Caste aside? Names, Networks and Justice in the courts of Bihar, India

Sandeep Bhupatiraju (World Bank)

Daniel L. Chen (Toulouse School of Economics)

Shareen Joshi (Georgetown University)

Peter Neis (CERDI, Université Clermont Auvergne)*

January 2, 2024

This research analyzes over a million cases from Bihar's Patna High Court (2009-2019), revealing social identity dynamics in judicial proceedings. A key focus is the use of family names as markers of caste identity. The study finds a notable absence of scheduled castes in influential judicial roles, indicating a divergence between the judiciary's composition and Bihar's demographic makeup. A trend of adopting 'caste-neutral' names is identified across court participants, yet there is no significant caste-based 'matching' between judges and litigants. In the outcomes of cases however, a distinct pattern emerges: cases involving petitioners and advocates with caste-neutral names show a slightly higher chance of dismissal relative to other low-status litigants. Conversely, respondents with caste-neutral names paired with similarly named advocates face a reduced likelihood of case dismissal. This pattern persists in appeals. These findings are consistent with litigants discriminating against higher-quality lawyers in favor of matching with a lawyer of their own group identity and highlight the enduring impact of social networks in perpetuating systemic disparities within the judicial system.

JEL Codes: Q53, Q56, C36, K32, K40

Keywords: inequality, social identity, justice

1 Introduction

A lot rides on a person's name. When printed on top of a resume, application form or petition in an institutional setting, a name can provide clues about an applicant's gender, race, religion, ethnicity and socioeconomic background. This can become the basis of institutional discrimination (Small and Pager 2020). In the United States, names that appear to be distinctively black have been associated with higher mortality (Cook, Logan, and Parman 2016), fewer responses to job-search applications (Bertrand and Mullainathan 2004) and less mentorship in educational institutions (Milkman, Akinola, and Chugh 2012). In Germany, unpopular or "negative" names are associated with neglect in online dating platforms, low achievement and weak self-esteem (Gebauer, Leary, and Neberich 2012).

Names are known to trigger deep psychological processes that shape a person's sense of self throughout their life, shaping their preferences and influencing their decisions in subtle ways (Pelham, Mirenberg, and Jones 2002). Names are known to correlate with the structure of social groups and networks that shape an individual's identity in childhood (Akerlof and Kranton 2000; Akerlof and Kranton 2002). Networks of people who have common core beliefs, values and norms however, can get locked into stable equilibria that persist over time (Collier 2016; McPherson, Smith-Lovin, and Cook 2001). Wealth and power associated with such networks can thus lead to the persistence of wealth over generations (Clark and Cummins 2015). Names can also be the basis of taste-based and statistical discrimination in markets (Small and Pager 2020).[1]

The power of names, and their role in shaping biases, is even more pronounced in post-colonial and poorer societies. In many parts of India, for example, a name may reveal a person's religion, region of birth and inherited caste (Gidla 2017; Parmar 2020; Deshpande 2011; Mosse 2018; Banerjee, Bertrand, et al. 2009).[2] Though discrimination on the basis of caste is illegal under the Indian constitution, names are routinely scrutinized for clues about a candidate's family background (Deshpande 2011). Banerjee, Bertrand, et al. (2009)

---

[1]Standard economic models of discrimination typically assume that actors are rational and make optimal decisions based on their preferences and information. The economic model of "taste-based discrimination" posits that people discriminate when they have a taste or preference for engaging with a certain group and are willing to pay a price to limit their engagement or interactions with another group (Becker 2010). An alternate model – "statistical discrimination" – posits that people who lack information about a specific individual use group characteristics to make inferences about them (Arrow 1998). In both these models, names can be a critical component of an information set about a person.

[2]Caste is a form of rigid social stratification that is unique to the South Asian context and has been a central organizing tenet of South Asian society for hundreds of years (Bayly 2001; Jodhka 2017).

find that in a field experiment, similar resumes with different caste-linked surnames receive large and significant differences in callback rates in call-center jobs. Thorat and Attewell (2007) find that college-educated lower-caste and Muslim job applicants fare less well than equivalently-qualified applicants with high-caste names when responding to "help wanted" advertisements in English language newspapers. Similarly, Deshpande and Newman (2007) report that recruiters on college campuses interpret names to draw inferences about class, family background and professional networks to assess the suitability of candidates for jobs at India's elite technical colleges.

Recent research from India suggests that names can be scrutinized for caste-markers even without the conscious awareness of the participants (Munshi 2019; Gidla 2017). Fisman, Paravisini, and Vig (2017) for example, examine dyadic data on religion and caste for officers and borrowers from an Indian bank and find that when both parties come from the same caste or religious group, loan amounts are higher, collateral requirements are lower, and repayment rates are also better after the officer leaves, suggesting that "cultural proximity" is valued in these settings. Even children who are made aware of their caste status change their behavior to conform to caste norms (Hoff and Pandey 2006; Hoff, Kshetramade, and Fehr 2011).

One approach towards mitigating the effects of caste-based discrimination has been to change names. In the United States, immigrants have often chosen non-ethnic names for their children as a deliberate effort to assimilate into the new culture and avoid the discrimination towards ethnic minorities (Abramitzky, Boustan, and Eriksson 2020). In many large companies and organizations, algorithms are being increasingly favored to screen resumes to eliminate the conscious and unconscious biases of humans in the screening process (Rambachan et al. 2020). The practice of "name blinding", i.e. removing names from resumes and applications at educational institutions or firms to reduce bias in hiring and recruitment, has also gained considerable popularity. The British government, for example, adopted "name blind" applications in 2017 for admissions to university as well as employment in the civil service, local government and various other government agencies.[3] While this research is still in its early stages and the long-term impact is largely unknown, evidence suggests that these practices may indeed lead to more minorities being hired in labor markets, but may have other perverse consequences. In France, an optional service offered to firms to join a new anonymous screening system actually increased inequality along other dimensions of identity and socio-economic status (Behaghel, Crépon, and Le Barbanchon 2015).

---

[3]https://www.bbc.com/news/magazine-34636464.

In India, efforts to conceal and change names have been gaining popularity since Independence in 1947. Indians may change their last names and, indeed, have been doing so to conceal their caste, religion and background (Jayaraman 2005; Parmar 2020).[4] Though the prevalence of this practice remains unknown, it is widely recognized in Indian society that some of India's most famous celebrities have changed their names or entirely renounced their last names.[5] Some elected leaders have renounced their last names to disassociate from their caste identities in a polarized political system.[6] New evidence suggests that in recent years, Indian Muslims have been increasingly adopting caste-neutral Hindu names to conceal their identity in urban labor markets and social engagements, but they refrain from official name changes due to growing anti-Muslim sentiment within the ranks of institutions (Menon 2022).

To what extent do names influence outcomes in institutional settings? Moreover, does the adoption of a caste-neutral name reduce the considerations of caste in these settings? There is almost no large-scale non-experimental study on this issue in a post-colonial setting like India. We study name-based networks and their impacts on cases in the judiciary of Bihar, a single state within India. The justice system is an ideal arena to study the power of a name to shape actual outcomes. In most instances, it is the *only* marker of identity that is observable through the pipeline of justice. Since it is an arena of power where decisions have high stakes, people are required to adhere to strong guidelines (Djankov et al. 2003). They are likely to use their official names. Courts are also ideal settings to study unconscious bias (Chen, Moskowitz, and Shue 2016; Berdejo and Chen 2017).

Our analysis focuses on the state of Bihar. With a population of over 100 million people, Bihar compares in size to Vietnam and is larger than the UK, France or Germany. It is also one of India's most socially stratified states. Caste and religion are major dividing

---

[4]According to the Gazette of India, the process takes three steps. First, an individual must sign an affidavit with the old and new names, as well as the reasons for changing the name. Next, they must publish the name change in two local newspapers (one English newspaper and one regional newspaper published in the official language of the State). Finally, the applicant must notify the Central Gazette of India, which is located in Delhi, either through an in-person visit or else via registered mail. Once the change of name is approved, a citizen can change their name in all documents (other than past educational documents) and use their new name. This information can be found at https://www.deptpub.nic.in/sites/default/files/Change-of-Name-and-Gender_1.PDF.

[5]India's most successful actor, Amitabh Bachhan was originally named Inquilab Srivastava, Akshay Kumar was named Rajiv Hari Om Bhatia, Govinda was named Govind Arun Ahuja, Madhubala was Begum Mumtaz Jehan Dehelvi, Dilip Kumar was originally named Muhammad Yusuf Khan, Meen Kumari was originally Mahjabeen Bano, and the renowned comedian who is widely known as Ajit was Hamid Ali Khan.

[6]In 2016, Manohar Lal Khattar, the Chief Minister of the state of Haryana, renounced his last name after caste-based violence led to the death of 30 people in his state (https://timesofindia.indiatimes.com/india/haryana-cm-castes-aside-surname-after-jat-agitation/articleshow/51520736.cms)

lines (Joshi, Kochhar, and Rao 2018b; Joshi, Kochhar, and Rao 2022). Bihar's economy is agrarian, social relations are feudal, and gender relations are patriarchal with some of the highest levels of gender inequality in India (Chakrabarti 2013). After Indian independence, agrarian movements, affirmative action programs, caste-based movements and more recently, women's movements have produced new leaders in the state (Shah 2004). Since 1991, alliances between minorities have resulted in novel arrangements of power-sharing and many new leaders have emerged from these groups (Jaffrelot 2010; Jaffrelot and Kumar 2012). Gains in political representation for these groups, however, have not always translated into greater economic or social opportunities for those who have been historically vulnerable. The age-old fault lines of caste, religion and gender continue to be salient in the state.

We use a novel dataset of more than one million cases, all listed in the public domain, filed at the Patna High Court between 2009 and 2019. We explore the role of social identity at the courts through the application of machine learning algorithms on databases of names from petitioners, respondents, judges and lawyers at the Patna High Court as well as other large state institutions in Bihar. We examine the frequency of last names, draw inference about caste, and then use an empirical regression framework to examine whether petitioners, respondents, and advocates from these groups are more likely to match with each other. We also examine how matching on the basis of social identity may affect the outcomes of cases in the justice system.

Our analysis begins with the study of names. We examine the top last names at the court and note that there is a significant concentration of names in the data. The top 10 last names (Singh, Kumar, Devi, Yadav, Kumari, Prasad, Jha, Rai, Sharma, Sah) account for 55% of petitioners. Next, we use algorithms to assign each name the likely caste, religion and gender. We find that almost all the top 10 names are caste-neutral, i.e. the caste identity of the person who uses the name cannot be clearly identified. The majority of all petitioners at the court – more than 50% use such caste-neutral names and the prevalence has been increasing between 2009 and 2019. The practice shows considerable variation within the state. We see the highest prevalence of caste-neutral names in the urbanized district of Patna and considerably lower prevalence in the northern areas of the state.

Next, we study the networks of names as well as caste and religion within the courts. We find very little evidence of "matching" between judges and litigants, but we do find that petitioners and their advocates match along the lines of caste. Advocates with caste-neutral names appear to be particularly effective. We find that matching on the basis of broad

measures of identity, such as belonging to "Scheduled Castes" or the use of caste-neutral names, can have modest but yet discernible impacts on both the processes of justice as well as outcomes. Low-status respondents who match with advocates from their own group appear to have shorter case-processing times. Matching can also affect the outcomes of appeal cases. We find that relative to respondents with SC-sounding names, respondents who match with lawyers with caste-neutral names appear to experience a 23 percentage point decrease in the likelihood of having their case "Dismissed," and a 16.2 percent increase in the likelihood of having their case "Withdrawn." This is likely for a variety of reasons that include better communication between advocates and their clients, stronger community support and also, the better positioning of name-blinded lawyers to negotiate the processes of justice and avoid the pitfalls of bias and discrimination within the system.

The rest of our paper is organized as follows. Section 2 provides an overview of the context, Section 3 provides an overview of data, Section 4 gives a summary over identities of petitioners, respondents and judges at the Patna HC, Section 5.3 focuses on the use of caste-neutral names, Section **??** analyses matching between judicial actors, Section 7 provides a discussion and the final section concludes.

2 Context: Politics, Society and Justice in Bihar

Bihar has long been regarded as India's poorest and most lawless state. In the words of a recent Chief Minister, Jatin Ram Manjhi, Bihar was until 2005 "seen as a perfect example of a failed state" where "law and order was abysmal, different caste-groups possessed private armies that would often kill innocent people, public and social infrastructure was in appalling condition and welfare supports for the poor rarely reached its target" (Manjhi 2014). This malaise is widely believed to have its roots in the colonial period, where regressive land tenure systems constrained opportunities for investment in infrastructure, industrialization, and other aspects of social development and persisted through institutions (Banerjee and Iyer 2005). In the subsections below, we review the important lines of social stratification and the system of justice in Bihar. These details are essential to understand the context in which an individual may change their name or seek representation to override concerns about their community background.

2.1 Caste

According to the most recent estimates, Hindus account for 82% of the population of Bihar (Verma 2023). Though all religious groups of India are stratified by caste, for Hindus it is a particularly important line of stratification. The concept of caste is understood in a variety of ways. In official records and policies, the Hindu population of Bihar is routinely defined in broad official categories such as Forward Caste (FC), Scheduled Caste (SC) and Scheduled Tribe (ST).[7] In everyday life however, for Hindus and even many Muslims in the state, identity is experienced and practiced as jāti (henceforth, jati) (Bayly 2001; Jodhka 2017). These are hereditarily formed endogamous groups whose identities are manifested in a variety of ways that include (but are not restricted to) naming conventions, geographic location, occupation, property ownership, diet, gender norms, social practices, and religious practices. The population of Bihar comprises hundreds of individual jati groups that have distinctive economic, political and social identities. The placement of jatis in broad government "caste" categories has always been complicated: a single category can include jatis that display considerable inter- and intra-level inequality (Joshi, Kochhar, and Rao 2022).

British administrators used caste as an official marker of identity in colonial governance practices. Recruitment into the army as well as government jobs was often made on the basis of jati-identity, which was determined from detailed census surveys that began in the late 1800s (Dirks 1989; Bayly 2001). In the aftermath of colonial rule, the Indian Constitution of 1947 sought to adress caste-inequality in a variety of ways. Most notably, Article 14 guarantees equality before the law and equal protection of the laws to all citizens, and also prohibits discrimination on the grounds of caste or religion (among other forms of identity). Articles 15(4) and 16(4) allow the state to make special provisions for the advancement of socially and educationally backward classes, including Scheduled Castes, Scheduled Tribes, and Other Backward Classes (OBCs). This provision enables the reservation of seats in educational institutions and public employment to address historical social and educational disadvantages.

Bihar's modern caste hierarchy conforms to this modern categorization. At the top of the system are upper-caste Hindus, officially known as Forward Castes. According to estimates from the recent controversial caste census of Bihar, this group accounts for about 15-20%

---

[7]The groups "Scheduled Caste (SC) and Scheduled Tribe (ST) are officially designated groups of people and among the most disadvantaged socioeconomic groups in India. Lists for each state, as specified by the government of India, are available here:https://socialjustice.gov.in/common/76750.

of the state's population (Kumar 2018, Verma 2023).[8] It includes mainly Brahmins, Rajputs, Bhumihars, and Kayasthas. These groups have wielded significant power in Bihar for much of the colonial and post-colonial periods of history, and have owned arable land and dominated state politics (Diwakar 1959). In the years before Indian independence, the Kayasthas were dominant and were recruited heavily into the colonial government. After Indian independence, the other upper-caste groups entered the political arena. Though the groups competed, together they dominated the Indian National Congress (INC) that ruled the state almost continuously until 1977. Today, they are about twice as likely to be literate and hold land than their lower-status counterparts (Joshi, Kochhar, and Rao 2018a). Recent estimates from Bihar's caste-based census suggest that this group has the highest levels of income and asset ownership in the status than any other caste-group (Tewary 2023).

The "backward classes" of Bihar have always been a significant fraction of the population of the state, though estimates depend on the precise definition that is used (Blair 1980). Today the group "Scheduled Caste" includes the jatis who were historically landless and engaged in manual labor. According to the most recent estimates, these groups account for 19.65% of the state's population (Verma 2023, Tewary 2023).[9] Additional categories such as Other Backward Castes (OBC), Backward Caste (BC) and Extremely Backward Caste (EBC) have also been recently defined to include certain groups within this broad category as well as some Muslim groups. The OBC group includes jatis who are educationally or socially disadvantaged due to historical marginalization, but not classified as SC or ST. This group now accounts for 27.13% of the population (Verma 2023, Tewary 2023). Scheduled Tribes, once a significant minority in the undivided state of Bihar, now account for just about 1.7% of the population; most of the tribal areas are part of the newly created state of Jharkhand.

The power struggle between the upper castes and the landless "backward" groups of Bihar has been a frequent driver of political instability, social unrest and outright violence in the state (Jaffrelot and Kumar 2012). The 1980s were marked by the emergence of private caste armies that defended and represented the interest of upper castes and landed backward castes (Chakrabarti 2013). In the 1990s, backward caste mobilization intensified under the leader Lalu Prasad Yadav, who openly declared war on the *Bhura Bal* (Bhumihars, Rajputs, Brahmins and Kayasthas) (Kumar 2012; Kumar 2018). For much of this period, law and

---

[8]We are unable to access the formal report and are thus relying on secondary sources for these estimates.

[9]Until this recent report, estimates of the population of the SC group was available from the Census of India, 2011. According to this source, SCs accounted for 15.7% of the state (Registrar General & Census Commissioner of India 2011).

order languished in Bihar, and most development policies were designed and implemented almost entirely in caste terms. The absence of a pan-state identity and sub-national solidarity presents a striking contrast with many other Indian states (Singh 2015).

These pressures, however, while eroding law and order and stifling development in the state, have also contributed to the emergence of some of the most ambitious affirmative action policies that were intended by Articles 15 and 16 of the constitution (Kumar 2018). In 1977, the first non-INC government instituted policies that reserved 20 percent of public sector jobs for Bihar's Other Backward Classes (Chakrabarti 2013). In 1991, reservation policies for OBCs were implemented all over India. The architect of the report that inspired the policies – Bindheshwari Prasad Mandal – was the former Chief Minister of Bihar and hailed from an OBC community. Today, lists of OBCs are maintained by both the National Commission for Backward Classes and the states of India. In Bihar, this group includes jatis such as Bania, Yadav, Kurmi, and Koiri. These OBCs are now widely understood to be agrarian communities that have acquired land, adopted improved agricultural technology, and benefitted from the reservation policies in public sector education and employment, particularly after 1991 (Kumar 2018). Unlike other states of India, in Bihar this group is large enough to change the balance of political power in the state. The Yadavs for example, account for 15% of the Bihari population, and have made a distinct mark on electoral politics and the composition of the legislature and leadership of the state (Jaffrelot and Kumar 2012; Kumar 2012; Kumar 2018).

Overall, caste has proved to be a particularly persistent form of inequality in Bihar: inequality between broad caste-groups remains significant, even though some jatis have improved their socio-economic position as a result of their efforts to mobilize in recent years (Joshi, Kochhar, and Rao 2022). The early 2000s brought momentous changes to Bihar – Nitish Kumar and the elites who enabled his ascendancy to power have overtly rejected caste-based politics and social identity and emphasized a distinct "Bihari" identity that transcends caste and religion (Chakrabarti 2013; Kumar 2018). Investments in the social sector of the state have intensified and Bihar has been the site of numerous innovative development programs, particularly aimed at women and minorities. The recent release of Bihar's caste-based survey has intensified these debates and raised the stakes for politics in the state.

While this paper largely focuses on caste, we note that gender identity intersects with both religion and caste. Bihar has some of the highest levels of gender inequality in India. According to the Census of 2011, only 63% of adult women were literate, and even though

9

estimates from surveys since then suggest that the ranks of literate women are growing, Bihar has the lowest percent of girls completing secondary education in India, and it ranks at the bottom of the education index used by the Government of India (Niti Aayog, 2020). With 916 females per thousand males, its sex ratio remains below the Indian average. With just 9 percent of women participating in the formal labor force, Bihar also has the lowest rate of female labor force participation in the country (World Bank, 2016). Though women's groups have gained considerable power in the state through livelihoods programs and self-help groups, and have even managed to achieve a prohibition of alcohol in 2016 (we say more about this later in this paper), women continue to be underrepresented in markets as well as formal institutions. Women from disadvantaged caste groups face the double disadvantage of caste- and gender-based inequality (Joshi, Kochhar, and Rao 2018a). In future work, we plan to examine women's access to the courts and their challenges of securing justice more broadly.

2.2   Justice System

The Patna High Court is about 100 years old. It was first established by the British in 1912 and began hearing cases in 1916, with a Chief Justice and six other judges. The sanctioned strength of the court has expanded several times over the years. There are currently 22 permanent judges, including the Chief Justice and 14 additional judges, while the sanctioned strength was a total of 53 judges in 2019. Bihar has sent more justices to the Supreme Court than any other Indian state (Chandrachud 2020).

Prior research on the Indian justice system has argued that the system's colonial roots continue to influence the courts in a variety of ways. Colonial courts were designed by colonial administrators and sought to secure Indian subjecthood rather than serve citizens (Menski 2006). After India's independence, the development of the Patna High Court has been constrained by weaknesses of state capacity, caste-based conflict and the episodic violence in the state (Chakrabarti 2013; Jaffrelot 2010; Kumar 2012; Kumar 2018). Political battles, often accompanied by complex allegations of corruption and criminality, have often found themselves being decided in the Patna High Court, straining the court's political neutrality (Roy 1997). In recent years, however, the challenges of the Patna High Court largely align with those of the Indian justice system more broadly (Sen 2017).

3 Data

3.1 Patna High Court Cases

Our data were drawn from the publicly accessible records of the Patna High Court. We scraped the data and processed it to construct a comprehensive database with attributes of cases that include the filing date, case type, and names of the judge, lawyer, petitioner and respondent. To get a comprehensive understanding of the functioning of the court and related ecosystems, we used auxiliary data from police stations, district courts, and judge biographies that are available on the court websites. We focus on the time period 2009-2019. This gives us a total of 1,071,068 unique cases in the high court. 364,163 (34%) of these are civil cases, and the remainder are criminal.

A trend of these cases is presented in Figure 1. Note that the number of criminal cases has been rising since 2015. This is likely a result of the controversial Bihar Prohibition law that was imposed by Nitish Kumar's government in 2015, declared unconstitutional by the Patna High Court in 2016 and then reimposed by the Supreme Court of India later that year, allowing the government to continue with the ban (Dar and Sahay 2018). Anecdotal evidence from the courts suggests that more than 200,000 people have been booked under the Bihar Prohibition and Excise Act 2016 and more than 50,000 bail applications are pending at the High Court of Patna.

Figure 2 displays the spatial variation of civil cases (Panel A) and the share of civil cases of all filed cases (Panel B) across Bihar's districts. While the district of Patna strikes out as having by far the highest number of civil filings, there is much less spatial heterogeneity in the share of civil cases.

3.2 Appeal Cases

Indian high courts serve as the highest courts of appellate jurisdiction in each state and union territory. Prior research has noted a steady increase in the number of appeals that have been filed at these courts (Ghosh 2018). We scraped data from the eCourts portal to obtain summary information on case types, filing dates, and decision dates (if a given case has been decided), as well as the names of the presiding judge, plaintiff, and respondent. These data

allow us to track the progress of appeals in the justice system.[10,11]

An appeal filed by a litigant moves through the pipeline of justice in several stages (see Figure 3). Tracking the status of appeal cases in the court system requires us to merge the eCourts databases at the district and high court level. We do this by leveraging information in the high court database. Specific variables include the decision date of the lower court case, the registration number, and the registration year (or a subset of these variables). This information, however, is not sufficient to uniquely identify the lower court case since names can be spelled differently, a middle initial can be included or fully spelled out, and/or details can be entered with error. We assess the quality of the match and choose the cases where the litigant's names in the district case closely match (i.e., above a threshold) the litigant's names in the high court case.

Once we obtain a set of appeal cases that we are confident that we have matched, we rely on the official language of the courts themselves to label these stages.[12] When the case is first filed at the high court, a case is marked as "Admitted" (the paperwork is in order and the case is deemed suitable for processing at the high court and listed to the court for hearing/decision) or "Rejected" (the application is not in the jurisdiction of the court - territorial or in content - or the paperwork has defects of non-curable nature). If a case was admitted, it can be heard by the high court and when the bench takes a (whatever kind of) decision the appeal is said to be "Disposed". A disposal can be of different natures: the appeal can be "Allowed" (accepted) or "Dismissed" by the bench or "Withdrawn" by the petitioner.[13] Figure 3 gives an overview over these different stages and potential outcomes of an appeal case.

3.3 Additional Data on Names

For the purpose of comparing the courts to other institutions in Bihar, we supplement the judicial data with additional data from several sources, all of which are fully publicly available:

---

[10]India's high courts have the jurisdiction to hear appeals from both civil and criminal cases from the lower (district) courts.

[11]Civil cases that were decided by district judges and Munsif courts can be appealed at the high court. Criminal cases that were decided by district judges, judges at Munsif courts and the sessions courts can also be heard as long as the sentence for imprisonment is seven years or more. High Courts can also hear appeals from the orders of the tribunal.

[12]These are official terms in the eCourts system that are applied by the e-filing administrator at the time of the review of the paperwork in the case (see https://ecourts.gov.in/ecourts_home/static/manuals/efiling-User-manual.pdf)

[13]Theoretically, the decision to withdraw can be made by the appellant at any time, even at the administrative stage before being admitted.

Socio-economic Caste Census (SECC) for Bihar: The SECC data has not been officially released by the Government of India. We rely on the replication files of existing research to access these data (Sood and Laohaprapanon 2018).[14]

Registered Farmers: We use a database of about 1.3 million registered farmers from the Bihar Cooperative Department[15] We have no information on the criteria that were used to include farmers in this database. It is quite likely, however, that the most well-connected and well-placed farmers were able to register and take loans in the early stages of this registration effort. This, however, makes the dataset particularly well-suited to understand who benefits from agrarian policies in Bihar.

Government Employees: We draw on a database of 210,389 employees of the Bihar state-government who have officially disclosed their financial status to comply with policies of the Government of India.[16]

Judges: Our research team constructed a database of all judges who have served at the Patna High Court, which includes not only their years of service at the court, but also their age, recruitment source, date of appointment as an additional judge, date of appointment as a permanent judge, and retirement date from handbooks from the years 2014, 2017 and 2020. This gives us a sample of 83 judges.

Each of these databases provides insights into the occupational choices and professional activities of groups that are defined by distinct caste and religious identities. These comparisons illustrate the extent to which the judiciary can be compared to other organizations or professional groups in the state of Bihar. The number of observations for each of these datasets is summarized in Table 4. Table 4 contains the breakdown of the populations in our datasets along the lines of caste, religion and gender. We look at the following categories: SC or ST, Muslim, "Hindu other" and women. Table 4 contains some striking findings. To the extent that the SECC is the most representative of the population of the state, it is noteworthy that no professional group we consider here perfectly represents the population of Bihar. Even in this unequal playing field, however, judges are quite distinctive. Almost all Judges fall into the category "Hindu Other", which includes all Hindus except SCs and STs. There are no

---

[14]These data were scraped from the public records of the MGNREGA website in 2020 (http://164.100.129.6/netnrega/secc_list.aspx) and made available on the Harvard University dataverse.

[15]The registered farmers' data is available on the Co-operative department website (http://cooperative.bih.nic.in/FarmerDB.htm). The data can be accessed district-wise, we scraped the data in June 2020.

[16]The asset data is available at http://bpsm.bihar.gov.in/assets/. It can be accessed via the employee's department, district, or the public sector undertaking that they are a part of. We scraped these data in June 2020.

ST judges and very few other minority judges. Of the 84 judges observed over the 11-year period, only 6% are women and less than 10% are Muslim. Women are underrepresented everywhere. They constitute half of the population, but their share in the different professions varies between 10% (judges) and 26% (government employees). The professional group that comes closest to representing the population of Bihar is the group of government employees. As noted in the earlier background section, several affirmative action programs have reserved seats at these institutions for SCs and STs.

For the sake of a historical comparison, we also draw on a newly compiled dataset of Indian surnames from Ancestry.com and familysearch.org, two leading websites that have attempted to gather detailed ancestral records for individuals who once served in the British Indian Army or the British Indian government that was present in India until 1947. Given that Bihar was one of the earliest regions that was placed under British administration, these records provide excellent insights into the prevalence of certain names and identities in the history of Bihar. We rely on these data simply for the prevalence of specific last names, which are discussed in the section below.

4   Who Contests Cases at the Patna High Court? A look at last names

We begin by examining the frequencies of names in our different sources of data for Bihar. Naming conventions in this region largely follow North Indian practices, i.e. an individual's given name and last name (surname) are typically assigned at birth and contain markers of their region of birth, caste and religion.[17] The process of changing names follows the process that was outlined earlier in this paper.

We begin by examining the names that occur most frequently in the data from the Patna High Court. We compare the incidence of specific last names to other populations in Bihar (Figure 4). We find that the top 10 last names across all the databases considered here (not including the judges data) are remarkably similar. Singh, Kumar, Kumari, Prasad, Yadav, Paswan,

---

[17]In their review of North Indian naming conventions, a leading British document explains this as follows: "As a Hindu family name can denote caste, a person may drop his/her last name to reject the caste system. It is then typical to use the middle name as a surname. In this case, the name used as a surname would not be common to all family members: e.g. Raj Lal VASANI would become Raj Lal and therefore Mr. Lal" (Financial Banking and Information Infrastructure Committee of the United Kingdom, 2006). Report is available at: https://www.fbiic.gov/public/2008/nov/Naming_practice_guide_UK_2006.pdf

Ram, Jha Sinha and Mishra occur the most frequently.[18] These specific names alone account for 58% of senior government employees, 59% of farmers, 40% of petitioners and 29% of respondents at the Patna High Court.

Within the group of the top 10 most popular names, some are more popular than others. The last name Singh, for example, accounts for 10% of senior government employees, 14% of farmers and 12% of petitioners in the Patna High Court. Kumar, however, accounts for 19% of all senior government employees but only 7% of farmers (Figure 4). The name Yadav is one of the less common in this group. In our data however, this name accounts for 3% of government employees, 7% of farmers, 4% of petitioners in the Patna High Court and just 0.2% of respondents. Yet this name accounted for 11% of the Bihari population in the 1931 Census of India, 14% in the 2011 Census of India and 7% in the SECC data. We observe no judges with this last name.

In contrast to Singh and Kumar, which are caste-neutral names, stand names that contain strong signals of a person's caste status. The name Yadav is a good example of this. According to the Yadav Welfare Association – a pan-India association that exists to serve this group – Yadavs are a caste-group who descended from an ancient royal clan, are "mentioned in ancient Dharmic scriptures" and are now classified by many state governments as "Other Backward Classes" because of "prevailing general economic and educational condition".[19] This group has gained considerable political power in recent decades (Kumar 2018; Chakrabarti 2013; Jaffrelot 2003). One of Bihar's most controversial and high-profile Chief Ministers, Laloo Prasad Yadav, hailed from this community. It is noteworthy that even though they have become prominent in the executive and legislative branches of government, they are still poorly represented in the judiciary. It is also noteworthy that we see no Muslim names in the top 10. This is largely due to their minority status in the state.

This high concentration of last names in all the different groups considered here also presents a striking contrast with other regions of the world and other periods in history. As reported

---

[18]We omit judges from consideration here because the process of becoming a High Court Judge is strongly affected by decision-makers outside the state. The process follows Article 217 of the Constitution of India. This article states that a judge is to be appointed by the President of India in consultation with the Chief Justice of a High Court in consultation with the Governor of the state. In the case of the Patna High Court, our analysis found that 43% of high court judges were appointed from the bar (i.e. they were practicing as advocates at the time of being appointed to the High Court), 25% had been serving as judges in the subordinate judiciary and the remainder had been transferred from outside the state in accordance with the preferences of the Chief Justice of India.

[19]https://www.yadavsamaj.co.in/, Accessed on March 27, 2023.

in Greif and Tabellini (2017), in 1541 only 7.85% of the population of London had the top 10 most common last names. Similarly, the English Census of 1881 shows that the mean proportion of the population with the most common last name in English towns was 0.079 and the mean proportion of the most common 10 last names was 0.39 (Greif and Tabellini 2017). These are significantly lower than what we see in any survey of Bihar today. Previously, literature has argued that the diversity of last names in a population can be indicative of migration, industrialization and the ease of impersonal exchange (Greif and Tabellini 2017; Clark and Cummins 2015). In the context of Bihar, however, there are many reasons why this is likely to be more complicated.

Overall, the high concentration of last names in our data highlights the caste-based inequality in modern Bihar. The study of name clustering in other societies suggests that this can be indicative of barriers to social mobility (Clark and Cummins 2015; Greif and Tabellini 2017).

Next we go beyond the list of most frequently occurring names, and decode caste in our data. The biggest challenge we face here is that legal data sources contain no information on an individual's caste or even their religion. To map names onto caste, we must rely on the SECC data. We also conducted a small qualitative survey in the state of Bihar to better understand the associations made by people in the state between caste and social status.

5 Analysis of Names for Markers of Social Identity

As discussed earlier, Indians routinely scrutinize names for markers of caste. In Bihar, like most other parts of northern India, a person's gender is decoded from a first name and their "group" based identity, i.e. caste, are decoded by last names. For Hindus, markers of caste are almost always contained exclusively by last names.[20] An individual can belong to multiple groups. Women can belong to any religious or caste group in the sample. And even for men, a name can signal affiliation with multiple groups since a certain name can be associated with several religions.[21] We rely on simple methods of machine learning to identify the relationship between caste and social status in our data. We also rely on a small survey conducted in the state of Bihar to check the validity of our findings. We perform additional validation of the results by examining online caste-identification discussion forums that have become common in recent years. These are described below.

---

[20]Caste among Muslims and Christians in India may also be decoded from last names, but in the context of Bihar, given the minority status of these populations, it is difficult to draw inference about these sub-groupings.

[21]The last names *Malik* and *Shah* for example, can belong to Hindus or Muslims.

5.1 Algorithmic Inference of Names

We first focus on identifying the religion of all individuals in our data. Since our SECC sample does not have identifiers for respondent's religion, we begin by using a sub-sample of our legal data to construct a comprehensive set of Hindu and Muslim names. We construct this by extracting names of litigants who contested cases at the Patna High Court related to the Muslim Women (Protection of Rights on Divorce) Act, 1986 and the Hindu Marriage Act, 1955. Under the assumption that only parties who hail from these religious backgrounds can file cases under these acts, this provides us with an extensive and representative set of names for both men and women from the two religions and genders that are relevant to the case of Bihar.

This database of names functions as training data for a machine learning algorithm whose objective is to predict whether any name that is given to it is Hindu, Muslim or "Other".[22] The machine learning algorithm extracts groups of 1–4 characters (including white spaces) from the names, stores this in the form of "feature vectors" and measures the co-occurrence of alphabets. Since the training data contains information on religion, the algorithm is able to note the distinctive features of Muslim and Hindu names. For example, for Muslim names (relative to Hindu names), it notes features such as the higher incidence of the alphabets 'z' and 'q', or the higher frequency of the co-occurrence of "mm" or "ee". Once the analysis of the training data is complete, the algorithm is used to make out-of-sample predictions in the court data.[23]

Next, we refine this algorithm to also predict caste on the basis of a last name. For training data for caste, we rely on the list of names in the SECC data. Since each household head in this survey was asked to report their full name and their caste affiliation, we can explore the distribution of caste groups that are associated with a single last name.[24] We thus split the SECC randomly into training data (90%) and testing data (10%). We use the training data to make a list of predicted caste affiliations for the top-100 names and then use out-of-sample

---

[22]Unfortunately, the number of litigants who hail from Christian, Budhist or other religious groups in our sample is too small to make appropriate matches. All those who are neither Hindu nor Muslim are included in the "Other" category.

[23]To make the final prediction out-of-sample, we used a variety of methods to evaluate various classifiers, including decision trees, random forests, logistic regression, linear SVM and character RNNs. We found that the simple linear SVM and logistic regression classifiers performed the best on the various metrics, producing an accuracy of around 91%. For this, and their simplicity, we relied on logistic regression to make the out-of-sample prediction on the basis of the feature vectors.

[24]We emphasize here that we do not have access to the full SECC data on caste and religion. We simply rely on the broad caste categorizations (SC, ST, and Other).

predictions on the testing data to determine the accuracy.

We flag a name as "caste-neutral" if at least 15% of respondents report a different status than the rest of the group. Using this threshold, of the 22,390,585 households in the SECC data, approximately 23% reported caste-neutral names. To get stability in our estimates across this heterogeneous sample, we use the trained algorithm to construct a prediction of caste for each name in the testing data. This process is repeated several times (N=10 for the purpose of this paper) on the training data to refine the quality of the out-of-sample accuracy within the SECC. Once this iterative process concludes, we calculate the share of predictions for each caste in the testing data. This set of predictions is then merged with the names of the judicial actors. In the end, the accuracy of our algorithm in the in-sample SECC data was 92% and the accuracy of our prediction on the testing group (also in the SECC data) was 87%.

At the end of this process, the caste status predicted by our algorithm is the outcome of a random variable whose distribution over the categories is given by the distribution of the last name over the social categories in the training data. When presented with the same last name (for a different individual) we make another independent prediction, i.e., use an independent random variable with the same distribution. The key assumption made here is that the statistical composition of the population is the same as that of the set of people appearing in the courts data. In order to reduce the generalization error, we trained multiple classifiers, including a logistic regression model and a random forest classifier, to make predictions of caste for every name in our sample. We then implemented a voting procedure on the outcomes of these models to generate a final prediction of the three aspects of social identity for each name. We are thus able to obtain a prediction of caste of each participant in the courts.

It is important to note that there is a stochastic component to this classifier. The same name could be predicted to be of a different category when the classifier is reapplied, however the probability of a specific category is dictated by the normalized name counts. The name Trivedi for example, has the normalized weights on [Other, SC, ST] given by [1,0,0], so it is always predicted to be of the 'Other' category. In contrast, the name Kumar has weights given by [0.88, 0.12, 0.01], so although there is a very high chance that the name is predicted as 'Other', there is also more than 12% chance that it is predicted as 'SC'.

To further validate the predictions of religion and caste that emerge from this method, we also conduct a small and informal survey of elderly residents of Panta to validate our findings.

5.2 Robustness Check with a Qualitative Survey

To check the accuracy of our findings, we also conducted an informal survey of a dozen elderly women in the city of Patna who had spent their entire lives in the state and had extensive knowledge of social structures in the state. The goal was to check what associations, if any, were made between specific last names and markers of caste and religion. We presented survey respondents, who were all grandparents who had successfully contracted caste-based arranged marriages within the past 10 years, with a list of names, followed by a series of questions about the caste, or religious background associated with the name.[25]

Although our survey was modest in scale and not designed for rigorous analysis, it provided valuable insights into the intricate interplay between caste and names in Bihar. Notably, our findings revealed that Muslim names were universally acknowledged as such, alongside upper-caste names like Bhumihar Brahmins (who have names such as Ojha, Pande or Upadhyaya). Names associated with dominant castes in Bihar's politics (such as Yadav) are also immediately understood to be from the OBC category. Moreover, while not universally recognized, at least some Scheduled Caste and Scheduled Tribe names were identified by significant proportions of our surveyed population. For instance, over three-quarters of respondents correctly associated Manjhi with Scheduled Caste last names, though opinions varied on names like Mandal, Prasad, and Rai. As one respondent noted, "these are supposed to be caste-neutral, but only lowest caste individuals adopt them, so they are trying to remove the stigma of belonging to the lowest caste and join in with the Dalits more broadly". This largely aligns with the broad discussions of caste-neutral names in the Indian media.[26]

Nearly all respondents said that within the Hindu community, certain names were completely caste-neutral and no clear inference could be made about a person's caste from these names. This category include Kumar, Kumari, Singh, Dev, Devi, Sinha (which is similar to Singh) and Mishra. Some also argued that an additional category of names was once clearly indicative of low status, but has become increasingly challenging to interpret and is increasingly viewed as caste-neutral. These respondents spoke of the blurring of lines between "caste-neutral" and "SC" names. Names like Paswan, Ram, Prasad, Mandal and Rai have gained

---

[25]The names we focused on included the top caste-neutral names that emerged from our algorithm, i.e. Kumar, Kumari, Prasad, Singh, Sinha, Jha, Mandal, Mishra, Baitha, Bharthi, Das, Dev, Devi, Safi, Ram, Rai, Roy and the many variants of the name Chaudhary as well as Bhumihar Brahmin names like Thakur, Sharma, Pandit, Dutt or typically SC names such as Mandal, or Paswan.

[26]See https://timesofindia.indiatimes.com/city/patna/using-surnames-to-conceal-identity/articleshow/4162892.cms.

popularity in recent years, not only because of their widespread adoption by Hindus from lower-status groups, but also because these are the surnames of politicians who have been staunch advocates for affirmative action programs at the national level. Individuals from deeply disadvantaged communities such as Musahars have adopted these names to conceal the depth of their marginalization in mainstream society. By adopting these names, they not only signal their eligibility for affirmative action programs but they also reduce the risks of discrimination towards them by other SC groups. In essence, some marginalized communities opt for caste-neutral names not to eradicate but to mitigate the extent of their disadvantage in Bihari society. All our respondents felt that in courts in particular, low-status petitioners would be more likely to use caste-neutral names because of the widespread belief that judges are typically drawn from high-status caste-groups in Bihari society.

Finally, we note that respondents in our survey also pointed out that caste-neutral names can also be adopted by higher castes or Other Backward Classes (OBCs) to obscure their privilege and minimize the risk of reverse discrimination when dealing with state institutions. Most of our respondents felt that in these settings even some high-caste groups would adopt caste-neutral names to ensure that their caste is not used against them in the court proceedings.

5.3   Caste-Neutral Names: Broad Patterns of Use

We combine information from the algorithmic caste assignment as well as well as our qualitative research described above to arrive at a list of about 16 caste-neutral names: Kumar, Kumari, Prasad, Singh, Sinha, Mandal, Mishra, Baitha, Bharthi, Das, Dev, Devi, Safi, Ram, Rai, and the many variants of the name Chaudhary (this includes Chowdhry, Chowdhury, Choudhary, Chaudhry, Chowdhry, Chodhry, etc.). We emphasize here that many of these names, such as Prasad, Mandal and Ram are widely regarded as low-caste names that have been increasingly adopted to conceal a deeper form of disadvantage than just SC. Not all the respondents in our sample would agree that a name like "Mandal" should be regarded as just SC. To address this issue, we will conduct our empirical analysis in two steps: we will group SC and neutral names together and call this group "Low status", and we will also study them separately.

We first explore the overall prevalence of caste-neutral names. Table 1, which presents the descriptive statistics of the full sample, shows that 50% of petitioners and 57% respondents use caste-neutral names. Only 12% of petitioners and 10% of respondents use SC-sounding names. Estimates of caste-neutral names among advocates are similar. 63% of petitioner's

advocates and 59% of respondent advocates use such names. Figure 5 illustrates these trends over the sample period. We note that the group of petitioners, respondents and judges who use caste-neutral names is larger than any other group in the courts, and this has remained stable over the sample period. In the case of judges we see a small decline, from approximately 50% in 2014 to 43% in 2018, but this has returned to 50% in the last year of our sample, 2019.

We find it noteworthy that all stakeholders in the courts appear to be relying on caste-neutral names. In the narrower sample of data on appeals, we continue to find that over half of all petitioners and 68% of petitioners' advocates who are contesting appeals in our sample are relying on caste-neutral names (Table 3).

In Figure 6 we explore the spatial distribution of the use of caste-neutral names by petitioners (panel A), petitioner advocates (panel B) and judges (panel C). Note that there is considerable variation by district, among all three sets of stakeholders. The prevalence appears to be lowest in the northern and eastern regions of the state. These are mostly rural areas where governance is weaker, lawlessness is greater and strife along the lines of caste and religion have been common in Bihar's recent history (Chakrabarti 2013; Kumar 2012; Kumar 2018). Conversely, the practice of using caste-neutral names seems to be most favored in Patna, the largest city in Bihar, which also has the highest proportion of civil cases (Figure 6). In Panels (D)–(F) of Figure 6 we explore the overlaps of these three categories. Specifically, we note that the likelihood of seeing caste-neutral petitioners, advocates and judges, in any combination, matched on a single case is the highest in the southern and relatively urbanized districts around Patna. These include Nalanda, Gaya, Sheikhpura, Bhojpur, Newada, Aurangabad and Bhojpur.

What factors contribute to the inclination of judges and petitioners in urbanized districts to adopt names that are devoid of caste connotations? In simpler terms, why are individuals in these more urban areas more prone to concealing their caste identity when engaging with the state? At first glance, this observation may align with the observation that rural regions in India exhibit a traditional caste-centric social stratification, whereas urban areas tend to manifest a stratification based on class and citizens in these areas may wish to conceal their caste as they engage with modern labor markets and state institutions (Singh 2018). Caste however, shapes urban life in a variety of ways, affecting housing, transportation and labor market opportunities (Srinivas 2000; Munshi 2019). It is thus entirely plausible however, that the adoption of caste-neutral names in urban areas is correlated with other aspects of citizens'

socio-economic background or else their experience of urban life, and these confounding factors drive a greater demand for caste-anonymity. Distinguishing between these two is beyond the scope of this paper, but an interesting avenue for future work. We emphasize here only that the practice of using caste-neutral names in engagement with the judiciary has become popular throughout Bihar, but particularly in urban and wealthier areas of the state.

A related question that emerges here is whether the rise of caste-neutral names in Bihar is a recent phenomenon, as was argued by some respondents in our qualitative survey. To explore this, we examine the incidence of caste-neutral names in the records of the British Indian army in the year 1912 (as seen on the website Ancestry.com). Of the more than 100,000 names, we find almost no records of any individuals with caste-neutral names such as Sinha and Kumar (N=65). We do, however, find many instances of the name Singh (N=8,562) – but the percentage of the population with this name is still lower than what we observe in any of the databases considered here today. The absence of caste-neutral names in these records is consistent with perspectives from historical sociology that emphasize the role of the British in reinforcing caste-based identities through administrative policies that favored the achievement of communal balance in the colonial administration through caste- and religion-based recruitment (Srinivas 1957; Gould 2007). Certain names are thus undoubtedly over-represented in colonial administrations due to policies that recruited specific castes for military positions. The near-complete absence of many names that are among the most common ones today is, however, still noteworthy. Urbanization, migration and differential fertility in the era after independence do not easily account for these trends since all these behaviors are affect the "grammar" of caste in contemporary India (Deshpande 2011).

5.4 Networks of Names

Networks between specific advocates and judges can emerge in groups where actors interact repeatedly (Jackson 2014). Previous research on Indian courts has documented that the most influential lawyers and judges continue to come from privileged groups, leading to a significant social distance between them and the populations they serve (Gadbois 2011; Galanter and Robinson 2017). In India's larger cities – "grand advocates" appear to offer clients opportunities to sidestep legal gridlock and expedite important cases (Galanter and Robinson 2017). And even though appointments to the Supreme Court have increasingly attempted to balance geographic representation, India's collegium based system of judicial

appointment suggests that informal networks among judges can be a powerful determinant of advancement within the ranks of the judiciary (Chandrachud 2020).

To explore this, we examine whether petitioners, judges and lawyers form networks within the judiciary, and whether these networks seem to be more likely to feature caste-neutral names. We examine the presence of networks by analyzing the incidence of co-occurrences of specific judges and specific lawyers, in their roles as representatives of petitioners and respondents in the court records in our sample. Figure 7 illustrates the network within the court. The edges depict pairs of last names of judges, litigants or advocates who appear at least 1000 times in the case data. The thickness of an edge is proportional to the number of times that pair of last names appears.[27] If last names did not matter at all, and judges and advocates were all similar in terms of case-loads and relationships, we would expect very little structure to emerge here. We see considerable structure – there are clearly some judges and some lawyers who appear together quite frequently.

Figure 8 presents a similar analysis for the set of appeals cases. As above, the weighted network is a visualization of the most frequent pairings of the last names along with the role. The thickness represents the number of cases where such a pairing has occurred. Recall that 34.42% of all petitioners and 46.6% of petitioners' advocates who are contesting appeals in our sample are relying on caste-neutral names. A look at the network structure suggests that the strongest connections are between petitioners and their advocates, as well as between advocates and judges. Since the majority of respondents are the state, the use of caste-neutral names by respondent's lawyers is negligible (just 6%). We note that the thickest edge is (P: Singh, A: Singh) where persons with the last name Singh as petitioner have worked with an advocate with last name also Singh. This occurs in 486 cases (2% of the sample).

In summary, the networks of names in our cases suggest that caste-neutral names are not only highly popular in the courts, but co-occur together in the complex network structure. Judges, petitioners, advocates and lawyers who use caste-neutral names are more likely to be represented in cases alongside each other. Next, we apply algorithms to decode last names by caste and religion to explore these issues further in our data.

---

[27]Here, the thinnest line indicates a pair appearing about a thousand times and the thickest indicates a pair (between advocates named Singh and litigants named Singh) appearing about 21 thousand times. The size of the node, representing a last name by category (judges, litigants or advocates), is proportional to the weighted degree of the node, i.e., the total number of pairs of last names by category containing the given last name and category.

The co-occurrence of judges and lawyers may be driven by a variety of factors. These include factors such as seasonality, the composition of judges at the court at a particular time, the preferences of a specific chief justice, and the specifics of the "not before me" lists that exclude certain judges from working with specific lawyers who they may be related to or have conflicts of interest with. Social identity may overlap with these. To better understand the robustness of the relationships described by Figure 7, we use a statistical framework to test some hypotheses about matching between (a) petitioners and judges; (b) advocates and judges; and (c) petitioners and advocates.

6   Matching on the Basis of Social Identity

To understand where petitioners, respondents, their advocates and judges can match with each other, or choose to work with each other on the basis of social identity, we must examine the processes of justice at the courts. At an abstract level, the pipeline of justice involves some distinct steps. A petitioner files a case against a respondent. Both have legal representation through advocates. The advocate who files the case is widely known as the filing lawyer. The case is assigned to a judge by the Chief Justice. After that point, any additional lawyers may provide services to argue the case in court. Arguing lawyers are often selected for their courtroom-specific skills and their performance in front of judges (Galanter and Robinson 2017).

This system gives petitioners and respondents the freedom to select their own lawyers. Judges are assigned to cases through the "roster system" – the system of case assignment to judges by the Chief Justice of a court – which is designed to ensure that case-assignment is as objective as possible. Unless a case is already at the final argument stage (after completion of evidence, etc.), a change in the roster results in a change in the judge hearing the case. The court strives to ensure that judges do not work with parties with whom they have had any familial or social connection. The lists of relatives and close contacts of the judge with whom the judge may have clear conflicts of interest are updated annually or on an as-needed basis.

With this background, we first examine matching between judges and petitioners. Given the stringent rules of the roster system, our first hypothesis, which we will call Hypothesis (A) is that the identity of the petitioner should not be associated with the identities of the judges

assigned to a case. Specifically, we consider the following model:

$$Litigant\ Identity_{cymdt} = \beta_0 + \beta_1 Matched\ Judge_{cymdt}$$
$$+ \Theta X_c + \alpha_y + \gamma_m + \nu_d + \phi_c + \epsilon_{cymdt} \tag{1}$$

Here $Litigant\ Identity_{cymdt}$ denotes the social status of either petitioners or respondents of case c of type t in year y month m and district d and $Matched\ Judge$ denotes the identity of the judge selected by the litigant. ΘX_c, α_y, γ_m and ν_d correspond to case-type, year, month and district fixed-effects respectively.

Next, we turn to the case of matching between advocates and judges. The rules allow for a petitioner to switch advocates during a case: one advocate can file the case and another advocate can argue a case. If a judge is known to have a strong working relationship with a specific lawyer, and the lawyer has been known to argue well in front of a specific judge, a petitioner is allowed to recruit that lawyer to represent them. If we make the further assumption that people with a similar group identity may have a greater ease of communicating in the courtroom, it is plausible that we may see some matching between judges and lawyers (with the caveat that these lawyers are not on the judge's official list of excluded people which contains the judge's familial and social network). However, in our data, for now, we only observe the filing advocates. As these advocates are chosen before the assignment of judges to a case, random assignment leads us to Hypothesis (B): Identity of petitioner advocates filing the case in the high court should not be associated with the identities of the judges assigned to the case. Specifically, we use the following model:

$$Advocate\ Identity_{cymdt} = \beta_0 + \beta_1 Matched\ Judge_{cymdt}$$
$$+ \Theta X_c + \alpha_y + \gamma_m + \nu_d + \phi_c + \epsilon_{cymdt} \tag{2}$$

Here $Advocate\ Identity_{cymdt}$ denotes the social status of either petitioner's or respondent's advocates of case c of type t in year y month m and district d and $Matched\ Judge$ denotes the identity of the judge selected by the litigant. ΘX_c, α_y, γ_m and ν_d are as in Equation 1.

Finally, we examine the matching between litigants and the lawyers who represent them. Here the official rules provide considerable freedom of choice. In some cases, such as bail applications (which are an increasing fraction of cases at the Patna High Court), a petitioner

seeking justice in the courts of India is permitted to file a case in the lower court as well as the high court, and to transfer a dismissed case between the two courts. In the event that the case needs to move between these institutions, a file must be transferred over. Given the enormous complexity of the courts, the backlog of cases, and the hierarchical structure of the courts in India, reliance on an advocate from a single community confers considerable advantages. Lawyers in close touch with their client can ensure that the file is transferred to the right person at the right time. This leads us to Hypothesis (C): Identity of the advocates representing petitioners should show strong association with the identities of the petitioners. To test this, we use the following model:

$$Litigant\ Identity_{cymdt} = \beta_0 + \beta_1 Advocate\ Identity_{cymdt}$$
$$+ \Theta X_c + \alpha_y + \gamma_m + \nu_d + \phi_c + \epsilon_{cymdt} \tag{3}$$

Here $Litigant\ Identity_{cymdt}$ denotes the social status of either petitioners or respondents of case c of type t in year y month m and district d and $Advocate\ Identity_{cymdt}$ denotes the social status of their advocates respectively. ΘX_c, α_y, γ_m and ν_d are as in Equation 1.

We perform the analysis in two steps. First we group together all low status names; this includes SC as well as caste-neutral names in a single category that we call "Low Status". We perform this analysis on the full sample of all Hindu litigants. Next we examine matching on the basis of caste neutral names relative to SC names in a restricted sample that includes just petitioners (or respondents) who use caste-neutral or SC names. Throughout, we restrict our sample only to first orders of any case and exclude any case where the government is either a petitioner or a respondent. When we examine matching between litigants and judges, we estimate the equation with a simple OLS regression that includes control variables some judge characteristics as controls, such as judge age, whether a judge was Chief Justice of any high court, the number of years a judge has been a permanent judge at the Patna High Court and whether a judge has been transferred to or from another court. Standard errors are clustered at a district-year level. We include year, month, district and case type fixed-effects.

Since we consider two main forms of social identity among Hindus, we drop Muslims from the sample of litigants in the specific regressions and focus our analysis on Hindus only. In future work, we plan to examine matching along the lines of religion much more closely.

Results for hypotheses (A) and (B) are presented in visual form in Figures 9 and 10. These visuals present just the relevant coefficients from the regressions. Note in the top panel of Figure 9 that for matching between caste-neutral and SC litigants, no coefficient here is statistically significant and the confidence intervals of all reported coefficients include 0. In other words, we find no clear matching patterns between individuals from these two groups and judges from their own groups. In the lower panel of this figure, we see that grouping together these two groups into the single "low-caste" (i.e. caste-neutral or Scheduled Caste) category does not change this result. Petitioners and their judge counterparts do not appear to match on the basis of this broader form of identity.

In Figure 10, we repeat this analysis for advocates and judges. Here we do find that caste-neutral advocates who represent petitioners are about 2% less likely to be matched with a Muslim judge relative to the omitted group of higher caste Hindus, but not more likely to match with a judge from their own group. In the lower panel, we repeat the regression once we combine the two groups and see that low-status advocates are about 1% less likely to match with a low status judge, but this result is barely significant at the 10% level. The key finding, here is that caste-neutral advocates are not more likely to match with judges from their own social group in our simple specification.

This result is broadly consistent with recent literature that has argued that judge assignment at the Indian courts appears to be as-good-as-random. Chandra, Kalantry, and Hubbard (2023) use more than a decade of data on cases at the Supreme Court to demonstrate that the Supreme Court randomly assigns cases to small benches. While these authors did not study the High Courts, the unified structure of the Indian justice system requires the protocols that are followed at High Courts to be aligned with the apex court. Ash et al. (2022) use a database of 5.5 million criminal cases in the entire Indian justice system to test for religious and gender bias in case assignment as well as case outcomes and report "tight zero effects of in-group bias" (Ash et al. 2022).[28]

Results to test hypothesis (C) are in Figure 11. In the upper panel, which presents estimates of matches between litigants and advocates from specific groups, we do find evidence of match-

---

[28]These authors also demonstrate that "the upper end of their 95% confidence interval rejects effect sizes that are one-fifth of those in most of the prior literature". This paper does not, however, examine caste bias in the Indian judiciary, largely because of the complexity of identifying caste in a pan-Indian dataset. The authors do use common last names to test for positive in-group bias and find that defendants assigned to judges with their same last name are 2 percentage points more likely to be acquitted, but that this effect is small in comparison with bias studies in other contexts.

ing for both petitioners and respondents who use caste-neutral last names. Petitioners and respondents are 3.7% and 6% more likely to match with advocates who share a caste-neutral last name, and these results are statistically significant at the 1% level. Interestingly, caste-neutral petitioners are about 0.7% *less* likely to match with an advocate with a distinctively SC name, and this too is significant at the 1% level. For respondents, the effect is stronger (coefficient -0.042) but the result is not statistically significant (standard error 0.026).

We see very different patterns for petitioners and respondents with SC sounding names. Petitioners with SC-names are about 0.3% less likely to match with an advocate with a caste-neutral name, about 0.7% more likely to match with an advocate with an SC name and about 1.7% more likely to match with an advocate with a Muslim name. These effects are significantly smaller than what was seen for the case of petitioners with caste-neutral names. For respondents, these coefficients are close to zero and not statistically significant.[29]

In the lower panel of Figure 11, we see that these results persist when we aggregate caste-neutral and SC names into a single category: litigants who are from low-status groups appear to match with advocates from these groups.

In summary, we see that both petitioners and respondents who use caste-neutral names appear to be more likely to show in-group matching than their counterparts with SC names. We infer from this that even though both neutral names and SC names may be regarded as low-status names in Bihar, they contain different markers of social identity at the courts. Caste-neutral petitioners and respondents are the most likely to match with advocates that also have caste-neutral names.

6.1 Outcomes of Justice: Regression Analysis

Next we examine the outcomes of the justice system. Does social identity of the petitioner and their lawyer affect the likelihood of the case being disposed, withdrawn or resolved? To analyze the impact of social identity on case outcomes, we consider the following model:

---

[29]All coefficients in Figure 11 are less precisely estimated for respondents because this group is much smaller than the petitioner group. Recall from 2 that 89% of respondents are, in fact, the government. This estimate is just 2% for petitioners.

$$y_{cymdt} = \beta_0 + \beta_1[PetitionerIdentity]_i + \beta_2[AdvocateIdentity]_j +$$
$$\beta_3[Petitioner \times Advocate_{(i,j)}] + \delta X_c + \alpha_y + \Gamma_m + \nu_d + \phi_c + \epsilon_{cymdt} \tag{4}$$

Here y_{cymdt} denotes the outcome of case c of type t in year y month m and district d. We consider several outcome measures: case status (Whether the case has been decided and the length of time taken to a decision) and the outcome of the case (rejected, dismissed, withdrawn and resolved). The sub-scripts i and j denote the types of social identity on the basis of names. We consider three types of groups: All Low-status (which includes both caste-neutral and SC names) and caste-neutral and SC names separately. Once again, we include year, month, district and case type fixed effects. Standard errors are clustered at a district-year level. We restrict our sample only to first orders of any case and exclude any case where the government is either a petitioner or a respondent.

In line with our approach in the prior findings, we exclude Muslims from the litigant sample, focusing exclusively on Hindus. Our analysis begins by exploring in-group matching effects among petitioners and respondents to their respective advocates. Subsequently, we narrow our focus to exclusively include low-status individuals, allowing us to gauge the influence of utilizing a caste-neutral name compared to one with a Scheduled Caste (SC) connotation.

We begin by examining the impact of low-status petitioner-advocate matches on case outcomes among Hindu litigants, with our findings presented in Figures 12 and 13, and detailed results in the Appendix (Tables A2 for petitioners and A3 for respondents).

From Figure 12, we observe that the interaction terms *Petitioner Advocate Low Status ×* *Petitioner Low Status* and *Respondent Advocate Low Status × Respondent Low Status* are not statistically significant in the regressions for either case decisions (top panel) or the time to decision (bottom panel). This suggests that litigant-advocate pairs who are matched on the basis of broadly classified low status names do not appear to have expedited review of their cases relative to their higher-status counterparts. The coefficient for *Respondent Low Status* in the case duration regression however, does take on a positive coefficient that is statistically significant at the 5% level, suggesting that low status respondents overall take longer to have their cases resolved. The combined effect for this variable is however, not statistically significant (Appendix Tables A2 for petitioners and A3 for respondents).

Figure 13 presents the case outcomes for these groups. We find here a marginal disadvantage for petitioners from these groups, evidenced by a 0.7 percentage points lower likelihood of a case being allowed in the court. This coefficient is significant at the 5% level. This however, is a small effect and we note that no other outcome considered here shows a significant effect. Even the joint tests of significance of the low status identities are not significant at even the 10% level (Table A2) in four out of the six regressions considered here.

Next, we examine the impact of caste-neutral names relative to SC names. We restrict the sample to the low status group alone, excluding not only Muslims but all other higher status Hindus as well. Results are shown in Figure 14. Complete results are shown in Appendix Table A4 (Petitioners) and A5 (Respondents). Here we note that the use of caste-neutral names does yield some impacts relative to petitioners with SC names. Petitioners with caste-neutral names are 1.5 percentage points more likely to have their cases dismissed relative to their counterparts with SC names and this result is significant at the 10% level. Respondents however, are 3.9 percentage points more likely to have their case decided and this result significant at the 1% level. Case decisions take less time, i.e. they are decided in 5.1 fewer months and this result is also significant at the 1% level. Cases are 0.4 percentage points less likely to be rejected (significant at the 1% level) and 11 percentage points less likely to be dismissed (significant at the 10% level). The joint test for caste-neutral advocate names *and* respondent names is significant in four of the six regressions. Given that dismissals are not the desired outcomes for litigants at the courts, it is noteworthy that when compared to SC names, caste-neutral names do not appear to be associated with better outcomes for both petitioners and respondents.

Our analysis of appeal cases further illustrates these patterns. As detailed in Figures 16 and 17 (full results in Appendix Tables A6 and A7), we find that low-status matches do not significantly affect the likelihood of the decision or the case duration. In Figure 17 (top panel) we see that for petitioners, matching on the basis of identity does not have any statistically significant impact on any of the case outcomes. Respondents, however, show a different pattern. Respondents who are matched to low-status advocates have a 23.7 percentage points higher likelihood of cases being allowed, a 19.2 percentage points lower likelihood of dismissal. These estimates are significant at the 1% level (Figure 17, Table A7). Given that respondents have an interest in having cases dismissed from the courts, we note that matching is disadvantageous for this group in appeal cases.

Figures 16 and 17 (full results in Appendix Tables A6 and A7) also present the estimates

from regressions with just the low-status group. This, as before, enables us to compare and contrast caste-neutral litigants with their counterparts who are SC. We note that petitioners who match with advocates on the basis of caste-neutral names are 19.8 percentage points less likely to be dismissed. This estimate is significant at the 1% level. Respondents are 32.9 percentage points less likely to have their cases pending. They are also 23.9 percentage points less likely to have their cases allowed, 22.0 percentage points less likely to have their cases dismissed and 20.5 percentage points more likely to have their cases withdrawn. These results are statistically significant at the 1%, 5% and 10% levels. Here, as before, we note that matching on the basis of caste-neutral names confers benefits on petitioners but disadvantages for respondents. This asymmetry is a key result of this paper.

Overall, our study reveals that litigants in the courts of Bihar tend to match with lawyers who share their own identity, and that this matching does seem to affect the outcomes of justice. When we group caste-neutral and SC names together, we see that they do not experience any significant advantages (or even significant disadvantages) in the judicial system. Within the group of low-status litigants, we also see that the adoption of caste-neutral names appears to affect petitioners and respondents differently, but that both types of names are associated with disadvantages in the processes and outcomes of justice.

7   Implications and Contributions to the Literature

An obvious question that emerges from these results on the adverse case outcomes when litigants match with lawyers is why they choose to do so, what are the perceived advantages in Bihar's courts to match with lawyers from their own community? We believe several mechanisms could be at play. First, lawyers from the same community may have a better understanding of the cultural nuances and social context relevant to the case. This can enhance communication between the respondent and the lawyer, ensuring that critical details are effectively conveyed and understood. A lawyer who has a strong bond with their client can potentially represent their client better in the arena of the court. This is consistent with ethnographic evidence that documents the many barriers faced by marginalized people in accessing the justice system (Krishnan et al. 2014).

Second, individuals from the same community may be easier to establish trust and rapport. Shared community ties often foster trust and rapport between respondents and lawyers. Evidence from the United States suggests that trust is crucial in legal representation, and

a strong attorney-client relationship can positively influence the client's confidence in the legal process. First, lawyers from the same community may have a better understanding of the cultural nuances and social context relevant to the case, enhancing communication and representation (Ryo 2018). Second, shared community ties can foster trust and rapport, which are crucial in legal representation (Young and Hassan 2020). However, this trust may be influenced by the perceived competence of the decision maker (Young and Hassan 2020). It's important to note that disadvantaged defendants may actively resist their court-appointed lawyers due to mistrust (Clair 2021). Lastly, people's evaluations of the justice of their experiences can be influenced by their motives for maintaining high status within their group (Tyler 1994).

The next question that emerges from our results is why matching on the basis of identity does not bring greater advantages to either petitioners or respondents from the low-status groups. It could be that litigants discriminate against higher-quality advocates in favor of advocates of their own group identity, resulting in worse outcomes for litigants in the court.

While these factors can influence case outcomes, it's essential to note that the effectiveness of such strategies may vary depending on the specifics of the case, and the individuals involved. Our analysis does not control for a lot of case-attributes or even the socio-economic status of the contesting parties, all of which can be powerful confounding variables. Nevertheless, we believe the results are important and interesting and open up many new questions about representation and justice in the courts of Bihar and India more generally.

Our research contributes to the expanding body of literature on India's judicial system, challenging the prevailing perception of its courts as isolated entities detached from societal dynamics (Sen 2017; Rudolph and Rudolph 2001). While it is widely acknowledged that judges and advocates often hail from privileged segments of society (Gadbois 2011), recent initiatives, such as those aimed at ensuring geographical representation and increasing inclusion of judges from marginalized regions (Chandrachud 2020), coexist with a growing presence of elite lawyers in major Indian cities (Galanter and Robinson 2017). Our study reveals an important nuance in the dynamics between litigants and legal representatives. While there may not be substantial alignment between litigants and judges, a discernible connection exists between litigants and lawyers, influencing case outcomes. Building on earlier observations, ethnographic evidence underscores the myriad challenges faced by Indian citizens in navigating bureaucratic complexities, limited infrastructure, inadequate legal information, pervasive corruption, and administrative indifference in their pursuit of justice (Krishnan

et al. 2014). Particularly for vulnerable individuals, having representation by a caste-neutral lawyer emerges as a pivotal factor shaping their journey toward justice.

Our research also comes as a reminder that courts are social spaces that are deeply embedded in society. The system ought to be looked at as a nonlinear superposition of a multitude of complex social networks, where nodes represent individuals and links capture a variety of different social relations between petitioners, respondents, advocates and judges. A deeper study of the complexity of such relationships, and their potential to shape outcomes of the system, is a promising area of future research (see, for example, Szell, Lambiotte, and Thurner 2010).

8 Conclusion

We study the names of petitioners, respondents, lawyers and judges in more than one million cases heard at the Patna High Court between 2009 and 2019. We rely on matching-learning algorithms to decode these names for markers of personal and social identity. We find that a small list of names (Singh, Kumar, Devi, Yadav, Kumari, Prasad, Jha, Rai, Sharma, Sah) account for more than half the cases filed at the Patna High court and most of these are caste-neutral. The prevalence of caste-neutral names however, varies considerably across the state with the highest prevalence in the urbanized district of Patna.

Next we test for three hypotheses for matching: (a) Between petitioners and judges; (b) Between advocates and judges; and (c) Between petitioners and their advocates. We find almost no evidence for any matching between petitioners and judges or between (filing) advocates and judges. Instead, we find strong evidence of matching between petitioners and their advocates. When petitioners and advocates match or when respondents and advocates match, the results in the court are on average worse for the litigants. These results suggest that the social movements that disrupted existing social structures in the past may have inadvertently created new social categories that reinforced networks and inequalities in the formal justice system.

We find almost no matching between judges and litigants, but considerable matching between litigants and advocates. This has consequences. We find that matching on the basis of broad measures of identity, such as belonging to "Scheduled Castes" or the use of caste-neutral names, can have modest but yet discernible impacts on both the processes of justice as well as outcomes. Low-status respondents who match with advocates from their own group appear

33

to have shorter case-processing times. Matching can also affect the outcomes of appeal cases.

This research serves as a poignant reminder that courts function as integral components of the broader societal fabric. Rather than viewing the legal system in isolation, it is more aptly perceived as a dynamic, nonlinear superposition of intricate social networks consisting of people with complex identities. Delving deeper into the complexity of these relationships and understanding their potential influence on the outcomes of the legal system is a promising avenue for future research.

References

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson (2020). "Do immigrants assimilate more slowly today than in the past?" In: *American Economic Review: Insights* 2.1, pp. 125–141.

Akerlof, George A and Rachel E Kranton (2000). "Economics and identity". In: *The quarterly journal of economics* 115.3, pp. 715–753.

— (2002). "Identity and schooling: Some lessons for the economics of education". In: *Journal of economic literature* 40.4, pp. 1167–1201.

Arrow, Kenneth J (1998). "What has economics to say about racial discrimination?" In: *Journal of economic perspectives* 12.2, pp. 91–100.

Ash, Elliott et al. (2022). "Measuring gender and religious bias in the indian judiciary". In.

Banerjee, Abhijit, Marianne Bertrand, et al. (2009). "Labor market discrimination in Delhi: Evidence from a field experiment". In: *Journal of comparative Economics* 37.1, pp. 14–27.

Banerjee, Abhijit and Lakshmi Iyer (2005). "History, institutions, and economic performance: The legacy of colonial land tenure systems in India". In: *American economic review* 95.4, pp. 1190–1213.

Bayly, Susan (2001). *Caste, society and politics in India from the eighteenth century to the modern age*. Vol. 3. Cambridge University Press.

Becker, Gary S (2010). *The economics of discrimination*. University of Chicago press.

Behaghel, Luc, Bruno Crépon, and Thomas Le Barbanchon (2015). "Unintended effects of anonymous resumes". In: *American Economic Journal: Applied Economics* 7.3, pp. 1–27.

Berdejo, Carlos and Daniel L Chen (2017). "Electoral cycles among us courts of appeals judges". In: *The Journal of Law and Economics* 60.3, pp. 479–496.

Bertrand, Marianne and Sendhil Mullainathan (2004). "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination". In: *American economic review* 94.4, pp. 991–1013.

Blair, Harry W (1980). "Rising kulaks and backward classes in Bihar: social change in the late 1970s". In: *Economic and Political Weekly*, pp. 64–74.

Chakrabarti, Rajesh (2013). *Bihar breakthrough: the turnaround of a beleaguered state*. Rupa Publications.

Chandra, Aparna, Sital Kalantry, and William HJ Hubbard (2023). *Court on Trial: A Data-Driven Account of the Supreme Court of India*. Penguin Random House India. ISBN: 9780670091584.

Chandrachud, Abhinav (2020). *The informal constitution: Unwritten criteria in selecting judges for the Supreme Court of India*. Oxford University Press.

Chen, Daniel L, Tobias J Moskowitz, and Kelly Shue (2016). "Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires". In: *The Quarterly Journal of Economics* 131.3, pp. 1181–1242.

Clair, Matthew (2021). "Being a disadvantaged criminal defendant: Mistrust and resistance in attorney-client Interactions". In: *Social Forces* 100.1, pp. 194–217.

Clark, Gregory and Neil Cummins (2015). "Intergenerational wealth mobility in England, 1858–2012: surnames and social mobility". In: *The Economic Journal* 125.582, pp. 61–85.

Collier, Paul (2016). "The cultural foundations of economic failure: A conceptual toolkit". In: *Journal of Economic Behavior & Organization* 126, pp. 5–24.

Cook, Lisa, Trevon Logan, and John Parman (2016). "The mortality consequences of distinctively black names". In: *Explorations in Economic History* 59, pp. 114–125.

Dar, Aaditya and Abhilasha Sahay (2018). "Designing policy in weak states: Unintended consequences of alcohol prohibition in Bihar". In: *Available at SSRN 3165159*.

Deshpande, Ashwini (2011). *The Grammar of Caste: Economic Discrimination in Contemporary India*. Oxford University Press.

Deshpande, Ashwini and Katherine Newman (2007). "Where the path leads: The role of caste in post-university employment expectations". In: *Economic and Political Weekly*, pp. 4133–4140.

Dirks, Nicholas B (1989). "The invention of caste: Civil society in colonial India". In: *Social Analysis: The International Journal of Social and Cultural Practice* 25, pp. 42–52.

Diwakar, Ranganath Ramachandra (1959). *Bihar through the Ages*. 44. Orient Longmans.

Djankov, Simeon et al. (2003). "Courts". In: *The Quarterly Journal of Economics* 118.2, pp. 453–517.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig (2017). "Cultural proximity and loan outcomes". In: *American Economic Review* 107.2, pp. 457–92.

Gadbois, George H (2011). *Judges of the Supreme Court of India: 1950–1989*. Oxford University Press.

Galanter, Marc and Nick Robinson (2017). "Grand advocates: The traditional elite lawyers". In: *The Indian legal profession in the age of globalization*, p. 455.

Gebauer, Jochen E, Mark R Leary, and Wiebke Neberich (2012). "Unfortunate first names: Effects of name-based relational devaluation and interpersonal neglect". In: *Social Psychological and Personality Science* 3.5, pp. 590–596.

Ghosh, Yashomati (2018). "Indian judiciary: an analysis of the cyclic syndrome of delay, arrears and pendency". In: *Asian Journal of Legal Education* 5.1, pp. 21–39.

Gidla, Sujatha (2017). *Ants among elephants: An untouchable family and the making of modern India*. Farrar, Straus and Giroux.

Gould, William (2007). "The dual state: The unruly "subordinate", caste, community and civil service recruitment in North India, 1930–1955". In: *Journal of Historical Sociology* 20.1-2, pp. 13–43.

Greif, Avner and Guido Tabellini (2017). "The clan and the corporation: Sustaining cooperation in China and Europe". In: *Journal of Comparative Economics* 45.1, pp. 1–35.

Hoff, Karla, Mayuresh Kshetramade, and Ernst Fehr (2011). "Caste and punishment: The legacy of caste culture in norm enforcement". In: *The economic journal* 121.556, F449–F475.

Hoff, Karla and Priyanka Pandey (2006). "Discrimination, social identity, and durable inequalities". In: *American economic review* 96.2, pp. 206–211.

Jackson, Matthew O (2014). "Networks in the understanding of economic behaviors". In: *Journal of Economic Perspectives* 28.4, pp. 3–22.

Jaffrelot, Christophe (2003). *India's Silent Revolution: The Rise of the Lower Castes in North India*. Columbia University Press.

— (2010). *Religion, caste, and politics in India*. Primus Books.

Jaffrelot, Christophe and Sanjay Kumar (2012). *Rise of the plebeians?: The changing face of the Indian legislative assemblies*. Routledge.

Jayaraman, Raja (2005). "Personal identity in a globalized world: Cultural roots of Hindu personal names and surnames". In: *The Journal of Popular Culture* 38.3, pp. 476–490.

Jodhka, Surinder S (2017). *Caste in contemporary India*. Routledge India.

Joshi, Shareen, Nishtha Kochhar, and Vijayendra Rao (2018a). *Are caste categories misleading? The relationship between gender and jati in three Indian states*. Oxford University Press.

— (2018b). "Jati inequality in rural Bihar". In: *World Bank Policy Research Working Paper* 8512.

— (2022). "Fractal inequality in rural India: class, caste and jati in Bihar". In: *Oxford Open Economics* 1.

Krishnan, Jayanth K et al. (2014). "Grappling at the Grassroots: Access to justice in India's Lower Tier". In: *Harv. Hum. Rts. J.* 27, p. 151.

Kumar, Avinash (2012). "Fallacy of the State in Bihar". In: *Economic and Political Weekly*, pp. 23–25.

Kumar, Sanjay (2018). *Post-Mandal Politics in Bihar: Changing Electoral Patterns*. Vol. 1. SAGE Publishing India.

Manjhi, Jitan Ram (Oct. 2014). *The Bihar Story: Resurrection of the State*. Blog. DOI: http://eprints.lse.ac.uk/81984/1/The\%20Bihar\%20Story_\%20Resurrection\%20of\%20the\%20State\%20-\%20IGC.pdf.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook (2001). "Birds of a feather: Homophily in social networks". In: *Annual review of sociology* 27.1, pp. 415–444.

Menon, Kalyani Devaki (2022). *Making Place for Muslims in Contemporary India*. Cornell University Press.

Menski, Werner F (2006). *Comparative law in a global context: the legal systems of Asia and Africa*. Cambridge University Press.

Milkman, Katherine L, Modupe Akinola, and Dolly Chugh (2012). "Temporal distance and discrimination: An audit study in academia". In: *Psychological science* 23.7, pp. 710–717.

Mosse, David (2018). "Caste and development: Contemporary perspectives on a structure of discrimination and advantage". In: *World development* 110, pp. 422–436.

Munshi, Kaivan (2019). "Caste and the Indian economy". In: *Journal of Economic Literature* 57.4, pp. 781–834.

Parmar, Rajnikant (2020). "Transacting caste in modern times: Changing social identity through surnames in urban Gujarat". In: *Contemporary Voice of Dalit* 12.2, pp. 220–233.

Pelham, Brett W, Matthew C Mirenberg, and John T Jones (2002). "Why Susie sells seashells by the seashore: implicit egotism and major life decisions." In: *Journal of personality and social psychology* 82.4, p. 469.

Rambachan, Ashesh et al. (2020). "An economic perspective on algorithmic fairness". In: *AEA Papers and Proceedings*. Vol. 110, pp. 91–95.

Registrar General & Census Commissioner of India, Office of the (2011). *Table A-11 Appendix: District wise scheduled tribe population (Appendix), Bihar - 2011*. Census of India.

Roy, AK (1997). "Can courts run the country?" In: *Economic and Political Weekly*, pp. 2579–2581.

Rudolph, Susanne and Lloyd Rudolph (2001). "Redoing the constitutional design: from an interventionist to a regulatory state". In: *The success of India's democracy*. Ed. by Atul Kohli. Vol. 6. Cambridge University Press. Chap. 6, pp. 127–163.

Ryo, Emily (May 2018). "Representing Immigrants: The Role of Lawyers in Immigration Bond Hearings". In: *Law amp; Society Review* 52.2, pp. 503–531.

Sen, Ronojoy (2017). "India's Democracy at 70: The Disputed Role of the Courts". In: *Journal of Democracy* 28.3, pp. 96–105.

Shah, Ghanshyam (2004). "Social movements in India: A review of literature". In.

Singh, Prerna (2015). *How solidarity works for welfare: Subnationalism and social development in India*. Cambridge University Press.

Singh, Yogendra (2018). "Modernization of Indian tradition". In: Rawat Publications.

Small, Mario L and Devah Pager (2020). "Sociological perspectives on racial discrimination". In: *Journal of Economic Perspectives* 34.2, pp. 49–67.

Sood, Gaurav and Suriyan Laohaprapanon (2018). *Socio-Economic and Caste Census 2011*. Version V8. DOI: 10.7910/DVN/LIIBNB. URL: https://doi.org/10.7910/DVN/LIIBNB.

Srinivas, Mysore Narasimhachar (1957). "Caste in modern India". In: *The Journal of Asian Studies* 16.4, pp. 529–548.

— (2000). *Caste: Its 20th century avatar*. Penguin UK.

Szell, Michael, Renaud Lambiotte, and Stefan Thurner (2010). "Multirelational organization of large-scale social networks in an online world". In: *Proceedings of the National Academy of Sciences* 107.31, pp. 13636–13641.

Tewary, Amarnath (2023). "Bihar caste survey, OBCs, EBCs comprise more than 63% of State's population". In: *The Hindu*.

Thorat, Sukhadeo and Paul Attewell (2007). "The legacy of social exclusion: A correspondence study of job discrimination in India". In: *Economic and Political Weekly*, pp. 4141–4145.

Tyler, Tom R (1994). "Psychological models of the justice motive: Antecedents of distributive and procedural justice." In: *Journal of personality and social psychology* 67.5, p. 850.

Verma, Ritesh (2023). "Full list of castes and their population". In: *Live Hindustan*.

Young, Kim A and Shahidul Hassan (2020). "How procedural experiences shape citizens' perceptions of and orientations toward legal institutions: Evidence from a household survey in Bangladesh". In: *International Review of Administrative Sciences* 86.2, pp. 278–294.

Tables And Figures

Table 1: Summary Statistics - Raw Data

Case Data	N	Mean	SD	Min	Max
Filing Year	1,071,068	2014.48	3.23	2009.0	2019.0
Civil Case	1,071,068	0.34	0.47	0.0	1.0
Petitioner is Government?	1,071,068	0.01	0.07	0.0	1.0
Respondent is Government?	1,071,068	0.89	0.31	0.0	1.0
Petitioner is Muslim?	1,032,838	0.12	0.32	0.0	1.0
Respondent is Muslim?	64,449	0.11	0.31	0.0	1.0
Petitioner's Advocate is Muslim?	1,068,991	0.05	0.22	0.0	1.0
Respondent's Advocate is Muslim?	959,450	0.12	0.32	0.0	1.0
Petitioner is SC?	1,001,830	0.12	0.32	0.0	1.0
Respondent is SC?	58,849	0.10	0.30	0.0	1.0
Petitioner's Advocate is SC?	1,040,380	0.06	0.24	0.0	1.0
Respondent's Advocate is SC?	841,301	0.07	0.25	0.0	1.0
Petitoner has caste-neutral name?	1,001,846	0.50	0.50	0.0	1.0
Responndent has caste-neutral name?	58,851	0.57	0.50	0.0	1.0
Petitoners's Advocate has caste-neutral name?	1,040,380	0.63	0.48	0.0	1.0
Respondent's Advocate has caste-neutral name?	841,304	0.59	0.49	0.0	1.0

Judge Data	N	Mean	SD	Min	Max
Judge has caste-neutral name?	73	0.53	0.50	0.0	1.0
Judge is SC?	52	0.02	0.14	0.0	1.0
Judge is Muslim?	83	0.06	0.24	0.0	1.0
Year of birth	79	1956.13	5.56	1947.0	1969.0
Year when became permanent	30	2009.00	7.38	1991.0	2019.0
Was Chief Justice	83	0.13	0.34	0.0	1.0
Promoted to Supreme Court?	83	0.08	0.28	0.0	1.0

Table 2: Summary Statistics - Merged Data

Case-Level	N	Mean	SD	Min	Max
Filing Year	345,239	2013.73	3.19	2009.0	2019.0
Civil Case	345,239	0.82	0.38	0.0	1.0
Petitioner is Government?	345,239	0.01	0.09	0.0	1.0
Respondent is Government?	345,239	0.79	0.41	0.0	1.0
Petitioner is Individual?	345,239	0.92	0.27	0.0	1.0
Respondent is Individual?	345,239	0.14	0.34	0.0	1.0
Petitioner is Muslim?	320,791	0.09	0.29	0.0	1.0
Respondent is Muslim?	47,731	0.10	0.30	0.0	1.0
Petitioner's Advocate is Muslim?	344,369	0.05	0.21	0.0	1.0
Respondent's Advocate is Muslim?	260,623	0.10	0.30	0.0	1.0
Petitioner is a Woman?	320,791	0.27	0.44	0.0	1.0
Respondent is a Woman?	47,731	0.64	0.48	0.0	1.0
Petitioner's Advocate is a Woman?	344,369	0.14	0.35	0.0	1.0
Respondent's Advocate is a Woman?	260,623	0.17	0.37	0.0	1.0
Petitioner is SC?	311,265	0.10	0.30	0.0	1.0
Respondent is SC?	43,129	0.10	0.30	0.0	1.0
Petitioner's Advocate is SC?	334,447	0.06	0.23	0.0	1.0
Respondent's Advocate is SC?	242,436	0.07	0.26	0.0	1.0
Petitoner has caste-neutral name?	311,277	0.56	0.50	0.0	1.0
Responndent has caste-neutral name?	43,131	0.57	0.50	0.0	1.0
Petitoners's Advocate has caste-neutral name?	334,447	0.62	0.49	0.0	1.0
Respondent's Advocate has caste-neutral name?	242,436	0.58	0.49	0.0	1.0
Total Number of Judges	345,239	2.88	3.66	1.0	152.0
Number of Distinct Judges	345,239	1.65	1.19	1.0	31.0
Fraction of Judges with Caste-Neutral Names?	342,370	0.51	0.44	0.0	1.0
Fraction of SC Judges?	303,503	0.03	0.15	0.0	1.0
Fraction of Muslim Judges?	343,063	0.05	0.19	0.0	1.0
Fraction of Female Judges?	343,063	0.05	0.19	0.0	1.0
How many Judges with Caste-Neutral Names?	345,239	0.80	0.78	0.0	5.0
How many judges female?	345,239	0.09	0.30	0.0	3.0
How many judges SC?	345,239	0.04	0.21	0.0	1.0
How many judges Muslim?	345,239	0.09	0.29	0.0	3.0
Any Judge with Caste-Neutral Name?	345,239	0.62	0.48	0.0	1.0
Any Judge is SC?	345,239	0.04	0.21	0.0	1.0
Any Judge is Muslim?	345,239	0.08	0.28	0.0	1.0
Any Judge is female?	345,239	0.09	0.29	0.0	1.0
Average Age of Judges	343,365	55.91	3.24	41.0	62.0
Average Years of being Permanent of Judges	345,239	0.38	0.89	-4.8	9.4
Bench includes Chief Justice	343,366	0.07	0.25	0.0	1.0

Table 3: Summary Statistics for Appeal Sample

Variable	N	Mean	SD	Min	Max
Pending	23,978	0.57	0.50	0.0	1.0
Months between Registration and Decision	10,401	6.36	9.91	0.0	84.0
Rejected	10,416	0.01	0.07	0.0	1.0
Dismissed	10,416	0.31	0.46	0.0	1.0
Allowed	10,416	0.41	0.49	0.0	1.0
Withdrawn	10,416	0.10	0.30	0.0	1.0
Petitioners and their Advocates					
Petitioner Low Status X Advocate Low Status	15,220	0.37	0.48	0.0	1.0
Petitioner Neutral X Advocate Neutral	15,885	0.32	0.47	0.0	1.0
Petitioner SC X Advocate SC	15,220	0.01	0.09	0.0	1.0
Petitioner Low Status	19,722	0.50	0.50	0.0	1.0
Petitioner Advocate Low Status	17,128	0.70	0.46	0.0	1.0
Petitioner Neutral	20,065	0.44	0.50	0.0	1.0
Petitioner SC	19,747	0.11	0.32	0.0	1.0
Petitioner's Advocate Neutral	17,928	0.66	0.47	0.0	1.0
Petitioner's Advocate SC	17,128	0.06	0.23	0.0	1.0
Respondents and their Advocates					
Petitioner Low Status X Advocate Low Status	15,220	0.37	0.48	0.0	1.0
Petitioner Neutral X Advocate Neutral	15,885	0.32	0.47	0.0	1.0
Petitioner SC X Advocate SC	15,220	0.01	0.09	0.0	1.0
Petitioner Low Status	19,722	0.50	0.50	0.0	1.0
Petitioner Advocate Low Status	17,128	0.70	0.46	0.0	1.0
Petitioner Neutral	20,065	0.44	0.50	0.0	1.0
Petitioner SC	19,747	0.11	0.32	0.0	1.0
Petitioner's Advocate Neutral	17,928	0.66	0.47	0.0	1.0
Petitioner's Advocate SC	17,128	0.06	0.23	0.0	1.0

Note: Summary Statistics of key dependent and independent variables for the set of appeals.
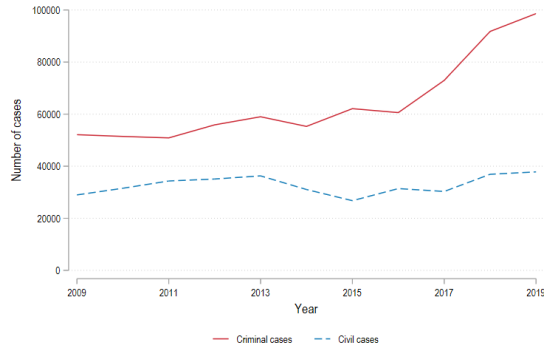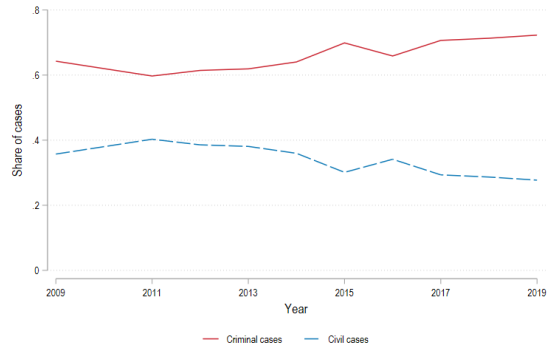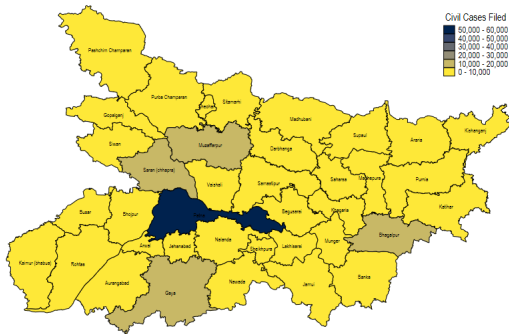
Figure 1: Time Trends of Criminal and Civil Cases Filed at the Patna High Court, 2009-2019

Note: The figure depicts time trends of the number (Panel A) and share (Panel B) of civil and criminal cases filed per year in the Patna High Court between 2009 and 2019. Calculations are based on the full sample of 1,071,068 cases filed in this time period.

(A) Number of Civil Cases Filed
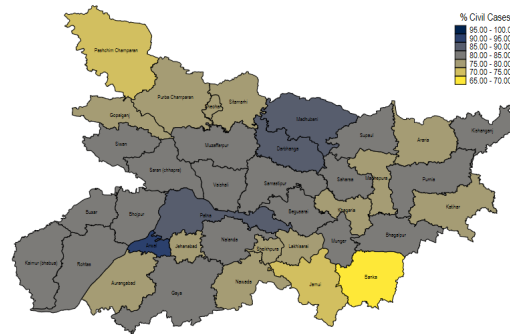
(B) Civil Share of Filed Cases

Figure 2: Spatial Distribution of Civil cases filed at the Patna High Court, 2009-2019

Note: Panel (A) displays the total number of civil cases filed per district in the Patna High Court between 2009 and 2019. Panel (B) plots the share of cases filed per district which are civil cases.
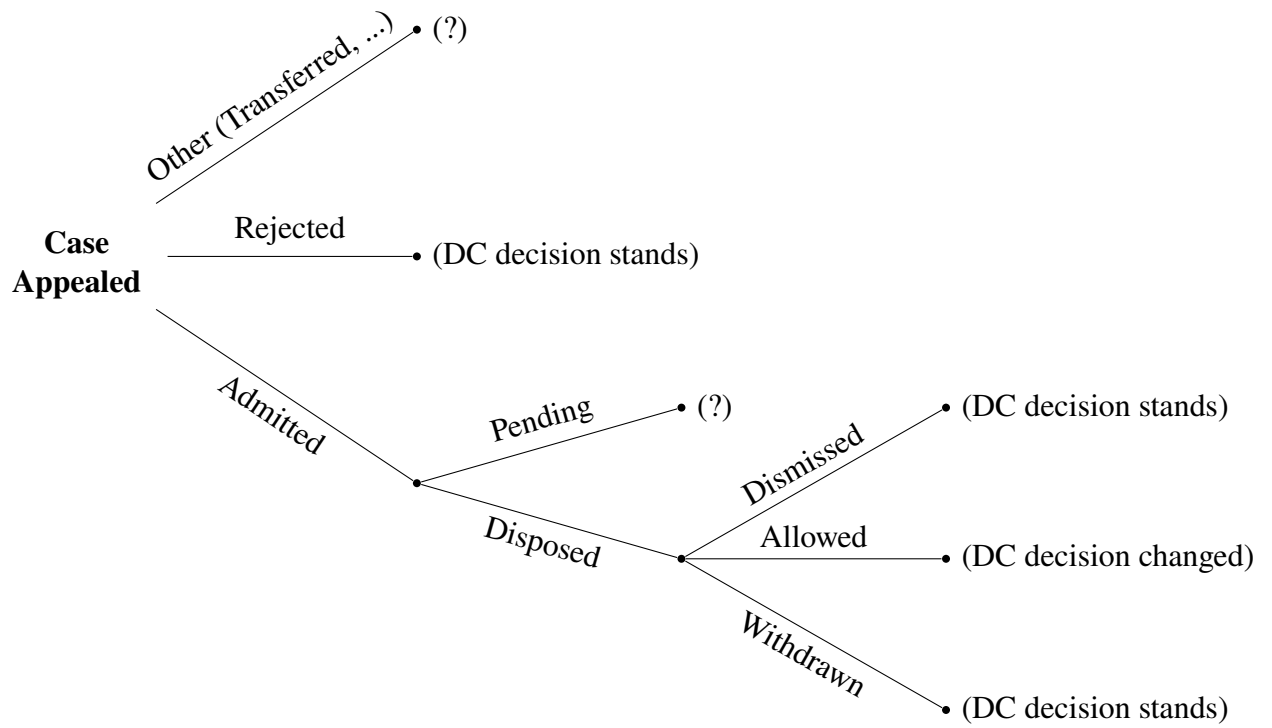
Figure 3: Decision Tree for DC cases appealed in the HC
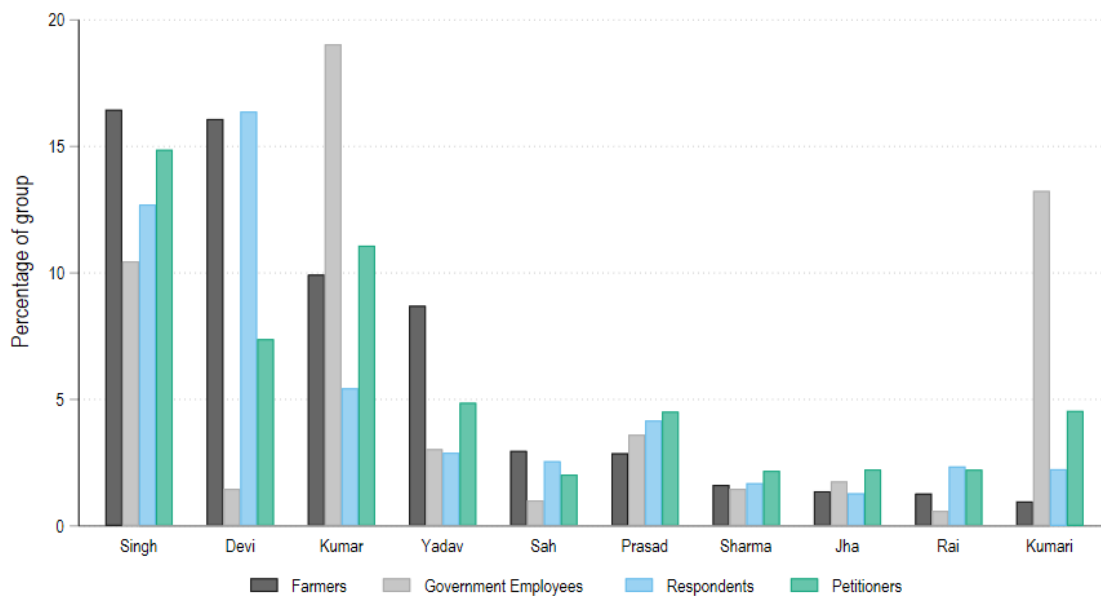
Figure 4: Proportion of sample, by top-10 last names

Note: Rewrite
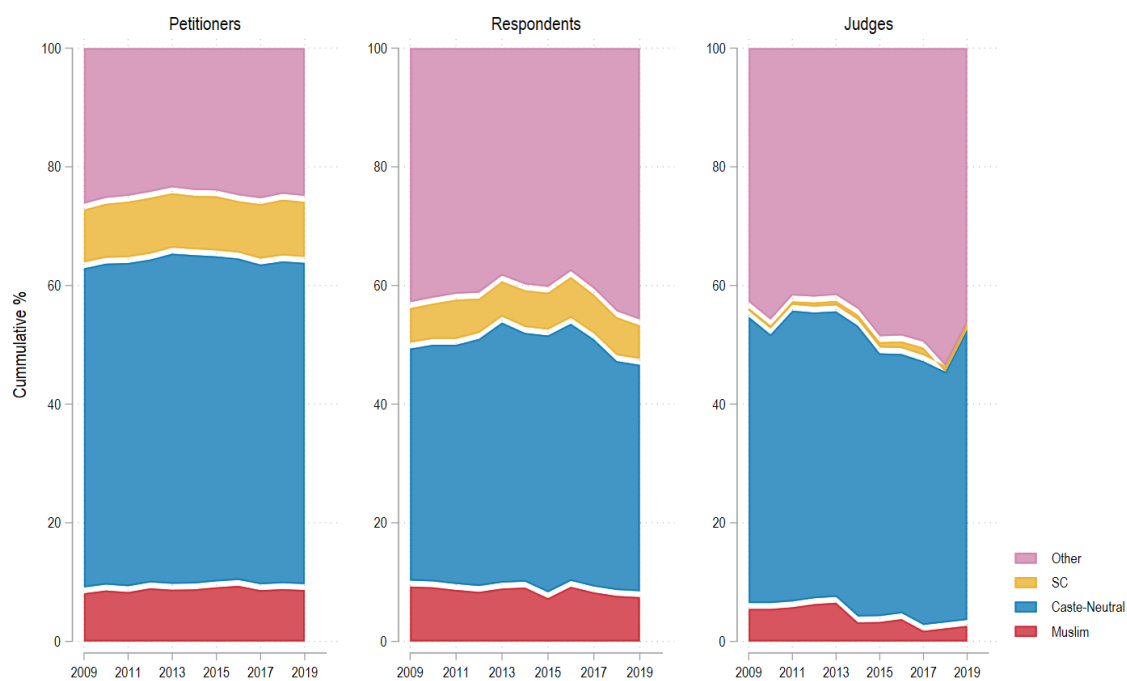
Figure 5: Trends of Petitioners, Respondents and Judges by Gender and Caste.

Note: The four panels for petitioners and respondents are at the case level. The top-left and top-center include all cases with information on. respectively, the petitioner and respondent. The bottom-left and bottom-center panel include only cases where respectively the petitioner and respondent are identified as individuals. The judge panels (top-right and bottom-right) include each judge with at least one case in the Patna HC in a given year exactly once.

Table 4: Breakdown of the Population by Gender, Religion and Caste

	N	Woman	Muslim	SC	ST	Other Hindu
Patna State Sources						
SECC	22,400,539	9.7	15.22	16.99	1.99	67.23
Farmers	1,341,181	27.62	6.17	10.28	1.05	82.74
Government Employees	210,389	29.01	9.35	13.28	1.67	76.00
Patna HC						
Judges	83	9.64	6.02	6.41	1.28	92.31
Petitioners	1,013,871	22.17	10.99	11.53	1.02	76.81
Respondents	63,374	34.89	10.53	9.93	1.20	78.80
Advocates	210,389	29.01	9.35	13.41	1.67	75.88

Notes: (i) Since the SECC was conducted by interviewing the designated head of the household, and only 9.7% of women in Bihar were coded as household heads, we do not present the estimates of gender from this survey; (ii) Estimates for Advocates, Petitioners, Respondents and Judges are calculated using our data from the Patna High Court, 2009–2019.
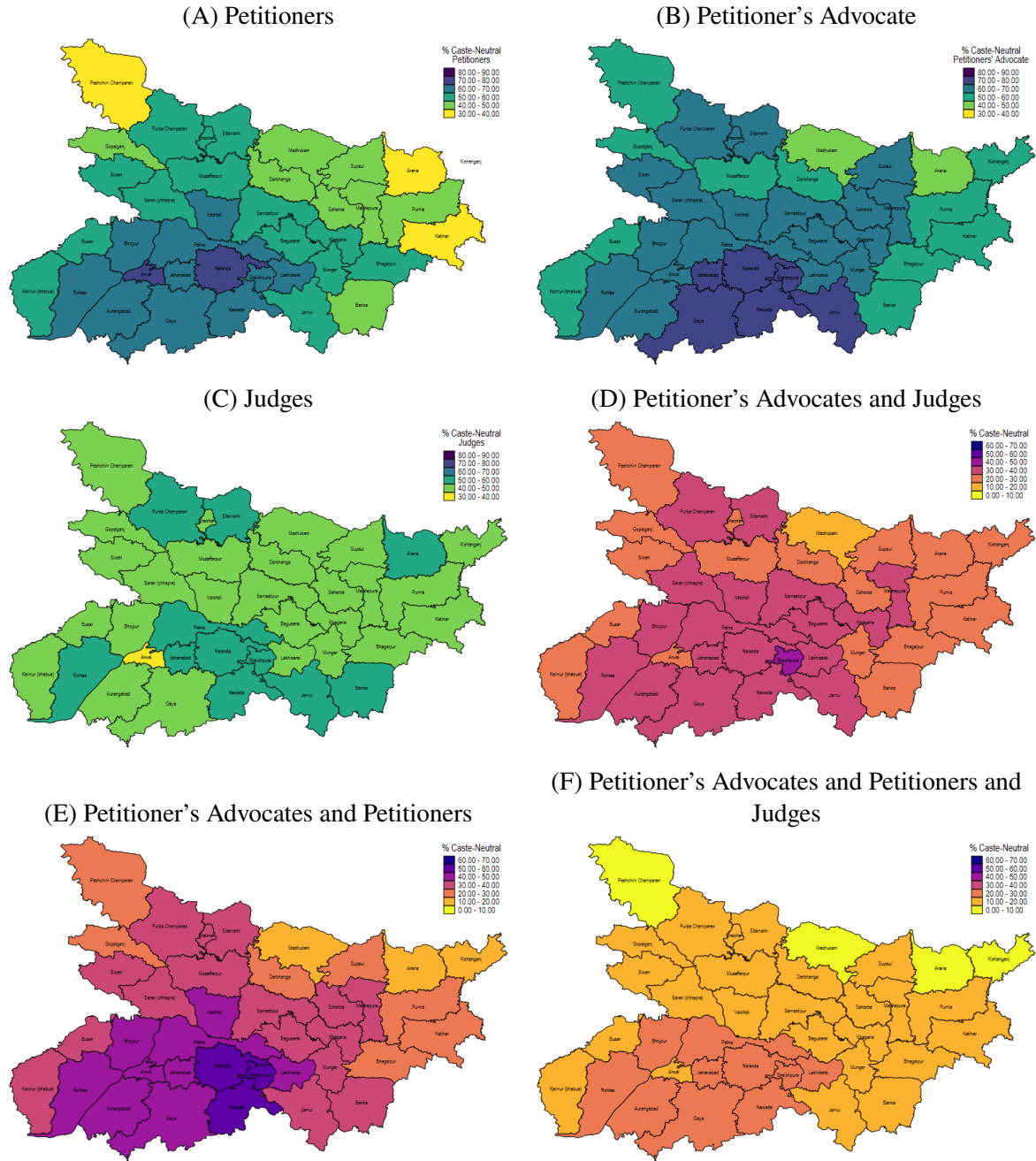
Figure 6: Spatial distribution of the use of caste-neutral names for cases filed at the Patna HC from 2008 to 2019.

Note: Districts with the fewest observations are dropped and marked in red. Bote also the different scale between Panels A to C (10-80%) and Panels D to F (0-50%).
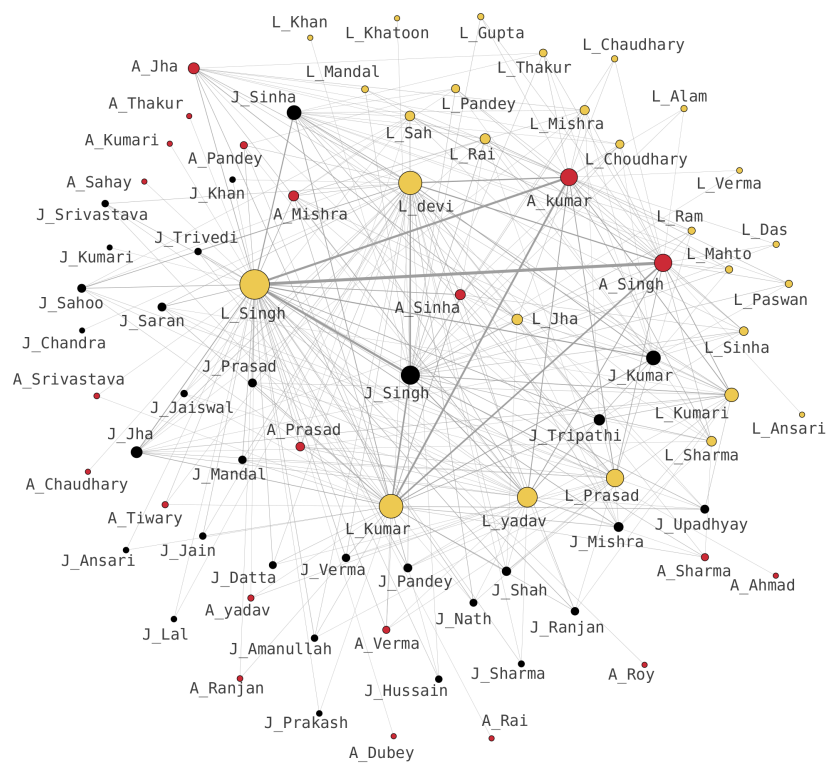
Figure 7: Networks between judges and lawyers, 2009–2019.

Notes: The data restricted to the names that appear at least 1000 times; "J: Judge", "A: Advocate", "L: Litigant"; size of the node is proportional to the weight of the edges that are associated with it; Lowest weight is 1000; Highest weight is 20,000.
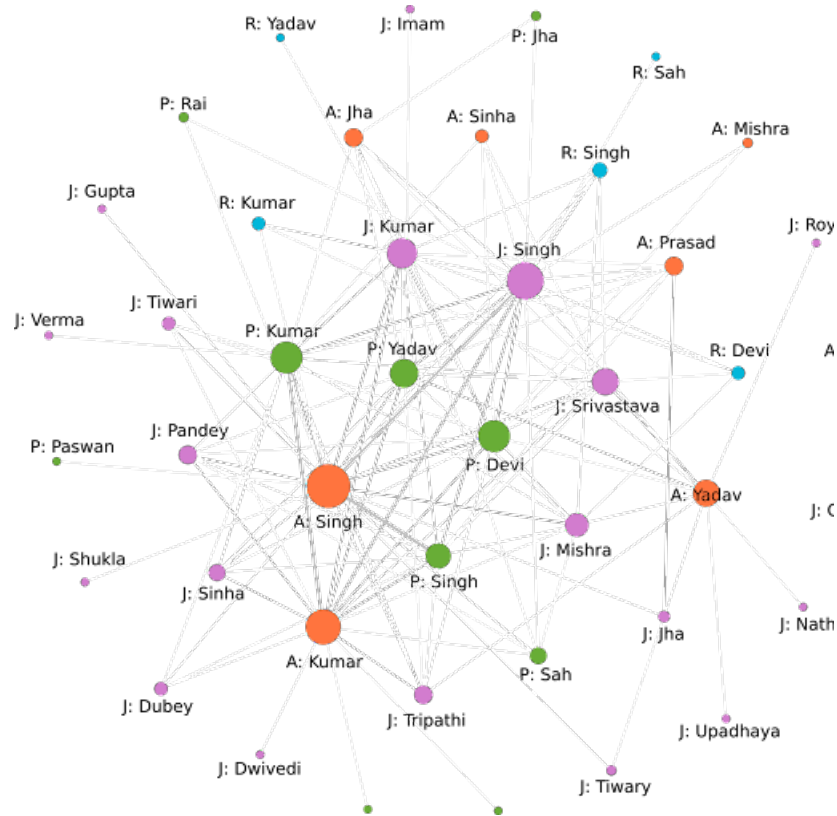
Figure 8: Networks between judges and lawyers in appeal cases, 2009–2019.

Notes: The data restricted to the set of appeals; "P: Petitioner", "R: Respondent", "A: Advocate" and "J: Judge". The color of the node represents the role, the size the degree and the thickness of the edge represents the number of cases with the respective matches in the given roles.
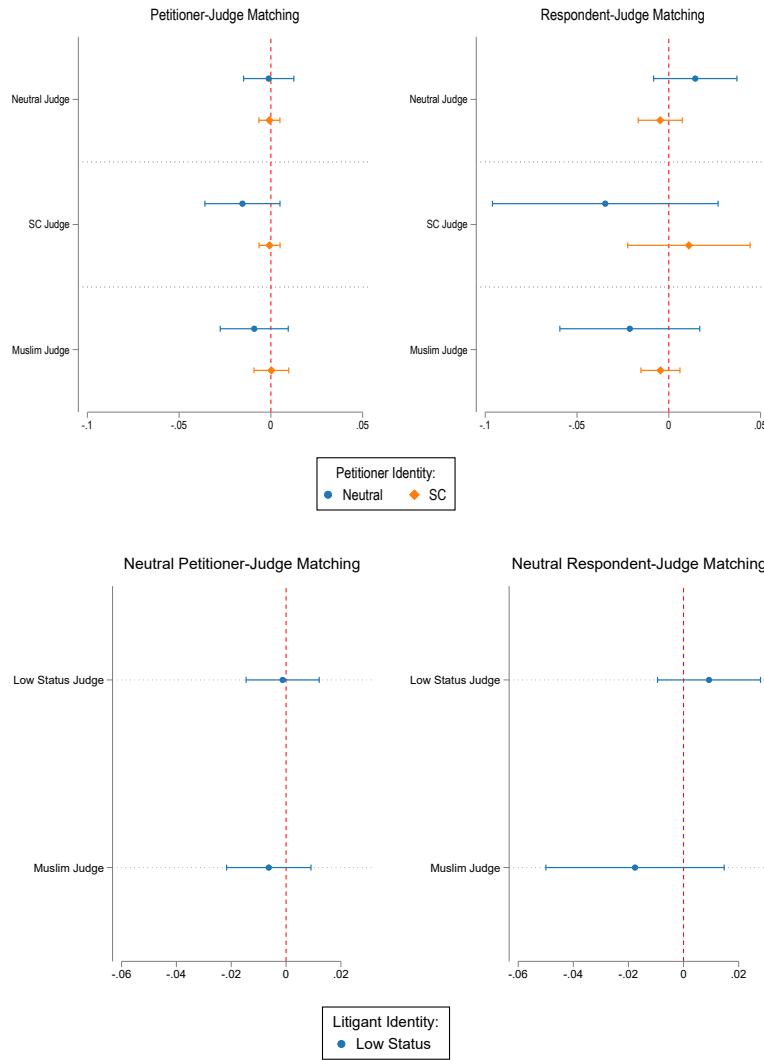
Figure 9: Test of Hypothesis (A): Likelihood of Low-Status Petitioners and Respondents matching with Judges with similar characteristics based on their attributes.

Note: Sample includes only judges from the first observable order in the regression. Panel (a) considers the two groups of names – caste-neutral and SC – separately, while Panel (b) groups them together. All regressions control for the age of the judge, if the judge pursued their career in the supreme court, the number of years the judge has a permanent position in the high court, district, year, month, the type of case and an interaction of the later four variables. Regressions are estimated separately across Panel A and Panel B and across petitioner's identities. Standard errors are clustered at district and year level. Confidence intervals correspond to 5% statistical significance.
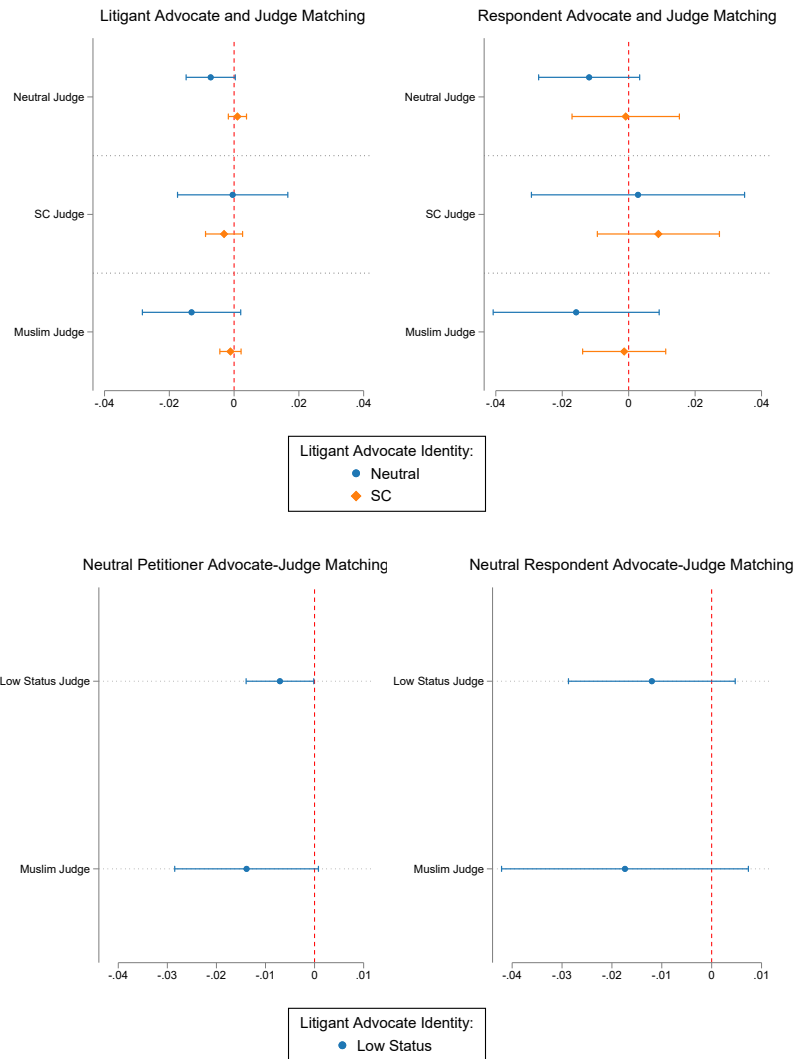
Figure 10: Test of Hypothesis (B): Likelihood of Advocates matching with Judges based on their attributes. Notes from Figure 9 apply.
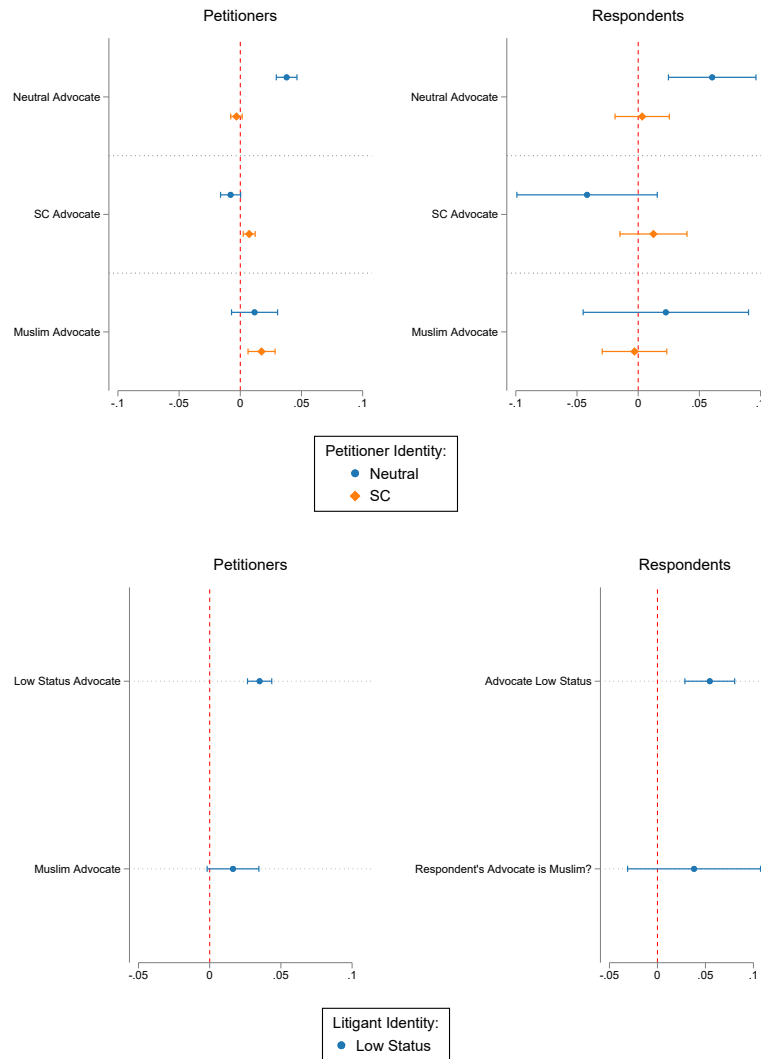
Figure 11: Test of Hypothesis (C): Likelihood of Petitioners matching with their Advocates based on attributes. Notes from Figure 9 apply.
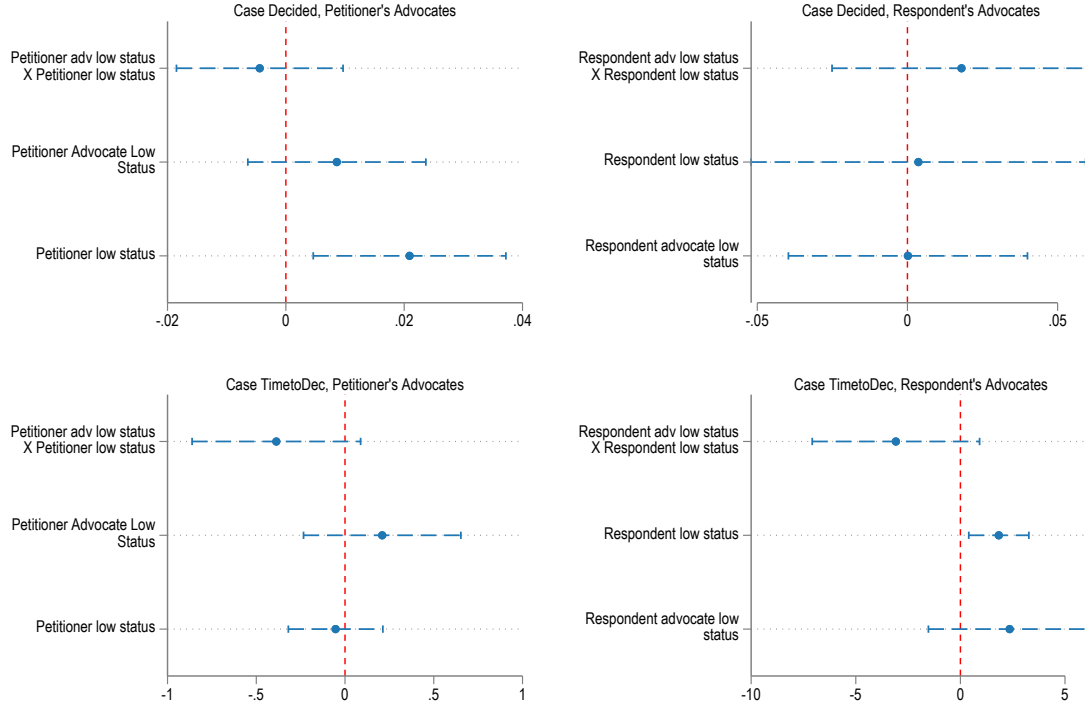
Figure 12: Litigants, advocates and case status.

Note: Regressions are based on the first observed order for a specific court case. Definition of judge identity is based on any judge on the bench with that identity. All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level. Confidence intervals correspond to 5% statistical significance.
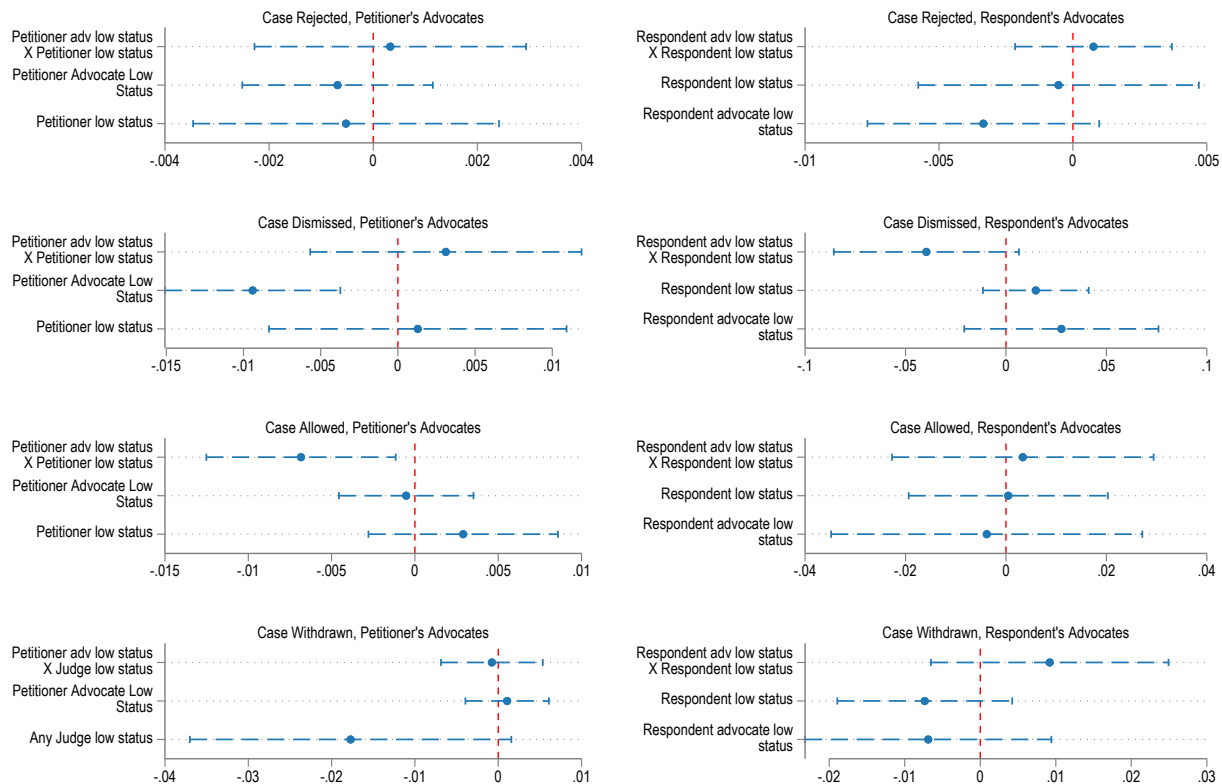
Figure 13: Litigants, advocates and case outcomes.

Note: Regressions are based on the first observed order for a specific court case. Definition of judge identity is based on any judge on the bench with that identity. All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level. Confidence intervals correspond to 5% statistical significance.
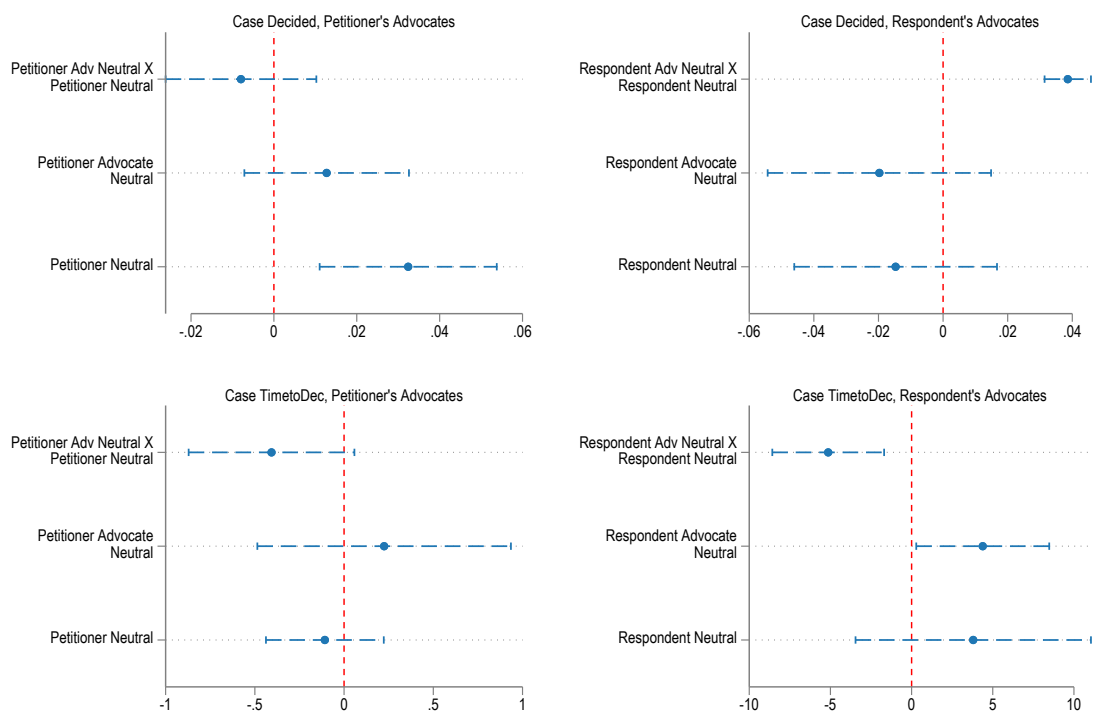
Figure 14: Litigants, advocates and case status. Sample includes only SC and caste-neutral petitioners. Notes from Figure 12 apply.
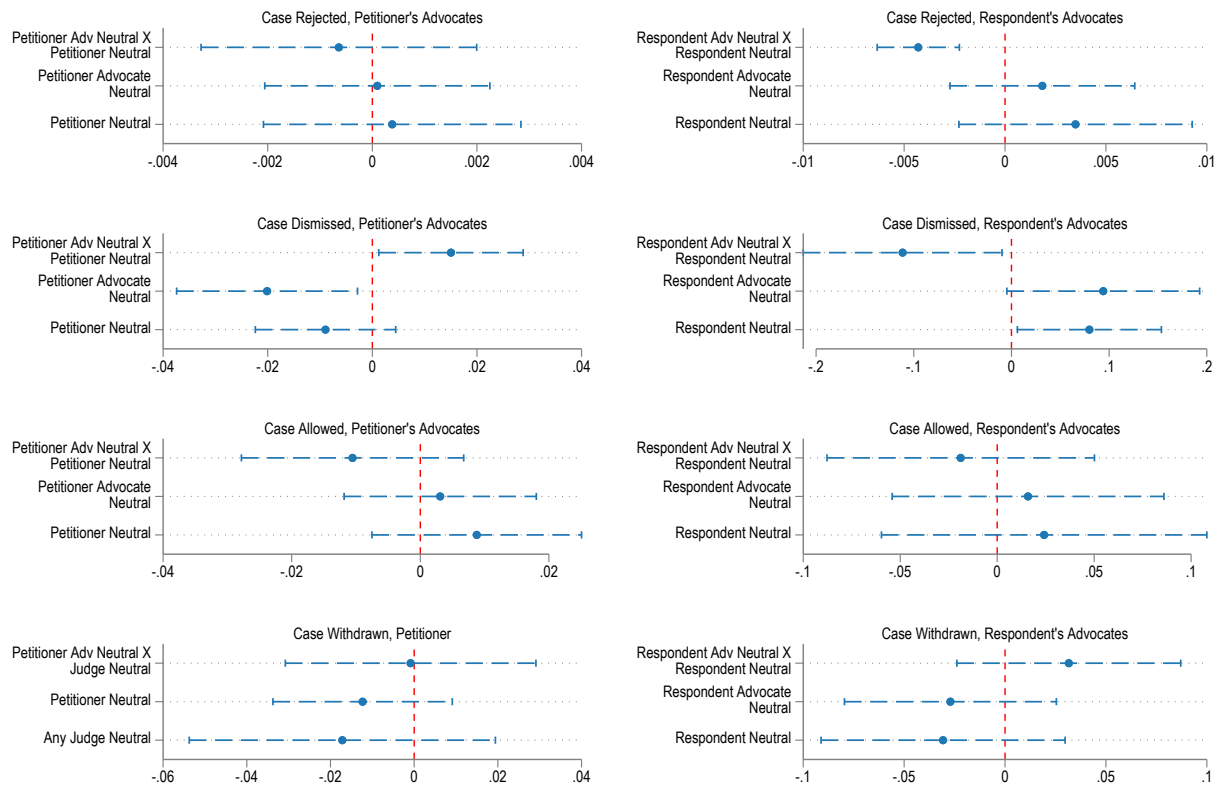
Figure 15: Litigants, advocates and case outcomes. Sample includes only SC and caste-neutral petitioners. Notes from Figure 13 apply.
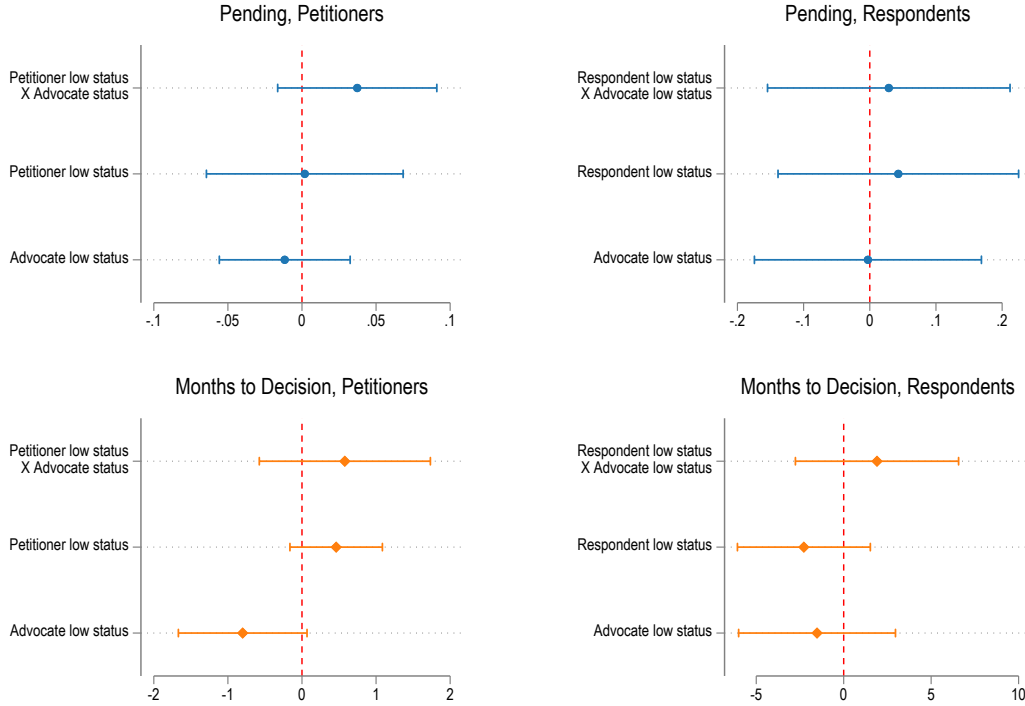
Figure 16: Status of Apppeals and Time to Decision, Petitioners and Respondents with Advocates.

Note: The regressions are based on appeal cases in the Patna High Court between 2012 and 2019. Panel (a) depicts the estimates of a dummy variable equal to one if a case was pending in the high court on a set of dummies of the petitioner's and petitioner's advocate's identity (at the district court level). Panel (b) is estimated on the subset of cases which were decided. The outcome variable is the number of months between the registration and the decision date in the high court. Both regressions include the year and month of the district court decision, the district, and a full set of interactions between the three variables as control variables. Standard errors are clustered at the district-year-level.

Figure 17: Appeal Case Outcome, Depending on Petitioner and Advocate Identities

Note: The regressions are based on appeal cases in the Patna High Court between 2012 and 2019. Top panel presents case outcomes for all petitioners. Bottom panel presents case outcomes for respondents.
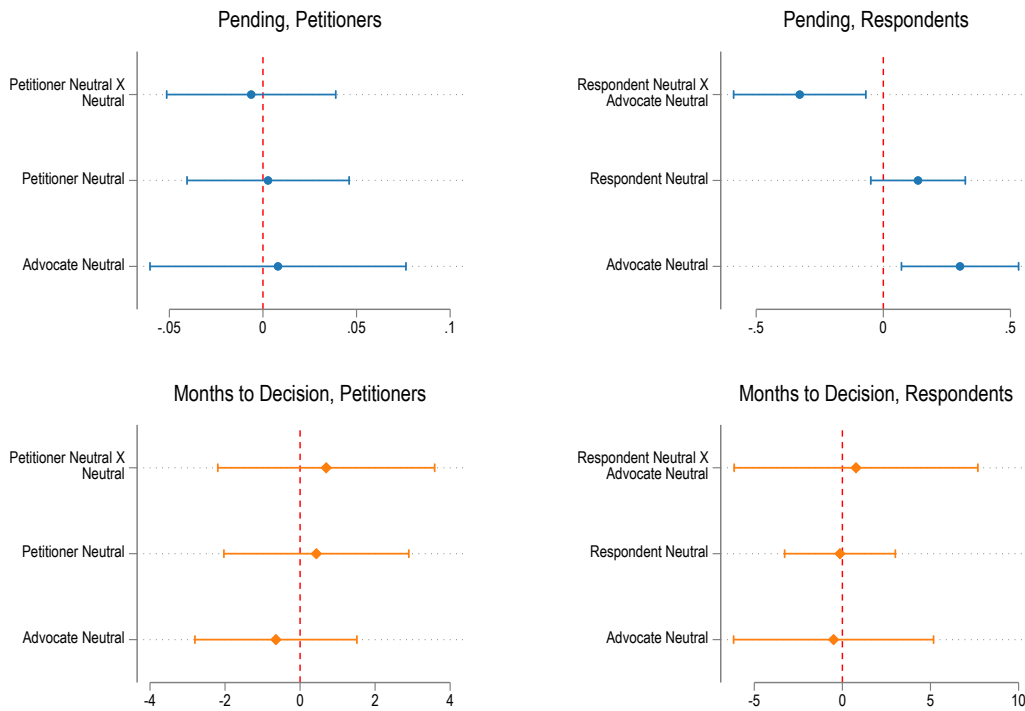
Figure 18: Status of Apppeals and Time to Decision, Petitioners and Respondents with Advocates, Restricted Sample

Note: Sample includes only litigants with caste-neutral and SC names who filed appeal cases in the Patna High Court between 2012 and 2019.Panel (a) depicts the estimates of a dummy variable equal to one if a case was pending in the high court on a set of dummies of the petitioner's and petitioner's advocate's identity (at the district court level). Panel (b) is estimated on the subset of cases which were decided. The outcome variable is the number of months between the registration and the decision date in the high court. Both regressions include the year and month of the district court decision, the district, and a full set of interactions between the three variables as control variables. Standard errors are clustered at the district-year-level.

Figure 19: Appeal Case Outcome, Depending on Petitioner and Advocate Identities, Restricted Sample

Note: Sample includes only litigants with caste-neutral and SC names who filed appeal cases in the Patna High Court between 2012 and 2019. Top panel presents case outcomes for all petitioners. Bottom panel presents case outcomes for respondents.

Appendix

Table A1: Summary Statistics - Merged Data Long

| Case-Judge-Level	N	Mean	SD	Min	Max
Filing Year	995,390	2013.45	3.05	2009.0	2019.0
Civil Case	995,390	0.79	0.40	0.0	1.0
Petitioner is Government?	995,390	0.01	0.11	0.0	1.0
Respondent is Government?	995,390	0.77	0.42	0.0	1.0
Petitioner is Individual?	995,390	0.89	0.31	0.0	1.0
Respondent is Individual?	995,390	0.16	0.36	0.0	1.0
Petitioner is Muslim?	898,991	0.09	0.29	0.0	1.0
Respondent is Muslim?	158,132	0.10	0.30	0.0	1.0
Petitioner's Advocate is Muslim?	990,626	0.05	0.21	0.0	1.0
Respondent's Advocate is Muslim?	754,812	0.08	0.28	0.0	1.0
Petitioner is a Woman?	898,991	0.26	0.44	0.0	1.0
Respondent is a Woman?	158,132	0.64	0.48	0.0	1.0
Petitioner's Advocate is a Woman?	990,626	0.14	0.35	0.0	1.0
Respondent's Advocate is a Woman?	754,812	0.16	0.37	0.0	1.0
Petitioner is SC?	867,788	0.10	0.30	0.0	1.0
Respondent is SC?	141,917	0.10	0.30	0.0	1.0
Petitioner's Advocate is SC?	957,218	0.06	0.24	0.0	1.0
Respondent's Advocate is SC?	693,944	0.06	0.24	0.0	1.0
Petitoner has caste-neutral name?	867,847	0.56	0.50	0.0	1.0
Responndent has caste-neutral name?	141,932	0.58	0.49	0.0	1.0
Petitoners's Advocate has caste-neutral name?	957,218	0.62	0.49	0.0	1.0
Respondent's Advocate has caste-neutral name?	693,944	0.58	0.49	0.0	1.0
Judge has caste-neutral name?	959,608	0.52	0.50	0.0	1.0
Judge is SC?	819,551	0.03	0.18	0.0	1.0
Judge is Muslim?	981,835	0.06	0.24	0.0	1.0
Judge is a Woman?	981,835	0.06	0.24	0.0	1.0
Judge's age	981,829	55.81	3.70	40.0	64.0
Years since Judge became permanent	995,390	1.20	3.25	-10.0	18.0
Was Chief Justice	981,835	0.05	0.22	0.0	1.0

Table A2: Petitioner identity, advocate identity and case status: Full Results

	(1) Decided b/se	(2) Months to Decision b/se	(3) Rejected b/se	(4) Dismissed b/se	(5) Allowed b/se	(6) Withdrawn b/se
Petitioner Advocate Low Status X Petitioner Low Status	-0.004	-0.387	0.000	0.003	-0.007*	0.001
	(0.006)	(0.213)	(0.001)	(0.004)	(0.003)	(0.002)
Petitioner Advocate Low Status	0.009	0.209	-0.001	-0.009**	-0.001	0.002
	(0.007)	(0.199)	(0.001)	(0.003)	(0.002)	(0.002)
Petitioner Low Status	0.021*	-0.053	-0.001	0.001	0.003	0.006
	(0.007)	(0.120)	(0.001)	(0.004)	(0.003)	(0.003)
Constant	0.678***	14.308***	0.006***	0.182***	0.098***	0.057***
	(0.004)	(0.076)	(0.001)	(0.002)	(0.002)	(0.002)
R-Squared	0.082	0.119	0.015	0.038	0.026	0.019
N	156293	108625	156293	156293	156293	156293
Petitioner Advocate LowStatus (b)	0.004	-0.178	-0.000	-0.006	-0.007***	0.002
Petitioner Advocate LowStatus (se)	(0.003)	(0.203)	(0.001)	(0.004)	(0.002)	(0.002)
Petitioner LowStatus (b)	0.016*	-0.440	-0.000	0.004	-0.004	0.007**
Petitioner LowStatus (se)	(0.009)	(0.291)	(0.001)	(0.003)	(0.003)	(0.003)
F:Petitioner Advocate LowStatus	1.270	1.659	1.294	7.130	6.813	1.226
p-value	0.322	0.239	0.316	0.012	0.014	0.334
F:Petitioner LowStatus	4.099	1.783	0.082	0.902	3.659	3.574
p-value	0.050	0.218	0.922	0.436	0.064	0.067

All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.

Table A3: Respondent identity, advocate identity and case status: Full Results

	(1) Decided b/se	(2) Months to Decision b/se	(3) Rejected b/se	(4) Dismissed b/se	(5) Allowed b/se	(6) Withdrawn b/se
Respondent Advocate Low Status X Respondent Low Status	0.018	-3.083	0.001	-0.040	0.003	0.009
	(0.019)	(1.795)	(0.001)	(0.021)	(0.012)	(0.007)
Respondent Advocate Low Status	0.000	2.356	-0.003	0.028	-0.004	-0.007
	(0.018)	(1.743)	(0.002)	(0.022)	(0.014)	(0.007)
Respondent Low Status	0.004	1.836*	-0.001	0.015	0.000	-0.007
	(0.025)	(0.644)	(0.002)	(0.012)	(0.009)	(0.005)
Constant	0.716***	19.358***	0.007***	0.163***	0.086***	0.036***
	(0.021)	(0.600)	(0.001)	(0.010)	(0.011)	(0.005)
R-Squared	0.282	0.129	0.068	0.108	0.110	0.074
N	6453	4656	6453	6453	6453	6453
Respondent Advocate LowStatus (b)	0.018	-0.727	-0.003	-0.012*	-0.000	0.002
Respondent Advocate LowStatus (se)	(0.011)	(0.464)	(0.002)	(0.006)	(0.010)	(0.005)
Respondent LowStatus (b)	0.022	-1.247	0.000	-0.025	0.004	0.002
Respondent LowStatus (se)	(0.013)	(2.087)	(0.002)	(0.021)	(0.010)	(0.002)
F:Respondent Advocate LowStatus	1.304	2.184	1.623	3.707	0.047	0.850
p-value	0.314	0.163	0.245	0.062	0.955	0.456
F:Respondent LowStatus	2.335	7.819	0.189	2.193	0.070	1.037
p-value	0.147	0.009	0.831	0.162	0.933	0.390

All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.

Table A4: Petitioner identity, petitioner's advocate identity and case status

	(1) Decided b/se	(2) Months to Decision b/se	(3) Rejected b/se	(4) Dismissed b/se	(5) Allowed b/se	(6) Withdrawn b/se
Petitioner Advocate Neutral X Petitioner Neutral	-0.008	-0.406	-0.001	0.015*	-0.011	-0.003
	(0.008)	(0.208)	(0.001)	(0.006)	(0.008)	(0.005)
Petitioner's Advocate Neutral	0.013	0.225	0.000	-0.020*	0.003	0.004
	(0.009)	(0.319)	(0.001)	(0.008)	(0.007)	(0.005)
Petitioner Neutral	0.032**	-0.108	0.000	-0.009	0.009	0.011
	(0.010)	(0.148)	(0.001)	(0.006)	(0.007)	(0.005)
Constant	0.670***	14.316***	0.005***	0.191***	0.093***	0.053***
	(0.008)	(0.249)	(0.001)	(0.007)	(0.006)	(0.005)
R-Squared	0.084	0.120	0.013	0.040	0.027	0.020
N	134184	93720	134184	134184	134184	134184
Petitioner Advocate Neutral (b)	0.005	-0.182	-0.001	-0.005	-0.007***	0.002
Petitioner Advocate Neutral (se)	(0.003)	(0.206)	(0.001)	(0.004)	(0.002)	(0.002)
Petitioner Neutral (b)	0.024**	-0.514*	-0.000	0.006*	-0.002	0.009***
Petitioner Neutral (se)	(0.009)	(0.273)	(0.001)	(0.003)	(0.003)	(0.003)
F:Petitioner Advocate Neutral	1.739	2.697	0.379	3.393	7.038	0.791
p-value	0.225	0.116	0.694	0.075	0.012	0.480
F:Petitioner Neutral	6.083	1.996	0.154	3.654	0.956	5.761
p-value	0.019	0.186	0.860	0.064	0.417	0.022

Sample restricted to SC and neutral petitioners only. All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.

Table A5: Respondent identity, respondent's advocate identity and case status

	(1) Decided b/se	(2) Months to Decision b/se	(3) Rejected b/se	(4) Dismissed b/se	(5) Allowed b/se	(6) Withdrawn b/se
Respondent Advocate Neutral X Respondent Neutral	0.039***	-5.137**	-0.004***	-0.111*	-0.019	0.032
	(0.003)	(1.546)	(0.001)	(0.046)	(0.031)	(0.025)
Respondent's Advocate Neutral	-0.020	4.383*	0.002	0.094	0.016	-0.027
	(0.016)	(1.838)	(0.002)	(0.044)	(0.031)	(0.024)
Respondent Neutral	-0.015	3.797	0.003	0.080*	0.024	-0.031
	(0.014)	(3.254)	(0.003)	(0.033)	(0.038)	(0.027)
Constant	0.736***	17.425***	0.003	0.101*	0.065	0.057*
	(0.020)	(2.921)	(0.002)	(0.034)	(0.036)	(0.025)
R-Squared	0.299	0.138	0.071	0.118	0.113	0.083
N	5548	4015	5548	5548	5548	5548
Respondent Advocate Neutral (b)	0.019	-0.755	-0.002	-0.017***	-0.003	0.005
Respondent Advocate Neutral (se)	(0.014)	(0.509)	(0.002)	(0.005)	(0.011)	(0.005)
Respondent Neutral (b)	0.024	-1.340	-0.001	-0.032	0.005	0.001
Respondent Neutral (se)	(0.015)	(1.815)	(0.002)	(0.021)	(0.011)	(0.003)
F:Respondent Advocate Neutral	88.248	11.390	11.235	7.651	0.204	0.952
p-value	0.000	0.003	0.003	0.010	0.818	0.418
F:Respondent Neutral	80.473	33.972	78.455	3.092	0.208	1.688
p-value	0.000	0.000	0.000	0.090	0.816	0.234

Sample includes SC and Neutral respondents only. All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.

Table A6: Appeal Regressions: Petitioner identity, petitioner's advocate identity and case status

	(1)	(2)	(3)	(4)	(5)	(6)
	Pending	Months to Decision	Rejected	Allowed	Dismissed	Withdrawn
	b/se	b/se	b/se	b/se	b/se	b/se
Petitioner Low Status X Advocate Low Status	0.037	0.578	-0.001	-0.009	0.000	-0.003
	(0.023)	(0.472)	(0.003)	(0.039)	(0.007)	(0.035)
Petitioner Low Status	0.002	0.461	0.000	-0.013	0.030**	-0.028
	(0.028)	(0.255)	(0.004)	(0.031)	(0.007)	(0.041)
Petitioner Advocate Low Status	-0.012	-0.800	0.004	0.006	-0.025	0.011
	(0.019)	(0.355)	(0.004)	(0.018)	(0.016)	(0.014)
Constant	0.556***	5.858***	0.002	0.429***	0.300***	0.116***
	(0.018)	(0.216)	(0.003)	(0.018)	(0.015)	(0.017)
R-Squared	0.165	0.451	0.190	0.351	0.277	0.245
N	11264	4645	4654	4654	4654	4654
Petitioner Advocate LowStatus (b)	0.026**	-0.222	0.003	-0.002	-0.024	0.008
Petitioner Advocate LowStatus (se)	(0.008)	(0.362)	(0.002)	(0.023)	(0.021)	(0.024)
Petitioner LowStatus (b)	0.039***	1.040*	-0.000	-0.022	0.030*	-0.031**
Petitioner LowStatus (se)	(0.009)	(0.470)	(0.003)	(0.023)	(0.014)	(0.011)
F:Petitioner Advocate LowStatus	4.747	2.590	1.047	0.129	1.882	1.175
p-value	0.050	0.155	0.407	0.881	0.232	0.371
F:Petitioner LowStatus	17.552	3.254	0.030	0.588	29.321	4.552
p-value	0.002	0.110	0.970	0.585	0.001	0.063

All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.

Table A7: Appeal Regressions: Respondent identity, respondent's advocate identity and case status

	(1) Pending b/se	(2) Months to Decision b/se	(3) Rejected b/se	(4) Allowed b/se	(5) Dismissed b/se	(6) Withdrawn b/se
Respondent Low Status X Advocate Low Status	0.029	1.904	-0.026	0.237***	-0.192*	0.003
	(0.075)	(1.906)	(0.011)	(0.033)	(0.053)	(0.074)
Respondent Low Status	0.043	-2.277	0.006	-0.112	0.091	0.016
	(0.074)	(1.553)	(0.008)	(0.059)	(0.071)	(0.067)
Respondent Advocate Low Status	-0.003	-1.523	0.027	-0.080**	0.137***	-0.012
	(0.070)	(1.832)	(0.012)	(0.017)	(0.016)	(0.075)
Constant	0.397***	7.943**	-0.002	0.441***	0.190***	0.111
	(0.063)	(1.427)	(0.008)	(0.012)	(0.018)	(0.057)
R-Squared	0.330	0.649	0.231	0.424	0.339	0.320
N	1059	550	552	552	552	552
Respondent Advocate LowStatus (b)	0.026*	0.381	0.001	0.157**	-0.055	-0.009
Respondent Advocate LowStatus (se)	(0.012)	(0.308)	(0.003)	(0.044)	(0.069)	(0.040)
Respondent LowStatus (b)	0.072*	-0.372	-0.020***	0.124**	-0.100**	0.020
Respondent LowStatus (se)	(0.032)	(0.959)	(0.005)	(0.043)	(0.036)	(0.020)
F:Respondent Advocate LowStatus	2.458	0.971	2.744	76.323	4005.714	0.031
p-value	0.166	0.431	0.142	0.000	0.000	0.969
F:Respondent LowStatus	2.503	1.222	7.332	34.710	13.935	0.624
p-value	0.162	0.359	0.024	0.001	0.006	0.567

All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.

Table A8: Appeal Regressions: Petitioner identity, petitioner's advocate identity and case status, Restricted Sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Pending	Months to Decision	Rejected	Allowed	Dismissed	Withdrawn
	b/se	b/se	b/se	b/se	b/se	b/se
Petitioner Neutral X Advocate Neutral	-0.006	0.697	-0.003	0.152	-0.198***	0.009
	(0.018)	(1.124)	(0.007)	(0.078)	(0.009)	(0.010)
Petitioner Neutral	0.003	0.434	0.000	-0.102	0.108**	-0.019
	(0.018)	(0.959)	(0.003)	(0.071)	(0.021)	(0.020)
Petitioner's Advocate Neutral	0.008	-0.644	0.007	-0.138*	0.159***	0.015
	(0.028)	(0.840)	(0.005)	(0.048)	(0.022)	(0.018)
Constant	0.574***	5.547**	0.001	0.521***	0.217***	0.101***
	(0.025)	(0.847)	(0.003)	(0.045)	(0.029)	(0.004)
R-Squared	0.199	0.490	0.265	0.396	0.346	0.271
N	6897	2624	2631	2631	2631	2631
Petitioner Advocate Neutral (b)	0.002	0.053	0.004	0.014	-0.039**	0.023
Petitioner Advocate Neutral (se)	(0.012)	(0.464)	(0.002)	(0.033)	(0.013)	(0.024)
Petitioner Neutral (b)	-0.004	1.131*	-0.003	0.050**	-0.089***	-0.011
Petitioner Neutral (se)	(0.018)	(0.467)	(0.006)	(0.016)	(0.012)	(0.016)
F:Petitioner Advocate Neutral	0.064	0.325	10.159	8.772	51.256	0.471
p-value	0.939	0.737	0.017	0.023	0.001	0.650
F:Petitioner Neutral	0.058	3.258	0.126	5.062	26.212	0.528
p-value	0.944	0.124	0.884	0.063	0.004	0.619

Sample includes only litigants with caste-neutral and SC names who filed appeal cases in the Patna High Court between 2012 and 2019.All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.

Table A9: Appeal Regressions: Respondent identity, respondent's advocate identity and case status, Restricted Sample

	(1) Pending b/se	(2) Months to Decision b/se	(3) Rejected b/se	(4) Allowed b/se	(5) Dismissed b/se	(6) Withdrawn b/se
Respondent Neutral X Advocate Neutral	-0.329*	0.775	-0.038	-0.239***	-0.220*	0.205**
	(0.106)	(2.826)	(0.031)	(0.037)	(0.064)	(0.048)
Respondent Neutral	0.137	-0.135	-0.001	0.236***	0.101	-0.151**
	(0.076)	(1.283)	(0.003)	(0.034)	(0.056)	(0.031)
Respondent's Advocate Neutral	0.302*	-0.498	0.034	0.312***	0.106	-0.178**
	(0.094)	(2.318)	(0.029)	(0.038)	(0.074)	(0.032)
Constant	0.324**	5.606***	0.005	0.216***	0.198***	0.228***
	(0.077)	(0.710)	(0.003)	(0.021)	(0.026)	(0.010)
R-Squared	0.346	0.658	0.518	0.465	0.374	0.329
N	606	296	296	296	296	296
Respondent Advocate Neutral (b)	-0.027	0.277	-0.004	0.072*	-0.115*	0.027
Respondent Advocate Neutral (se)	(0.026)	(1.541)	(0.003)	(0.037)	(0.054)	(0.047)
Respondent Neutral (b)	-0.192**	0.640	-0.039	-0.004	-0.119	0.055
Respondent Neutral (se)	(0.066)	(1.772)	(0.031)	(0.058)	(0.102)	(0.042)
F:Respondent Advocate Neutral	5.139	0.038	1.145	37.509	6.913	18.039
p-value	0.050	0.963	0.379	0.000	0.028	0.003
F:Respondent Neutral	5.287	0.087	0.891	66.502	12.281	13.963
p-value	0.047	0.918	0.458	0.000	0.008	0.006

Sample includes only litigants with caste-neutral and SC names who filed appeal cases in the Patna High Court between 2012 and 2019. All regressions control for high court, district, year, month, the type of case fixed-effects. Standard errors are clustered at district and year level.