having documented that early career experiences can have long-lasting impacts on judicial bias, we ask next: to what extent do interactions in the workplace amplify bias? To do so, we leverage quasi-random variation in the timing of judge movements across courts to study the causal effect of exposure to biased peers, as measured by pre-move cases, on incumbent judges' decisions in the receiving court. Our specifications compares within-judge variation in acquittal rates for out-group and in-group cases, before and after a new judge arrives, in a difference-in-differences framework. We find that exposure to a highly biased judge (top tercile of in-group bias) increases incumbent judges' own in-group bias: they become 3.21 percentage points less likely to acquit out-group defendants (18.6% of the mean) after 6 months of the move, with the effect persisting up to another year. This effect is economically and statistically meaningful. In contrast, we find that exposure to a judge who exhibits out-group favoritism (whose judge fixed-effect is in the bottom tercile) had a small and insignificant effect on judges' acquittal rates for defendants.

We proceed to test for potential mechanisms, finding support for the role of group-specific interactions, proximity, seniority, and incumbent judge characteristics. First, our estimates are entirely driven by judges' exposure to peers of the same gender. Exposure to a biased judge of the opposite gender has a small, insignificant effect. Second, we look into the size of courts, as smaller courts increase potential interaction between incumbent and incoming judges. We find that courts employing at most eight judges (corresponding to the sample mean) drive the effect. Third, examining seniority, we find that effects are more pronounced when the mover judge is more senior than the court average or the incumbent judge.

Most strikingly, we find asymmetric transmission by incumbent bias level: exposure to biased colleagues affects only incumbent judges who previously exhibited out-group favoritism (bottom tercile of the bias distribution), moving them toward neutral treatment. In contrast, incumbent judges who already exhibit in-group bias (top tercile) show no moderation when exposed to out-group-favoring colleagues.

Our study focuses on India, drawing on a new dataset of 5 million cases in India's criminal courts used by Ash et al. (2025). This is a setting of great policy relevance. First, as a multi-ethnic and cultural society, there are increasing concerns over unequal representation and segregation by ethnicity and religion in India (Asher et al. 2024; Ash et al. 2025). This makes it highly urgent and policy-relevant to understand whether, and how, the judicial system may aggravate existing biases. Second, studying the world's second-largest common law system enables us to obtain a judge-level estimate of in-group bias with precision and retain statistical power.

This paper contributes to multiple strands of related research areas. The first is the literature on discrimination and in-group bias, beginning with foundational work by Becker (1957) on taste-based discrimination and Arrow et al. (1973) on statistical discrimination. One of the first pieces of empirical evidence, by Tajfel et al. (1971), showed through a series of experiments that in-group favouritism exists in situations where groups are sorted arbitrarily. Further building on these foundations, judicial bias has been studied in economics across various contexts. The seminal paper by Shayo and Zussman (2011) was the first to exploit the quasi-random assignment of cases to judges in Israel to identify ethnic bias that intensified during terrorism periods, and in later work, found that this bias later persisted at the *local* level (Shayo and Zussman, 2017). Arnold et al. (2018) showed that inexperienced and part-time judges exhibit more racial bias against Black defendants in the US, where

Abrams et al. (2012) documented significant variation in the conviction rate of minorities between judges. McConnell and Rasul (2021) showed that judicial biases are malleable, finding that post 9/11 animosity toward Muslims spilled over to Hispanic defendants in federal courts.

In the Indian context, using all criminal cases across India, Ash et al. (2025) found a precisely estimated null effect for religious and gender bias in the cross-section. Bharti and Roy (2023) showed that judges in Uttar Pradesh, India, who were exposed to Hindu-Muslim communal riots in their early childhood (ages 0-6) exhibit more stringency in their bail decisions. We advance this literature by providing the first account for the lifecycle of judicial bias, from its formation to its transmission within the judicial system. While existing work has mainly documented when and where judicial bias exists, only a few studies have examined the mechanisms through which these discriminatory preferences are formed and can be influenced by policy.

Second, we contribute to the literature on intergroup contact and social interactions. Building on foundational contact theory of Allport (1954), which proposed that contact between groups can reduce prejudice under favorable conditions, empirical work has examined contact effects across various settings: hiring decisions (Battaglini et al. 2023), judicial rulings based on contemporaneous peer composition (Eren and Mocan 2020), workplace diversity initiatives (Darling-Hammond et al. 2021), and educational environments (Boisjoly et al. 2006; Corno et al. 2022; Carrell et al. 2018; Finseraas and Kotsadam 2017; Rao 2019). This literature has primarily focused on out-group interactions with those of a different gender, ethnicity, or socioeconomic status.

We make advances to this literature in three ways. First, we show that contact during critical early-career periods has lasting effects on professional behavior, underscoring the importance of timing for contact effectiveness. Second, we identify that contact works under specific conditions of social proximity, where demographically similar out-group peers (same gender, similar age) successfully mitigate bias formation. Third, and importantly, we document the flip side: interactions with biased in-group members can transmit prejudice by eliminating out-group favoritism.

Third, the literature on peer effects in the justice system. A central question in this work is whether judges influence one another's decisions, and through what channels. Anwar et al. (2012) provided early evidence that the racial composition of juries affects conviction rates, later showing that jurors' political ideology shapes verdicts in a predictable direction Anwar et al. (2019). Holden et al. (2021) documented ideological conformity among US Supreme Court justices, where exposure to colleagues shifts voting behavior. Eren and Mocan (2020) found that judges are influenced by their peers' sentencing decisions on the same day, suggesting real-time spillovers in judicial behavior.

We advance this literature in two ways. First, existing work focuses primarily on settings where judges interact directly: deliberating on panels, voting together, or observing each other's decisions in real time. In contrast, Indian district court judges decide cases individually, without consulting colleagues. Any peer effects we detect must therefore operate through informal workplace interactions rather than deliberation or deference (Epstein and Jacobi 2008; Holden et al. 2021). Second, following Dobbie et al. (2022), we leverage quasi-random case assignment to estimate each judge's revealed in-group bias, allowing us to study how exposure to a colleague with measured bias affects an incumbent's own decisions, rather than relying on demographic composition as a proxy for influence.

Lastly, another strand of literature examines how experiences in early-career stages shape long-term outcomes and preferences. Research has shown that various types of shocks during one's formative years affect: political attitudes (Margalit, 2019), social trust (Alesina and Fuchs-Schündeln, 2007), and economic values (Abramitzky et al., 2025), among others. Similarly, early career environments have been shown to have lasting effects on economic outcomes, such as wages and career progression (Kahn, 2010; Wachter, 2020). However, identifying causal effects on behavioral outcomes is understudied as it is empirically challenging. We contribute to this literature by showing that experiences also influence professional decision-making in individuals' early careers.

The remainder of this paper is organized as follows. Section 2 describes India's judicial system and the institutional setting. Section 3 introduces our three primary data sources used in our analyses. Section 4 documents the substantial heterogeneity in judicial bias within courts. Section 5 examines how bias forms through early career exposure to communal riots and how contact with out-group peers can prevent this formation. Section 6 explores how bias spreads through peer interactions when judges move across courts. Section 7 discusses our findings and concludes.

2. Institutional Setting

A. India's Judicial System

India has one of the world's largest judicial systems, which relies on British common law. The system has three tiers: the Supreme Court at the top, 25 High Courts at the state level, and, beneath them, approximately 700 district courts and over 7,000 subordinate courts. Our analysis focuses on criminal cases in the District and Sessions courts (district courts), which constitute the lower judiciary and serve as the first point of contact for litigants dealing with civil (property/debt), criminal (violent/property/financial crimes), and commercial (contractual enforcement) disputes. Unlike jury-based systems, judges in India make decisions individually.

Each district has only one district court, despite significant variation across districts in court size, as indicated by the number of courtrooms and active judges in a given period (Rao, 2024).[3] This lower judiciary operates under severe case backlogs and overwhelming caseloads.

B. Recruitment of Judges

The Indian judicial system follows a particular approach to recruitment and career structure. The lower judiciary varies by state but broadly consists of three hierarchical positions: (i) Civil Judge (Junior Division), (ii) Senior Civil Judge, and (iii) District Judge. The corresponding state High Court appoints all judges serving in the lower judiciary.

There are two primary pathways into the judiciary. The first is direct entry-level recruitment, where candidates take written and oral exams administered by the High Court. For this pathway, candidates must be between 22 and 35 years old. A significant policy change occurred in 2002 when the Supreme Court eliminated the previously mandatory three-year practice requirement for entry-level positions, allowing law graduates to apply directly after graduation. This reform substantially lowered entry ages, with judges in states such as Punjab now averaging under 30 years old upon recruitment (Ratra et al., 2023).[4]

The second pathway allows experienced lawyers to enter at higher levels. District Judges can be recruited directly from practicing lawyers with at least seven years of experience, though such appointments are limited to 25% of District Judge positions. These entrants cannot be older than 45 years and join at a more senior level than direct recruits (Allahabad High Court, 2001).

For those entering through direct recruitment, two critical milestones mark the early stages of their careers. The two-year point marks the end of the "probationary period," which coincides with the completion of their first court appointment. The five-year milestone is significant as judges become eligible for promotion to Senior Civil Judge and receive their first pay increase. After these initial years, judges serve for life until the mandatory retirement

[3]The number of district courts varies by state, with the largest being in the state of Uttar Pradesh, with 75 districts.

[4]Recently, on May 20, 2025, the Supreme Court restored the minimum legal practice requirement: https://www.thehindu.com/news/national/supreme-court-makes-three-years-legal-practice-mandatory-to-apply-for-judicial-service/article69596456.ece.

age of 60, with removal possible only by the governor upon agreement with the corresponding High Court (Ash et al., 2025).

C. Judges Appointments

While the High Court determines judicial appointments, the allocation of judges to specific courts follows a systematic rotation policy, where judges move to new court locations approximately every two to three years. The total number of judges in each district is generally decided by district-level population counts based on the most recent census, though judge vacancy rates have been increasing over time.

The mechanism through which judges are allocated across courts is based on a seniority-first, serial dictatorship mechanism, subject to no-repeat and no-home-district assignments. The most senior judges rank their locations and are assigned to their first choice, followed by the second-most-senior judge, in an iterative process. All judges are subject to a non-home-district restriction, under which the Assistant Registrar compiles a list of places where the judge has relatives or received education (Bharti and Roy, 2023). For newly recruited judges, assignment to courts with vacancies is effectively random, given this constraint. For more senior judges, after an average service period of two years, allocation to courts with vacancies includes an additional non-repeat rule that prohibits assignment to an entire "zone," which is a collection of nearby districts.

The stated motive of High Courts in designing such rotation policies is to fight corruption and collusion. Judges move every two to three years on average. In Uttar Pradesh, for example, the average reallocation distance is 325 km, with a standard deviation of 165 km (Bharti and Roy, 2023). The reallocation process can sometimes occur rapidly, with judges required to relocate on short notice to fill urgent vacancies. This allocation policy is crucial for our study as it induces quasi-random variation in court composition, exposure to local inter-group conflict, and judge-level exposure to biased peers. We provide formal statistical tests supporting the quasi-random assignment of judges to courts in Section 6.

D. Case Assignment

Our identification strategy also relies on the quasi-random assignment of judicial cases to judges. Cases are allocated in the Indian judicial system based on the premise of restricting individuals from "judge-shopping," a practice where defendants seek to be allocated to a specific judge. This practice was explicitly condemned by the Indian Supreme Court in *M/s Chetak Construction Ltd. v. Om Prakash & Ors., 1998(4) SCC 577*, and this decision has been frequently cited in subsequent rulings.

Once a crime is reported at a police station, a First Information Report (FIR) is filed. As each police station corresponds to a specific district court jurisdiction, the case is assigned to the court. Within courts with multiple judges, cases are allocated based on courtroom rules, where each courtroom is designated to handle cases from specific police stations and charge types.

Several institutional features support the quasi-random allocation of cases to judges. First, judges are routinely rotated across courts every two to three years and across different courtrooms within the same court. This frequent rotation limits the predictability of case assignments. Second, the mapping between police stations and courtrooms remains fixed, but the rotation of judges across these courtrooms makes it difficult to predict which judge will handle any given case (Ash et al., 2025). We provide formal statistical tests of the correlation between demographic characteristics of assigned judges and defendants in both Sections 5 and 6.

E. Hindu-Muslim Relations in India

Since the partition of the Indian subcontinent into India and Pakistan in 1947, tension between the Hindu and Muslim communities living in India has persisted in various forms. The partition itself resulted in the displacement of roughly twelve million people and massive massacres along the border (Khan, 2017). Post-independence, violence has been transformed into more sporadic yet regular occurrences. There have been, since then, approximately 10,000 casualties from more than 1,000 Hindu-Muslim communal conflicts across India (Mitra and Ray, 2014). Despite their localized nature, these incidents tend to be highly socially disruptive. The government typically reacts with harsh measures such as mass arrests and curfews affecting several villages or entire districts, even though few are usually directly involved in the violence itself (Bharti and Roy, 2023).

The Muslim minority in India comprises around 14% of the population (approximately 200 million people, making them the world's largest Muslim minority) and has faced persistent marginalization. The 2006 Sachar Committee report documented that the Muslim community performs substantially worse than the Hindu population across education, employment, and other dimensions of life. Recent research has reinforced these findings: Asher et al. (2024) showed that Muslim boys have worse upward income mobility than any other marginalized community in India, with a steady decline over the years, while Asher et al. (2024) documented that Muslim communities concentrate in poorer cities with less access to schools, hospitals, and public infrastructure.

Since the Bharatiya Janata Party (BJP) came to power in 2014, these disparities have become increasingly politicized. Legislative efforts emphasizing Hinduism as central to Indian identity have brought the Muslim divide into public discourse, raising concerns about the rights of the Muslim population. These tensions, intertwined with political discourse and policymaking around issues such as interfaith marriages and anti-conversion sentiments (Jaffrelot, 2021), make the judicial system's role in mediating this religious divide particularly important for understanding how these divisions influence Indian institutions.

3. Data Description

Our empirical analysis draws on three primary data sources: criminal case records from India's e-Courts platform, detailed judge biographies from the Allahabad High Court, and data on Hindu-Muslim communal riots across India. We describe each source in turn, noting which analyses utilize particular data elements. The specific sample construction procedures for each analysis are detailed in their relevant empirical sections.

A. Cases

Our primary data source comes from the Indian e-Courts platform, compiled by the Development Data Lab (Ash et al., 2025). This e-Courts system is a publicly available data source, implemented in 2013 by the Indian judicial system, which includes information on the entirety of cases held at the courts in the jurisdiction of the District and Sessions courts across India.

The data includes all case information concerning the offense, including the specific statutory provisions (act and section) under which charges were filed, all filing, hearing, and decision dates, and the names of petitioners, defendants, and judges. Moreover, each case includes its corresponding charge section, the court and courtroom in which the case was decided, and the final decision. Another important data source we utilize is judges' appointments, which are also from the e-Courts system. The appointments describe each judge's tenure start and end dates at the court for a given position.

To infer the religion and gender of both defendants and judges, we use the resulting classification from Ash et al. (2025), which comes from a trained machine classifier. This classifier utilizes a bidirectional Long Short-Term Memory (LSTM) model, which processes the sequence of characters in names directly. The gender classifier is trained on a dataset comprising 13.7 million names, categorized by gender from Delhi voters. Similarly, the religion classifier is based on 1.4 million entries from data labeled by religion among National Railway exam takers. The outcomes of these processes categorize gender as “Female,” “Male,” or “Unidentified,” and religion as “Hindu,” “Muslim,” “Christian,” “Buddhist,” or “Other”. Our analysis does not differentiate between the non-Muslim population. As reported in Ash et al. (2025), the classifier demonstrates strong performance on standard metrics, achieving a balanced accuracy and harmonic mean of precision and recall (F1) of 0.975 and 0.976 for gender and 0.98 and 0.99 for religion, respectively.

B. Judges Biographies

For the analysis of bias formation in Section 5, we collected detailed biographical information on judges serving in Uttar Pradesh from the Allahabad High Court website.[5] The website maintains comprehensive biographies of both past and present judges serving under the court’s jurisdiction. The data includes each judge’s name, date of birth, home district and state, along with dates of entering judicial service and subsequent promotions. Moreover, there are educational qualifications detailing all levels of education a judge has completed, with their corresponding grades and years.

Importantly, each judge also has a complete list of current and previous postings, including each court a judge has served in, with the position held, location, and start and end dates.[6] We manually standardize all judge home districts and posting locations to match the 2011 Indian Census district boundaries, ensuring consistent geographic matching with other data sources. As the gender and religion of each judge are not provided directly, we rely on the same classification method described above.

C. Communal Riots

For the analysis of conflict exposure in Section 5, we construct comprehensive data on Hindu-Muslim communal riots in India. We begin with the datasets compiled by Varshney (2006) and Mitra and Ray (2014) covering 1950 to 1995 and 1996 to 2000, respectively. To extend coverage through 2018, we use the Factiva website by Dow Jones, which includes digital articles from common newspapers.

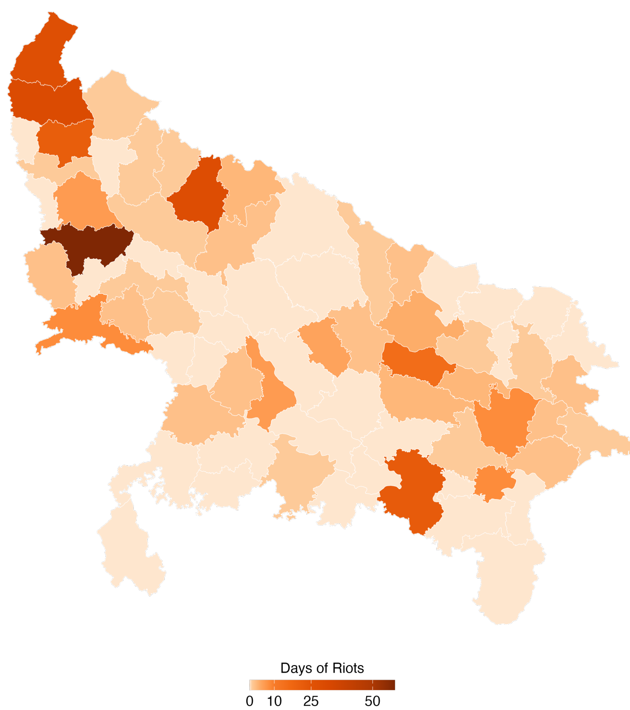
After filtering for all articles from the *Times of India* in the relevant timespan, we limit the search to articles with the subject “Civil Unrest” and require relevant keywords to appear in the text.[7] Following the “Data-Entering Protocol for Riot Database” from Varshney (2006),

[5] We collect information on judges only in Uttar Pradesh given the availability of these biographies.
[6] For an example of such a biography see Appendix Figure A1.
[7] The following restriction was applied: articles must contain two words, one from each category. Category

we detect events of inter-religion conflict. The definition of such an event follows Soule (1992), where there is required to be: (i) violence, and (ii) at least two or more communally identified groups confronting each other.

We review the filtered articles and identify riots between communities, documenting location, timing, and casualties (killed, arrested, or injured), and whether a curfew was imposed. We then restrict all riots to those classified as "Strong Likelihood" or "Certain" according to the protocol described above. While communal riots occur throughout India, Uttar Pradesh is the most riot-prone state, making it an ideal setting for studying the effects of conflict exposure. Appendix Table A1 describes the frequency and intensity of communal riots across India and in Uttar Pradesh specifically by decade. Beyond Uttar Pradesh's prominence, the wide spatial variation in Hindu-Muslim conflict is evident in Figure 1.

Figure 1: Hindu-Muslim Communal Riots in Uttar Pradesh, 1950-2018



Notes: This figure shows the spatial distribution of communal riots in the state of Uttar Pradesh between 1950 and 2018. The map is colored by a gradient corresponding to the total days of communal riots, with darker colors indicating higher riot intensity. Geographic boundaries correspond to 2011 census district definitions. Source of the shape-file: India boundaries by DataMeet India community (CC BY 4.0)

one refers to conflict: riot, riots, rioting, violent, violence, clash, clashes, killed, arrested, arrests, arrest, injured, murder, quarrel, attacked, attack. Category two refers to communities involved: communities, communal, Hindu, Muslim, religious, mosque, cow slaughter, pray, prayer, group, mob.

D. Matching and Sample Construction

Cases

To define our analysis sample, we impose several restrictions. We begin by limiting our data to only those criminal cases filed under the Indian Penal Code or the Code of Criminal Procedure, following Ash et al. (2025). This approach serves two primary purposes. Firstly, it ensures that each case involves a single defendant and a single judge, allowing us to clearly assign gender and religious labels to each party for our analysis. Secondly, criminal cases ensure clear and objective outcomes, as defendants are either acquitted or not. We further make the restriction of omitting any available offenses, which have different legal procedures and may follow different assignment mechanisms.

Courts and Judges

For our purposes, we define courts as distinct complexes within a district, as they are situated in various locations. With the presence of both district and subdistrict courts, our sample includes over 2,898 courts across 29 states and union territories, encompassing approximately 597 districts. For instance, in the state of Maharashtra, the district of Nandurbar houses more than three different subdistrict courts, including the "*Chief Judicial Magistrate*," "*Civil Court Senior Division*," and the "*District and Session Courts*."

In our methodology for identifying judges over time, we take the following steps. First, from the text strings for judge designation, we discard any inconsistent information, including prefixes, educational qualifications, and similar details. Next, we eliminate terms that typically pertain to the court or the position of a judge—phrases like "morning court," "vacant," or any empty entries are excluded from our sample. We then consolidate the names of judges within each state into a single judge identifier. This focus on individual states is grounded in the fact that judges operate under the jurisdiction of that state's High Court, with the majority of their career movements occurring within the same state. Finally, to ensure the reliability of judge identification over time, we restrict our identification to those judges who, on average, held no more than two positions per year. For judges who averaged more than two positions per year, each position was classified as a separate judge.

We define judge appointments by considering all subsequent appointments of a judge in a court as a singular appointment. This approach accounts for instances where judges have their appointments renewed within the same court. We define a subsequent appointment as one occurring within six months of a prior appointment in the same court, corresponding to the default duration stipulated by the *Maternity Benefit Act of 1961*.[8]

[8]Judges that return to a court after a more extended hiatus (without an appointment in another court)

E. Summary Statistics

Panel A of Appendix Table A2 details our full sample, which is comprised of 12,784 judges who decided on more than 4 million criminal cases between 2010 and 2019.[9] Among these judges, 7% are Muslim and 29% are female. The defendant composition shows that approximately 15% of defendants are Muslim and 10% are female. The mean acquittal rate in our sample is 12%. Cases where the defendant and judge share the same-religion constitute 81% of our sample, while 67% involve defendants and judges of the same gender.

As our empirical analyses employ different sub-samples constructed from these data sources to address each research question. For our analyses of heterogeneity in judicial bias (Section 4) and bias transmission through peer effects (Section 6), we use all criminal cases from across India, mentioned above. For our analysis of bias formation (Section 5), we focus on Uttar Pradesh, matching the judge's biographical data with the case records. The specific sample construction procedures, including the restrictions we impose and the identification strategies we employ, are detailed in their respective empirical sections to maintain clarity about the assumptions and variation underlying each analysis.

4. Heterogeniety in Judge Bias

In this section, we use the all-India sample described in Section 3 to document the substantial heterogeneity in judicial bias among judges.

A. Measuring Judge Bias

We begin by measuring the in-group bias (same-religion) of judges using quasi-random assignment to cases, a method first developed by Shayo and Zussman (2011), and argued by Ash et al. (2025) to be applicable in the context of Indian lower-level courts. We differ from their approach, which focuses on documenting the average level of in-group bias, by estimating judge-level disparities in acquittal rates between in-group and out-group defendants, similar to Dobbie et al. (2022).

To quantify the in-group bias, we estimate the judge fixed effect, which is captured by the differential probability of acquitting a defendant who shares the judge's religion, relative to acquitting a defendant of a different-religion. For case i, handled by judge j, in court c,

are discarded from our analysis.

[9]For consistency, we report the summary statistics for the cases among religions of the judge and the defendants, and charge type is not missing. These are the cases that are included in Section 4 analysis.

first filed in time period t, we model the probability of acquittal as:

$$Acquitted_{ijct} = \alpha_{ct} + \sum_k \theta_k Charge_i + \beta_{1,j} + \beta_{2,j} \times I[\text{InGroup}]_{ij} + \epsilon_{ijct}, \tag{1}$$

where $Acquitted_{ijct}$ represents the probability of acquitting the defendant, α_{ct} is the court-year-month fixed effect, and $Charge_i$ represents fixed effects for the specific charge under which the defendant was accused (e.g., IPC Section 302 for murder).

By controlling for the court-year-month fixed effect and the criminal charge of each case, we weaken the assumption of quasi-random assignment to consistently identify $\beta_{1,j}$ and $\beta_{2,j}$. Formally, we assume that cases within the same criminal charge category, processed by the same court, within the same month are randomly assigned to judges. As described in Section 2.D, while courtrooms are linked to specific police stations, judges regularly rotate across courtrooms within courts, limiting the predictability of which judge handles cases from any given police station.

While $\beta_{1,j}$ estimates judge effects on the acquittal probability, for cases where the judge and defendant do not share the same-religion, our coefficient of interest, $\beta_{2,j}$, measures the judge fixed effect on the differential probability of acquitting an in-group (same-religion) defendant, relative to an out-group defendant.

To validate the assumption of conditional quasi-random assignment of cases to judges, we follow Ash et al. (2025), and perform the following balance test:

$$\begin{aligned} Muslim\_Judge_i = \gamma_1 \times Muslim\_Defendant_i + \gamma_2 \times Female\_Defendant_i \\ + \sum_k \theta_k Charge_i + \alpha_{ct} + \epsilon_i \end{aligned} \tag{2}$$

$$\begin{aligned} Female\_Judge_i = \delta_1 \times Female\_Defendant_i + \delta_2 \times Muslim\_Defendant_i \\ + \sum_k \theta_k Charge_i + \alpha_{ct} + \epsilon_i, \end{aligned} \tag{3}$$

with the variables corresponding to those in Equation 1.

The coefficients of interest are γ_1, γ_2, δ_1, and δ_2, which represent the differential probability that cases by defendants with a given characteristic are assigned to judges of a given demographic characteristic. As seen in Table 1, the likelihood of a case being assigned to a Muslim or female judge isn't predicted by either of the defendant's characteristics. These results provide evidence for our identifying assumption that cases are assigned randomly to judges within a court.

Table 1: Assignment of Cases to Judges

	Muslim Judge	Female Judge
	(1)	(2)
Muslim defendant	0.001	0.001
	(0.001)	(0.002)
Female defendant	0.001	-0.000
	(0.001)	(0.001)
Mean Dep.	0.066	0.288
Observations	3,821,644	3,705,991

Notes: The dependent variable is a binary indicator of whether the case was assigned a Muslim judge in columns 1 or a female judge in 2. All regressions include charge section fixed effects and court-year-month fixed effects. Standard errors are clustered at the court level and appear in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sample Restriction

We impose two restrictions on our sample of judges for measuring bias, which are consistent with the sample restriction we impose in Section 6. First, we include only judges who have ruled on at least 100 cases and served for at least six months in a position they've held. Second, we focus on the judges who've served in at least two courts during the sample. This restriction allows us to observe the same judge operating in multiple institutional environments, which helps separate judge-specific bias from court-specific factors.

Empirical Bayes Shrinkage

We apply empirical Bayes shrinkage (EB) to our fixed effect estimates to address measurement error, following Bonhomme and Weidner (2022). For judges with fewer cases, estimates are noisier; the EB procedure shrinks these toward the sample mean while leaving precise estimates largely unchanged. We obtain standard errors for shrunk estimates via Bayesian bootstrap with 100 replications. Appendix Figure A2 shows that the shrunk and unshrunk distributions are nearly identical, confirming sufficient case volumes for most judges. All subsequent analyses use the shrunk estimates.

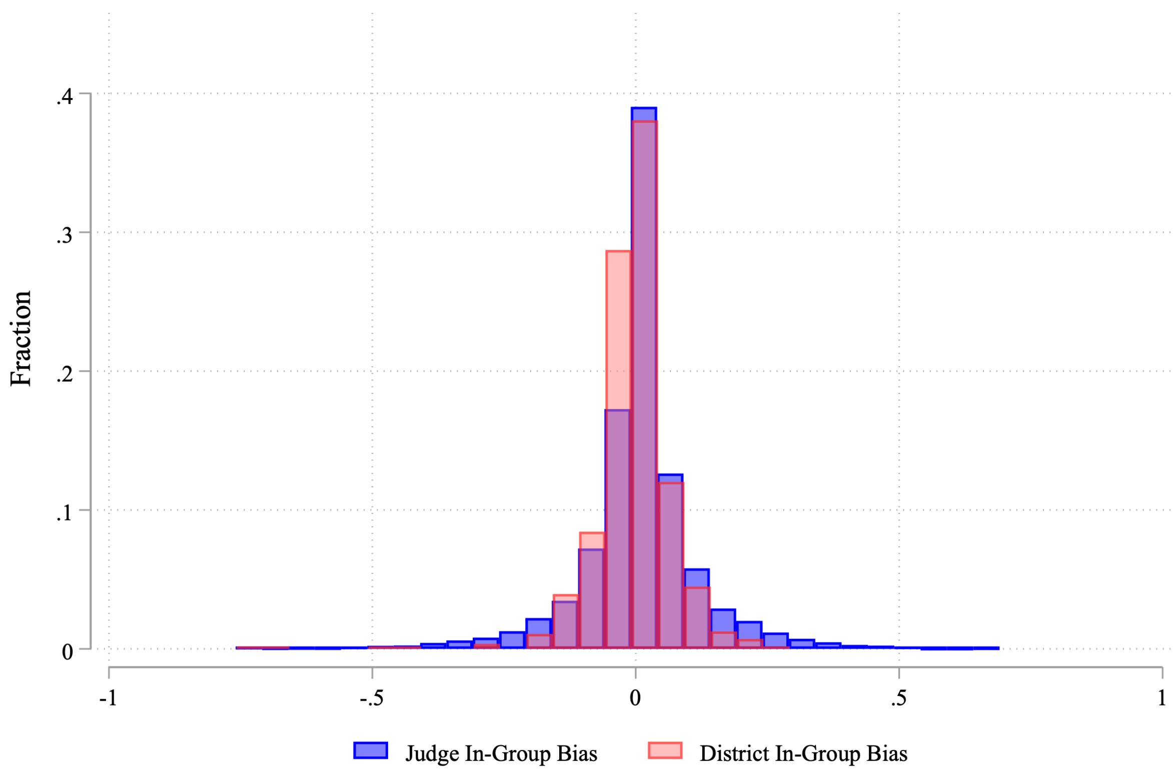
B. Distribution of Judge In-Group Bias

Figure 2 presents two distributions of in-group bias measures. The blue histogram shows our judge-level fixed effects ($\beta_{2,j}$) estimates from Equation 1, representing the differential likelihood each judge acquits same-religion versus different-religion defendants. The red histogram shows district-level measures, calculated as the simple difference in mean acquittal rates between in-group and out-group cases across India's districts.

Comparing these distributions reveals substantial within-court heterogeneity in judicial bias. The judge-level distribution is considerably wider than the district-level distribution, indicating that variation across judges within the same court exceeds variation across geographic units. Reassigning an in-group defendant from a judge at the 25th to the 75th percentile of the bias distribution increases their differential probability of acquittal by 7.5 percentage points, which is 45% of the mean acquittal rate.

The distribution is approximately symmetric around zero, indicating that both in-group bias and out-group bias exist among judges. Some judges systematically favor same-religion defendants (positive values) while others systematically favor different-religion defendants (negative values) by similar magnitudes. This heterogeneity in both the direction and magnitude of bias motivates our analysis of the drivers of these different patterns across judges.

Figure 2: Histograms of Judge and District Level In-Group Bias



Note: This figure plots the judge-level in-group bias fixed effects ($\beta_{2,j}$), as estimated in Equation 1. The district-level in-group bias measure is the difference in the average acquittal rate for in-group and out-group cases. To improve readability, we exclude extreme outliers beyond the 1st and 99th percentiles plus two standard deviations from this figure.

5. Formation of Bias

In this section, we examine how judicial bias forms through early career experiences using data from Uttar Pradesh, India's most riot-prone state. We leverage the quasi-random assignment of judges to courts and cases to judges to identify how exposure to communal riots and contact with out-group peers during the early stages of judges' careers shape their long-term discriminatory practices.

A. Identification Strategy

Our empirical model follows the modelling approach from Shayo and Zussman (2011) and uses both the quasi-random assignment of cases to judges and the assignment of judges to courts. We estimate the effect of early-career conflict exposure, on the probability of acquitting a defendant, and the differential probability that a judge acquits a defendant who shares their religion, with the equation:

$$\begin{aligned} Acquitted_{ijct} = \beta_1 I[\text{InGroup}]_i + \beta_2 Riot_j + \beta_3 I[\text{InGroup}]_i Riot_j \\ + \alpha_{ct} + \sum_k \theta_k Charge_i + \Omega X_j + \Phi W_i + \epsilon_{ijct}, \end{aligned} \tag{4}$$

where $Acquitted_{ijct}$ indicates whether judge j acquitted case i in courtroom c at time t.

In this specification, we include courtroom-by-year fixed effects (α_{ct}) and Act_i fixed effects to account for the conditional random assignment, alongside a set of judge characteristics (X_j) and case characteristics (W_i).[10] Baseline judge controls include gender and religion. We show robustness to richer controls, including the year of birth, year of entry, highest educational qualification, a home district dummy, and a same-gender-with-judge indicator. Defendant controls include the defendant's gender and religion. Charge fixed effects include 598 specific offense categories across the Indian Penal Code, Code of Criminal Procedure, and other applicable statutes.[11] Therefore, in our estimation β_1 will identify overall in-group bias in the sample, that is, the differential likelihood that a judge acquits a defendant with whom they share characteristics. β_2 would be the overall change in leniency of the judge due to exposure to a riot, and β_3 the additional in-group bias a judge exhibits due to that exposure.

While random assignment theoretically eliminates the need for additional judge controls,

[10] In the robustness subsection we show that our results, with judge level fixed effects are robust to any of the following set of controls: courtroom-year-month, court-year-month, and court-year.

[11] Examples include IPC Section 302 (murder), IPC Section 376 (rape), IPC Section 390 (robbery), and charges under special statutes like the Protection of Women from Domestic Violence Act.

we include them in our main estimates table to show robustness. The core coefficients remain stable across specifications while standard errors decrease, suggesting the controls absorb residual variation rather than addressing selection bias. We also present results with judge fixed effects, which cannot identify the direct leniency effect (β_2) but provides a stringent test of the bias estimates by comparing the same judge's treatment of different defendants.

In defining our outcome variable, we follow Ash et al. (2025) and broadly define favourable outcomes to include acquittals, dismissals, and other case dispositions that benefit defendants, coded as 1, versus all other outcomes, coded as 0. While this coding includes some ambiguous cases due to data entry inconsistencies, this measurement error should not systematically correlate with the interaction of judge and defendant's religion unless court clerks exhibit the same bias patterns as judges. Results are robust to the exclusion of ambiguous cases.

Validating Identifying Assumptions

We validate the quasi-random assignment of cases to judges at the courtroom-year level, extending the balance tests from Section 4 to include additional judge characteristics gathered from the judges' biographies: experience, age at entry, and exposure to riots during childhood and early career (Equations 2 and 3).

Appendix Table A3 presents these validation results. The courtroom-level specification yields precisely estimated null effects for 11 of 12 coefficients, all economically small. These results provide reassurance that, at this level of conditioning, cases are assigned randomly to judges within courtrooms in Uttar Pradesh.

We also validate that early career riot exposure is quasi-random by testing whether predetermined judge characteristics predict the probability of experiencing riots during the first five years of service. Specifically, we run the following validation test:

$$Riot_j = \beta_0 + \beta_1 MuslimJudge_j + \beta_2 FemaleJudge_j + \beta_3 AgeEnter_j + \epsilon_j, \qquad (5)$$

where $Riot_j$ is an indicator for riot exposure in the first five years of a judge's practice.

The results show that predetermined judge characteristics do not systematically predict exposure to riots. The joint F-statistic stands at 0.33 (p-value = 0.89), indicating that we cannot reject the null hypothesis of no association for all estimates. This provides suggestive evidence that early career riot exposure is effectively random, supporting the identification assumption.

Finally, another potential threat to interpretation is a systematic judicial system response to riots. If the administration strategically alters judicial personnel during riots, for example, by reassigning certain types of judges or changing court composition, the estimated effects could reflect institutional responses rather than direct bias formation from conflict exposure.

To test for this possibility, we examine whether riots trigger immediate or lagged changes in judicial personnel. Appendix Figure A3 presents results from an event study testing whether riots affect hiring and separation of judges across multiple dimensions: overall hiring and separation patterns, average experience levels, and pre-assignment leniency and bias characteristics of judges.[12] Specifically, we estimate the following stacked event study:

$$Personnel_{rdt} = \sum_{t\neq -1} \tau_t \cdot I[\text{EventTime} = t] + \rho_r + \phi_t + \varepsilon_{rdt}, \quad (6) \tag{6}$$

where $Personnel_{rdt}$ represents various personnel outcomes for riot r, district d, at time t. EventTime is measured in months relative to the riot occurrence, with $t = -1$ as the omitted reference period. The specification includes riot-specific fixed effects (ρ_r) and month-year fixed effects (ϕ_t).

The results show no statistically significant personnel changes in response to riots across any margin tested. Although the confidence intervals are wide, there appears to be no strategic reallocation of judges in response to these conflicts. Joint F-tests of the post-riot coefficients cannot reject the null hypothesis of no systematic response across all personnel dimensions tested.[13] This provides additional support for the assumption of quasi-random judge assignment to courts. Moreover, even if some personnel responses occurred, they would need to be both substantial and systematically biased in the direction of the conflict effects to drive the main findings.

B. Sample Construction

For the analysis of bias formation, we match the judge biographical data from the Allahabad High Court website with the criminal case data. After implementing a cleaning algorithm for both the criminal cases and biographies, and restricting to only informative names of judges,[14] we match the two datasets on exact names, ensuring each judge is matched to a unique judge in both datasets (following practices from Abramitzky et al. 2021).

[12]Leniency and bias are estimated on the judge's cases before hire or separation, restricting to those with more than 30 cases. The estimation of the fixed effects comes from the specification in Equation 1 with a judge fixed effects capturing leniency and judge fixed effects interacted with $I[\text{InGroup}]_i$ capturing bias.

[13]The following p-values are associated with F-tests of joint significance of post-riot coefficients: Hired Judges (p-value = 0.36), Separated Judges (p-value = 0.80), Experience of Hired (p-value = 0.81), Experience of Separated (p-value = 0.80), Leniency of Hired (p-value = 0.56), Leniency of Separated (p-value = 0.29), Bias of Hired (p-value = 0.48), Bias of Separated (p-value = 0.24).

[14]We impose the following restriction for names of judges to be considered: the string must contain more than one word, and one of the words needs to be more than two letters.

We then apply several sample restrictions to ensure data quality and proper identification. First, we retain only judges who have completed at least one full appointment, that is, those with recorded end dates for their first postings. This restriction addresses concerns that incomplete appointments may reflect non-random attrition, where judges who leave the system prematurely might differ systematically from those who remain. Second, we keep only cases with complete information on all key variables: case outcome, court and courtroom identifiers, dates, charges, and demographic characteristics of both defendants and judges. This reduces the sample to 653,075 cases. Third, we restrict the analysis to cases in which the filing judge and the deciding judge are the same individual, as cases in which they differ introduce measurement error into our bias estimates. Finally, to ensure consistency across specifications, we restrict the sample to observations that appear in our most stringent specification with judge fixed effects. This restriction automatically drops judges who would be singletons once judge fixed effects are included, ensuring all coefficient estimates are based on the same sample of cases. This results in a final sample of 839 judges and 316,885 cases.[15]

Panel B of Appendix Table A2 presents summary statistics for this sample. The majority of cases are decided on well after the early career phase of the judge, with 87% after the initial 5 years, and 38% more than a decade after enter. The defendant composition broadly reflects Uttar Pradesh's Muslim demographics, with 16% being Muslim, while women are substantially under-represented at only 8% of defendants. In terms of judicial outcomes, the acquittal rate is half of that of the all-India sample and stands at 5%. In terms of the judges' demographics, the Muslim share stands at 7% and women represent only 24%. The judges are also well educated and high achieving: while all judges hold a Bachelor of Laws (LL.B.), 38% graduated with First Class Honors and 34% earned a Master of Laws (LL.M.). The average judge was born in 1976 and entered the judicial system at age 32, approximately seven years after completing their most recent degree.

Early career experiences vary substantially across judges. During their first five years of service, 17% of judges experience at least one communal riot in their district. Professional interactions are frequent: judges average 20 colleagues with whom they work, yet out-group contact remains limited. Only half of judges work alongside out-group colleagues, 45% work with out-group peers of the same gender, and just 25% work with out-group colleagues within two years of age, highlighting the potential limitation of close relationships across religious lines.

[15] When considering all cases of judges, irrespective of missing non-essential demographics or those that didn't necessarily have an end date for their first appointment, the main estimates remain robust (includes 344,270 cases, and 939 judges).

C. Exposure to Conflict

Table 2 presents the main results: the effect of exposure to intergroup conflict for judges matters depending on the stage of their careers. We test whether riot exposure during three distinct career periods affects judicial bias: the first five years (Panel A), years 5-10 (Panel B), and years 10-20 (Panel C). To isolate the effects of exposure during each period, Panel A uses all cases decided by judges in our sample, Panel B restricts to cases decided after judges completed 5 years of service, and Panel C restricts to cases decided after 10 years of service.

Panel A shows that judges exposed to conflict in their early career are 0.8 percentage points more likely to acquit a defendant who shares their religion. This corresponds to a 17.5% increase in relative terms to a 4.8% acquittal rate. The effect is significant at the 1 percent level and is robust across all specifications. We find no effect of exposure on the average leniency of a judge. That is, the increase in in-group bias comes from increased acquittal rate for in-group defendants and not from a reduction of the overall acquittal rate.

In contrast to these results, Panels B and C show that exposure to these riots later in their careers does not lead judges to adjust their behavior. Mid-career exposure (Panel B), those with between 5 and 10 years of experience, shows no effect on either bias or leniency. Similarly, late-career exposure does not affect the judges' workplace practices. The null effect we find is not due to the lack of power, as the standard errors are comparable and mostly smaller than in Panel A. The 95% confidence intervals from Panels B and C also rule out any large effects. Riots in judges' mid-career would not result in more than 2% increase in ingroup bias, and in their late career, no more than 8.2%.

These findings demonstrate that experiences in judges' early careers shape their discriminatory behavior. While the same conflictual events shock the judge's local environment at different stages, it matters when they occur. The timing importance lends itself to the question of what other experiences in these formative years can offset the formation of bias or exacerbate the effect.

D. Contact with Out-group Judges

While early-career riot exposure creates persistent bias, we next examine how contact with out-group colleagues can counteract this formation process. Specifically, we examine how working alongside Muslim colleagues during the first five years of a Hindu judge's career affects their responses to communal riots.[16]

[16] We focus on Hindu judges because Muslim judges constitute only 7% of the judiciary, making exposure to Hindu colleagues nearly universal (95%).

Table 2: Exposure to Riots at Various Stages of Career

	Dependent Variable: Acquitted		
	(1)	(2)	(3)
Panel A: Early Career (0-5 years)			
316,899 observations; 839 judges			
$I[\text{InGroup}]_i$	0.001	0.001	0.000
	(0.002)	(0.002)	(0.002)
Riot (0-5 years)	0.000	-0.021	
	(0.010)	(0.016)	
$I[\text{InGroup}]_i$ × Riot (0-5 years)	0.008	0.009	0.010
	(0.003)	(0.003)	(0.003)
Panel B: Mid Career (5-10 years)			
275,882 observations; 733 judges			
$I[\text{InGroup}]_i$	0.001	0.001	0.000
	(0.002)	(0.002)	(0.002)
Riot (5-10 years)	-0.004	0.005	
	(0.007)	(0.010)	
$I[\text{InGroup}]_i$ × Riot (5-10 years)	-0.003	-0.003	-0.002
	(0.002)	(0.002)	(0.002)
Panel C: Late Career (10-20 years)			
121,143 observations; 412 judges			
$I[\text{InGroup}]_i$	-0.000	-0.000	-0.000
	(0.002)	(0.002)	(0.002)
Riot (10-20 years)	0.031	0.017	
	(0.015)	(0.039)	
$I[\text{InGroup}]_i$ × Riot (10-20 years)	-0.005	-0.004	-0.004
	(0.004)	(0.004)	(0.004)
Fixed Effects:			
Charge	✓	✓	✓
Courtroom-by-year	✓	✓	✓
Additional Controls		✓	✓
Judge FE			✓

Notes: The dependent variable is a dummy equal to 1 if the case was acquitted. The coefficient on In-group indicates whether the judge and defendant share the same religion. The coefficient on Riot (0-5 years) is a dummy equal to 1 if the judge experienced a communal riot during their first 5 years of judicial service (similarly for 5-10 and 10-20 years). The interaction term captures the "in-group bias" induced by riot exposure. All specifications include courtroom-by-year and charge fixed effects. Column 1 includes defendant gender and religion, and judge gender and religion controls. Column 2 adds year of birth, year of entry, highest educational qualification, a home district dummy, and a judge and defendant shared gender indicator. Column 3 adds judge fixed effects. Panel B restricts to cases decided after the judge had more than 5 years of experience, and Panel C to cases after 10 years. The mean acquittal rate for Panels A, B, and C is 4.8%, 4.7%, and 5.3%, respectively. Standard errors are robust, adjusted for clusters by judge, and appear in parentheses.

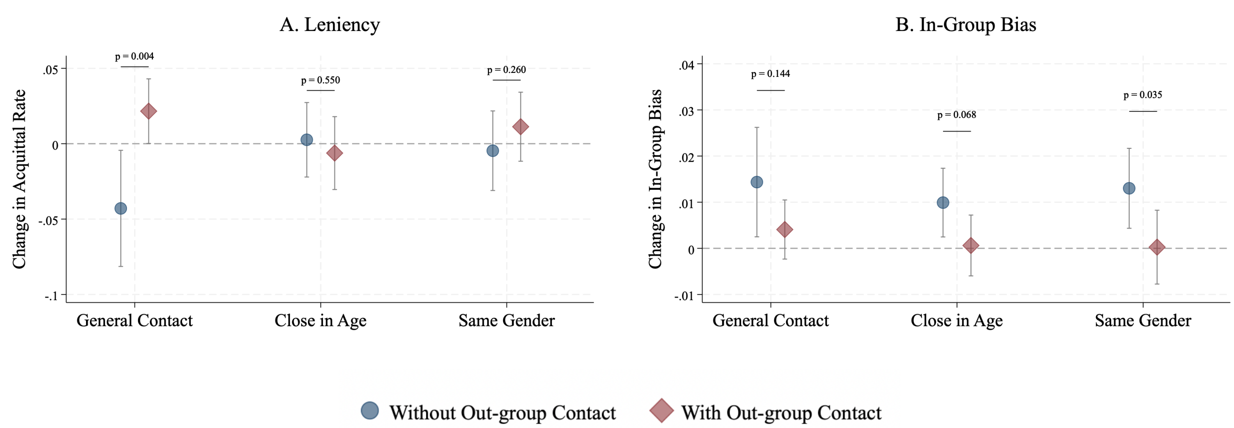
To examine these moderating effects, we extend Equation 4 by interacting all main coefficients with indicators for different types of out-group contact during judges' first five years:

$$\begin{aligned} Acquitted_{ijct} =& \beta_1 I[\text{InGroup}]_i + \beta_2 Riot_j + \beta_3 I[\text{InGroup}]_i \times Riot_j \\ &+ \beta_4 Contact_j + \beta_5 I[\text{InGroup}]_i \times Contact_j \\ &+ \beta_6 Riot_j \times Contact_j + \beta_7 I[\text{InGroup}]_i \times Riot_j \times Contact_j \\ &+ \alpha_{ct} + \sum_k \theta_k Charge_i + \Omega X_j + \Phi W_i + \epsilon_{ijct}, \end{aligned} \tag{7}$$

where $Contact_j$ indicates whether judge j worked with Muslim colleagues during their first five years.

Figure 3 show that this exposure is associated with a significant difference. Hindu judges who experienced early-career riots but had no Muslim colleagues developed extreme bias and stringency: they reduced acquittal rates for Muslim defendants to effectively zero while maintaining an average differential acquittal rate for Hindu defendants of 1.43 percentage points, significant at the 5 percent level (Appendix Table A4 shows the results in a table format).

Figure 3: The Effect of Early Career Riot Interacted with Out-Group Contact, Non-Muslim Judges



Notes: This figure shows how contact with Muslim colleagues during the first five years of a Hindu judge's career moderates riot exposure effects, based on the specification in Equation 7. Panel A plots the effect on leniency: β_2 for judges without contact, $\beta_2 + \beta_6$ for judges with contact. Panel B plots the effect on in-group bias: β_3 for judges without contact, $\beta_3 + \beta_7$ for judges with contact. Contact is measured three ways: any Muslim colleague (General Contact), within two years of age (Close in Age), and same gender (Same Gender). Standard errors are clustered by judge, with 95% confidence intervals shown.

In contrast, Hindu judges who did come in contact with a Muslim colleague behaved differently. Despite being exposed to the same communal riots, these judges showed no significant bias in their treatment of defendants. However, the moderating effect, while substantial, is not statistically different from the bias effect. This motivates us to examine whether specific types of contact, such as with colleagues of a similar age or the same gender, who may be more likely to form meaningful professional relationships, provide more effective protection against bias formation.

To interpret the magnitudes: among Hindu judges exposed to riots without Muslim colleagues (blue dots in Panel B, left), the estimated in-group bias effect is 1.43 percentage points (p=0.044). Combined with the negative leniency effect shown in Panel A (left), this indicates these judges became more stringent overall with seniority while developing strong in-group favoritism. In contrast, Hindu judges who worked with Muslim colleagues during riot exposure (red dots in Panel B) show near-zero in-group bias effects across all contact measures, with the strongest protection from same-gender contact (point estimate of 0.001, p=0.903).

Importantly, contact with out-group colleagues alone, without riot exposure, has no direct effect on either acquittal rates or religious bias. This suggests that intergroup contact operates specifically as a protective measure against bias formation during critical periods of conflict exposure, rather than promoting general changes in judicial behavior. As we will show in Section 6, the nature of peer influence depends crucially on group identity: while contact with out-group peers can prevent bias formation, exposure to biased in-group peers can transmit discriminatory practices.

Figure 3 further examines whether the social proximity of the inter-group contact matters by focusing on colleagues who share demographic characteristics. The results reveal that contact with similar peers provides the strongest protection against bias formation.

Hindu judges who work with Muslim colleagues of similar age (within two years) show virtually no riot-induced bias. The point estimate of 0.0017 percentage points is precisely estimated, with 95% confidence intervals ruling out any effect larger than 7% of the baseline acquittal rate. This corresponds to a complete elimination of the riot-induced bias effect and is marginally statistically different from judges without age-similar contact (p = 0.054).

The strongest results show for same-gender contact. Hindu judges who work with Muslim colleagues of the same gender develop no in-group bias whatsoever. The point estimate is exactly zero, and we also rule out large magnitude of bias. This contrasts sharply with riot-exposed judges lacking same-gender Muslim colleagues, who exhibit a riot-induced bias of a 27% increase in ingroup bias from the baseline acquittal rate mean.

These findings align with Allport (1954) theoretical work on the effectiveness of equal-status interactions on eliminating prejudice. In the judicial workplace, colleagues of similar age or same gender are more likely to have meaningful contact, discuss cases, and potentially develop a professional relationship that breaks down intergroup barriers. These results suggest that institutions should consider not mere diversity, but also strategic composition of early-career training and composition of equal-status peers that would have meaningful cross-group interaction.

E. The Reinforcement Mechanism

The nature of early career experiences lends itself to the question of how childhood experiences affect one's vulnerability to conflict. Motivated by the findings from Bharti and Roy (2023), we examine the judges who experienced these communal riots in their early childhood.

To examine this pattern, we extend Equation 4 by interacting all main coefficients with an indicator for childhood riot exposure during different developmental periods:

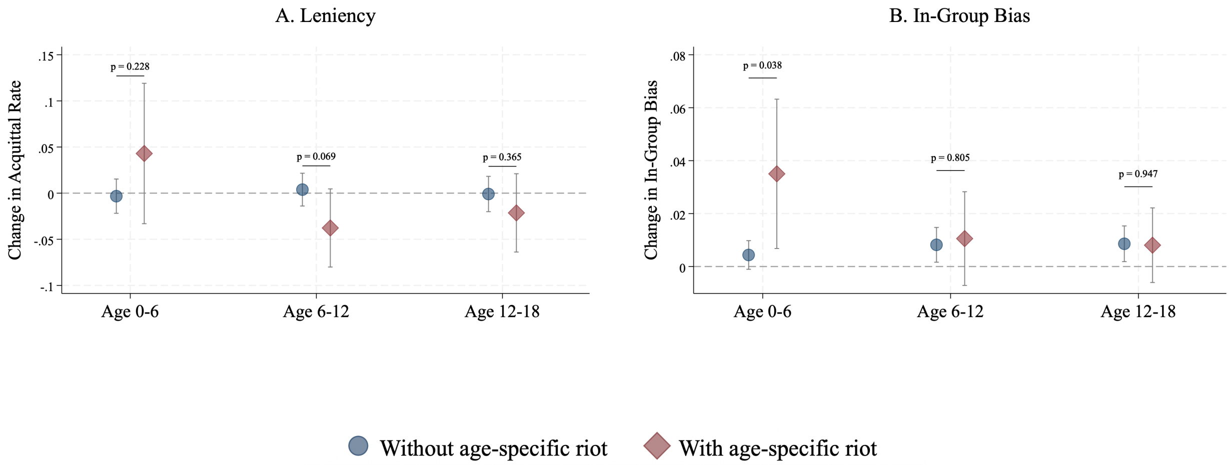
$$\begin{aligned} Acquitted_{ijct} =& \beta_1 I[\text{InGroup}]_i + \beta_2 Riot_j^{career} + \beta_3 I[\text{InGroup}]_i \times Riot_j^{career} \\ &+ \beta_4 Riot_j^{child} + \beta_5 I[\text{InGroup}]_i \times Riot_j^{child} \\ &+ \beta_6 Riot_j^{career} \times Riot_j^{child} + \beta_7 I[\text{InGroup}]_i \times Riot_j^{career} \times Riot_j^{child} \\ &+ \alpha_{ct} + \sum_k \theta_k Charge_i + \Omega X_j + \Phi W_i + \epsilon_{ijct}, \end{aligned} \tag{8}$$

where $Riot_j^{career}$ indicates riot exposure during the first five years of judicial service and $Riot_j^{child}$ indicates riot exposure during specified childhood periods (ages 0-6, 6-12, or 12-18). We estimate this specification separately for each childhood age range.

Figure 4 shows what we refer to as a "reinforcement mechanism." Judges who experienced communal riots in both early childhood (ages 0-6) and their first five years as judges show dramatically amplified bias compared to those exposed only during their careers. These twice-exposed judges increase their in-group bias by 3.06 percentage points, a 75% increase from the baseline mean. This effect is significant at the five percent level and statistically different from judges exposed to riots only in their early careers. The estimates remain unchanged when including judges' home district fixed effects, indicating that the reinforcement comes from the double exposure itself rather than growing up in a generally conflictual environment (Appendix Table A5 shows the results in a table format).

This reinforcement pattern provides additional evidence that these formative years play a significant role in the development of later workplace discriminatory practices. As early career experiences serve a professional formative period, early childhood appears to provide the underlying susceptibilities. These susceptibilities can be triggered and interact with

Figure 4: Childhood Riots Exacerbate the Reaction to Early Career Riots



Notes: This figure shows how childhood riot exposure reinforces early career riot effects, based on Equation 8. Panel A plots the effect on leniency: β_2 for judges without childhood exposure, $\beta_2 + \beta_6$ for judges with childhood exposure. Panel B plots the effect on in-group bias: β_3 for judges without childhood exposure, $\beta_3 + \beta_7$ for judges with childhood exposure. P-values shown are those corresponding to the statistical test examining the difference between the two coefficients. Separate estimates for three childhood periods: ages 0-6, 6-12, and 12-18. Standard errors are clustered by judge, with 95% confidence intervals shown.

later professional experiences. Important to note, judges who grew up in places that solely experienced riots when they were young do not seem to be more biased. Aligned with the findings of Bharti and Roy (2023), which used the bail decisions of these judges, judges who experienced early-childhood violence appear to be more stringent (see Appendix Table A6).

F. Persistence, Heterogeneity, and Distributional Consequences

Having established that early-career riot exposure increases in-group bias, we examine whether these effects persist throughout judges' careers, whether they vary by judge characteristics, and how they manifest across the distribution of the judiciary.

An important question is whether early-career riot effects persist over time. As 87 percent of cases in our sample are decided more than five years after judges enter the system, and 38 percent after more than a decade, most cases provide out-of-sample tests of persistence. Column 1 of Appendix Table A7 examines this by interacting our main specification with an indicator for cases decided more than 10 years after entry. The bias effect remains unchanged and statistically significant, showing no fade-out over time. Judges appear to become more stringent overall while maintaining their biased patterns.

Columns 2-5 of Appendix Table A7 explore whether judge characteristics moderate the riot exposure effect. Column 2 shows that judges who entered younger (below median age at entry) exhibit similar patterns to those entering later, suggesting that early professional experience rather than age drives the results. Column 3 examines academic achievement: judges with First Class honors show a marginally significant decrease in overall acquittal rates when exposed to riots ($p = 0.059$), but their religious bias remains unchanged. Female judges (Column 4) show no significant difference from male colleagues. Column 5 tests whether riots triggering government curfews have different effects. The bias coefficient is larger for curfew-imposing riots (0.014 versus 0.007) but not statistically different, though limited statistical power may explain this, given that only 10 of 62 riots since the year 2000 resulted in curfews. These results suggest that early-career riot exposure affects judges through similar mechanisms regardless of individual traits.

Beyond the individual-level persistence and heterogeneity in exposure to conflict, an important aspect is how these events reshape the distribution of judicial bias across the system. Appendix Table A8 compares the distribution of judge-level bias estimates between riot-exposed and unexposed judges, indicating an asymmetric effect that varies across the distribution. The center (25th-75th percentiles) shows minimal movement, with the median remaining at 0 for both groups. The moderate tails exhibit rightward shifts: the 10th and 90th percentiles both increase by approximately 3 percentage points. The extreme tails, however, reveal pronounced asymmetry: the 5th percentile shifts by 7.7 percentage points, while the 95th percentile shifts by 1.5 percentage points. This differential movement transforms the distribution from near-symmetric to severely right-skewed. These patterns suggest that riot exposure operates primarily by eliminating extreme out-group favoritism while also increasing in-group bias at the tails, producing a judicial system with systematic asymmetries favoring same-religion defendants.

G. Robustness

Our main results are robust to alternative specifications and sample definitions. In the Appendix Section A.1 we show that: (1) effects are similar when defining exposure by first appointment rather than first five years; (2) riot exposure does not affect gender bias, confirming the effect is specific to the Hindu-Muslim dimension; (3) estimates remain stable across alternative fixed effect structures; and (4) results are not driven by individual outliers. Full details of these robustness checks appear in Appendix Tables A9-A10 and Appendix Figure A4.

6. Transmission of Bias

In this section, we test whether judicial bias spreads through peer interactions within courts. Using pre-move judge-level bias measures, we leverage the quasi-random timing of judge movements across courts to identify how exposure to biased peers affects incumbent judges' decision-making.

A. Identification Strategy

After recovering the judge fixed effects, as in Section 4, we exploit the quasi-random movements of judges across courts to estimate the spillover effects on the in-group bias of non-mover judges. Our mover design follows closely the specifications in labor and organisational economics (Card et al., 2013; Fenizia, 2022). However, one crucial difference in our study is that we are interested in the judicial bias of non-mover judges who became colleagues of mover judges. This approach in estimating peer effects in judicial bias overcomes common challenges, such as the Reflection Problem, when exploiting group-level variation (Manski, 1995).

We compare case outcomes of in-group cases versus out-group cases for judges in courts exposed to new, high-bias judges relative to those who are in courts that were not exposed to a new judge. Our parameters of interest are the differential effect of exposure to a highly biased judge on the incumbent judges' probability of acquitting a defendant when they are quasi-randomly assigned to an in-group (same-religion) case, relative to when they are assigned to an out-group (different-religion) case.

To flexibly estimate changes in judge propensity to acquit defendants over time, and examine dynamics of the potential treatment effects of exposure to high-bias judges, we estimate the following event-study model:

$$\begin{aligned} Acquitted_{ijct} =& \rho_t + \rho_t \times I[\text{InGroup}]_i + \sum_k \theta_k Charge_i + \sum_k \theta_k Charge_i \times I[\text{InGroup}]_i \\ &+ \sum_g \tau_{1,g} I[\text{HighBias}]_c \times I[\text{EventTime} = g] \\ &+ \sum_g \tau_{2,g} I[\text{HighBias}]_c \times I[\text{EventTime} = g] \times I[\text{InGroup}]_i \\ &+ \delta_j + \delta_j \times I[\text{InGroup}]_i + \epsilon_{ijct}, \end{aligned} \tag{9}$$

where $Acquitted_{ijct}$ indicates whether judge j acquitted case i in court c at time t (month-year). $I[\text{HighBias}]_c$ is an indicator for whether the mover judge into court c had high religious bias (top tercile of the bias distribution). $I[\text{InGroup}]_i$ indicates whether the defendant shares their religion with the judge.

The triple differences setup allows us to flexibly control for various potential confounding shocks. First, we include both month-year fixed effects and month-year by in-group case fixed effects (ρ_t, and $\rho_t \times I[\text{InGroup}]_i$), to control for potential differential trends in case outcomes in general and across in-group cases. Second, we control for the type of charge and charge by in-group ($Charge_i$, and $Charge_i \times I[\text{InGroup}]_i$), to allow for differences in treatment of the same offenses by whether the defendant is of the same-religion. Lastly, we control for judge fixed effects and judge by in-group case fixed effects (δ_j and $\delta_j \times I[\text{InGroup}]_i$), allowing for permanent differences, in leniency and bias, across judges to vary in general and by whether the case involves a same-religion defendant.

We also summarize the treatment effect of exposure to a judge with a particular level of in-group bias (in a given tercile) in the following static difference-in-differences model:

$$\begin{aligned}
Acquitted_{ijct} = &\rho_t + \rho_t \times I[\text{InGroup}]_i + \sum_k \theta_k Charge_i + \sum_k \theta_k Charge_i \times I[\text{InGroup}]_i \\
&+ \beta_1 Pre_{it} + \beta_2 Pre_{it} \times I[\text{InGroup}]_i \\
&+ \beta_3 PostShort_{it} + \beta_4 PostShort_{it} \times I[\text{InGroup}]_i \\
&+ \beta_5 PostMedium_{it} + \beta_6 PostMedium_{it} \times I[\text{InGroup}]_i \\
&+ \beta_7 PostLong_{it} + \beta_8 PostLong_{it} \times I[\text{InGroup}]_i \\
&+ \delta_{jc} + \delta_{jc} \times I[\text{InGroup}]_i + \epsilon_{ijct}, \qquad (10)
\end{aligned}$$

where $PostShort_{it}$, $PostMedium_{it}$, and $PostLong_{it}$ correspond to the first, second, and third 6 months following the move, respectively. These coefficients capture the immediate, short- and medium-term effects of exposure to biased judges.

B. Sample Construction

Our analysis relies on cross-court movements of judges. We impose four main sample restrictions for our analysis sample.

First, we only include cross-court movements where the mover judge had ruled on at least 100 cases before relocation and was appointed for at least 6 months.[17] This restriction ensures that we have a sufficiently large sample to estimate the in-group bias of mover judges with precision. Judges with too few cases would have noisy bias estimates that could attenuate our results.

[17] This aligns with cut-off choices made by Dobbie et al. (2022) and Bharti and Roy (2023).

Second, as our case-level data achieves universal coverage of all states from 2014 and continues until December 2019, we focus on judge moves from January 2015 to December 2018. This ensures we observe case outcomes in treated courts before and after a given move when estimating event study regressions, allowing us to establish baseline trends and track post-treatment dynamics.[18]

Third, we exclude courts that received over 13 mover judges over four years, which is the 90th percentile of the number of moves. These courts receive a new mover judge on average every 3 months; as such, we are unable to define a long enough event-time window. The interpretation of treatment effects is also complicated by frequent moves and lagged effects of past exposure (for a detailed discussion, see Borusyak et al. 2024). In our sample, both the mean and median number of incoming judges are one per court every year.

Fourth, we restrict our sample to cases handled by incumbent judges who have ruled cases in the treated courts before a given treatment event. This restriction ensures that our estimates capture peer effects among judges with court presence, thereby excluding new incoming judges who arrive at after the mover judge. Finally, we include cases in courts that never received a mover judge in our sample period as a never-treated/pure control group in our dataset. Appendix Table A11 shows that treated and never-treated courts are statistically similar across all observable characteristics, validating our use of never-treated courts as a control group.

Stacked Dataset

We construct a stacked dataset to address the presence of multiple treatments for each court and alleviate issues in two-way fixed effects regressions due to the staggered treatment of courts over time (Borusyak et al., 2024). Each stack of the dataset includes cases prior to and after each court move/treatment event as a separate dataset, and cases in never-treated courts. Then, we combine all stacks, where each stack corresponds to a specific treatment event. This accommodates the presence of courts that were treated more than once. For instance, if a court had two mover judges, they would appear twice in the dataset. Each time, only the cases within the defined event window appear.

Panel C of Table A2 presents summary statistics for our case-level dataset utilized in the main analysis. Our sample comprises 1,939 judges who transferred between courts (providing identifying variation in judge-level exposure to biased judges) and 7,935 judges who were either exposed to a biased judge or remained untreated during the sample period. The mean acquittal rate in our sample is 17 percentage points, where both defendants' and judges'

[18]Our results are robust to focusing on judge moves (treatments) over shorter periods, such as 2016-2018 and 2017-2018.

shares consistent with the all-India sample (Panel A).

Panel A of Appendix Figure A5 presents the distribution of the month-year in which courts received new judges (which are our events of interest). Around 30% of events in our sample happened before June 2017, with the remaining 75% happening between July 2017 and January 2019. Complementing this, Panel B shows the distribution of cases relative to the judges' arrival date, confirming that case composition remains stable across event time and alleviating concerns that compositional changes drive our results.

Panel C of Appendix Figure A5 presents the distribution of the number of months a mover judge stayed in the court they moved to before making another move. The mean and median number of months is 20 months. While there is some variability in the amount of time mover judges stayed in a given court, and we restrict it to 6 months, our results are robust to a lower cutoff.

Lastly, Panel D of Appendix Figure A5 presents the distribution of the number of times a given court was exposed to a mover judge. The mean and median numbers are both three treatments, with 30% of courts in our sample only experiencing one move.

C. Validation

Next, we validate that our pre-move bias measures predict post-move judicial behavior. For each mover judge, we separately estimate in-group bias fixed effects using only pre-move cases and only post-move cases, following Equation 1. Regressing post-move bias on pre-move bias yields a slope of 0.921 (SE = 0.004), showing strong persistence of judicial bias across court assignments.[19] This validates our use of pre-move bias measures to classify mover judges in the transmission analysis that follows.

D. Dynamic Effects of Exposure

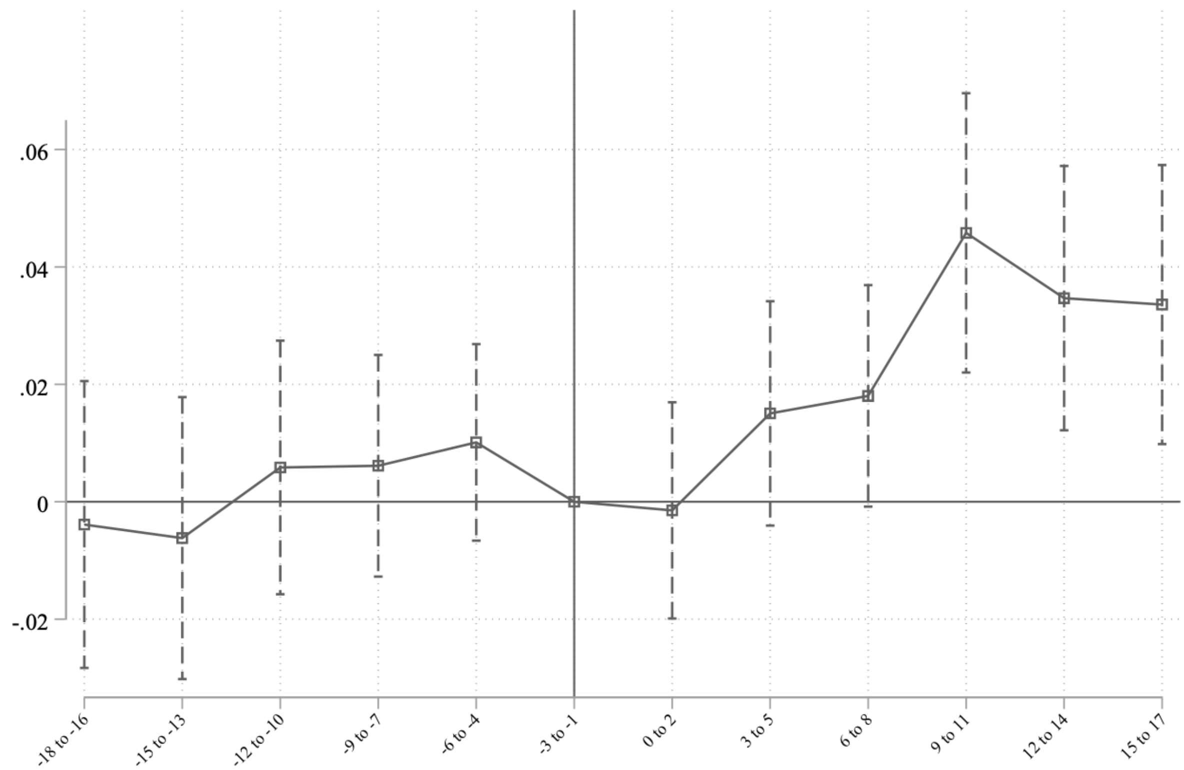
In this section, we present our estimates of the effects of exposure to biased mover judges on non-mover/incumbent judges' behaviors. We begin by presenting the dynamic effects of exposure to mover judges on the incumbent judges' decision-making. To do so, we examine changes in judges' in-group bias among judges exposed to a judge in each tertile of the judge-bias fixed-effects distribution, compared with judges in courts without a mover judge.

Figure 5 shows the coefficients of the triple interaction terms from Equation 9 for incumbent judges in courts receiving a judge in the top tercile for bias. We observe a strong,

[19] Restricting the sample to judges who ruled on more than 100 cases prior to and after a respective move increases the association to 0.967 (SE=0.003).

positive effect of exposure to a high-bias judge (top tercile) on in-group bias, beginning 3 months after the incoming judge's move. Between 3 and 8 months after exposure, in-group bias increases by around 2 percentage points; between 9 and 17 months after exposure, in-group bias increases by around 4 percentage points. This indicates that judges exposed to highly biased colleagues increasingly favor defendants who share their religion, compared to those who do not.

Figure 5: Dynamic Effect of Exposure to Highly Religious-Biased Mover Judges on Religious In-Group Bias



Note: This graph displays the coefficients of the triple interaction terms from Equation 9. Each coefficient represents the change in the differential probability of acquitting a defendant of a judge's own religious group compared to a defendant from a different religious group. The figure shows the dynamic effect for courts receiving a judge in the top tercile for bias. Standard errors are clustered at the level of the incumbent judge, and the graph includes the corresponding 95% confidence intervals.

The magnitude is economically meaningful: the mean acquittal probability in our sample is around 0.173, so the effect indicates that judges were almost 20% more likely to acquit their in-group defendants than out-group defendants after being exposed to a high-bias judge. Notably, the effect seemed to persist up to 17 months after the mover judge arrived in the court.

Next, we summarize the treatment effect by estimating regressions that examine the effect of exposure to a mover judge in three post-move periods. Table 3 presents the regression results, which qualitatively align with our previous findings. The specification follows Equation 10, where the base coefficients (Post) capture the baseline effect on out-group defendants, while the interaction terms (Post × $I[\text{InGroup}]_i$) directly measure the differential acquittal rate for in-group defendants.

Specifically, in column (1), we can see that exposure to a high-bias judge (top tercile) reduces acquittal rates for out-group defendants, but no change in the acquittal rate for in-group defendants. The effect intensifies over time: there is no significant effect in the first six months. In the subsequent period (6-11 months post-treatment), out-group defendants experience a 3.9 percentage point decline in acquittal rates. The interaction coefficient of 3.2 percentage points indicates that in-group defendants are largely insulated from this decline, and the 3.2 percentage point increase in bias (18.5% of the baseline mean acquittal rate) comes entirely from reduced acquittals for out-group defendants. In the final period (12-17 months after treatment), the pattern persists: a 3.1 percentage-point decline for out-group defendants, offset by a 3.6 percentage-point interaction term for in-group defendants.

Looking at columns (2) and (3) of Table 3, we see no significant changes in differential acquittal probability after exposure to a middle tercile judge (whose judge-fixed effects are, on average, close to zero). We also observe no significant effect of exposure to a bottom tercile judge (whose judge fixed effect would be negative, suggesting they were less likely to acquit same-religion defendants compared to different-religions).

Non-linearities in Treatment Effects

Table 3 suggests that incumbent judges only experienced a change in their decision-making when exposed to the most biased judges (in the top tercile of the in-group bias fixed effects distribution), suggesting non-linearities in treatment effects. To further explore whether this is the case, we estimate the effect of exposure to a mover judge whose bias level was at different top fractions of the fixed effects distribution.

In Appendix Figure A6 we plot the differential change in the in-group bias rate for judges 0-8 months and 9-17 months following the arrival of a judge who belongs to different top shares of the religious bias distribution, as specified in Equation 10. The figure presents coefficients for different top fractions of the distribution. The estimates show that the magnitude of the effect increases as we focus on mover judges with higher in-group bias, suggesting potential nonlinearity in the effects of exposure to biased peer judges.

Focusing on the mover judges who fall among the highest top shares of the bias distribution suggests a more pronounced long-term effect. Specifically, the increase in in-group

Table 3: Difference-in-Differences Estimates, by Terciles Mover Judge of Religion Bias

	Dependent Variable: Acquitted		
	Top Tercile	Middle Tercile	Bottom Tercile
	(1)	(2)	(3)
Pre	-0.010	0.006	0.012
	(0.010)	(0.014)	(0.009)
Pre $\times I[\textbf{InGroup}]_i$	0.008	-0.009	-0.009
	(0.009)	(0.014)	(0.009)
Post			
(0-5 months)	-0.009	-0.008	0.008
	(0.010)	(0.014)	(0.008)
(6-11 months)	-0.039	0.029	0.005
	(0.012)	(0.019)	(0.011)
(12-17 months)	-0.031	0.009	-0.016
	(0.012)	(0.018)	(0.012)
Post $\times I[\textbf{InGroup}]_i$			
(0-5 months)	0.010	-0.005	-0.009
	(0.009)	(0.014)	(0.009)
(6-11 months)	0.032	-0.023	-0.008
	(0.011)	(0.018)	(0.010)
(12-17 months)	0.036	-0.013	0.017
	(0.011)	(0.017)	(0.011)
Mean Dep.	0.173	0.128	0.165
Clusters	5,157	4,803	5,174
Observations	1,498,455	1,499,533	1,504,708

Notes: The dependent variable is a binary indicator of whether the defendant in a case was acquitted. The pre-treatment period includes all cases a year and a half before a moving judge arrives in court. The control variables used are those specified in Equation 10. Standard errors are clustered at the incumbent judge level and appear in parentheses.

bias during the last nine months of our analysis (9-17 months) following the judge's move is larger in magnitude, increasing from a non-significant decline of 2 percentage points for judges above the median to 5 percentage points for those in the top quintile of judicial bias.

E. Mechanisms

In this section, we investigate the mechanisms underlying our main estimates. We focus on three complementary questions: First, under what conditions does bias transmission occur? We examine proxies for increased social interaction opportunities, such as court size and shared gender. Second, how do different types of judges respond to biased colleagues? We examine whether transmission varies by incumbents' pre-existing bias levels. Third, does seniority structure the flow of influence? We test whether behavioral spillovers flow from senior to junior judges.

Proximity and Shared Characteristics

We begin by examining conditions that facilitate peer transmission. Table 4 presents difference-in-differences estimates split by court characteristics and judge relationships.

Columns 1 and 2 compare courts above and below the median size of eight judges. Judges in smaller courts experience large and significant reductions in differential acquittal rates following exposure to biased colleagues: after 6 to 11 months, they become 4.1 percentage points less likely to acquit in-group defendants than out-group defendants, and after 12 to 17 it increases to 4.8 percentage points. In contrast, judges in larger courts show no statistically significant effect at any period.

Columns 3 and 4 examine whether shared gender affects transmission. The effects concentrate among same-gender pairs: after 6 to 17 months, judges exposed to same-gender biased colleagues become 3.5 percentage points less likely to acquit out-group defendants, while different-gender pairs show small and statistically insignificant effects.

These results provide suggestive evidence that interactions between the judges are those that drive the observed peer effects. Both the size of the court and the shared gender of the judges increase opportunities for meaningful interaction, amplifying behavioral transmission.

Heterogeneous Transmission by Incumbent Bias

Next, we examine how transmission varies by incumbent judges' pre-existing bias levels. Columns 5 to 7 of Table 4 split the sample by incumbent judges' in-group bias tercile, measured using their judge-level fixed effects estimated from cases that were filed before the arrival of the mover judge.[20]

The results reveal asymmetric transmission operating through two distinct channels. Column 5 shows that incumbent judges in the top tercile, those who already exhibited in-group bias (average bias of 8.7 pp), respond by solely reducing acquittal rates for all

[20]These judge-level fixed effects are not Bayesian shrunk.

Table 4: Suggestive Mechanisms, Top Tercile Difference in Differences Estimates

	Dependent Variable: Acquitted						
	Court Size		Gender		Incum. Bias Tercile		
	Small	Large	Same	Different	Top	Middle	Bottom
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post							
(0-5 months)	-0.009	-0.006	-0.014	0.009	-0.010	-0.010	-0.007
	(0.014)	(0.013)	(0.012)	(0.016)	(0.017)	(0.014)	(0.018)
(6-11 months)	-0.027	-0.042	-0.043	-0.043	-0.037	-0.008	-0.068
	(0.018)	(0.015)	(0.015)	(0.019)	(0.019)	(0.019)	(0.021)
(12-17 months)	-0.026	-0.024	-0.042	-0.024	-0.038	0.010	-0.056
	(0.020)	(0.016)	(0.015)	(0.020)	(0.021)	(0.022)	(0.020)
Post $\times I[\text{InGroup}]_i$							
(0-5 months)	0.022	-0.003	0.010	-0.006	0.008	0.014	0.004
	(0.014)	(0.013)	(0.012)	(0.015)	(0.015)	(0.014)	(0.018)
(6-11 months)	0.041	0.019	0.035	0.014	0.003	0.020	0.058
	(0.016)	(0.014)	(0.013)	(0.018)	(0.016)	(0.017)	(0.019)
(12-17 months)	0.048	0.016	0.037	0.017	0.018	0.010	0.063
	(0.020)	(0.015)	(0.013)	(0.019)	(0.019)	(0.020)	(0.019)
Mean Dep.	0.159	0.180	0.165	0.161	0.178	0.119	0.162
Clusters	2,051	3,404	4,155	3,313	3,258	2,724	3,127
Observations	493,671	1,004,494	1,120,514	844,858	844,846	735,811	818,992

Notes: The dependent variable is a dummy variable of whether the defendant was acquitted or not. Small courts correspond to those with fewer than 8 judges (corresponding to the median size court). Controls included are those that appear in Equation 10. Standard errors are clustered at the incumbent judge level and appear in parentheses.

defendants by 3.7 percentage points. That is, these biased judges do not change their level of bias but become more stringent in their rulings.

Column 6 shows that judges in the middle tercile, those with near-zero pre-existing bias, exhibit no significant response in their judicial behavior. In contrast, Column 7 shows that incumbent judges in the bottom tercile, who previously exhibited out-group favoritism (average bias of -9.4 pp), are those driving the effect observed. After 6 to 11 and 12 to 17 months, these judges reduce their out-group acquittal rates by 6.8 and 5.6 percentage points, respectively. Actively eliminating any prior out-group favoritism.

Thus, the mechanism through which transmission operates is asymmetrical: biased peers eliminate out-group favoritism among out-group-favoring judges, while out-group-favoring peers fail to moderate biased judges. This norm-transmission mechanism mirrors the distributional effects documented in Section 5. Just as riot exposure primarily eliminates out-group favoritism (truncating the lower tail) while moderately affecting in-group bias, peer transmission constrains out-group favoritism. Both the formation (riots) and transmission (peers) channels operate through establishing and enforcing boundaries on acceptable judicial behavior: boundaries that systematically disadvantage out-group defendants on average.

Career Stages and Seniority

Next, we examine whether incumbent judges respond differently to incoming colleagues based on seniority, a test of the mechanism through which judicial peer effects might operate. We proxy experience using the year of entry of a judge to our dataset, splitting judges within each court as junior (entered after the court's mean entry year) or senior (entered before the court's mean entry year).

In Table 5, columns 1 and 2, we split the sample by incumbent judge seniority when receiving highly biased colleagues. Junior incumbents (column 1) show substantial and immediate increases in differential acquittal rates: 4.6 and 3.6 percentage points at 6-11 and 12-17 months, respectively. Although marginally significant, senior incumbent judges (column 2) also show an increase in the long term (3.1 pp at 12-17 months).

We then examine relative seniority between incumbent and incoming judges. Columns 3 and 4 split the sample based on whether the incoming judge is more or less experienced than the incumbent. When incumbents are less senior than the incoming judge (column 3), we observe significant reductions in differential acquittal rates of 3.9 and 3.6 percentage points in the medium and long term. In contrast, column 4 shows that when incumbents are more senior than the mover, effects are smaller and only marginally significant in the long term (3.0 percentage points). Together, these results suggest asymmetry in judicial peer effects, with behavioral spillovers flowing primarily from senior to more junior judges.

F. Robustness

Our main results on bias transmission are robust to several potential concerns. In Appendix Section A.2 we show that: (1) judge moves are likely not to be systematically related to pre-move trends in court case outcomes or characteristics; (2) case assignment patterns do not change systematically when biased judges arrive; (3) results are unchanged when controlling for state-by-year fixed effects; (4) permutation tests confirm effects are not due to chance assignment of bias levels to courts; and (5) our findings are robust to potential violations of parallel trends using recent sensitivity analysis methods (Roth, 2022; Rambachan and Roth, 2023). Full details of these robustness checks appear in Appendix Tables A12-A14 and Appendix Figures A7-A8.

Table 5: Career Stages of the Judges, Top Tercile Difference in Differences Estimates

	Dependent Variable: Acquitted			
	By Incumbent Seniority		By Relative Seniority *Incum. relative to Incom.*	
	Junior (1)	Senior (2)	Junior (3)	Senior (4)
Post				
(0-5 months)	-0.008	-0.009	-0.010	-0.008
	(0.016)	(0.012)	(0.014)	(0.013)
(6-11 months)	-0.061	-0.028	-0.050	-0.037
	(0.018)	(0.015)	(0.016)	(0.018)
(12-17 months)	-0.052	-0.018	-0.038	-0.032
	(0.017)	(0.017)	(0.016)	(0.019)
Post $\times I[\text{InGroup}]_i$				
(0-5 months)	0.002	0.014	0.014	0.005
	(0.016)	(0.011)	(0.013)	(0.013)
(6-11 months)	0.046	0.016	0.039	0.020
	(0.017)	(0.013)	(0.015)	(0.016)
(12-17 months)	0.036	0.031	0.036	0.030
	(0.016)	(0.016)	(0.015)	(0.018)
Mean Acquittal Rate	0.163	0.161	0.172	0.150
Judges	3,364	3,611	3,970	3,497
Cases	922,483	1,014,417	1,031,584	905,267

Notes: The dependent variable is a dummy variable of whether the case's defendant was acquitted or not. Columns 1-2 split incumbent judges by seniority: Junior (Senior) incumbents are those who entered the court after (before) the mean entry year within that court. Columns 3-4 split cases by relative seniority: column 3 includes incumbent judges who entered after the incoming judge (incumbent is junior relative to incoming), while column 4 includes cases where the incumbent judge entered the judicial system before the incoming judge (incumbent is senior relative to incoming). Controls included are those that appear in Equation 10. Standard errors are clustered at the incumbent judge level and appear in parentheses.

7. Concluding Remarks

In this paper, we provide a comprehensive account of judicial bias across its lifecycle: from its formation through early-career experiences to its transmission through peer interactions. Using data from India's criminal judicial system, we leverage the quasi-random assignment of cases to judges and the rotation of judges across courts to trace how discriminatory practices emerge and spread in professional institutions.

Our analysis reveals three main findings. First, we document substantial heterogeneity in judicial bias: within the same court, between-judge variability in religious in-group bias is larger than across-district variability. Second, we identify the early-career period as critical for bias formation. Exposure to Hindu-Muslim communal riots during judges' first five years significantly increases religious in-group bias throughout their careers. However, working alongside out-group peers during these formative years counteracts riot exposure effects, especially through meaningful interactions with colleagues of similar age or the same gender. Third, we show that bias spreads through peer interactions. Exposure to highly biased colleagues significantly increases incumbent judges' own in-group bias, with transmission concentrated in smaller courts and flowing primarily from senior to junior judges.

Together, these findings reveal an important asymmetry in how social interactions shape judicial behavior: contact with out-group peers during formative years provides protection against bias formation, while exposure to biased in-group peers transmits discriminatory practices. This asymmetry depends fundamentally on group identity. When judges develop bias from riot exposure, their out-group colleagues can counteract this formation without themselves becoming biased: a Muslim colleague can prevent a Hindu judge from developing pro-Hindu bias precisely because they don't share that group identity. In contrast, bias transmission from senior colleagues occurs through in-group channels: judges are influenced by biased peers with whom they share religion and gender, suggesting that role modeling and persuasion operate within group boundaries. The moderating peers (out-group) are distinct from the transmitting peers (in-group), explaining why prevention and transmission operate through different social channels rather than reinforcing each other.

Beyond these different social channels, both formation and transmission mechanisms operate through asymmetric pruning of the distribution. Riot exposure and peer influence share a common feature: they systematically eliminate out-group favoritism among judges, moving them toward neutrality. This one-directional correction likely stems from out-group favoritism being an effortful state that requires institutional support, which collapses under normative pressure.

These findings have important implications that go beyond judicial settings. The mechanisms we identify, such as critical periods of professional development, the protective role of equal-status out-group contact, and the spread of bias through in-group peer networks, likely play a role in many high-stakes decision-making areas where discrimination has been found, including education (Lavy et al., 2022; Alesina et al., 2024), healthcare (Angerer et al., 2019), and law enforcement (Goncalves and Mello, 2021).

For workplace design more broadly, our results suggest several concrete principles. First, the early-career period is a critical window during which professional norms and practices are established. Organizations should view this period not merely as skills development but as identity formation. Exposure to organizational shocks or crises during this critical window can have lasting effects on worker behavior, making the composition of early-career environments particularly consequential.

Second, diversity alone is insufficient. Our findings show that meaningful cross-group interaction requires social proximity: colleagues of similar status, age, or other demographic characteristics are more likely to form relationships that break down intergroup barriers. This suggests that effective diversity initiatives should focus on creating opportunities for favorable interactions (Allport, 1954). For example, cohort-based training programs, mentorship structures that pair demographically similar cross-group colleagues, or team compositions that facilitate genuine collaboration may be more effective than merely targeting representation.

Third, organizational hierarchy creates pathways for both positive and negative cultural transmission. Since influence flows primarily from senior to junior workers, the composition of senior peers becomes critical, particularly in units with many early-career employees. Concentrations of biased senior workers can create self-reinforcing environments where discriminatory practices become normalized and transmitted across cohorts. This highlights the importance of both selection and development of senior staff, not just entry-level hiring.

This paper opens an agenda to study social interactions in shaping workplace discrimination more broadly. Our analytical framework, exploiting quasi-random assignment of tasks and worker rotation, may apply directly to other professions with similar institutional structures (such as health, law enforcement, and education), along with the broader mechanisms we identify, which are likely to be present in other settings. Further research should examine how these formation and transmission mechanisms interact with organizational features such as promotion systems, performance evaluation, and explicit anti-discrimination policies. Understanding these interactions is important for the future of designing effective interventions to reduce bias in high-stakes professional decisions across contexts.

References

Abramitzky, R., N. Ben-Porath, V. Lavy, and M. Palgi (2025). Economic crisis and disillusionment from socialism: Evidence from a quasi-natural experiment. *Journal of the European Economic Association*, jvaf057.

Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. Pérez (2021). Automated linking of historical data. *Journal of Economic Literature 59*(3), 865–918.

Abrams, D. S., M. Bertrand, and S. Mullainathan (2012). Do judges vary in their treatment of race? *The Journal of Legal Studies 41*(2), 347–383.

Alesina, A., M. Carlana, E. La Ferrara, and P. Pinotti (2024). Revealing stereotypes: Evidence from immigrants in schools. *American Economic Review 114*(7), 1916–1948.

Alesina, A. and N. Fuchs-Schündeln (2007). Good-bye lenin (or not?): The effect of communism on people's preferences. *American Economic Review 97*(4), 1507–1528.

Allahabad High Court (2001). The uttar pradesh judicial service rules, 2001. Notification No. 1696/II-4-2001-36(1)-85. Published in consultation with the Uttar Pradesh Public Service Commission and the High Court of Judicature at Allahabad.

Allport, G. W. (1954). *The Nature of Prejudice*. Addison-Wesley.

Angerer, S., C. Waibel, and H. Stummer (2019). Discrimination in health care: a field experiment on the impact of patients' socioeconomic status on access to care. *American Journal of Health Economics 5*(4), 407–427.

Anwar, S., P. Bayer, and R. Hjalmarsson (2012). The impact of jury race in criminal trials. *The Quarterly Journal of Economics 127*(2), 1017–1055.

Anwar, S., P. Bayer, and R. Hjalmarsson (2019). Politics in the courtroom: Political ideology and jury decision making. *Journal of the European Economic Association 17*(3), 834–875.

Arnold, D., W. Dobbie, and C. S. Yang (2018). Racial bias in bail decisions. *The Quarterly Journal of Economics 133*(4), 1885–1932.

Arrow, K., O. Ashenfelter, and A. Rees (1973). Discrimination in labor markets. *The Theory of Discrimination*, 3–33.

Ash, E., S. Asher, A. Bhowmick, S. Bhupatiraju, D. Chen, T. Devi, C. Goessmann, P. Novosad, and B. Siddiqi (2025). In-group bias in the indian judiciary: Evidence from 5 million criminal cases. *Review of Economics and Statistics*, 1–45.

Asher, S., K. Jha, P. Novosad, A. Adukia, and B. Tan (2024). Residential segregation and unequal access to local public services in india: Evidence from 1.5 m neighborhoods. Working Paper.

Asher, S., P. Novosad, and C. Rafkin (2024). Intergenerational mobility in india: New measures and estimates across time and social groups. *American Economic Journal: Applied Economics 16*(2), 66–98.

Bagues, M., M. Sylos-Labini, and N. Zinovyeva (2017). Does the gender composition of scientific committees matter? *American Economic Review 107*(4), 1207–1238.

Battaglini, M., J. M. Harris, and E. Patacchini (2023). Interactions with powerful female colleagues promote diversity in hiring. *Journal of Labor Economics 41*(3), 589–614.

Bau, N. and A. Matray (2023). Misallocation and capital market integration: Evidence from india. *Econometrica 91*(1), 67–106.

Becker, G. (1957). The economics of discrimination.

Bharti, N. K. and S. Roy (2023). The early origins of judicial stringency in bail decisions: Evidence from early childhood exposure to hindu-muslim riots in india. *Journal of Public Economics 221*, 104846.

Boisjoly, J., G. J. Duncan, M. Kremer, D. M. Levy, and J. Eccles (2006). Empathy or antipathy? the impact of diversity. *American Economic Review 96*(5), 1890–1905.

Bonhomme, S. and M. Weidner (2022). Posterior average effects. *Journal of Business & Economic Statistics 40*(4), 1849–1862.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, rdae007.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics 128*(3), 967–1015.

Carrell, S. E., M. Hoekstra, and J. E. West (2018). The impact of intergroup contact on racial attitudes and revealed preferences. Working Paper.

Corno, L., E. La Ferrara, and J. Burns (2022, December). Interaction, stereotypes, and performance: Evidence from south africa. *American Economic Review 112*(12), 3848–75.

Cullen, Z. and R. Perez-Truglia (2023). The old boys' club: Schmoozing and the gender gap. *American Economic Review 113*(7), 1703–1740.

Darling-Hammond, S., R. T. Lee, and R. Mendoza-Denton (2021). Interracial contact at work: Does workplace diversity reduce bias? *Group Processes & Intergroup Relations 24*(7), 1114–1131.

Dobbie, W., P. Hull, and D. Arnold (2022). Measuring racial discrimination in bail decisions. *American Economic Review 112*(9).

Epstein, L. and T. Jacobi (2008). Super medians. *Stan. L. Rev. 61*, 37.

Eren, O. and N. H. Mocan (2020). Judge peer effects in the courthouse. Working Paper.

Fenizia, A. (2022). Managers and productivity in the public sector. *Econometrica 90*(3), 1063–1084.

Finseraas, H. and A. Kotsadam (2017). Does personal contact with ethnic minorities affect anti-immigrant sentiments? evidence from a field experiment. *European Journal of Political Research 56*(3), 703–722.

Goncalves, F. and S. Mello (2021). A few bad apples? racial bias in policing. *American Economic Review 111*(5), 1406–1441.

Haegele, I. (2022). Talent hoarding in organizations. Working Paper.

Hester, R. (2017). Judicial rotation as centripetal force: Sentencing in the court communities of south carolina. *Criminology 55*(1), 205–235.

Holden, R., M. Keane, and M. Lilley (2021). Peer effects on the united states supreme court. *Quantitative Economics 12*(3), 981–1019.

Jaffrelot, C. (2021). *Modi's India: Hindu nationalism and the rise of ethnic democracy*. Princeton University Press.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour economics 17*(2), 303–316.

Khan, Y. (2017). *The great partition: The making of India and Pakistan*. Yale University Press.

Lavy, V., E. Sand, and M. Shayo (2022). Discrimination between religious and non-religious groups: Evidence from marking high-stakes exams. *The Economic Journal 132*(646), 2308–2324.

Manski, C. F. (1995). *Identification problems in the social sciences.* Harvard University Press.

Margalit, Y. (2019). Political responses to economic shocks. *Annual Review of Political Science 22*(1), 277–295.

McConnell, B. and I. Rasul (2021). Contagious animosity in the field: Evidence from the federal criminal justice system. *Journal of Labor Economics 39*(3), 739–785.

Mitra, A. and D. Ray (2014). Implications of an economic theory of conflict: Hindu-muslim violence in india. *Journal of political economy 122*(4), 719–765.

Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies 90*(5), 2555–2591.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in delhi schools. *American Economic Review 109*(3), 774–809.

Rao, M. (2024, June). Front-line courts as state capacity: Evidence from india. Working Paper.

Ratra, V., D. Gupta, and R. Kumar (2023). Hiring reforms and the composition of lower judiciary a study of punjab's district courts. *Economic and Political Weekly 58*(32), 27.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights 4*(3), 305–322.

Sachar Committee (2006). Social, economic and educational status of the muslim community of india: A report. Technical report, Prime Minister's High Level Committee, Cabinet Secretariat, Government of India, New Delhi.

Sandefur, J. and B. Siddiqi (2013). Delivering justice to the poor: Theory and experimental evidence from liberia. In *World Bank Workshop on African Political Economy*, Volume 20, pp. 1–61. World Bank Washington, DC.

Shayo, M. and A. Zussman (2011). Judicial ingroup bias in the shadow of terrorism. *The Quarterly journal of economics 126*(3), 1447–1484.

Shayo, M. and A. Zussman (2017). Conflict and the persistence of ethnic bias. *American Economic Journal: Applied Economics 9*(4), 137–165.

Soule, S. A. (1992). Populism and black lynching in georgia, 1890–1900. *Social Forces 71*(2), 431–449.

Tajfel, H., M. G. Billig, R. P. Bundy, and C. Flament (1971). Social categorization and intergroup behaviour. *European journal of social psychology 1*(2), 149–178.

Varshney, A. (2006). Varshney–wilkinson dataset on hindu–muslim violence in india, 1950–1995, version 2. inter-university consortium for political and social research. *Ann Arbor, MI 17*.

Wachter, T. V. (2020). The persistent effects of initial labor market conditions for young adults and their sources. *Journal of Economic Perspectives 34*(4), 168–194.

1. Appendix

A. Robustness of Section 5

To validate our findings of Section 5, we conduct several robustness checks addressing concerns about specification, sample composition, and measurement.

First Appointment

There are trade-offs between defining the treatment as a fixed window and scheduling it by appointment. The first appointment in a judge's career is of particular interest, as it marks the first interaction with the judicial system and typically corresponds to the judge's probation period. A downside of considering this period is the limited power associated with a low fraction of judges exposed to riots, and the asymmetry in the length of the appointments.[21]

The main advantage of this definition is that it enables us to learn about whether it's the place that affects the judges' practices or the conflictual environment. Looking at Column 1 of Appendix Table A9, the estimates on the bias are substantially large, and stand at a 2.2 percentage point increase in bias, a 46% increase from the mean. This larger estimate also indicates that the earlier a judge is exposed to a conflict, the stronger the effect is, alluding to the importance of the early career aspect. Column 2 adds a fixed effect of the first court that the individual served in. The associated coefficient remains virtually unchanged, indicating that it is not the particular location driving the results.

Another test that can be incorporated into the analysis is looking at pre-appointment riots at the local level. Column 3 of Appendix Table A9 includes the same specification with an indicator for riots that occurred up to two years before the arrival of the judge. The results show that the during-appointment riot remains stable, and although half as large in magnitude, the prior-to-appointment riot has an effect. This suggests that while location itself is not the determining factor, the local tension associated with the conflict is important.

Gender Dimension

While the communal riots in question are between Hindus and Muslims, it raises a question of whether the effect spills over to another unrelated dimension. Column 4 of Appendix Table A9 changes Equation 4 to include an interaction term of $I[\text{ingender}]_i$ which is one if the judge and the defendant are of the same gender, and zero otherwise. The results show

[21] While the judge's average first appointment lasts a year and 8 months, we see empirically that it can be as long as almost four years.

that these judges show no difference in gender in-group bias due to the early career violence, validating that the effect is purely related to the Hindu-Muslim dimension.

Alternative Specifications and Outliers

Appendix Table A10 shows the robustness of the results to other alternative fixed-effect structures. While the baseline specification includes courtroom-by-year fixed effects, we show that all estimates remain significant and unchanged under: (1) court and year-month, (2) court-by-year, and (3) courtroom and year-month specifications. All are considered under the most stringent specification, which includes a judge fixed effect.

Lastly, as we have only 839 judges, we run a leave-one-out judge specification, estimating Equation 4 and excluding one judge at a time. Appendix Figure A4 presents the results of this estimation. The bias estimate ranges from 0.67 to 0.91 percentage points, with a standard deviation of those estimates of 0.01 percentage points. All estimates remained statistically significant at the five percent level. This robustness test suggests that these results reflect a general pattern rather than a result from the influence of a single judge.

B. Robustness of Section 6

We conduct several robustness checks to address potential concerns and alternative explanations for our findings on the transmission of judicial bias.

Assignment of Judges to Courts

We employ two strategies to check whether judge mobility is endogenous to changes in the recipient courts. First, we estimate event-study regressions to examine pre-move trends in case outcomes of courts that received judges who relocated. If there are differential pre-trends between courts receiving a more biased judge and those receiving a less biased one, this would indicate systematic patterns in judge mobility based on the characteristics of the receiving courts. Such a finding would suggest that the assumption of parallel trends for our event-study analysis is likely violated.

Further, we examine whether pre-move changes in observable court characteristics can predict whether a court received a more/less biased judge. Restricting the sample to cases before each move, we run the following regressions at the court level, in similar spirits to Bau and Matray (2023), who used lagged industry-level characteristics to predict reform timing.

$$Y_c = \Gamma \Delta X_c + \epsilon_c, \tag{11}$$

where Y_c is a set of indicators for whether the mover judge's fixed effects were in each of the three terciles. ΔX_c is the average monthly change of a vector of court-level characteristics, including the number of cases, number of same-religion and same-gender cases, mean acquittal rates, and decision time for cases and same-religion/gender cases. Suppose the receiving courts were experiencing different shocks (such as an increase in caseload) that prompted the need for an additional judge to relieve the case burden, and the judge who moved tended to be more/less biased towards their in-group defendants. In that case, we will obtain significant correlations between some of these characteristics and the type of judge moved.

Appendix Table A12 shows the estimated correlations between month-by-month pre-move changes in court characteristics, a year prior to the move, and the estimated judicial bias tercile of incoming mover judges. We find consistently little evidence of significant correlations between pre-move changes in court characteristics and whether the mover judge is in each of the three terciles of the fixed effects distribution. This provides systematic evidence against selection on pre-existing court changes or past shocks that drove differential moves of more/less biased judges to certain courts.

Assignment of Cases to Judges

Second, we test for potential changes in how cases were assigned to judges in courts that received judges of different bias levels. For instance, if the arrivals of high-bias judges led to incumbent judges receiving significantly different composition of cases, this could mechanically generate significant changes in their differential acquittal rates.

To check for this, we run the following regression:

$$y_{ijcmy} = \alpha_{j,m} + \rho_{my} + Charge_i \times I[\text{OutGroup}]_i + \beta_1 Pre_{icmy} + \beta_2 Post_{icmy} + \epsilon_{ijcmy} \tag{12}$$

In estimating this regression, we examine the effect of exposure to a mover judge within each tercile of the judge bias on the probability that a judge in court c receives a case involving a defendant who shares the same-religion, gender, or last name (used as a proxy for caste).

Appendix Table A13 presents the triple difference regression results. Across all three outcomes, we find no significant effect (with small point estimates), suggesting that courts did not respond to the arrival of more or less biased judges by reallocating cases among judges. This indicates that the effects we observe are not driven by changes in the types of cases specific judges handle after new judges arrive.

Confounding Shocks

We address potential confounding shocks and contemporaneous policy changes that may be correlated with judges' cross-court movements by re-estimating our main specifications, including state-by-year fixed effects. These control for state-level judicial reforms or other shocks that may confound our estimates. Appendix Table A14 shows our results remain largely unchanged, suggesting they are unlikely to be driven by state-specific shocks or policy changes that triggered judge movements.

Permutation Test

We conduct a placebo test to compare the effects of actual cross-court movements by judges with the effects of simulated movements. We randomly shuffle the assignment of incoming mover judge's bias measures across courts and re-estimate Equation 10 for each permutation. Appendix Figure A7 presents the distribution of coefficients from 1,000 random permutations alongside our actual estimates. The true coefficient for the second and third six months post-treatment falls at the 100th and 97th percentiles, respectively. These results suggest that the magnitude of our estimated effects is unlikely to be due to random assignment of judge bias levels, indicating a genuine relationship rather than one generated by chance.

A.2.1 Sensitivity to Potential Violations of Parallel Trends

We further complement the visual evidence of limited pre-trends in our event study graphs with two robustness checks that assess sensitivity to potential violations of parallel trends.

First, we examine the implications of our observed pre-trend. Panel (A) of Appendix Figure A8 displays the adjusted expectation of treatment effects assuming the time path would follow the linear extrapolation of our pre-period coefficients, which exhibit a slope of 0.0013. Following Roth (2022), we compute this adjusted path accounting for potential pre-testing bias arising from the analysis being conditional on passing the pre-trend test. Our observed treatment effects fall outside the bounds implied by this linear extrapolation, suggesting our results are not likely to be driven by the continuation of pre-existing trends.

Second, we follow Rambachan and Roth (2023) and estimate robust confidence intervals for our post-treatment event-study coefficients, allowing for non-linear deviations from the parallel trends assumption. Rather than assuming pre-trends continue linearly, this approach permits flexible violations bounded by a maximum change between consecutive periods. We set this bound (M) equal to 0.006, corresponding to the pre-trend slope detectable with 80% power given our estimation precision. Panel (b) of Appendix Figure A8 presents these robust confidence intervals, which continue to exclude zero even when allowing for such violations.

Taken together, these results provide additional supportive evidence that our findings are unlikely to be driven by differential pre-trends.

C. Tables and Figures

Appendix Table A1: Riots Descriptive, by Decade

Decade	Occurrences	Curfews	Killed	Injured	Arrested
Panel A: India					
1950	96	27	217	1,212	5,877
1960	141	46	1,096	3,011	15,270
1970	123	25	497	2,973	9,260
1980	420	67	1,867	7,182	38,594
1990	295	61	2,814	5,967	13,195
2000	54	5	1,190	3,202	30,008
2010	172	21	295	2,029	6,400
Panel B: Uttar Pradesh					
1950	29	16	45	298	2,344
1960	23	13	76	404	2,342
1970	17	6	72	367	1,124
1980	57	11	338	1,015	13,903
1990	54	11	379	325	6,787
2000	11	1	35	138	304
2010	41	9	106	668	3,583

Notes: This table presents descriptive statistics of communal riots by decade from 1950-2018. Panel A shows all-India statistics, while Panel B focuses on Uttar Pradesh. Occurrences count the total number of distinct riot events, where curfews measure the total number of curfew days imposed during the riots. Killed, injured, and arrested represent cumulative counts across all riot incidents within each decade. Missing values for curfews, killed, injured, and arrests are coded as zero.

Appendix Table A2: Summary Statistics

	Judges	Mean	SD	Cases
Panel A: All-India Sample				
Muslim Defendant	12,784	0.15	0.36	4,147,372
Female Defendant	12,745	0.10	0.30	3,828,282
In-Religious Case	12,784	0.81	0.39	4,147,372
In-Gender Case	12,361	0.67	0.47	3,712,571
Acquitted	12,784	0.12	0.33	4,147,372
Muslim Judge	12,784	0.07	0.25	4,147,372
Female Judge	12,399	0.29	0.45	4,019,335
Panel B: Uttar Pradesh				
Muslim Defendant	839	0.16	0.37	316,899
Female Defendant	839	0.08	0.28	316,899
Acquitted	839	0.05	0.21	316,899
Muslim Judge	839	0.07	0.26	316,899
Female Judge	839	0.24	0.43	316,899
Year of Birth	839	1976	6.85	316,899
Age Joined JS system	839	32.2	4.36	316,899
LL.B. (First Hons.)	839	0.38	0.49	316,899
LL.M.	839	0.34	0.47	316,899
Ph.D.	839	0.01	0.11	316,899
Years After Graduation Joined JS system	835	6.92	4.32	316,899
Riot at Early-Childhood (ages 0–6)	839	0.19	0.39	316,899
Riot at Middle-Childhood (ages 6–12)	839	0.23	0.42	316,899
Riot at Adolescence (ages 12–18)	839	0.22	0.41	316,899
First Appointment Length (years)	839	1.69	1.10	316,899
First 5 Years				
Early-Career Exposure to Riot	839	0.17	0.38	316,899
Number of Interactions	839	19.5	24.22	316,899
Contact with Out-group Judge	839	0.55	0.50	316,899
Contact with Same Gender Out-group Judge	839	0.45	0.50	316,899
Contact with Close in Age Out-group Judge (± 2y)	839	0.25	0.43	316,899
Panel C: Stacked Dataset				
Acquitted	7,935	0.17	0.37	3,997,464
In-Religious Case	7,935	0.82	0.38	3,997,464
Decision Time (Days)	7,805	252	301	2,526,745
Mover Muslim Judge	1,939	0.05	0.22	
Mover Female Judge	1,939	0.29	0.46	
New Appointment Length (months)	1,939	20.5	9.14	
Cases Prior to Move	1,939	474	601	
Ingroup Bias FE (shrunk)	1,939	0.00	0.10	

Notes: This table presents summary statistics for all three analyses samples. Panel A shows statistics for all criminal cases in India's district courts. Panel B shows statistics for Uttar Pradesh's criminal courts, corresponding to Section 5. Panel C shows attributes for our "stacked" mover design sample, corresponding to Section 6. "First 5 Years" variables capture judge' experiences during their initial five years in the judicial service system. Riot exposure variables are defined based on communal riots occurring in the judge's district of residence during specified age ranges or career periods. Contact variables measure whether judges had professional interactions with out-group colleagues (different religion) during their first five years.

Appendix Table A3: Validation Test of Random Assignment of Cases to Judges

Judge Characteristic:	(1) Muslim	(2) Female	(3) Experience	(4) Age at Entry	(5) Childhood Riot	(6) Career Riot
Muslim Def.	-0.000	0.002	0.000	-0.005	-0.001	-0.001
	(0.0003)	(0.001)	(0.003)	(0.009)	(0.001)	(0.001)
Female Def.	0.001	0.001	0.001	-0.003	-0.002	0.001
	(0.0004)	(0.001)	(0.003)	(0.008)	(0.001)	(0.001)
Mean Dep.	0.073	0.209	8.908	31.708	0.391	0.504
Observations	316,899	316,899	316,899	316,899	316,899	316,899

Notes: This table tests the random assignment assumption by examining whether defendant characteristics predict judge characteristics. Control variables include the courtroom-by-year and charge section fixed effects. Dependent variables are indicators for a Muslim judge (column 1), a female judge (column 2), and continuous measures of judge experience in years (column 3), age at entry (column 4), an indicator for exposure to a Hindu-Muslim riot in childhood (column 5), and in their career (column 6). The key regressors are indicators for a defendant being Muslim and female. Standard errors appear in parentheses, are robust, and clustered at the courtroom level.

Appendix Table A4: Early Career Hindu Judges Exposure to Muslim Judges

	Dependent Variable: Acquitted		
	(1)	(2)	(3)
Riot (0-5 years)	-0.043	0.003	-0.005
	(0.020)	(0.013)	(0.014)
$I[\text{InGroup}]_i \times$ Riot	0.014	0.010	0.013
	(0.006)	(0.004)	(0.004)
General Contact:			
Outgroup Contact (0-5 years)	-0.008		
	(0.006)		
$I[\text{InGroup}]_i \times$ Outgroup Contact	-0.000		
	(0.002)		
Riot $\times$ Outgroup Contact	0.065		
	(0.023)		
$I[\text{InGroup}]_i \times$ Riot $\times$ Outgroup Contact	-0.010		
	(0.007)		
Age-based Outgroup Contact:			
Outgroup Contact ($\pm$2 Age)		0.003	
		(0.010)	
$I[\text{InGroup}]_i \times$ Outgroup Contact ($\pm$2 Age)		0.001	
		(0.003)	
Riot $\times$ Outgroup Contact ($\pm$2 Age)		-0.009	
		(0.015)	
$I[\text{InGroup}]_i \times$ Riot $\times$ Outgroup Contact ($\pm$2 Age)		-0.009	
		(0.005)	
Gender-based Outgroup Contact:			
Outgroup Contact (Same Gender)			0.002
			(0.008)
$I[\text{InGroup}]_i \times$ Outgroup Contact (Same Gender)			-0.000
			(0.002)
Riot $\times$ Outgroup Contact (Same Gender)			0.016
			(0.014)
$I[\text{InGroup}]_i \times$ Riot $\times$ Outgroup Contact (Same Gender)			-0.013
			(0.006)
Mean of Dependent Variable	0.048	0.048	0.048
Number of Judges	777	777	777
Observations	293,777	293,777	293,777

Notes: This table examines how contact with outgroup defendants mitigates bias formation among non-Muslim judges exposed to riots early in their careers. Column (1) uses any outgroup contact, Column (2) focuses on age-based outgroup contact (defendants within 2 years in age difference), and Column (3) examines same-gender outgroup contact. Standard errors clustered by judge and appear in parentheses.

Appendix Table A5: Reinforcement Effects: Early Career and Childhood Riot Exposure

	Dependent Variable: Acquitted		
	(1) Age 0-6	(2) Age 6-12	(3) Age 12-18
$I[\text{InGroup}]_i$	0.000	0.001	0.001
	(0.002)	(0.002)	(0.002)
Riot (Career)	-0.003	0.004	-0.001
	(0.010)	(0.009)	(0.010)
$I[\text{InGroup}]_i \times$ Riot (Career)	0.004	0.008	0.009
	(0.003)	(0.003)	(0.003)
Childhood Exposure Effects:			
Riot (Age 0-6)	-0.008		
	(0.008)		
$I[\text{InGroup}]_i \times$ Riot (Age 0-6)	0.001		
	(0.002)		
Riot (Career) $\times$ Riot (Age 0-6)	0.046		
	(0.038)		
$I[\text{InGroup}]_i \times$ Riot (Career) $\times$ Riot (Age 0-6)	0.031		
	(0.015)		
Riot (Age 6-12)		-0.005	
		(0.006)	
$I[\text{InGroup}]_i \times$ Riot (Age 6-12)		-0.000	
		(0.003)	
Riot (Career) $\times$ Riot (Age 6-12)		-0.042	
		(0.023)	
$I[\text{InGroup}]_i \times$ Riot (Career) $\times$ Riot (Age 6-12)		0.002	
		(0.010)	
Riot (Age 12-18)			-0.023
			(0.007)
$I[\text{InGroup}]_i \times$ Riot (Age 12-18)			0.001
			(0.003)
Riot (Career) $\times$ Riot (Age 12-18)			-0.021
			(0.023)
$I[\text{InGroup}]_i \times$ Riot (Career) $\times$ Riot (Age 12-18)			-0.000
			(0.008)
Mean of Dependent Variable	0.048	0.048	0.048
Number of Judges	839	839	839
Observations	316,899	316,899	316,899

Notes: This table examines reinforcement effects of childhood riot exposure on early career bias formation. Each column tests whether exposure to riots during different childhood periods (ages 0-6, 6-12, or 12-18) reinforces the bias-inducing effects of riots experienced during judges' early careers (0-5 years). The triple interaction terms capture whether childhood exposure amplifies the in-group bias from early career riot exposure. Standard errors clustered by judge in parentheses.

Appendix Table A6: Exposure to Riots During Childhood

	Dependent Variable: Acquitted		
	(1)	(2)	(3)
Panel A: Early Childhood (0-6 years)			
$I[\text{InGroup}]_i$	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
Riot (Age 0-6)	-0.006	-0.002	
	(0.007)	(0.016)	
$I[\text{InGroup}]_i \times$ Riot (Age 0-6)	0.004	0.004	0.003
	(0.003)	(0.003)	(0.002)
Panel B: Middle Childhood (6-12 years)			
$I[\text{InGroup}]_i$	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)
Riot (Age 6-12)	-0.007	-0.022	
	(0.006)	(0.013)	
$I[\text{InGroup}]_i \times$ Riot (Age 6-12)	-0.001	0.000	-0.001
	(0.002)	(0.003)	(0.003)
Panel C: Adolescence (12-18 years)			
$I[\text{InGroup}]_i$	0.002	0.002	0.001
	(0.002)	(0.002)	(0.002)
Riot (Age 12-18)	-0.024	-0.012	
	(0.007)	(0.012)	
$I[\text{InGroup}]_i \times$ Riot (Age 12-18)	0.001	0.001	0.000
	(0.003)	(0.002)	(0.002)
Fixed Effects:			
Charge	✓	✓	✓
Courtroom-by-year	✓	✓	✓
Additional Controls		✓	✓
Judge FE			✓
Mean of Dependent Variable	0.048	0.048	0.048
Number of Judges	839	839	839
Observations	316,899	316,899	316,899

Notes: This table examines the effect of childhood exposure to communal riots on judicial in-group bias. The dependent variable is a dummy equal to 1 if the case was acquitted. $I[\text{InGroup}]_i$ indicates whether the judge and defendant share the same religion. Riot variables indicate whether the judge experienced a communal riot in their home district during the specified age range. Panel A examines early childhood (ages 0-6), Panel B middle childhood (ages 6-12), and Panel C adolescence (ages 12-18). All specifications include courtroom-by-year and charge fixed effects, as well as defendant controls (gender and religion) and judge controls (gender and religion). Column 2 adds year of birth, year of entry, highest educational qualification, a home district dummy, and a judge and defendant shared gender indicator. Column 3 adds judge fixed effects. Standard errors are robust, adjusted for clusters by judge.

Appendix Table A7: Heterogeneity in Early Career Riot Exposure Effects

	Dependent Variable: Acquitted				
	Post 10+ (1)	Entered Young (2)	LLB First (3)	Female (4)	Severity (5)
$I[\text{InGroup}]_i$	0.001	0.001	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Riot (0-5 years)	0.005	-0.008	0.016	0.002	0.003
	(0.010)	(0.012)	(0.011)	(0.010)	(0.010)
$I[\text{InGroup}]_i \times$ Riot	0.008	0.008	0.007	0.007	0.007
	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)
Char.	0.023	0.015	0.002		-0.023
	(0.008)	(0.006)	(0.006)		(0.017)
Riot $\times$ Char.	-0.038	0.005	-0.035	-0.010	
	(0.015)	(0.018)	(0.018)	(0.018)	
$I[\text{InGroup}]_i \times$ Char.	-0.000	-0.001	0.000	-0.000	0.014
	(0.003)	(0.002)	(0.002)	(0.003)	(0.015)
$I[\text{InGroup}]_i \times$ Riot $\times$ Char.	0.006	0.000	0.004	0.008	
	(0.009)	(0.006)	(0.007)	(0.009)	
Mean of Dependent Variable	0.048	0.048	0.048	0.048	0.048
Number of Judges	839	839	839	839	839
Observations	316,899	316,899	316,899	316,899	316,899

Notes: Each column tests whether the in-group bias effect from early career riot exposure varies by judge characteristics. The base interaction shows the effect of riot exposure (0-5 years) on in-group bias. The triple interaction tests heterogeneity: Column (1) persistence after 10+ years of experience; Column (2) judges who entered young (entered judiciary at young age); Column (3) academic achievement (LLB First Class); Column (4) female judges; Column (5) riot severity (curfew imposed). Standard errors are robust, adjusted for clusters by judge, and appear in parentheses.

Appendix Table A8: Distribution of Judge-Level Bias by Early Career Riot Exposure

	No Exposure	Riot Exposure	Change
Mean bias	-0.004	0.012	0.016
Standard deviation	0.118	0.085	-0.033
Percentiles of bias distribution:			
5th	-0.138	-0.061	0.077
10th	-0.055	-0.028	0.027
25th	-0.013	-0.006	0.007
50th	-0.000	0.000	0.000
75th	0.016	0.018	0.002
90th	0.048	0.080	0.032
95th	0.101	0.116	0.015
Skewness	-1.384	3.142	4.526
Kolmogorov-Smirnov (p-value)			0.039
Judges	679	143	

Appendix Table A9: Riot Exposure at First Appointment and the Gender Dimension

	Dependent Variable: Acquitted			
	(1) 1st Court	(2) + 1st Court FE	(3) + Pre-app.	(4) Gender
$I[\text{InGroup}]_i$	0.002	0.002	0.001	
	(0.002)	(0.002)	(0.002)	
Riot (1st Court)	0.006	0.030	0.016	
	(0.028)	(0.045)	(0.036)	
$I[\text{InGroup}]_i \times$ Riot (1st Court)	0.022	0.023	0.022	
	(0.012)	(0.012)	(0.012)	
Riot (Pre-appointment)			-0.083	
			(0.043)	
$I[\text{InGroup}]_i \times$ Riot (Pre-appointment)			0.013	
			(0.007)	
$I_{[ingender]_i}$				0.006
				(0.003)
Riot (0-5 years)				0.008
				(0.011)
$I_{[ingender]_i} \times$ Riot (0-5 years)				-0.002
				(0.006)
Mean of Dependent Variable	0.048	0.048	0.048	0.048
Number of Judges	839	839	839	839
Observations	316,899	316,899	316,899	316,899
Fixed Effects:				
Charge	✓	✓	✓	✓
Courtroom-by-year	✓	✓	✓	✓
1st Court FE		✓	✓	

Notes: This table examines the effect of riot exposure related to judges' first court assignments and a placebo test. The dependent variable is a dummy equal to 1 if the case was acquitted. $I[\text{InGroup}]_i$ indicates whether the judge and defendant share the same religion. $I_{[ingender]_i}$ indicates whether they share the same gender. Column (1) shows the baseline specification with riot exposure at the first court location. Column (2) adds the first court fixed effects to control for court-specific characteristics. Column (3) examines riots that occurred in the judge's home district before their appointment to the judiciary. Column (4) presents a placebo test examining whether exposure to riots affects in-gender bias. Standard errors are robust, adjusted for clusters by judge, and appear in parentheses.

Appendix Table A10: Alternative Specifications

	Dependent Variable: Acquitted			
	(1)	(2)	(3)	(4)
$I[\text{InGroup}]_i$	0.000	0.000	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)
$I[\text{InGroup}]_i$ × Riot (0-5 years)	0.009	0.0010	0.009	0.010
	(0.003)	(0.003)	(0.003)	(0.003)
Fixed Effects:				
Controls	Court & month-year	Court-by-year	Courtroom & month-year	Courtroom-by-year
Charge	✓	✓	✓	✓
Judge FE	✓	✓	✓	✓
Mean of Dependent Variable	0.048	0.048	0.048	0.048
Number of Judges	839	839	839	839
Observations	316,897	316,899	316,897	316,899

Notes: This table presents alternative specifications examining the effect of riot exposure on judicial bias. The dependent variable is a dummy equal to 1 if the case was acquitted. $I[\text{InGroup}]_i$ indicates whether the judge and defendant share the same religion. Standard errors are robust, adjusted for clusters by judge, and appear in parentheses.

Appendix Table A11: Balancing of Treated and Never Treated Courts

	Treated			Never-Treated			Test Year-Month & Charge		
	Mean	SD	N	Mean	SD	N	Coef.	SE	P-value
Acquittal	0.15	0.36	1743953	0.13	0.34	477797	0.01	0.01	0.15
Delay	247.87	300.98	1070013	262.71	313.90	276096	-0.69	10.74	0.95
Muslim Judge	0.07	0.25	1743953	0.07	0.25	477797	0.01	0.01	0.50
Female Judge	0.27	0.44	1692635	0.32	0.47	456836	-0.01	0.03	0.63
In-Religious Case	0.81	0.39	1743953	0.82	0.38	477797	-0.02	0.01	0.19
In-Gender Case	0.68	0.46	1570937	0.64	0.48	415222	0.02	0.03	0.49
In-religious Acquittal	0.15	0.36	1417899	0.13	0.33	392717	0.01	0.01	0.10
In-gender Acquittal	0.14	0.35	1074839	0.14	0.35	265258	0.00	0.01	0.74
N(Courts)		1019			902				

Notes: This table provides summary statistics for the treated and never-treated courts. The first three columns give the mean, SD, and number of unique case-level observations for the treated group. The next three columns do the same for the never-treated courts. The last three columns report the coefficient on treatment, its standard error, and p-value for the test of difference from zero, controlling for month-year and the case's corresponding charge, clustered at the court level.

Appendix Table A12: Predict Pr(Mover Judge in each Tercile) with Pre-Move Changes

Dependent Variable:	Pr(Mover Judge Tercile)		
	Bottom	Middle	Top
	(1)	(2)	(3)
Δ N(Case)	0.004	-0.001	-0.003
	(0.007)	(0.008)	(0.006)
Δ N(Same Relig Case)	-0.005	-0.002	0.008
	(0.009)	(0.011)	(0.007)
Δ N(Same Gender Case)	0.000	0.005	-0.005
	(0.003)	(0.004)	(0.003)
Δ Pr(Acquitted)	-0.074	-0.072	0.146
	(0.244)	(0.222)	(0.201)
Δ Pr(Same Relig Acquitted)	0.031	-0.117	0.086
	(0.239)	(0.189)	(0.169)
Δ Pr(Same Gender Acquitted)	0.120	0.204	-0.324
	(0.196)	(0.189)	(0.205)
Δ Decision Time (Days)	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Δ Same Relig Decision Time	0.000	0.001	-0.001
	(0.000)	(0.000)	(0.000)
Δ Same Gender Decision Time	-0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)
Joint F-stat	0.339	1.708	1.389
p-value	0.962	0.083	0.188
N (Clusters)	950	950	950
N (Obs)	1,824	1,824	1,824

Notes: This table presents court-level OLS regressions. Outcome variables are dummy variables for whether the mover judge's religious bias fixed effects is in the first/ second/ third tercile of the distribution. Independent variables are pre-move average changes in characteristics of courts that received mover judges. The sample is restricted to observations in 12 to 1 month before each move. Standard errors are clustered by court level in parentheses.

Appendix Table A13: Triple Differences, Alternative Outcomes, In-Religion Bias of Mover Judges

Dependent Variable:	Same Gender Case			Same Caste Case			Same Religion Case		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Terc. 1	Terc. 2	Terc. 3	Terc. 1	Terc. 2	Terc. 3	Terc. 1	Terc. 2	Terc. 3
Pre	0.004	-0.002	0.001	-0.002	0.001	0.000	0.002	0.003	0.000
(-18 to -2 months)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.004)	(0.004)
Post	0.002	-0.004	-0.001	-0.001	0.001	-0.002	0.004	0.006	-0.000
(0 to 17 months)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.004)	(0.004)
Control Mean Dep. Var.	0.666	0.692	0.659	0.042	0.036	0.034	0.832	0.812	0.833
Clusters	5,040	4,665	4,997	4,741	4,352	4,673	5,207	4,831	5,188
Observations	1,337,141	1,332,153	1,315,244	625,787	593,663	620,869	1,505,269	1,500,082	1,499,012

Notes: Every three columns are a different outcome variable, split by tercile of mover judges religion bias. The pre-treatment period includes all cases a year before the moving judge arrives in court. The post period includes all cases a year and a half after the move. The control variables include the charges, year-month, and judge fixed effects. Standard errors are clustered at the incumbent judge level and appear in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table A14: Difference-in-Differences Estimates Including State-by-Year Fixed Effects, by Terciles Mover Judge of Religion Bias

	Dependent Variable: Acquitted		
	(1) Top Tercile	(2) Middle Tercile	(3) Bottom Tercile
Pre	-0.023	0.001	0.002
	(0.010)	(0.013)	(0.008)
Pre $\times I[InGroup]_i$	0.015	-0.010	-0.006
	(0.009)	(0.013)	(0.008)
Post			
(0-5 months)	-0.006	-0.000	0.012
	(0.010)	(0.013)	(0.008)
(6-11 months)	-0.027	0.039	0.020
	(0.011)	(0.017)	(0.010)
(12-17 months)	-0.014	0.027	0.009
	(0.012)	(0.016)	(0.012)
Post $\times I[InGroup]_i$			
(0-5 months)	0.009	-0.010	-0.010
	(0.009)	(0.013)	(0.008)
(6-11 months)	0.030	-0.028	-0.014
	(0.010)	(0.016)	(0.010)
(12-17 months)	0.026	-0.021	0.004
	(0.011)	(0.016)	(0.011)
Control Mean Dep. Var.	0.165	0.128	0.173
Clusters	5,174	4,803	5,157
Observations	1,504,708	1,499,533	1,498,455

Notes: The dependent variable is a binary indicator of whether the defendant in a case was acquitted. The pre-treatment period includes all cases a year and a half before a moving judge arrives in court. The control variables used are those specified in Equation 10, with additional state-by-year fixed effects. Standard errors are clustered at the incumbent judge level and appear in parentheses.

Appendix Figure A1: Example of Judges' Biographies from the Allahabad High Court

[redacted]

[redacted]

Id No:
Father's/Husband's Name:
Date of Birth:
Date of Retirement:
Home District:
Home State:
Date of Appt. in Judicial Service as Munsif/CJ(JD):
Date of Promotion in CJ/CJ(SD):
Date of Conf. on Munsif/CJ(JD):
Date of Promotion to Class-I:
Date of Joining in HJS under Rule 22(3) of UP HJS Rule 1975:
Date of Joining in HJS under Rule 22(1) of UP HJS Rule 1975:
Date of Conf. in HJS:

[redacted]

Qualification Details

SL	DETAILS	DIVISION	YEAR
1.	High School	First	2011
2.	Intermediate	First	2013
3.	B.A. LL.B.(Hons.)	First	2018
4.	LL.M.	Passed	2020

Training Details

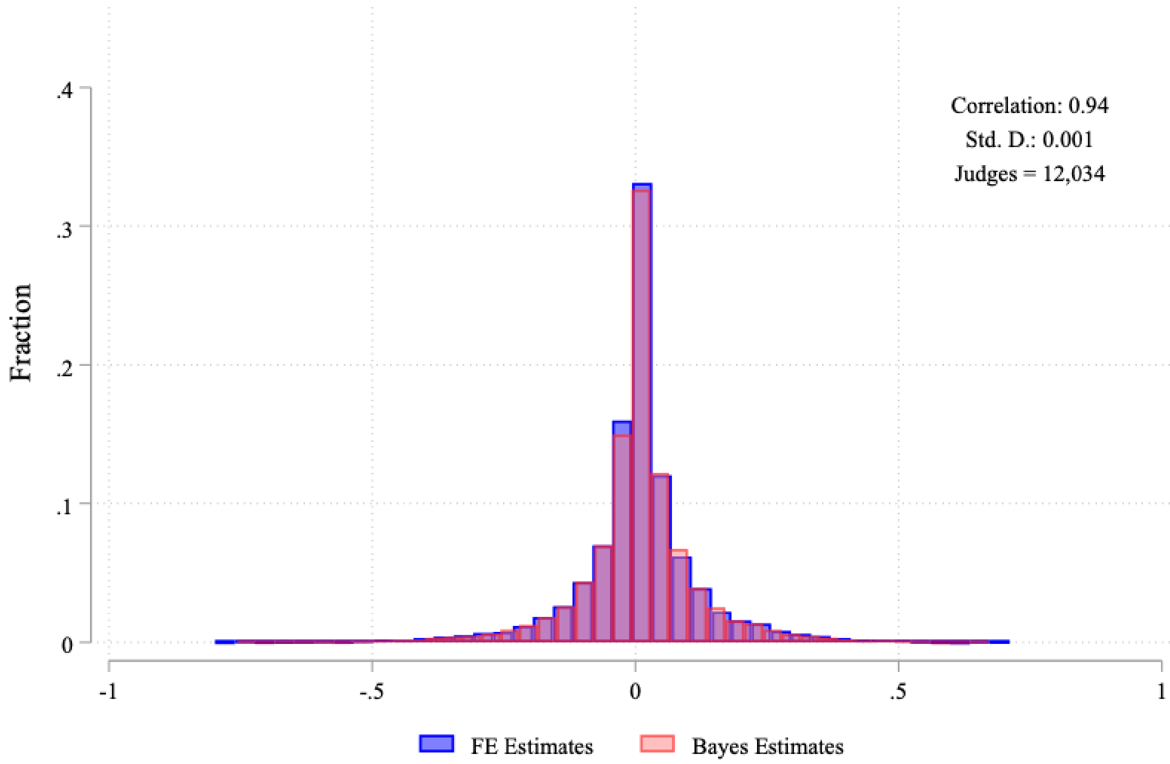
SL	SUBJECT	INSTITUTE	FROM-TO
1.	Induction Training Programme (Phase-1 Institutional Training) for Newly Appointed Civil Judges (JD)	I.J.T.R., U.P. Lucknow	22-06-20 - 05-08-20
2.	Induction Training Programme (Phase-II Institutional Training) for Newly Appointed Civil Judges (JD)	I.J.T.R., U.P. Lucknow	07-12-20 - 30-01-21

Posting Details

SL	DESIGNATION PLACE	FROM TO	REMARKS
1.	Civil Judge (Junior Div.) Varanasi	[redacted]	Fast Track
2.	Addl. Civil Judge (Junior Div.) Varanasi		
3.	Addl. Civil Judge (Junior Div.) Budaun		

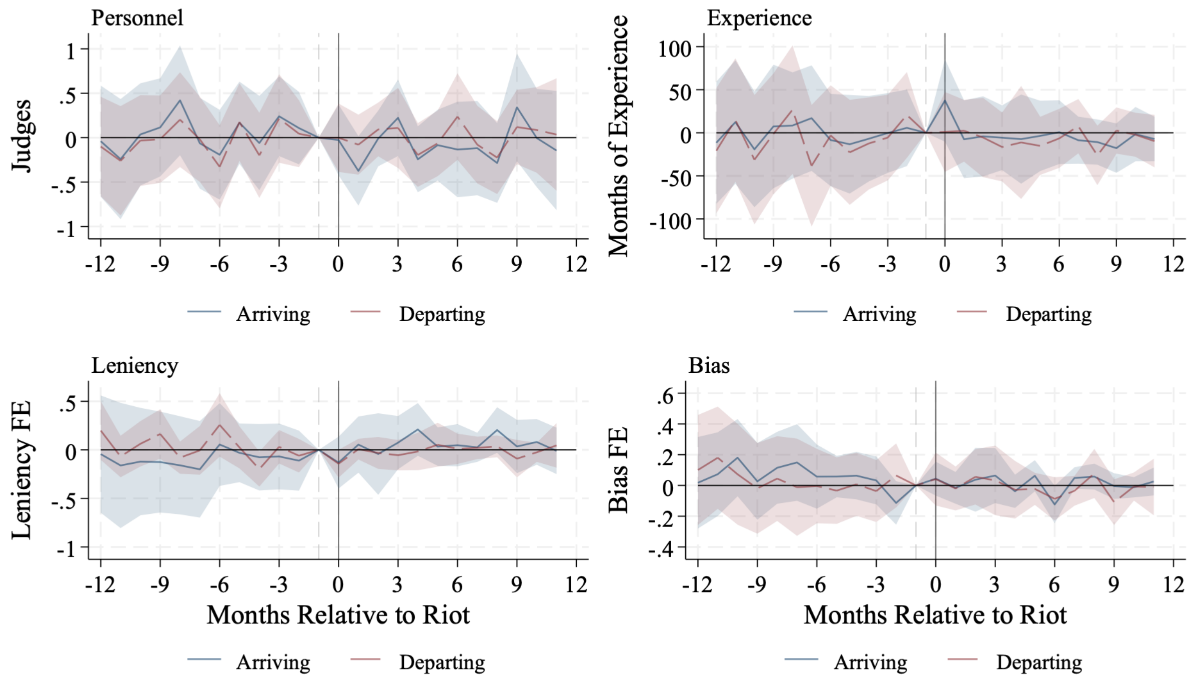
Notes: This figure shows a representative example of biographical information available for judges on the Allahabad High Court website. The data includes personal details (name, designation, date of birth), educational background, professional experience, and career progression within the judicial system.

Appendix Figure A2: Oridnary Least Squares and Bayesian Shrunk Estimates



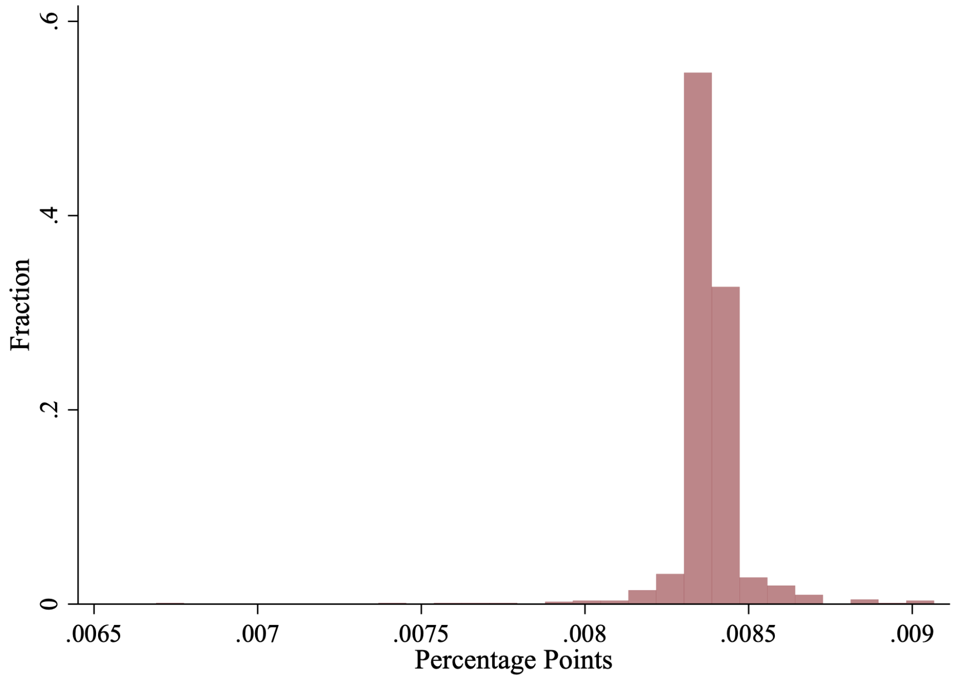
Note: This figure plots the judge-level in-group bias fixed effects, as estimated in Equation 1, and their corresponding Bayesian shrunk estimates. To improve readability, we exclude extreme outliers beyond the 1st and 99th percentiles plus two standard deviations from this figure.

Appendix Figure A3: No Strategic Allocation of Judges in Response to Riots



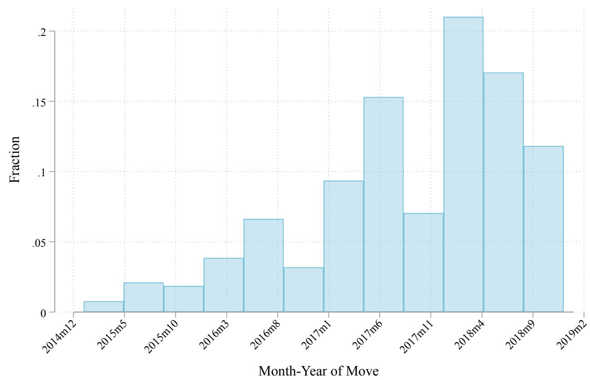
Notes: This figure shows event study results of judicial personnel changes 12 months before and after communal riots in a district, following Equation 6. Each Panel shows the point estimate with the associated 95% confidence intervals for arriving (solid navy line) and departing (dashed maroon line) judges. Panel A shows judge hiring and separation. Panel B shows the average experience of personnel changes. Panel C and D show the leniency and religious bias fixed effects of the incoming and departing judges, estimated from Equation 1.

Appendix Figure A4: Distribution of the In-group Bias Estimates from a Leave-One-Judge-Out Exercise

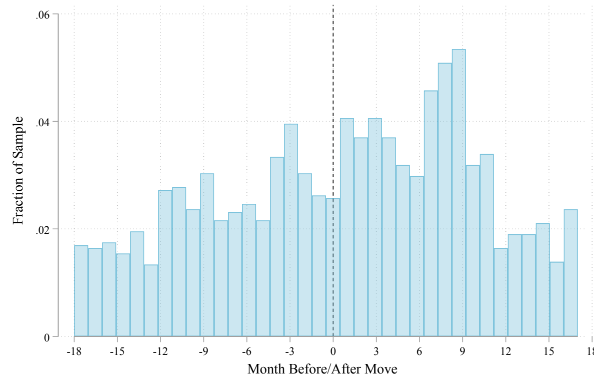


Notes: This histogram shows the distribution of the in-group bias estimate for Equation 4, each time excluding a single judge from the estimation.

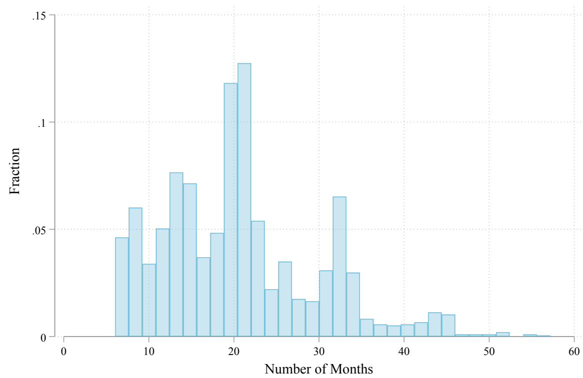
Appendix Figure A5: Stacked Dataset Descriptives



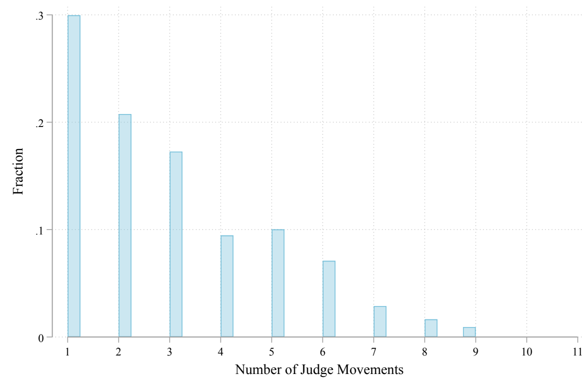
(A) Distribution of Month-Year of Judge Movement



(B) Distribution of Event-Time

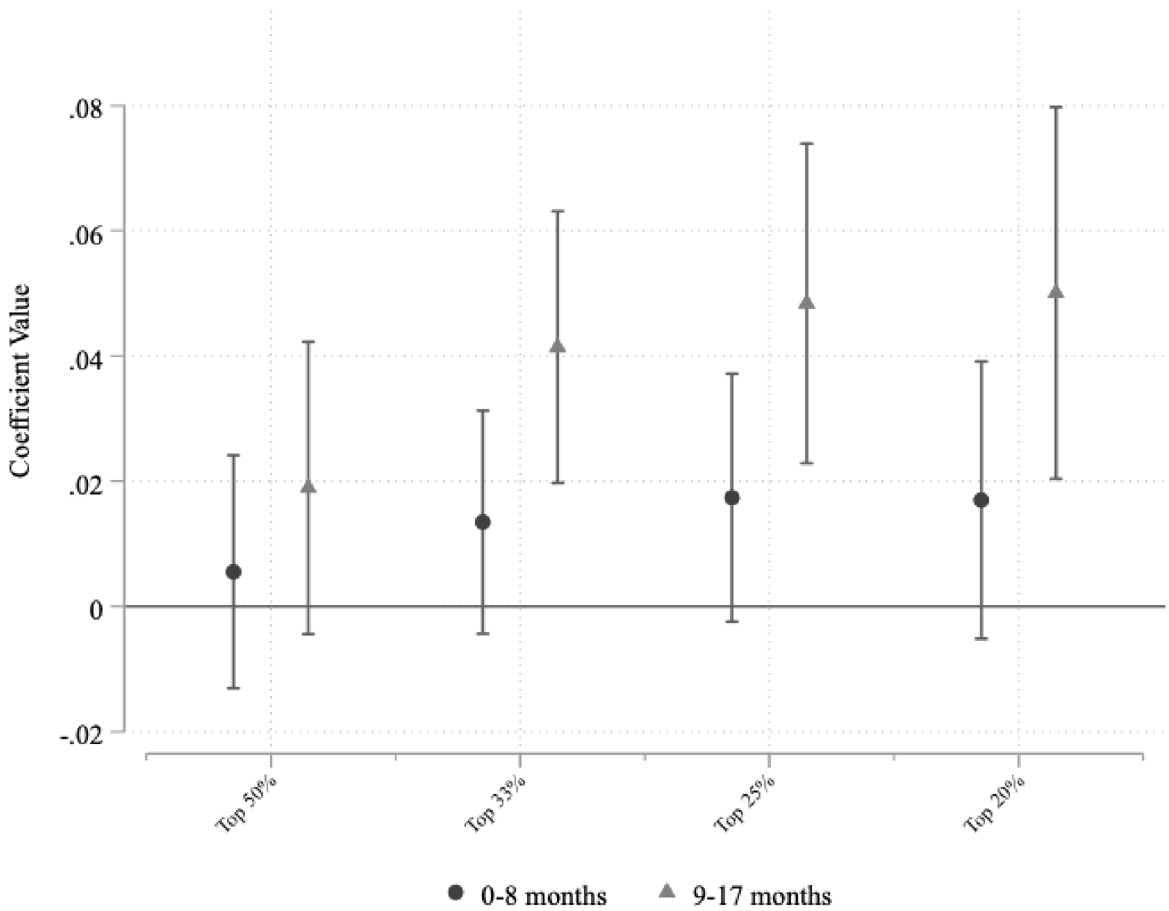


(C) Distribution of Length of Stay for Mover Judges



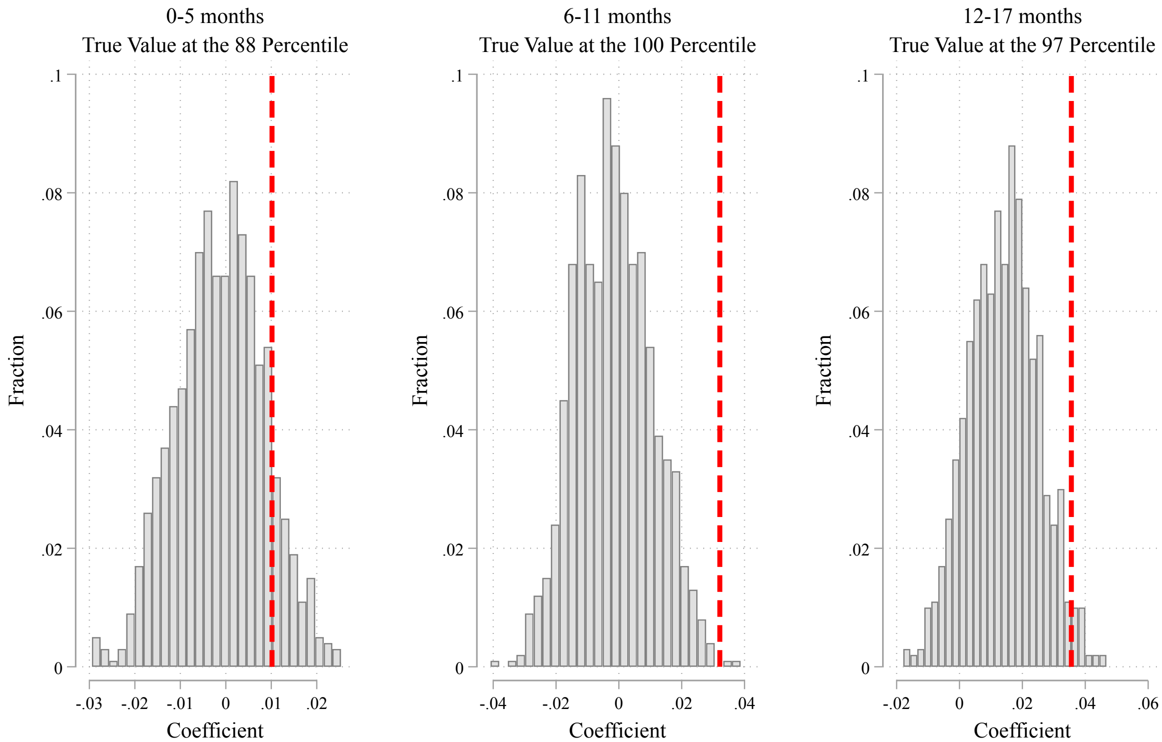
(D) Distribution of Number of Movers by Court

Appendix Figure A6: Change in Relative In-Group Bias in the Medium- and Long-Run, by Top Fraction of Mover Judges' Religion Bias



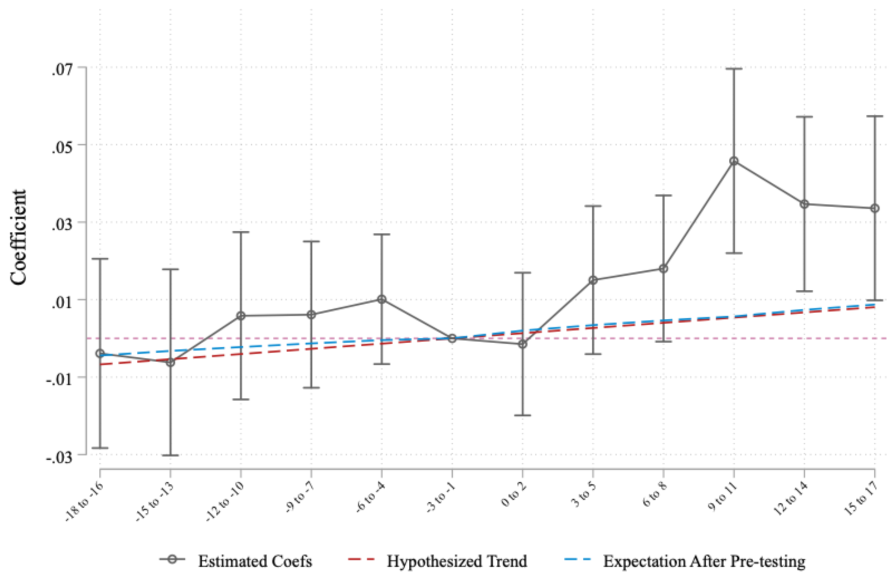
Note: This figure plots the two triple interaction terms from Equation 10. Standard errors are clustered at the incumbent judge level, and the plot shows their corresponding 95% confidence intervals.

Appendix Figure A7: Permutation Tests

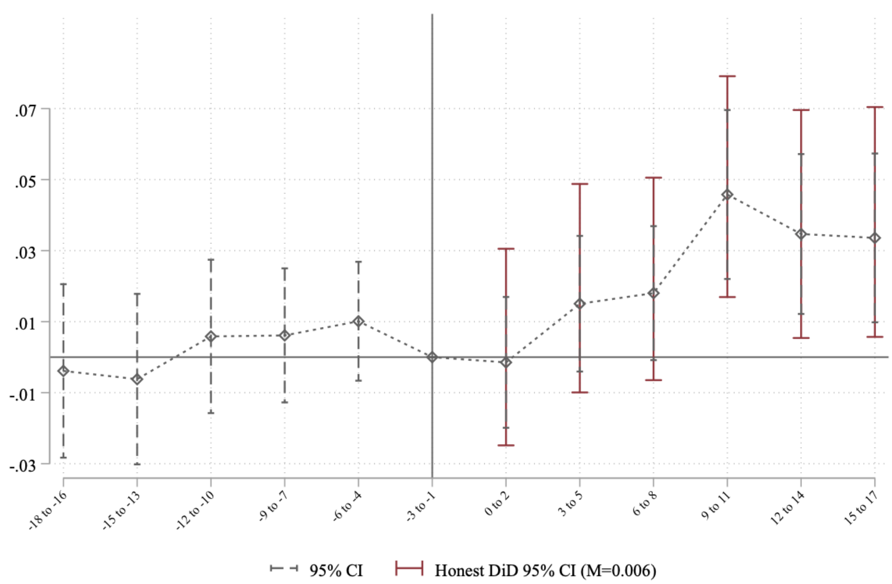


Notes: These figures present results from permutation tests in which we randomly shuffle the incoming mover judge's bias measure and re-estimate Equation 10 for each permutation. The red dashed line indicates the coefficient obtained using the actual (non-permuted) data, while the histogram displays the distribution of coefficients from 1,000 random permutations. Standard errors are clustered at the incumbent judge level.

Appendix Figure A8: Robustness Tests for Parallel Trends Assumption



(A) Linear Extrapolation of Pre-trend



(B) Robust Confidence Intervals

Notes: These figures show robustness tests relating to the parallel trends assumption. Panel A plots the adjusted expectation of treatment effects assuming the time path would follow the linear extrapolation of our pre-period coefficients (slope = -0.0013). Following Roth (2022), we also compute this adjusted path, accounting for potential pre-testing bias arising from conditioning the analysis on passing the pre-trend test. Panel B estimates robust 95% confidence intervals for post-treatment coefficients, allowing for non-linear deviations from parallel trends with maximum change in consecutive periods (M) equal to the 80% power pre-trend slope (M = 0.006) (Rambachan and Roth, 2023).