

# Does Victim Gender Matter for Justice Delivery? Police and Judicial Responses to Women's Cases in India

NIRVIKAR JASSAL *London School of Economics and Political Science, United Kingdom*

**A**re women disadvantaged whilst accessing justice? I chart, for the first time, the full trajectory of accessing justice in India using an original dataset of roughly half a million crime reports, subsequently merged with court files. I demonstrate that particular complaints can be hindered when passing through nodes of the criminal justice system, and illustrate a pattern of “multi-stage” discrimination. In particular, I show that women's complaints are more likely to be delayed and dismissed at the police station and courthouse compared to men. Suspects that female complainants accuse of crime are less likely to be convicted and more likely to be acquitted, an imbalance that persists even when accounting for cases of violence against women (VAW). The application of machine learning to complaints reveals—contrary to claims by policymakers and judges—that VAW, including the extortive crime of dowry, are not “petty quarrels,” but may involve starvation, poisoning, and marital rape. In an attempt to make a causal claim about the impact of complainant gender on verdicts, I utilize topical inverse regression matching, a method that leverages high-dimensional text data. I show that those who suffer from cumulative disadvantage in society may face challenges across sequential stages of seeking restitution or punitive justice through formal state institutions.


## INTRODUCTION

**A**re women disadvantaged when accessing justice, and if so how? In the largest democracy of India, journalists regularly report stories that women (and minorities) are discriminated against when seeking help from the state. Still, it remains unclear whether any disparities, if they did exist, are attributable to the case *types* registered by particular groups or their *identity*. If women are discriminated against, is it because of their gender or the nature of their complaints, for example, challenges associated with proving cases involving violence against women (VAW)?<sup>1</sup>

Not only is there limited research on policing and courts in political science, but also few discussions about inequities in accessing justice (Grossman et al. 2016). Further, scholarship on VAW in economics (Jayachandran 2015), sociology (Armstrong, Gleckman-Krut, and Johnson 2018), or criminology (Khan et al. 2020), is often carried out through the prism of sexual assault (McDougal et al. 2021). In political science, questions about VAW have focused on rape in conflict or post-conflict settings

(Agerberg and Kreft 2020; Cohen 2013), rather than how the state takes cognizance of everyday harassment and abuse (Khan et al. 2020). Moreover, while emerging scholarship has sought to re-focus attention toward law-and-order, existing studies primarily test the impact of interventions (e.g., police training or community engagement [Blair, Karim, and Morse 2019; Blair et al. 2021]), rather than spotlight the extensive system of justice delivery.

I ask whether women are less likely than men to access justice upon turning to the state (police and judiciary). The article is situated in India, a site dubbed the most unsafe country for women (Goldsmith and Beresford 2018), where surveys show that 28%, 6.6%, and 78.4% of women report physical violence, sexual assault, and fear of their spouse (sometimes or always), respectively (DHS 2017). The study extends research on gender disparities in South Asia—that has included scholarship on education (Beaman et al. 2012), politics (Chattopadhyay and Duflo 2004), health (Dupas and Jain 2021), and property rights (Brulé 2020)—to include justice delivery. I document how certain complaints are filtered while funneling through a tiered system. This filtration, evident at specific junctures in bureaucratic processing, compounds existing inequalities, including those rooted in gender. The results foreground how discrimination is iterative such that inequity in one agency can be reproduced in another. To illustrate, I create an individual-level dataset of crime, and merge it with court files, thereby tracing cases from the minute a victim enters a police station until (potentially years) later following a judicial verdict. I combine several research topics—for example, from courtroom gender biases to police responsiveness toward VAW—into one holistic study. By linking all

Corresponding author: Nirvikar Jassal , Assistant Professor, Department of Government, London School of Economics and Political Science, United Kingdom [n.jassal@lse.ac.uk](mailto:n.jassal@lse.ac.uk).

Received: January 19, 2022; revised: October 20, 2022; accepted: August 15, 2023.

<sup>1</sup> According to the UN (1994), VAW is, “any act of gender-based violence that results in, or is likely to result in, physical, sexual or mental harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life.” Sexual assault is one component of VAW.

arms of the system, I establish a series of facts, for example, cases of VAW are delayed vis-à-vis police registration and court verdict compared to non-VAW. Strikingly, even accounting for VAW, women are significantly more likely to have their cases dismissed or result in a suspect's acquittal rather than conviction compared to men. Using methodological advancements in text matching, I attempt to provide credible evidence that the discrepancies can indeed be attributed to the complainant's gender identity.

The paper makes additional contributions. For instance, scholarship has pointed to social impediments hindering vulnerable groups from registering crime (Green, Wilke, and Cooper 2020; Iyer et al. 2012), with an implicit assumption that if only they can be encouraged to report, the state will take action. The findings herein suggest that anxieties about reporting can be a rational or strategic response to the low probability of justice at the conclusion of an arduous process. Initial "gatekeeping" by police in terms of case filing, while important, does not fully capture the state's accountability toward victims of crime and abuse (Spohn and Tellis 2019).

The study also supplements work on bureaucratic discrimination (Emeriau 2022), much of which has focused on race rather than gender or involved audit experiments as opposed to administrative data (Butler and Broockman 2011; White, Nathan, and Faller 2015). With official records, I point to certain direct and indirect challenges that women face as their complaints are being processed, and simultaneously quantify the duration of police investigations, court sessions, station wait-times, and bail hearings, that is, granular points of interest to scholars of state capacity, bureaucracy, gender, policing, and judicial politics in the contemporary Global South.

Another novelty of the study is that it applies unsupervised machine learning to police reports, each of which contain  $\approx$  500-word first-person testimonies (Roberts, Stewart, and Airoidi 2016; Roberts, Stewart, and Tingley 2019). While such methods have been used to understand Islamic *fatwas* (Lucas et al. 2015), Indian rural deliberation (Parthasarathy, Rao, and Palaniswamy 2019), or British parliamentary debate (Sanders, Lisi, and Schonhardt-Bailey 2017), they have not frequently been applied to the study of crime or VAW. The benefits of a text-as-data approach are three-fold. First, it amplifies *victims'* voices. Second, topic modeling can disentangle VAW carried out *in* and *out* of the household and summarize real-life cases, for example, marital rape or abuse related to women's extortion for dowry. Third, text matching can adjust for confounding so as to make an attempt at causal inference using text (Roberts, Stewart, and Nielsen 2020).

The paper is structured as follows: I define "multi-stage" discrimination, contextualize India's criminal justice system, and explain the merging of records from two distinct agencies. I present tests of the argument, utilizing descriptive and OLS analyses, topic modeling, and text matching. I conclude by highlighting a new, broader research agenda that the findings illuminate.

## MULTI-STAGE DISCRIMINATION

Most work on discrimination<sup>2</sup> examines isolated stages or "episodic disparities" rather than the reproduction of unequal treatment from one setting to the next (Kurlychek and Johnson 2019; Kutateladze et al. 2014). The limited theorization and focus on dynamic processes of discrimination suggests that social science is underestimating the true levels of disadvantage that citizens face, including when interacting with linked agencies in a system like criminal justice (Bohren, Hull, and Imas 2022; Reskin 2012). Such neglect may yield inaccurate conclusions in scholarship; for instance, if police mishandle or carry out biased investigations, judges have limited evidence to prosecute, and so a single-stage analysis of judicial rulings alone could lead to an imprecise argument that judges are to blame (Lang and Kahn-Lang Spitzer 2020).

Unlike political science, the disciplines of economics and especially sociology have approached inequality from what Reskin (2012) calls "a systems perspective," yet there remains scant dialogue between the fields as to how to conceptualize non-episodic forms of discrimination (Small and Pager 2020). Often, terms such as "cumulative discrimination," "structural discrimination," "cumulative disadvantage," and "*über* discrimination" are interchangeably used to convey similar ideas (Bohren, Hull, and Imas 2022; DiPrete and Eirich 2006; Kurlychek and Johnson 2019; Reskin 2012). Blank (2005, 99) notes that, "cumulative discrimination is the measurement of discriminatory effects over time and across domains," where one disadvantage feeds into the next. Over-time discrimination may cut across domains (or systems), reinforcing disparities that affect future generations (Lang and Kahn-Lang Spitzer 2020). Discrimination against people of color in, say, the U.S. housing market can contribute to residential segregation, in turn negatively affecting health or educational opportunities downstream (Korver-Glenn 2018; Reskin 2012).

I focus on a specific category of "cumulative discrimination" outlined by Blank (2005, 99), that is, "discrimination that arises from multiple interactions within a single social domain over time." I refer to this as "multi-stage" discrimination because it occurs at sequential nodes within a system or domain (e.g., criminal justice) whereby the *process* and *outcomes* associated with a complaint, application, or request are affected at one or more decision-points or "stages." Despite serving in distinct agencies or sub-systems, administrators that mediate the system's stages may be inter-connected by rules, routines, or norms. Complaints, applications, or requests travel via stages vertically (*through* sub-systems, e.g., chain of police custody) or horizontally (*across* sub-systems, e.g., law enforcement to courthouse). In criminal justice, stages or decision-points might include: (a) *police registration*,

<sup>2</sup> Lang and Kahn-Lang Spitzer (2020, 68) define discrimination as, "treating someone differently based on characteristics such as gender, race, or religion."

for example, police may turn away or dissuade citizens from case filing; (b) *police investigation*, for example, officers might delay inquiries or cajole the complainant into withdrawing the report; (c) *court trial*, for example, judges can stall arbitration or postpone hearings; and (d) *court verdict*, for example, judges may vary sentences or acquit rather than convict suspects.

Tracing multi-stage discrimination has a defined scope. Broader ideas of cumulative discrimination encapsulate injustices across countless domains which could coincide or overlap at a single time point, or even refer to the experiences of individuals that have not had direct contact with a system. For instance, Black Americans may be victims of discrimination by a U.S. criminal justice system without ever even having officially interacted with the police (Soss and Weaver 2017). Instead, multi-stage discrimination refers to a specific channel by which inequities propagate across time, *conditional* on entering a system bounded by rules, routines, or norms. Those that enter this system may have more in common with one another than others that have had indirect or no contact. Predictable stages within the system allow for transitions or serve as leverage points at which administrators have discretion to influence pathways or outcomes.

In a single system, administrators might issue decisions with expectations of how other officials will react (e.g., police may inadequately investigate cases that they know, or perceive, judges will simply dismiss). Feedback loops are particularly salient when administrators are linked by rules, have a joint stake in future decisions, or share resource constraints (Kurlychek and Johnson 2019). For instance, if an administrator is biased, but is unable to discriminate in the first stage, they may rely on officials downstream to do so. Alternatively, because administrators are constrained by time or resources, they could depend on previous officials in the chain to assist in lightening system load. The mere fact of having multiple stages where human discretion can be applied, absent any coordination by administrators, could lead to greater opportunities for a complaint, application, or request to fail.<sup>3</sup> Broadly, multi-stage discrimination is only traceable longitudinally rather than at a single stage or time point, and is conceptually distinct from overlapping indirect forms of discrimination that citizens face via disparate domains.

Studying a single system's stages may shed light on *drivers* of discrimination. There are numerous mechanisms by which inequalities persist, for example, administrators may have a "taste" for discrimination (Becker 2010), they may lack information (statistical discrimination) (Phelps 1972), or the implementation of rules or procedures could "implicitly" or indirectly disadvantage some over others (Bertrand and Duflo 2017; Bohren, Hull, and Imas 2022). System decomposition could

underscore whether certain mechanisms are more likely than others. As an example, suppose biased administrators want to curtail minorities' cases in early stages but can find no grounds to do so because of the overwhelming evidence. Then, later stages might see a progressively comparable distribution of valid cases. Nevertheless, if minority groups' requests at a final stage of a system are *still* more likely to be rejected than the majority—despite having crossed a higher bar of filtration at earlier nodes—we may be better placed in pointing to, say, taste-based versus statistical discrimination on the part of administrators as drivers of injustice (Arrow 2015). In fact, we may even learn about the behaviors of those *outside* a system too. For example, finding an absence of police discrimination at the same time as citizen anxieties about approaching law enforcement for help could be a puzzle that is clarified by tracing stages; it could be that district attorneys in the *middle* rungs of a process are discriminatory,<sup>4</sup> which ultimately deters citizens from the *first* node of case filing with the police. Consequently, exploring within-system discrimination sequentially may reveal more than the sum of its parts at individual stages.

There are other benefits of charting multi-stage discrimination. While experiments causally identify the incidence of discrimination, they do not highlight the process by which disparities compound (Bertrand and Duflo 2017),<sup>5</sup> or spotlight where bottlenecks and leakage manifest (Holland 2016). Without this fuller picture, any policy intervention applied at one stage might not only fail to reduce inequality, but also have the unintended consequence of exacerbating it. Let us imagine an experiment that reveals statistical discrimination against minorities by judges. Such an empirical finding may drive policymakers to target resources at that node by, say, providing information to administrators so as to minimize future judicial discrimination (Bohren, Imas, and Rosenberg 2019; Small and Pager 2020). Nevertheless, in tiered systems connected by multiple stages, such one-stop solutions are likely to have limited impact. If judges are provided information, they might update their beliefs or be lenient with related cases in their docket, but patterns of disadvantage will ultimately remain static because of courts' inability to overturn prior decision-points or even observe cases that they were not sent by law enforcement. Suppose, instead, that it is law enforcement that is discriminatory in terms of preventing cases filed by minority groups from reaching court. Any policy intervention targeted at the node of case filing that reduces or constrains police officer discretion could simply shift discrimination downstream to the judicial stage. Because courts would now be overwhelmed with

<sup>3</sup> Suppose one survives a stage with probability  $p$ , but if there are  $n$  stages, the probability of survival becomes  $p^n$  (which is smaller than  $p$  and tends to zero as  $n \rightarrow \infty$ , assuming that each stage's probability is independent).

<sup>4</sup> Spohn and Tellis (2019) show how numerous sexual assault cases for which the Los Angeles Police Department have probable cause never yield arrest but are rejected by the District Attorney prior to felony charges.

<sup>5</sup> Dynamic processes of discrimination may even *mitigate* initial disparities rather than magnify them (Bohren, Imas, and Rosenberg 2019; Stolzenberg, D'Alessio, and Eitle 2013).



complaints that they relied on being filtered at earlier nodes, exasperated judges might become *more* dismissive of minorities' cases in their docket, thereby intensifying discrimination. In this way, charting a multi-stage process spotlights *where* anti-discrimination policy interventions should be targeted such that they can have a positive ripple effect through the chain.

Political science scholarship has pointed to inadequacies with the study of (racial) discrimination with administrative (police) data due to selection and incomplete information, for example, the stage at which police stop citizens in the United States (Zhao et al. 2022). As Knox, Lowe, and Mummolo (2020) astutely point out, analysts miss information on individuals that *police observe but do not investigate*, thereby introducing bias in the study of (racial) discrimination by the police. However, decomposing the criminal justice system could be useful vis-à-vis understanding discrimination at later nodes, for example, access to police records may at least shed light on individuals that *district attorneys and judges observe but do not prosecute*. Still, a multi-stage process induces selection across every node, and studying discrimination within a bounded system does not make it a less challenging statistical task to gain unbiased estimates of the cumulative effect or magnitude of total discrimination (e.g., those with stronger cases might select-in, only those with overwhelming evidence advance, and so on). In fact, if a higher burden of proof is applied to a minority group in an early undocumented stage (e.g., when cases are being reported), the unusually strong complaints reaching subsequent stages could be masking discrimination, especially in administrative records, because officially it might appear that all groups are being treated similarly or even that the disadvantaged group is being treated better.

In this study, I test just one plausible implication of multi-stage discrimination in justice delivery, that is, that disadvantaged groups will see a diminished *speed* and *likelihood* of their complaints, applications, or requests successfully crossing the system's stages. As Kurlychek and Johnson (2019) note, research has faced challenges in charting criminal cases across time (and space), even in the United States (Rehavi and Starr 2014). I simply follow each administrator decision sequentially in a criminal justice system from the stage of entry (police registration) to exit (judicial verdict) for every complaint. I document whether women face non-episodic unequal outcomes (*exclusion*) or a disproportionately trying process (*burdens*) as requests for help transit through the state apparatus of justice provision (Olsen, Kyhse-Andersen, and Moynihan 2022).

## GENDER AND THE INDIAN CRIMINAL JUSTICE SYSTEM

Research on criminal justice—often restricted to the United States, and focused on the police *or* judiciary—shows that Black Americans are discriminated against in terms of bail, sentencing, and incarceration (Abrams, Bertrand, and Mullainathan 2012; Alesina and La

Ferrara 2014; Arnold, Dobbie, and Yang 2018). Aside from related scholarship on ethnicity (Shayo and Zussman 2011), there are few analyses of gender-based disparities. I address this lacuna by focusing on India; here, women may have limited support such as access to lawyers (Roychowdhury 2021), and investigating VAW might be perceived as a strain on police resources. Culturally, officials might construe punitive justice for women's complaints, including dowry,<sup>6</sup> as a threat to societal norms or marriage. Further, because gender disparities are pronounced in areas like education or labor force participation, we may see imbalances between men and women vis-à-vis accessing justice even absent any malign intent by administrators, for example, women could be fearful about retribution and humiliation or lack autonomy from the household to travel to courthouses (Chhibber 2002).

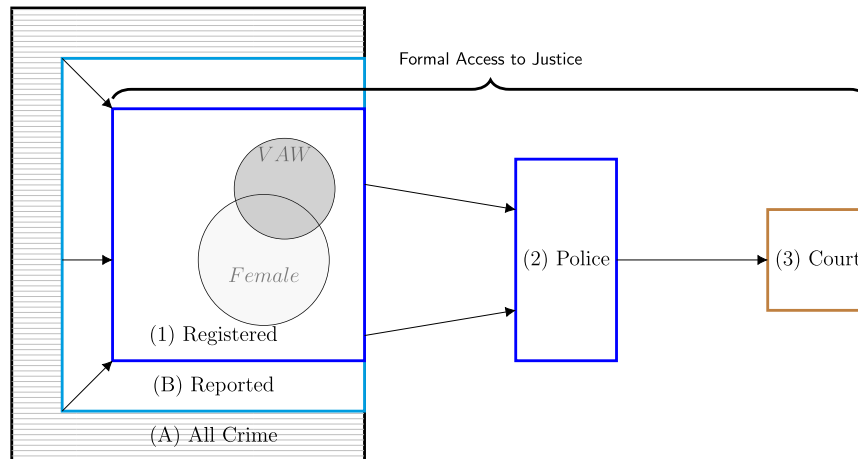
The Hindi-speaking belt of north and central India is a crucial site to study inequality (Jayachandran 2017); by global standards, this region, and especially Haryana state, retains among the most skewed sex-ratios in favor of men (834 girls to 1,000 boys in 2011 [Chhibber, Jensenius, and Ostermann 2021]) as a function of sex-selective abortion and female infanticide. Haryana may thus present not only an upper bound of gender-based disparities in accessing justice, but also illustrate how inequalities reproduce in the formal system of grievance redressal.

When seeking help from the state to address a wrong, crime registration is a primary step. It occurs at police stations run by a station officer, who is supported by staff (e.g., sub- and assistant sub-inspectors). Police are supposed to file all complaints whether they believe them to be truthful or not, but in practice have leeway as to which are formally registered. When filed, the station officer assigns a case to a deputy, and, depending on crime-type, investigations have to be completed within a window (e.g., 60 days). If not canceled, or withdrawn, the case is sent to court. Every station is located within a jurisdiction of a district court; police reports, and any evidence collected during investigation, are assigned to a judge.<sup>7</sup> In a major study that introduces a fine-grained dataset of Indian court decisions, Ash et al. (2021) find a null effect of (judicial) bias with regard to in-group *defendants*. Here, I focus not those who have been accused of crime per se, but on those who initiated the legal action in the previous stage: the *plaintiffs*.

Figure 1 presents a stylized illustration. Level A represents the abstract concept of all crime. Level B signifies those who came forward to *report*, for example, at a station or help-desk (Sukhtankar, Kruks-Wisner, and Mangla 2022). Level 1 is smaller than “B” because not all reported cases are registered. This study focuses on those cases that were filed with the

<sup>6</sup> Unlike bride-price, dowry involves a wife being coerced into providing resources to her spouse. The widespread practice has been linked to domestic violence, murder, and “missing girls” (Bhalotra, Chakravarty, and Gulesci 2020; Rao 1997; Srinivasan and Bedi 2007).

<sup>7</sup> On appeal, a case may travel to a High Court or Supreme Court.

**FIGURE 1. Levels of Accessing Formal Justice in India**

Note: Light and dark blue represent police jurisdiction; yellow represents the judiciary. The analyses in this study cover all levels from 1 to 3, and the corresponding in-between stages.

state. Within level 1, there are two sub-categories: women's complaints and VAW.<sup>8</sup> This is illustrated in a Venn diagram because not all VAW is reported by women.<sup>9</sup> Cases in level 2 represent those that, after registration, survive cancellation. The remaining cases, once investigated, enter the judiciary in level 3. There, unless stalled or dismissed, cases result in a verdict following trial that (dis)favors the complainant in the original filing from level 1 (who eventually became the plaintiff in court).

Judges arguably have greater discretion when handling cases than the police. For officers, there are explicit rules that mandate registration of "cognizable" or serious crimes,<sup>10</sup> some introduced after a gruesome 2012 gang-rape of a Delhi college student. Police are required to register all VAW complaints—including acid attacks, criminal force, trafficking, and rape—with the threat of 1-year jail time or fine for the officer, and rape investigations are to be completed within 2 months of filing.<sup>11</sup> Aside from being pressured "from above" via such guidelines and orders, the police are also constrained "from below" where, for example, feminist groups and NGOs can assist victims in filing cases, especially VAW (Htun and Weldon 2012). Judges are disassociated from such pressures<sup>12</sup> or from juries, which were formally abolished in 1973.<sup>13</sup>

<sup>8</sup> The gap between levels 1—"A" is called "the dark figure of crime" (Biderman and Reiss 1967).

<sup>9</sup> VAW can be further subdivided, for example, abuse *inside* the household may involve the spouse, family, or in-laws.

<sup>10</sup> Section 154 of the Code of Criminal Procedure.

<sup>11</sup> Section 166A of the Penal Code and Section 173 of the Code of Criminal Procedure.

<sup>12</sup> Law enforcement is also constrained by the civil bureaucracy (or Administrative Service) and, in practice, answerable to local politicians who hold sway over promotions or transfers.

<sup>13</sup> Jury trials existed since the Raj until roughly the mid-twentieth century. See 1973 Code of Criminal Procedure.

During registration, police officers stamp Penal Codes to cases to signal the probable laws violated. VAW Penal Codes (and related "acts") include Section 326-A (acid throwing), Section 498-A (dowry harassment),<sup>14</sup> Section 376 (rape),<sup>15</sup> and others.<sup>16</sup> Politicians have argued that women exaggerate when filing such cases, even noting, "Many families are destroyed or ruined under such [gendered] provisions, and the legal proceedings go on for years. Men's rights organizations are working to raise awareness ... in opposition to women ... men should be arrested after proper inquiries rather than on the basis of the woman's complaint" (Verma 2017).

These sentiments are not restricted to politicians. (All-male) benches of the Supreme Court have ruled that domestic violence provisions are, "a license for unscrupulous persons to wreck personal vendetta or unleash harassment [against men]," and a form of "legal terrorism [by women]."<sup>17</sup> The court has noted,

<sup>14</sup> In 1983, "cruelty" by a husband (or in-laws) against a wife was made a crime. Some criticized the law since it only covers married women (Kothari 2005). It was followed with Section 304-B or "dowry death" that outlaws dowry violence culminating in murder. In 2005, the Protection of Women from Domestic Violence Act expanded domestic violence's definition, but also prioritized "counseling" abused women. Agnes and D'Mello (2015, 80) argue that, "... counseling is based on a patriarchal premise and is laden with anti-women biases... [women are] advised to 'save the marriage' even at the cost of danger to her life."

<sup>15</sup> See Supplementary Table A1 for a list. I classify all official "gendered" sections as VAW, including Section 497 (adultery).

<sup>16</sup> There are implicit distinctions between "heinous" and "non-heinous" violations. Non-heinous cases include "compoundable" sections where police are not forced to take action if the victim settles. VAW cases such as Section 497 (adultery) or Section 312 (causing miscarriage) are compoundable. Bailable, compoundable, and non-cognizable laws are considered the least serious. Section 320 of the Code of Criminal Procedure.

<sup>17</sup> *Sushil Kumar Sharma v. Union of India*, No. 141, 2005.

“... complaints under Section 498-A are filed in the heat of the moment over trivial issues without proper deliberations. The learned members of the Bar have enormous social responsibility and obligation to ensure that the social fiber of family life is not ruined or demolished,”<sup>18</sup> and that women should be deterred from filing cases to, “satisfy the ego and anger of the complainant.”<sup>19</sup> These statements imply that cases of VAW are (a) frivolous, (b) reported without delay, (c) submitted by those with an agenda, or (d) best resolved through reconciliation (Basu 2012; Jassal 2021). I scrutinize these assumptions using two sources of data.

## THE FIRST-INFORMATION-REPORT DATASET AND JUDICIAL RECORDS

In a push for transparency, India made police reports, called First-Information-Reports (FIRs), accessible to citizens. Over several years, I collected and parsed millions of these records; this study utilizes 418,190 police files in Haryana from January 2015 to November 2018.<sup>20</sup> I focus on this state for which I translated reports into English, and was able to collaborate with the local police to gather information about officers and otherwise withheld cases.<sup>21</sup> Aside from details about victims and suspects, FIRs contain first-person descriptions of crime, less affected by social desirability.<sup>22</sup> I then merged FIRs with judicial records. India has made (semi-)public the universe of these files on a platform called E-Courts, similar to a domain established by China (Liebman et al. 2020). It contains particulars about the date of filing/first appearance in court for a police report, judges assigned, verdict (if any), and other details. With assistance from the Development Data Lab<sup>23</sup>—that compiled and released 77 million records from 2010—I connected the two databases via particulars of the station, complainant name,<sup>24</sup> and other identifiers. Out of 418,190 crime reports, I

merged precisely 251,804 or 60.2% to court files, a figure that appears to accurately reflect those police files that were ultimately sent to the judiciary.<sup>25</sup>

## Research Design

I chart the *process* and *outcomes* associated with each level of accessing formal justice (Figure 2). In level 1 (registration), I examine the duration of time to file a crime report. Crime records include when the complaint was filed, as well as the dates that a complainant told an officer that the crime *began* or *ended*. *Registration duration* reflects the difference (in days) between police filing and incident. For outcome, I examine the likelihood of a registered case being sent to court; non-merged cases are categorized as canceled, indicating that the police files were not present in the judiciary. For levels 2–3 (investigation and preliminary hearing), I create two measures. *Investigation duration*—the difference (in days) between police filing and the case’s first appearance in court—signifies how long the police inquiries took. *Dismissal* refers to whether the case was ejected by a judge at the initial hearing. At the final stage (level 3), I create a numeric variable corresponding to the number of days from the preliminary hearing until the most recent one on file (*duration in court*). I focus on two indicators for judicial review, that is, whether the verdict issued by a judge resulted in a suspect’s *conviction* and *acquittal*.

Broadly, I am interested not only in the primary question as to whether women are disadvantaged while accessing justice compared to men (including for generic or non-VAW cases), but also whether male complainants are less likely to face burdens and exclusion when registering cases of VAW on behalf of women. In the equation below,  $Y_i$  is a binary or numeric outcome for crime report  $i$  filed in police station  $s$  at time  $t$ :

$$Y_i = \alpha + \beta_1 \text{Female}_i + \beta_2 \text{VAW}_i + \beta_3 (\text{Female} \cdot \text{VAW})_i + X_i \gamma + \zeta_s + v_t + \epsilon_i \quad (1)$$

$\text{Female}_i$  indicates whether the case was filed by a woman, and  $\text{VAW}_i$  is a binary variable for any Penal Code classified a “crime against women” being affixed to the report. Since VAW may be registered not only by women but also by male family/friends, the interaction term allows for differences between male and female complainants for VAW and non-VAW crime.  $X_i$  represents covariates, for example, distance in kilometers that the crime took place from the station, rank of the officer that investigated the case, and, ultimately, rank of the judge that heard it in court. I also include fixed effects for the station that the case was registered in ( $\zeta_s$ )

<sup>18</sup> *Preeti Gupta & Anr. v. State of Jharkhand*, Appeal No. 1512, Criminal Appellate Jurisdiction, 2010.

<sup>19</sup> *Rajesh Sharma v. State of Uttar Pradesh*, Appeal No. 1265, Criminal Appellate Jurisdiction, 2017.

<sup>20</sup> Records on government databases raise ethical questions that relate to privacy debates in India (Bhatia 2018). If citizens are aware that the state is storing their details, may this preclude individuals from coming forward in the future? If police reports are accessible to citizens, may they come to have less evidentiary value in court (Dam 2016)? The government believes privacy concerns are outweighed by lowering opportunities for police graft, disincentivizing officers from “burking” or ignoring registration, giving suspects details about accusations, and generating transparency in a historically opaque institution (Kumar 2017; Raghavan and Sivanandhan 2016). We anonymize data and remove personally identifiable details for officers, citizens, and suspects to uphold privacy and prevent it from being used to track any individual or organization.

<sup>21</sup> The police are exempt from releasing “sensitive” cases involving certain forms of VAW and terrorism.

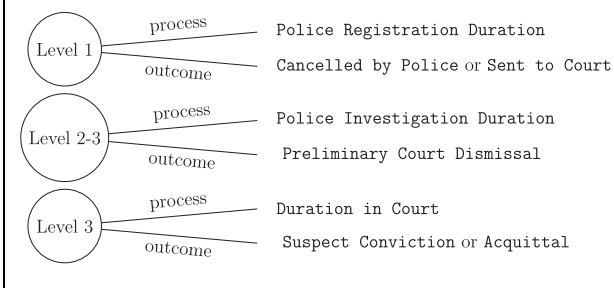
<sup>22</sup> Citizens would have had to provide as much detail to officers to initiate investigation.

<sup>23</sup> <https://www.devdata.org/>.

<sup>24</sup> Gender was classified based on 150,359 unique names that were manually coded. For a few ambiguous names, the testimonies (e.g.,

grammar or, say, reference to oneself as a housewife) made classification straightforward.

<sup>25</sup> As a validation exercise, I show that a third of cases of VAW could *not* be matched to court, reinforcing research based on internal police memos demonstrating  $\approx 30\%$  of such crime as canceled pre-Court (Jassal 2020).

**FIGURE 2. Process and Outcome Measures of Accessing Justice in India**

and the month–year of police filing ( $v_t$ ). In the Supplementary Material, when excluding  $VAW_i$ , I include primary Penal Code fixed effects.<sup>26</sup> Standard errors are clustered at the district-level.

I also estimate structural topic models (STMs) that, in a regression-type framework, can predict whether cases devoted to a topic (e.g., rape) are functions of covariates (Roberts, Stewart, and Airoidi 2016; Roberts, Stewart, and Tingley 2019; Roberts et al. 2014).<sup>27</sup> The method de-emphasizes Penal Codes by textually disaggregating what citizens told the police happened to them using statistical associations between words to separate, say, domestic violence from domestic violence that also happened to involve attempted murder (e.g., an inexperienced officer could have incorrectly classified a case or have forgotten to append a relevant Penal Code). I compiled and parsed text from each police report into a machine-readable format, and translated the 418,190 cases (two hundred million words or  $\approx$  450,000 A4-size single pages) using Google Translate.<sup>28</sup>

When comparing men and women's attempts at accessing justice, there may be concerns about omitted variable bias and distinct underlying case distributions (Arnold, Dobbie, and Yang 2018), for example, even *within* categories of crime (theft), women may report distinct sub-types (bag-snatching) compared to men (motorcycle robbery). Consequently, I utilize a *third* method: topical inverse regression matching (TIRM), introduced by Roberts, Stewart, and Nielsen (2020), which allows for the matching of complaints based on text, thereby adjusting for confounding. To implement

TIRM, I estimate a STM with an indicator for a woman's crime report as a content covariate. This estimates the relationship between having a female complainant and words in the corpus, as well as how crime reports registered by women discuss topics differently (Roberts, Stewart, and Airoidi 2016). Following the procedure outlined in Roberts, Stewart, and Nielsen (2020), I create indicators for whether the police report ultimately resulted in a suspect's conviction and acquittal in court. Put differently, I test if *final* stage outcomes are distinct for men and women who brought forward topically similar complaints at the very *first* stage. In this setting, the method likely serves as the closest one may attain in terms of having a valid counterfactual, short of randomizing citizens to be victims of crime or report at police stations.

## DESCRIPTIVE STATISTICS

Women registered 38,828 or 9% of all FIRs. Supplementary Figure A2 displays the top Penal Codes in women's cases, and Supplementary Figure A1 highlights those for men. For male complainants, the top substantive<sup>29</sup> Penal Codes relate to theft, rash driving, burglary, and public intoxication/bootlegging. For women, Section 498-A or domestic violence/dowry-related abuse perpetrated by a spouse (or in-laws) was present in 15% of their registrations.<sup>30</sup> Other common VAW Penal Codes include abduction (e.g., kidnapping a woman "to compel her into marriage"),<sup>31</sup> "obscene acts/songs,"<sup>32</sup> "criminal force against a woman,"<sup>33</sup> rape, "insulting the modesty of a woman,"<sup>34</sup> stalking, "intent to disrobe," sexual harassment, and "unnatural" (anal) sex.<sup>35</sup>

Table 1, which depicts the first quantification of Indian cases, presents summary statistics. *Distance* reveals that crime takes place, on average, 5.5 kilometers from a station. When identified by the complainant, cases are likely to have two suspects. Crimes registered by women, and VAW, are more likely to have a female suspect (*female suspects*). (Dowry may involve a mother-in-law.) While officers do not always record victim ages, non-missing data suggest that

<sup>26</sup> Most FIRs are combinations of multiple Penal Code clauses, with the first listed often (but not always) indicating the case type or its severity. There are approximately one thousand unique Penal Codes and special acts, and even greater combinations of cases/charges.

<sup>27</sup> For most analyses, I specify 20–40 topics. As seen in Supplementary Figures A1 and A2, most crimes can be slotted into roughly two-dozen Penal Code classifications. I see more repeat topics for values greater than 40.

<sup>28</sup> I analyze translations to ease pre-processing, that is, stemming, lemmatization, and the ejection of stop-words. While a majority of crime reports are in Hindi, a subset are in English or Punjabi which Google Translate standardizes. Vries, Schoonvelde, and Schumacher (2018) show that Google Translate is an excellent tool for comparative bag-of-word text models. The documents were translated on February 6, 2021.

<sup>29</sup> Most codes relate to concrete violations, for example, theft and murder. However, there are clauses that are often attached as supplements, for example, Section 323 (causing hurt) can be appended to rash driving, extortion, and so forth.

<sup>30</sup> Many Penal Codes are registered in conjunction with Section 498-A, for example, "unnatural"/anal sex (for marital rape), or dowry death (when domestic violence culminates in suicide or murder).

<sup>31</sup> Invoked in cases ranging from abductions to young women eloping or running away with boyfriends.

<sup>32</sup> Invoked in cases that may include lewd behavior in front of, or toward a woman, as well as "obscenity."

<sup>33</sup> Invoked in cases ranging from acting aggressively to attempted rape.

<sup>34</sup> Invoked in a range of cases, including exhibitionism and invasion of privacy.

<sup>35</sup> Invoked in cases of sodomy or, potentially, marital rape; this clause was "read-down" in 2018 as being discriminatory toward the gay community.



**TABLE 1. Descriptive Statistics on Select Variables: The First-Information-Report (FIR) Dataset**

	Complainant	N	Mean	SD	Crime type	N	Mean	SD	All Crime			
									N	Mean	SD	Median
Testimony word count	Female	38,828	577.41	421.49	VAW	20,869	722.44	526.30	418,189	452.30	257.86	381.00
	Other	379,361	439.49	230.98	Non-VAW	397,320	438.11	226.72				
Distance	Female	36,868	5.96	12.45	VAW	19,585	6.96	14.01	400,345	5.52	13.83	3.00
	Other	363,477	5.48	13.96	Non-VAW	380,760	5.45	13.82				
Female suspects	Female	22,022	0.70	1.07	VAW	17,676	0.75	1.08	220,943	0.18	0.72	0.00
	Other	198,921	0.12	0.65	Non-VAW	203,267	0.13	0.66				
Total suspects	Female	22,022	2.74	2.35	VAW	17,676	2.78	2.26	220,943	2.00	2.42	1.00
	Other	198,921	1.92	2.41	Non-VAW	203,267	1.93	2.42				
No. of sections	Female	38,828	2.57	1.59	VAW	20,869	3.31	1.65	418,190	2.11	1.42	2.00
	Other	379,362	2.07	1.39	Non-VAW	397,321	2.05	1.37				
Urban	Female	38,141	0.60	0.49	VAW	20,028	0.53	0.50	417,322	0.59	0.49	1.00
	Other	379,181	0.59	0.49	Non-VAW	397,294	0.59	0.49				
Registration duration	Female	33,766	68.94	341.50	VAW	17,269	112.88	440.69	381,836	27.78	225.87	1.00
	Other	348,070	23.79	210.88	Non-VAW	364,567	23.75	209.46				
Hours waited at PS	Female	38,690	7.51	62.08	VAW	20,775	9.32	79.25	416,045	7.06	52.72	0.68
	Other	377,355	7.01	51.67	Non-VAW	395,270	6.94	50.95				
Victim age	Female	17,953	35.93	10.39	VAW	9,131	34.35	10.29	192,939	38.40	9.07	38.00
	Other	174,986	38.65	8.88	Non-VAW	183,808	38.60	8.96				
Hour registered	Female	38,828	17.37	4.80	VAW	20,869	17.00	5.07	418,190	17.20	5.32	19.00
	Other	379,362	17.19	5.38	Non-VAW	397,321	17.21	5.34				
Hour arrived	Female	38,828	16.50	4.83	VAW	20,869	16.14	5.11	418,190	16.35	5.43	18.00
	Other	379,362	16.34	5.48	Non-VAW	397,321	16.36	5.44				
R:Head constable	Female	36,959	0.29	0.46	VAW	19,621	0.16	0.36	400,086	0.43	0.49	0.00
	Other	363,127	0.44	0.50	Non-VAW	380,465	0.44	0.50				
R:Assistant sub-inspector	Female	36,959	0.52	0.50	VAW	19,621	0.58	0.49	400,086	0.44	0.50	0.00
	Other	363,127	0.43	0.50	Non-VAW	380,465	0.44	0.50				
R:Sub-inspector	Female	36,959	0.16	0.36	VAW	19,621	0.22	0.41	400,086	0.10	0.30	0.00
	Other	363,127	0.10	0.30	Non-VAW	380,465	0.10	0.30				
R:Inspector	Female	36,959	0.03	0.17	VAW	19,621	0.04	0.21	400,086	0.02	0.16	0.00
	Other	363,127	0.02	0.15	Non-VAW	380,465	0.02	0.15				
No record/not sent to court	Female	38,828	0.42	0.49	VAW	20,869	0.32	0.47	418,190	0.40	0.49	0.00
	Other	379,362	0.40	0.49	Non-VAW	397,321	0.40	0.49				

*Note:* Descriptive statistics for variables in the FIR dataset, split by female/other complainants, as well as VAW/non-VAW crime. The term ‘Other’ is used because a small fraction of cases may be brought forward by organizations or institutions rather than individuals. VAW crime may be brought forward by male or female complainants. Variables prefixed with ‘R:’ represent investigator ranks.



complainants are, on average, in their 30s. VAW is likely to have more Penal Codes (*no. of sections*). Strikingly, victims of VAW wait longer at the station in anticipation of registration (9 vs. 7 hours).<sup>36</sup> Women's cases are infrequently assigned to junior officers, for example, constables.

The median days between crime and registration (*registration duration*) is 1, with a mean of 28. Women's cases, and VAW, have means of 69 and 113, respectively. A complainant may have visited a station to register a case but asked to drop it, or be forced to return at a later date.<sup>37</sup> *Prima facie*, this challenges the assumption that women's cases are filed, "in the heat of the moment."<sup>38</sup> *No record* shows that women's cases are 2 percentage points more likely to be canceled at the police-level.<sup>39</sup> Table 2 highlights variables post-merging with court files. *Investigation duration* shows that a police investigation takes 128 days, on average. Cases spend about 336 days in the judiciary (*duration in court*); yet, women's cases, and VAW, spend even longer awaiting a decision. The variable *duration in court* is noteworthy when juxtaposed with the comparable number of court meetings for male and female complainants (*no. of hearings*). While most cases are assigned to Judicial Magistrate 1st Class, women's cases, and VAW, are more likely to be assigned to senior judges, for example, Additional District Sessions Judge.<sup>40</sup>

Figure 3 illustrates judicial outcomes, which fall into roughly seven categories in the E-Courts database. *Acquitted* refers to whether the suspect is absolved; *allowed* denotes if the case is admitted but a trial has not been set; *convicted* represents suspect conviction, whereas *dismissed* underscores a judge's ejection of the case, typically at a bail hearing.<sup>41</sup> The basic cross-tabulations in Figure 3 show that—whether a function of all FIRs (panels a and b) or simply those in the court docket (panels c and d)—women's complaints (and

VAW) are more likely to be on-going (stalled), dismissed, or result in a suspect's acquittal, and less likely to see a suspect sent to prison. For women's cases, only 2.9% of the implicated suspects are convicted, unlike 10.8% for men's cases. For cases that ultimately make their way to court, the gap between female and male complainants' cases is even wider (5% and 17.9%, respectively).

## OLS RESULTS

While there appear to be gender gaps along a variety of measures, women are also more likely to register cases involving VAW which—for political, economic, and cultural reasons—could be treated specially by the criminal justice system. And so, I examine how women fare in each stage of justice delivery, taking into account this confounding factor.

Columns 1 and 2 of Table 3 show that women's cases have, on average, a lag of over a month longer than men between incident and registration, suggesting significant delays between crime occurrence and when the state takes cognizance of the case. Columns 3–6 show that cases of VAW account for a substantial portion of this gap. While this may be reflective of women's hesitancy in coming forward to complain of such crime, it is important to note that at the stage prior to cases formally entering the books, law enforcement has discretion in asking citizens to return at a later date or forwarding complainants of VAW to counseling and mediation centers. Still, controlling for VAW, columns 5 and 6 reveal significant discrepancies between men and women for generic cases; for non-VAW, the number of days from when a crime occurs and police officers initiating the case for investigation is roughly a week longer for women than men.

In terms of outcome in level 1, columns 7 and 8 of Table 3 reveal that women's cases are significantly less likely than men's to be sent to court. However, conditional on registration, cases of VAW are likely to transition to the next wing compared to non-VAW (columns 9 and 10), especially when a woman is the complainant (columns 11 and 12). Cases of VAW are approximately 7–8 points more likely to be sent to the judiciary than non-VAW crime. To reiterate, police officers are bound by rules to ensure (registered) cases of VAW transition or are investigated quickly. The gap between men and women in terms of a case being canceled by law enforcement (after registration) is noteworthy for non-VAW, that is, those complaints for which officers have discretion in influencing outcomes. This dynamic is reflected in terms of police investigations too. Specifically, columns 3 and 4 of Table 4 show that cases of VAW are investigated and sent to court, on average, sooner than non-VAW crime. And, while women's cases are investigated slower by police officers than men's complaints (column 2 of Table 4), this is especially true when the cases are restricted to non-VAW where women's complaints have investigative delays of  $\approx 19$  days (columns 5 and 6).

<sup>36</sup> The establishment of all-women police stations (AWPSs) posts the limited staff of female officers to segregated units (Jassal 2020). This may clash with rules mandating female officers be present for recording women's testimonies, forcing policewomen to be then called-in from other stations, inadvertently increasing wait-times.

<sup>37</sup> See Supplementary Figure A4. The Supplementary Material also provides details on *pre-registration duration*, which reflects the difference between registration and when a crime first *began* or started (e.g., domestic violence may have begun at the start of marriage) rather than when the last incident related to the crime took place.

<sup>38</sup> *Preeti Gupta & Anr. v. State of Jharkhand*, Appeal No. 1512, Criminal Appellate Jurisdiction, 2010.

<sup>39</sup> While it is possible some cases have not had time to move to the judiciary, the FIRs cover 2015–18. Investigations are supposed to be at most 90 days, and the E-Court's database was downloaded in 2020. The analyses thus "allow" a 2-year window for FIRs to appear in court, that is, far longer than the time allotted for investigation.

<sup>40</sup> Cases can have multiple successive judges; if judges are transferred, cases are overtaken by successors of identical rank.

<sup>41</sup> *Untraced* represents whether the suspect could not be brought to court. The remaining outcomes are classified as *disposed*, indicating that a decision was taken (e.g., fine) but further details are unavailable.

**TABLE 2. Descriptive Statistics: The First-Information-Report (FIR) Dataset Merged with Court Records**

	Complainant	N	Mean	SD	Crime type	N	Mean	SD	All Crime			
									N	Mean	SD	Median
Investigation duration	Female	22,500	133.93	206.49	VAW	14,039	113.73	185.78	249,930	127.77	204.13	55.00
	Other	227,430	127.16	203.88	Non-VAW	235,891	128.60	205.14				
Dismissed	Female	22,648	0.07	0.26	VAW	14,134	0.09	0.28	251,804	0.05	0.21	0.00
	Other	229,156	0.04	0.20	Non-VAW	237,670	0.04	0.20				
Ongoing	Female	22,648	0.44	0.50	VAW	14,134	0.43	0.50	251,804	0.38	0.49	0.00
	Other	229,156	0.38	0.48	Non-VAW	237,670	0.38	0.49				
Acquitted	Female	22,648	0.23	0.42	VAW	14,134	0.26	0.44	251,804	0.17	0.38	0.00
	Other	229,156	0.17	0.37	Non-VAW	237,670	0.17	0.37				
Convicted	Female	22,648	0.05	0.22	VAW	14,134	0.04	0.21	251,804	0.17	0.37	0.00
	Other	229,156	0.18	0.38	Non-VAW	237,670	0.17	0.38				
Duration in court	Female	22,522	377.37	368.07	VAW	14,120	378.43	362.50	250,287	336.18	365.50	205.00
	Other	227,765	332.10	364.99	Non-VAW	236,167	333.65	365.52				
No. of hearings	Female	20,077	9.82	9.04	VAW	12,852	10.41	9.49	195,480	9.84	9.15	7.00
	Other	175,403	9.84	9.17	Non-VAW	182,628	9.80	9.13				
R:Civil judge junior division	Female	22,634	0.06	0.25	VAW	14,124	0.06	0.24	251,629	0.07	0.25	0.00
	Other	228,995	0.07	0.25	Non-VAW	237,505	0.07	0.25				
R:Judicial magistrate 1st class	Female	22,634	0.43	0.50	VAW	14,124	0.39	0.49	251,629	0.46	0.50	0.00
	Other	228,995	0.47	0.50	Non-VAW	237,505	0.47	0.50				
R:Sub-divis. judicial magistrate	Female	22,634	0.08	0.27	VAW	14,124	0.07	0.26	251,629	0.09	0.29	0.00
	Other	228,995	0.09	0.29	Non-VAW	237,505	0.09	0.29				
R:Addl. chief judicial magistrate	Female	22,634	0.09	0.29	VAW	14,124	0.08	0.26	251,629	0.11	0.31	0.00
	Other	228,995	0.11	0.31	Non-VAW	237,505	0.11	0.31				
R:Chief judicial magistrate	Female	22,634	0.13	0.33	VAW	14,124	0.09	0.29	251,629	0.14	0.35	0.00
	Other	228,995	0.14	0.35	Non-VAW	237,505	0.14	0.35				
R:Addl. district sessions judge	Female	22,634	0.17	0.37	VAW	14,124	0.29	0.45	251,629	0.11	0.31	0.00
	Other	228,995	0.10	0.30	Non-VAW	237,505	0.10	0.30				
R:District sessions judge	Female	22,634	0.03	0.16	VAW	14,124	0.01	0.10	251,629	0.02	0.14	0.00
	Other	228,995	0.02	0.13	Non-VAW	237,505	0.02	0.14				
Duration in CJ system	Female	22,495	573.44	383.63	VAW	14,113	568.99	381.93	249,847	508.25	392.43	435.04
	Other	227,352	501.80	392.70	Non-VAW	235,734	504.62	392.75				

*Note:* Descriptives statistics for select variables in merged dataset of crime and judicial records, split by female/other complainants, as well as VAW/non-VAW crime. The term 'Other' is used because a small fraction of cases may be brought forward by organizations or institutions rather than individuals. VAW crime may be brought forward by male or female complainants. Variables prefixed with 'R:' represent judge ranks.

**FIGURE 3. Crime Report Statuses [Split by Complainant Gender and Crime Type]**

Note: Judicial outcomes for cases (% on Y-axis). Panels a and b reflect outcomes conditional on police registration, that is, including “no record” cases or police files that did not make their way to court. Panel a is separated by female ( $N=38,828$ ) and male/other complainants ( $N=379,362$ ). Panel b reflects VAW ( $N=20,869$ ) and non-VAW crime ( $N=397,321$ ). Panels c and d reflect outcomes as a function of cases just in the court docket. Panel c is separated by female ( $N=22,648$ ) and male/other complainants ( $N=229,156$ ), and Panel d by VAW ( $N=14,134$ ) and non-VAW crime ( $N=237,670$ ). 95% confidence intervals included.

Figure 4 presents marginal effects in a descriptive plot. Panel a shows that cases of VAW (registered by women) have the longest lag between incident and filing. However, having crossed the node of filing, VAW is generally allowed to pass through level 1 (panels b and c).

Nonetheless, discrimination appears more consistent—whether in contexts of VAW or not (columns 7–12 of Table 4 and Figure 4d)—when cases selected to leave the police jurisdiction enter the court for the first time. Specifically, even though women’s non-VAW cases are between 1 and 3 points more likely to be dismissed at a preliminary court hearing than cases brought by men, a gap persists for VAW. Figure 4d–f suggests that not only are women discriminated against across crime type, but also male complainants who register cases *on behalf* of female friends or relatives are less likely to face burdens or exclusion than if a woman was listed as the primary complainant. Columns 1 and 2 in Table 5 reveal that women’s cases spend longer in the judiciary by over a month. Graphically, Figure 4e shows that cases involving VAW registered *by* women spend significantly long stalled. To probe whether punitive justice was meted out or that the individual who sought help in the original report received a favorable ruling, I pay attention to the suspect’s conviction or acquittal who allegedly wronged the complainant. Columns 1, 2, 5, and 6 of Table 6 demonstrate that cases brought

forward by women are significantly more likely to yield a suspect’s acquittal. Women’s cases are associated with more than a 10-point reduction in convictions (columns 7, 8, 11, and 12). Figure 4f shows conviction for suspects that male complainants accuse of VAW (e.g., for female family/friends) decline from a non-VAW base, but not to the same level as when women file cases themselves. Indeed, women complainants seeking justice from the state have a lower chance of a suspect that wronged them being sent to prison for either type of complaint, VAW or not.

While this paper does not delve into *drivers* of discrimination, the findings are suggestive. On the one hand, administrators may be acting rationally but with imperfect information. Administrators could hold inaccurate beliefs about women’s tendency to exaggerate. Or, female complainants may be statistically less likely to afford lawyers or cope with in-person follow-ups necessary for trial, thereby precluding judges from making informed decisions. (In India, complainants have to be present in court at least twice, e.g., once during a magistrate’s “cognizance” of a case, and again during cross-examination.) Describing crime in open court can be difficult, especially for victims of VAW. Further, traveling long distances repeatedly to a district court for multiple hearings may pose distinct challenges than simply filing a one-time police report at a neighborhood police station. Low levels of development can

**TABLE 3. Police Process and Outcome for Female Complainants: Level 1**

	Registration duration						Canceled after registration					
	1	2	3	4	5	6)	7	8	9	10	11	12
Female	45.148*** (6.353)	40.670*** (7.241)			9.100*** (2.263)	6.866*** (2.491)	0.021*** (0.008)	0.024*** (0.008)			0.068*** (0.010)	0.063*** (0.009)
VAW			89.135*** (13.777)	85.411*** (16.035)	27.459*** (7.740)	24.344*** (9.078)			-0.079*** (0.014)	-0.065*** (0.011)	-0.046*** (0.017)	-0.034** (0.014)
Female:VAW					111.601*** (17.815)	115.105*** (18.129)					-0.120*** (0.018)	-0.113*** (0.018)
Constant	23.788*** (2.385)	8.993*** (2.896)	23.749*** (2.344)	6.961** (2.993)	23.128*** (2.350)	5.196* (2.799)	0.396*** (0.018)	0.388*** (0.013)	0.402*** (0.018)	0.397*** (0.013)	0.397*** (0.018)	0.392*** (0.013)
No. of obs.	381,836	360,666	381,836	360,666	381,836	360,666	418,190	383,033	418,190	383,033	418,190	383,033
R <sup>2</sup>	0.003	0.016	0.007	0.020	0.010	0.023	0.0002	0.112	0.001	0.112	0.003	0.113
Adj. R <sup>2</sup>	0.003	0.015	0.007	0.019	0.010	0.022	0.0001	0.111	0.001	0.111	0.003	0.112
Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
PS FE	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Month-year FE	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y

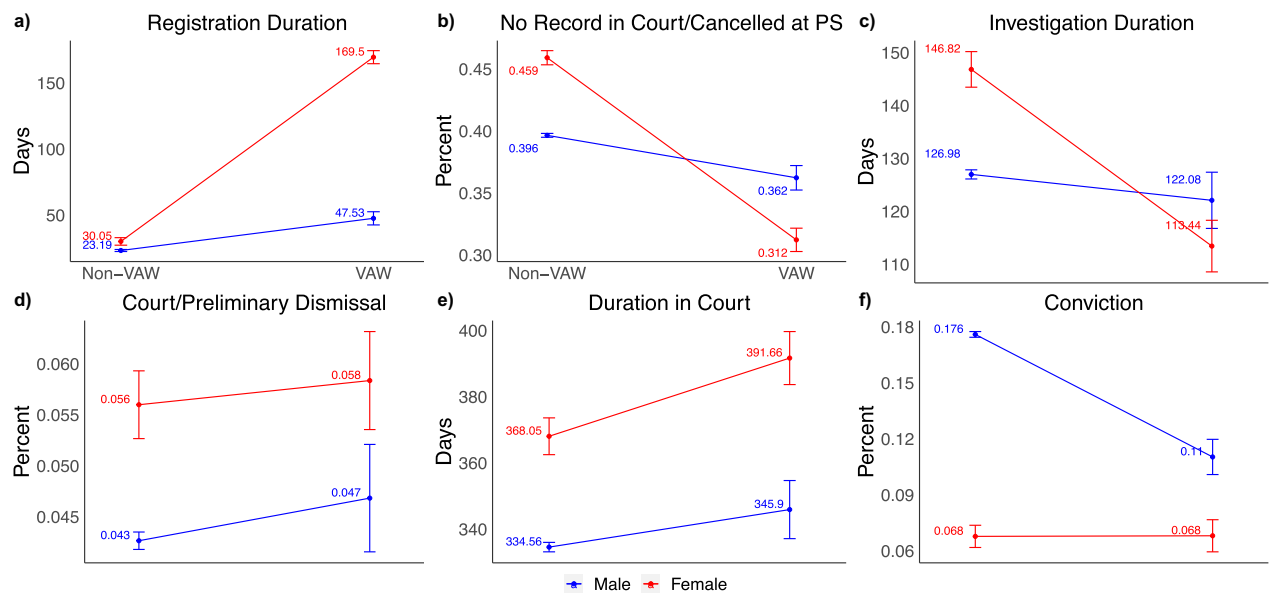
Note: Female indicates whether the complainant in the police report was a woman. Controls include a numeric variable for distance of crime from station and investigator rank. Standard errors clustered by district. For full model, see Section 7 of the Supplementary Material. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.



**TABLE 4. Police/Judicial Process and Outcome for Female Complainants: Level 2**

	Investigation duration						Court dismissal					
	1	2	3	4	5	6	7	8	9	10	11	12
Female	6.768 (4.885)	9.336*** (3.283)			19.250*** (4.203)	19.837*** (2.590)	0.029*** (0.003)	0.014*** (0.004)			0.024*** (0.003)	0.013*** (0.003)
VAW			-14.870** (6.956)	-11.066* (5.880)	-9.153 (7.132)	-4.898 (5.994)			0.043*** (0.006)	0.010 (0.006)	0.049*** (0.007)	0.004 (0.006)
Female:VAW					-27.336*** (6.213)	-28.485*** (6.544)					-0.032*** (0.006)	-0.002 (0.007)
Constant	127.158*** (5.957)	115.251*** (16.760)	128.602*** (5.953)	117.213*** (16.793)	127.407*** (6.042)	115.997*** (16.731)	0.043*** (0.003)	0.006 (0.008)	0.043*** (0.003)	0.007 (0.008)	0.041*** (0.003)	0.006 (0.008)
No. of obs.	249,930	228,823	249,930	228,823	249,930	228,823	251,804	230,555	251,804	230,555	251,804	230,555
R <sup>2</sup>	0.0001	0.070	0.0003	0.070	0.001	0.070	0.002	0.084	0.002	0.084	0.003	0.084
Adj. R <sup>2</sup>	0.0001	0.068	0.0003	0.068	0.001	0.069	0.002	0.083	0.002	0.083	0.003	0.083
Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
PS FE	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Month-year FE	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y

*Note:* Female indicates whether the complainant in the police report was a woman. Controls include a numeric variable for distance of crime from station, investigator rank, and judge rank. Standard errors clustered by district. For full model, see Section 7 of the Supplementary Material. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

**FIGURE 4. Marginal Effects for Male and Female Complainants across Stages (for VAW and Non-VAW)**

Note: Marginal effects based on regressions in columns 6 or 12 in Tables 3–6.

result in misgovernance without administrators behaving with repressive intent or harboring animus (Slough and Fariss 2021).

Then again, administrators may have a *taste* for discrimination or, say, eager to “protect” victims from the complex and public process of accessing formal justice (Bindler and Hjalmarsson 2020). The empirical findings suggest that statistical discrimination is unlikely to be the *only* mechanism, since it is plausible that the gaps would likely *decrease* across later stages as a function of the initial filtration. For example, Stolzenberg, D’Alessio, and Eitle (2013) note that racial discrimination in the U.S. justice system is less likely to manifest at final decision-points because initial processing decisions ultimately decrease variation, that is, cases in later stages become progressively similar vis-à-vis the validity of evidence. Here, gaps amplify, persisting in the final node of a judge’s verdict, a point at which plaintiffs do not have to be physically present or travel long distances. Relatedly, if police were dismissing complaints without animus by factoring in expected judicial rulings for what they perceived would be statistically hard-to-prove cases in court, then officers might presumably apply discretion to VAW.<sup>42</sup>

Instead, gaps at the police-level are stark for non-VAW crime, that is, cases for which officers have no prima facie reason to anticipate negative consequences in court to warrant differential treatment. And so, it is likely that taste-based discrimination operates in conjunction with other drivers of inequality.

In Supplementary Table A2, I re-run the analyses controlling for over a thousand Penal Codes/acts. I also analyze the outcomes for levels 2 and 3 as a function of all registrations (as opposed to just the court docket) to account for the selection that occurs when only 60% of police files are sent to the judiciary (Supplementary Tables A7 and A8 and Supplementary Figure A27). The coefficient on female remains significant in almost every model. A positive implication of the findings is that, conditional on registration, cases of VAW have better outcomes in in the realm of the police (e.g., plausibly as a function of the rules and guidelines that check officer behavior), but there is also suggestive evidence that, because of the multi-stage process, discrimination is delayed rather than mitigated. Discrimination does not appear to be restricted to one agency, but is iterative, and present in the mid- to late-stages of

<sup>42</sup> To explore variation by VAW, I focus on four sub-types (a) dowry harassment, (b) female kidnapping, (c) criminal force, and (d) rape. I select these codes because they are least likely to overlap. For example, dowry (Section 498-A) involves the spouse/in-laws, but this does not apply to rape (Section 376) which is stamped when a non-spouse commits assault. Female kidnappings (Section 366) are often registered by relatives rather than the victim. Of the four sub-types, rape is investigated the quickest (Supplementary Figure A9 and Supplementary Table A5). Female kidnapping cases are likely to be dismissed prior to entering court (Supplementary Table A4), and have significantly long investigative delays (Supplementary

Table A5). See Footnote 58 in Supplementary Material. Rape has high rates of suspect acquittal (Supplementary Figures A18 and A23 and Supplementary Table A6). Dowry harassment/marital violence is an exception by almost any measure: it has the longest delay (e.g., 300 days or more) between crime occurrence and registration, suggesting that abuse carried on for an extended period of time and/or police diverted or “counseled” complainants prior to filing (Supplementary Figure A5 and Supplementary Table A3). Dowry takes unusually long stalled in court (Supplementary Table A6 and Supplementary Figures A11 and A17), and is least likely to result in a suspect’s conviction compared to most other crimes (Supplementary Table A6).

**TABLE 5. Judicial Process for Female Complainants: Level 3**

	Duration in court					
	1	2	3	4	5	6
Female	45.264*** (11.042)	40.668*** (8.807)			27.042** (12.088)	33.481*** (10.230)
VAW			44.781*** (10.738)	33.912*** (10.873)	4.677 (10.194)	11.335 (10.942)
Female:VAW					47.649*** (8.335)	12.284 (10.663)
Constant	332.103*** (12.660)	549.451*** (32.856)	333.650*** (12.347)	551.414*** (32.793)	331.976*** (12.755)	548.655*** (32.761)
No. of obs.	250,287	229,143	250,287	229,143	250,287	229,143
$R^2$	0.001	0.201	0.001	0.200	0.002	0.201
Adj. $R^2$	0.001	0.200	0.001	0.199	0.002	0.200
Controls	N	Y	N	Y	N	Y
PS FE	N	Y	N	Y	N	Y
Month-Yr FE	N	Y	N	Y	N	Y

Note: Female indicates whether the complainant in the police report (or plaintiff in court) was a woman. Controls include a numeric variable for distance of crime from station, investigator rank, and judge rank. Standard errors clustered by district. For full model, see Section 7 of the Supplementary Material. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

justice delivery, “the last mile” at which complainants have not only spent time and resources to reach those stages but also decision-points at which there are few constraints on administrators from either “above” or “below.”<sup>43</sup>

### Structural Topic Modeling (STM)

There remain at least two methodological concerns with the analyses thus far. First, case categorizations have relied on Penal Codes. The assignment of codes is a strategic decision by officers who may, consciously or unconsciously, underweight a case’s seriousness by deciding which and how many to apply.<sup>44</sup> Second, even if one were to accept that there is an imbalance between men and women, perhaps women *are* more likely to register baseless cases, which the justice system happens to be efficiently weeding out. To explore the validity of this assumption, I utilize the unfiltered first-person testimony that citizens provide to law enforcement prior to case investigation, thereby setting aside the administrators’ classifications or any official coding schema.

I apply STMs to reports. Such modeling can estimate relationships between meta-data and topics from the

corpus and facilitate hypothesis testing (Roberts, Stewart, and Tingley 2019). Are there *topics* in the victims’ testimonies—including inside women’s complaints—that yield low convictions for suspects? To reiterate, the goals of STM are three-fold: (1) give voice to victims by utilizing their own words, (2) highlight the severity of claims, especially VAW, and (3) coarsen high-dimensional data to allow for text matching techniques.

I begin by summarizing crime in north India. In Supplementary Figure A29, there are crimes that are likely to be familiar to most readers including public intoxication/bootlegging or “alcohol” (Topic 19), “burglary” (Topic 16), “auto theft” (Topic 22/23), and “kidnapping” (Topic 27). However, there are South Asia-specific cases such as Topic 5 “cattle,” which represents the smuggling of cows or their slaughter, and Topic 6 “resource mafia” that signifies the sand or mining cartel (Asher and Novosad 2023). Figure 5 utilizes an indicator for suspect conviction as a predictor, and shows correlations (when topics likely co-occur). Across the full corpus, there are topics that appear more likely to result in suspect conviction, for example, fraudulent currency, gambling, and alcohol-related cases (Topics 2, 19, 28), compared to others like auto-theft, missing persons, and kidnapping (Topics 22, 7, 27), for which suspects infrequently go to prison or are even found. In Figure 5b, the machine estimated correlations where, for instance, Topics 10 and 13 (“accident” and “injury”) are unsurprisingly related, as are “cattle” and “minorities,” suggesting that members of the Muslim community are disproportionately victimized (or, more precisely, invoked in a police report) for alleged bovine-related offenses. More research is needed to explain the heterogeneity, and why, for instance, topics related to “arms” (Topic 18)

<sup>43</sup> While district judges may have greater leeway in influencing the trajectory of a case than law enforcement, like police officers, they are answerable to supervisors. Judges in district courts have a supervisor in the High Court. District judges may get “points” that influence promotions, for example, the number of cases settled by mediation. This is an informal and opaque system that warrants additional research.

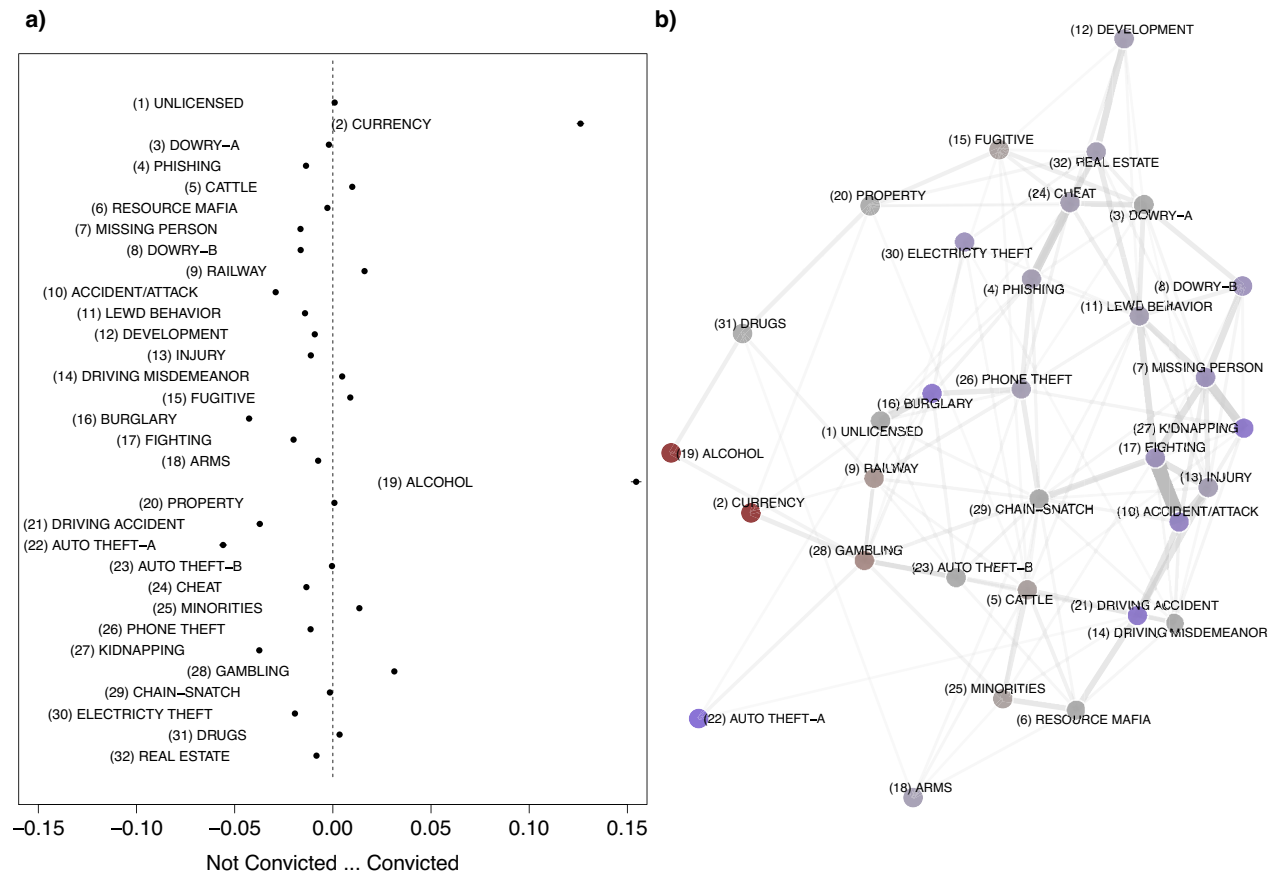
<sup>44</sup> For instance, a police officer may be inclined *not* to register the Penal Code for rape (Indian Penal Code, Section 376), despite being told of sexual assault by the complainant, if the officer fears added scrutiny from the media or by bureaucratic superiors.

**TABLE 6. Judicial Outcomes for Female Complainants: Level 3**

	Suspect acquittal						Suspect conviction					
	1	2	3	4	5	6	7	8	9	10	11	12
Female	0.064*** (0.009)	0.054*** (0.007)			0.056*** (0.008)	0.054*** (0.006)	-0.129*** (0.012)	-0.105*** (0.010)			-0.123*** (0.011)	-0.108*** (0.010)
VAW			0.088*** (0.012)	0.068*** (0.008)	0.100*** (0.010)	0.080*** (0.008)			-0.130*** (0.014)	-0.081*** (0.012)	-0.122*** (0.014)	-0.066*** (0.012)
Female:VAW					-0.071*** (0.010)	-0.067*** (0.009)					0.095*** (0.011)	0.066*** (0.011)
Constant	0.168*** (0.017)	0.394*** (0.018)	0.169*** (0.017)	0.395*** (0.018)	0.165*** (0.017)	0.391*** (0.018)	0.179*** (0.014)	0.266*** (0.028)	0.175*** (0.014)	0.261*** (0.028)	0.182*** (0.015)	0.268*** (0.028)
No. of obs.	251,804	230,555	251,804	230,555	251,804	230,555	251,804	230,555	251,804	230,555	251,804	230,555
R <sup>2</sup>	0.002	0.124	0.003	0.123	0.004	0.125	0.010	0.102	0.006	0.097	0.012	0.102
Adj. R <sup>2</sup>	0.002	0.122	0.003	0.122	0.004	0.123	0.010	0.100	0.006	0.096	0.012	0.101
Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
PS FE	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Month-year FE	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y

Note: Female indicates whether the complainant in the police report (or plaintiff in court) was a woman. The dependent variable refers to whether the suspect that was implicated in the crime was acquitted or convicted, respectively. Controls include a numeric variable for distance of crime from station, investigator rank, and judge rank. Standard errors clustered by district. For full model, see Section 7 of the Supplementary Material. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.



**FIGURE 5. Suspect Conviction and Correlation of Topics Associated with Full Crime Corpus**

*Note:* Panel a: Coefficients and standard errors for a structural topic model of all police complaints filed in Haryana with suspect conviction/non-conviction in court as the predictor. Right of the dashed vertical line represents positive coefficients. The stemmed words making up the topics appear in Supplementary Figure A29. Panel b: Figure depicts the network of correlated topics. Colors indicate the magnitude of the coefficient; red underscores positive coefficients and blue negative for the suspect conviction indicator. The gray widths of the edges are proportional to the strength of correlation between topics.

and organized crime (Topic 6) have better indicators on suspect acquittal (Supplementary Figure A31).

More importantly, the exercise sheds light on gradations of abuse that women face.<sup>45</sup> Figures 6 and 7 present the highest probability as well as FREX (frequent and exclusive) words for women's complaints and VAW, respectively. A top topic that emerges involves "fighting" (Topic 14), usually domestic violence. The word clouds in Supplementary Figures A37–A39 underscore terms such as: *wife, hospital, kill, beaten, domest, husband, hurt, and blunt*. Figure 7 breaks down VAW.<sup>46</sup> Topics range from the blackmail of women with compromising photographs/videos (Topic 18) to "trafficking" or being sold into prostitution (Topic 12). While certain topics such as sexual assault (by a non-spouse) or child abuse have better outcomes (vis-à-vis

suspect conviction) (Figure 9), a theme that emerges is the prioritization of sons over daughters. Specifically, in Figure 7, Topic 7 refers to abandoning or killing infants ("killing the girl child"), Topic 14 refers to (illegal) sex selective diagnostic technologies, and Topic 5 includes unlicensed doctors performing abortions. As highlighted in the word clouds of the Supplementary Material (Supplementary Figures A42 and A44), common words in these categories include: *children, child, medic, drug, abort, kill, patient, ultrasound, and pregnant*.

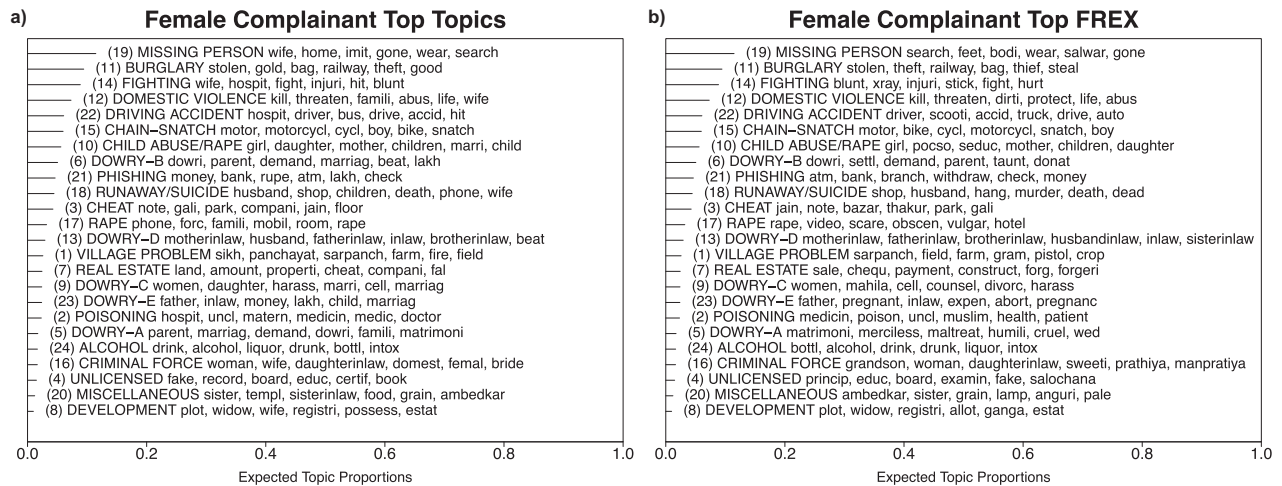
STM depicts how dowry underlies various forms of VAW. In both Figures 8 and 9, we see clusters of topics related to violence involving dowry-based harassment and domestic abuse. As the blue shading suggests, almost all these forms of crime are unlikely to see suspect conviction, for example, Topics 5, 6, 9, 13, and 23 in Figure 9.<sup>47</sup> Yet, based on the words that complainants use to describe this abuse, the cases do not appear to be frivolous. Supplementary Figures A37–

<sup>45</sup> Table 1 shows that complaints brought forward by women are longer. See Supplementary Figure A28.

<sup>46</sup> I use Penal Codes to subset VAW. STM is used here only to illustrate criminal activity in the region.

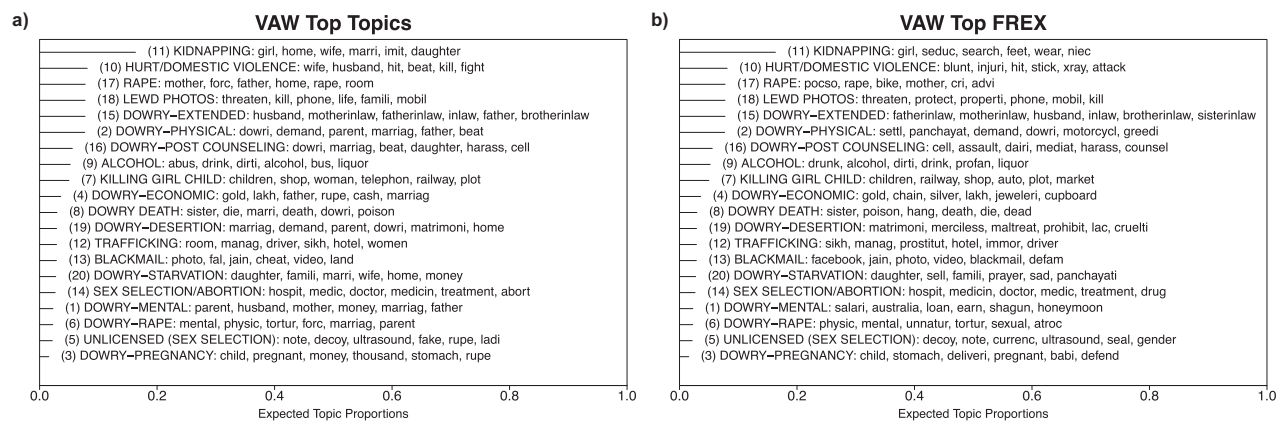
<sup>47</sup> Topics 5, 6, 9, 13, and 23 in Figure 8.

**FIGURE 6. Top Topics (Female Complainants, N=38,828)**



Note: Panel a: Top topics associated with women’s complaints and highest probability words in the topic. Panel b: FREX words (frequent and exclusive) or distinguishing words of the topics.

**FIGURE 7. Top Topics (VAW Crime, N=20,869)**



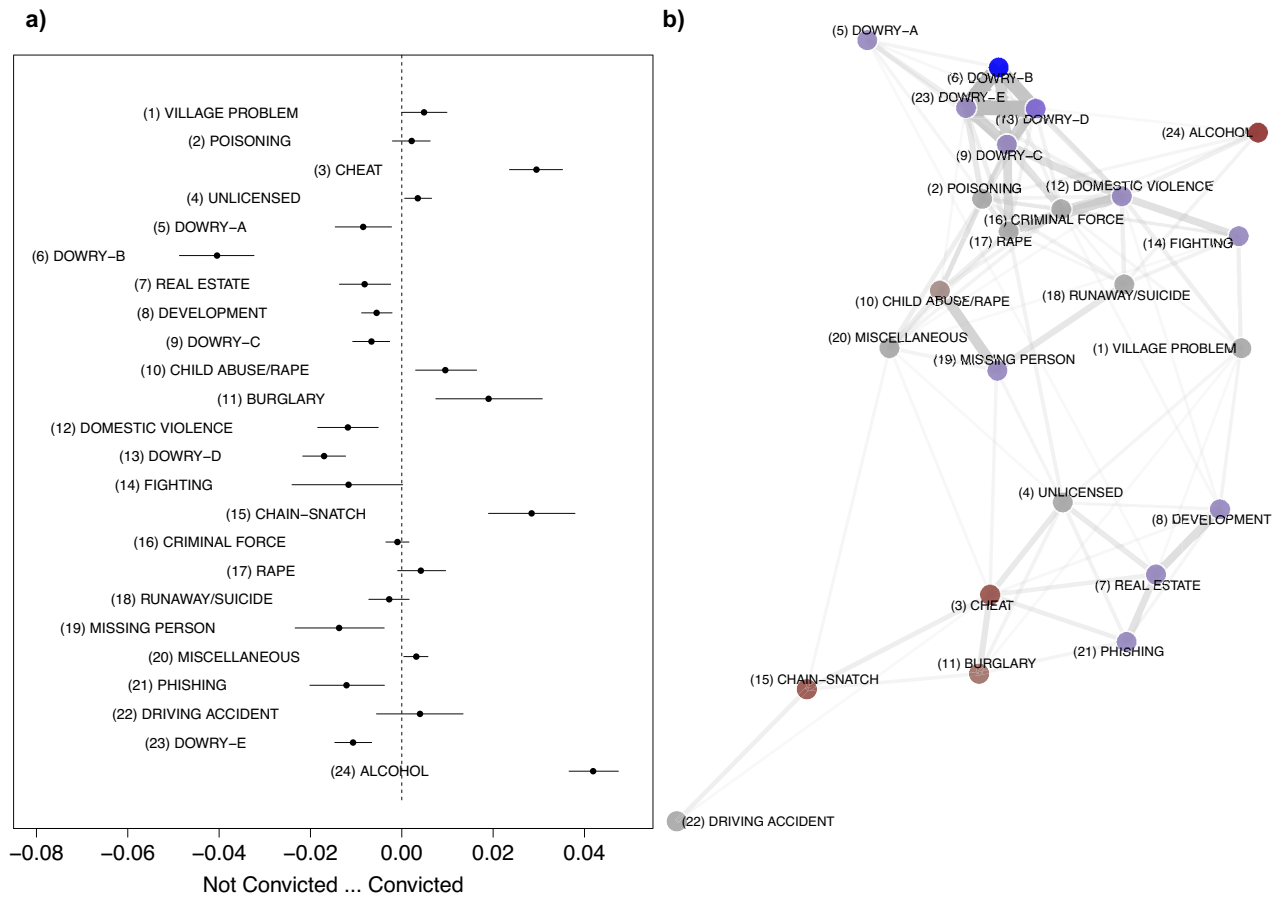
Note: Panel a: Top topics associated with cases of violence against women (VAW) and highest probability words in the topic. Panel b: FREX words (frequent and exclusive) or distinguishing words of the topics.

A39 show that common words include: *dowry, tortur, parent, cash, daughter, greed, kill, demand, cruelty, in-law, and assault*. “Mother-in-law” appears repeatedly, indicating that abuse often involves the extended family as opposed to just an intimate partner. The machine can separate dowry relating to mental or physical torment in Figure 7 (Topics 1 and 2) from others involving, for instance, harassment in conjunction with spousal rape (Topic 6). This is highlighted in the FREX words in Figure 7b that accentuate terms such as *unnatur* (i.e., “unnatural” or anal) and *sexual*.<sup>48</sup>

Topic 16 represents instances in which complainants explain that they *tried* to register a dowry or domestic violence case before but were asked to reconcile or participate in counseling with a spouse instead. This is seen in the FREX words *mediate* or *counsel* in panel b (Jassal and Barnhardt 2023). Topics 19 and 20 refer to abusers either deserting their wives or absconding, for example, possibly to extract dowry from another. The only type of dowry-related topic that *is* associated with higher levels of suspect conviction is Topic 8, that is, when harassment has culminated in suicide or the death

<sup>48</sup> Like the preceding analyses, STM reveals variation in VAW committed in and out of the household. Rape by a non-spouse has

better chances of suspect conviction (Topic 10), than spousal assault (Topic 6).

**FIGURE 8. Suspect Conviction and Correlation of Topics Associated with Women’s Cases**

Note: Panel a: Coefficients and standard errors for a structural topic model of police complaints filed by women with suspect conviction/non-conviction in court as the predictor. Right of the dashed vertical line represents positive coefficients. The stemmed words making up the topics appear in Figure 6. Panel b: Figure depicts the network of correlated topics. Colors indicate the magnitude of the coefficient; red underscores positive coefficients and blue negative for the suspect conviction indicator. The gray widths of the edges are proportional to the strength of correlation between topics. Gradations of VAW appear highly correlated with each other (top of panel b), while driving accidents/hit-and-runs have limited connections to other types of crime.

of a victim (equivalent to murder); this is evoked in the FREX words of Figure 7, for example, *dowri*, *die*, *marri*, *death*, *hang*, *poison*. The severity of cases which crossed the first stage of registration involving physical, mental, and emotional abuse that spouses (and in-laws) perpetrate, often to extort resources from victims’ natal homes, go against assertions that such complaints are family disputes unworthy of formal punishment by the state.<sup>49</sup>

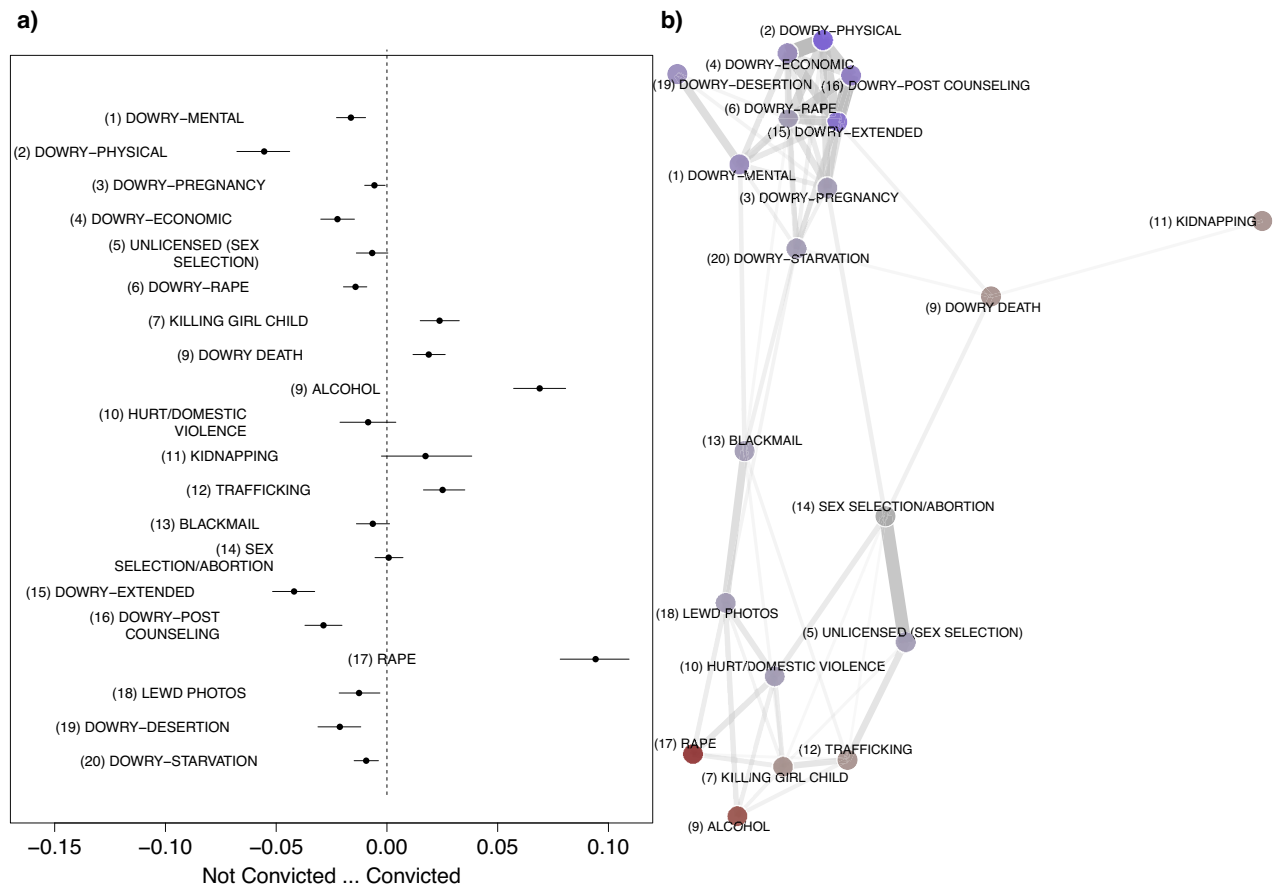
### Topical Inverse Regression Matching (TIRM)

STM underscores yet another challenge with interpreting the regression analyses and selection, for example, men and women’s cases could be distinct in a way that

controlling for Penal Codes or other variables across the system cannot account for. For women, a common form of theft is “chain-snatching” (Topic 15 in Figure 6), as opposed to vehicle robbery for men (Topic 22 in Supplementary Figure A29). While an officer would have simply classified both as “theft” (e.g., Section 379), one could assume that the criminal justice administration does not discriminate against women per se but merely takes stolen vehicles more seriously than burgled jewelry. One way to tackle this problem is via text matching utilizing all original police reports; then, after qualitatively ensuring that the technique was successful, compare final outcomes (Grimmer and Stewart 2013). Of course, even thinking about gender conceptually as a “treatment” is difficult; it is a non-manipulable “bundled” category encapsulating numerous factors (Holland 1986; Knox, Lowe, and Mummolo 2020). Nevertheless, I hold features of the testimony provided to law enforcement constant

<sup>49</sup> *Rajesh Sharma v. State of Uttar Pradesh*, Appeal No. 1265, Criminal Appellate Jurisdiction, 2017.

**FIGURE 9. Suspect Conviction and Correlation of Topics Associated with Violence against Women (VAW) Crime**



Note: Panel a: Coefficients and standard errors for a structural topic model of police complaints involving VAW (filed by women or male friends/family of victim) with suspect conviction/non-conviction in court as the predictor. Right of the dashed vertical line represents positive coefficients. The stemmed words making up the topics appear in Figure 7. Panel b: Figure depicts the network of correlated topics. Colors indicate the magnitude of the coefficient; red underscores positive coefficients and blue negative for the suspect conviction indicator. The gray widths of the edges are proportional to the strength of correlation between topics. Gradations of dowry/domestic violence cases appear highly correlated with each other. Sexual assault by a non-spouse is (relatively) more likely to lead to suspect conviction (Topic 17) than marital rape (Topic 6).

but for the complainant’s gender in a novel but imperfect attempt at underscoring a plausible causal link between complainant identity and justice.

I utilize registered complaints for topical inverse regression matching. Figure 10 is the first balance-test. The gray bars—which highlight the difference between female minus male complainants’ topics in the unmatched data—reveal differences by gender. Women are more likely to register dowry violence (Topic 24), whereas men cases of bootlegging and alcohol (Topic 4). Figure 10 shows that TIRM (black dots) is highly successful in minimizing these differences.

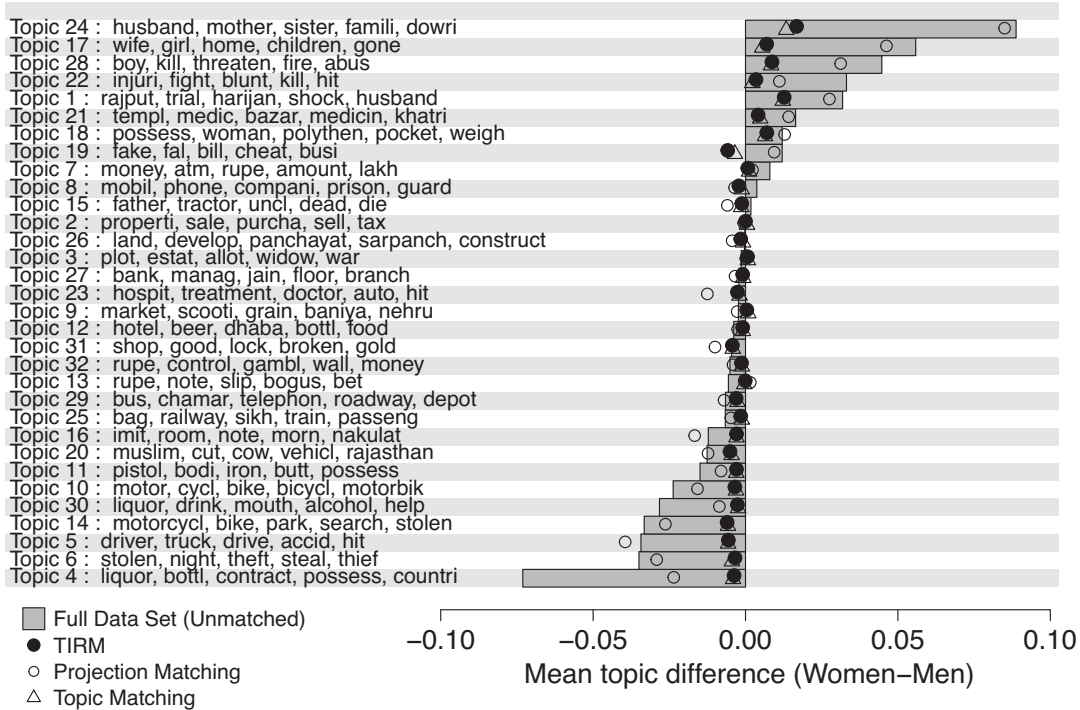
As a second balance test, I randomly select and present 12 matched testimonies in Table 7. This is a hard test for balance because the machine matched cases without any reference to Penal Codes; and still, post-TIRM, we see similarities in codes based on

content. In fact, the machine is more successful at categorizations than police officers.<sup>50</sup> In rows 1, 2, 5, and 6 of Table 7, we see non-VAW cases registered by either a male or female complainant (identifying information redacted). Row 1 depicts scooter theft, and row 2 a hit-and-run. In row 2, we see matched cases where a crash occurred. Still, despite being similar, there remain dissimilarities that the machine cannot (and, in fact, should not) match on. In row 6, for both men and women, the cases involve confidence-tricksters, but the type of con is distinct. Relatedly,

<sup>50</sup> In Table 8 (row 2), the officer did not attach Section 338. Administrators may now use algorithms to ensure correct Penal Codes are being utilized, instead of relying on memory or manuals. An online tool under development, called the Indian-Penal-Code Classifier, may benefit (a) citizens by ensuring accurate charges are applied, and (b) police officers by reducing cognitive load.



FIGURE 10. Balance Check I



Note: Gray bars indicate cases associated with male and female complainants. For instance, women complainants are likely to bring forward dowry cases (e.g., Topic 24), while male complainants are involved in or file alcohol and bootlegging (Topic 4). TIRM tries to achieve balance on estimated topics (black dots).

the woman's case in the dowry murder involves a killing, but in the matched column, a woman *and* her child have been found deceased. In the abduction cases, the complainants belong to distinct religious communities (Hindu and Muslim, respectively).

The language in the documents is rich, and allows for a brief interpretative exercise. In row 3 of Table 7, we see (relatively less violent) dowry cases wherein victims have been extorted, and in the left column of the table, beaten. Consider the way in which class is foregrounded. In the right column, the father—who is registering a case on behalf of his child—notes that his daughter is well-educated. He also underscores his social humiliation. The complainant in the left column of Table 7 is filing a case against a lawyer and judge, which suggests not only that the perpetrators have influence, but also that they are well-educated; yet, the suspects allegedly believe that they are owed luxury vehicles in view of their “status.” Similarly, in row 4, the complainant in the left column notes that the in-laws (in an arranged marriage) were given gifts in accordance “with their status.”<sup>51</sup> These dynamics suggest that class plays a role in mediating interactions; even

with officers, citizens signal their background, potentially to be taken seriously. Often, citizens do not have cultural capital, and are forced to plead for help (as in the right column for the “abduction” row where the complainant stresses his poverty). A puzzle arises as to how justice would vary across these contexts: would the system provide re-distributive justice (financial compensation), especially for losses in the dowry or cheating cases? Would those with political connections or social capital be more successful in seeing their requests cross the system's stages?

In the left column of row 4 in Table 7, the perpetrators previously went to prison. This raises concerns about the failure to deter perpetrators such that a woman was allegedly murdered despite the suspects' prior incarceration. The reports shed light on criminal impunity, where individuals may be abducted in broad daylight, or killed in defiance of the authorities. Many victims are threatened with further violence if they dare to reveal their oppression (e.g., row 5). Clearly, victims face challenges for breaking their silence, thereby not only hinting at the courage required to approach the first stage of police filing, but also the likely number of unreported cases that never enter the formal multi-stage process. The testimonies of the dowry murders (crimes that coincidentally have the highest percentage of suspect acquittal, Supplementary Table A21), illuminate the quantitative insights by demonstrating how real individuals are impacted.

<sup>51</sup> More well-to-do individuals might demand luxury vehicles as dowry—which for a less upwardly mobile group could involve a motorcycle instead of car—in addition to the mandatory jewelry and household effects.

**TABLE 7. Balance Check II, Matched Cases and Corresponding Charges/Penal Codes [Identifying Information Redacted]**

Female First-Information-Report	Matched Male First-Information-Report	
<p>I am Ankita ██████, daughter of ██████ ██████ ██████ ██████, Punjab. I live in ██████ ██████ Gurgaon. I work in ██████ ██████ company sector ██████. On date ██████ at 10 am I came to company for duty on my scooty. I parked my scooty in the parking lot, and I went to office. When I came back at around 6:00 pm, my scooty could not be found. My scooty color was Gray Model 20████, License ██████ ██████ Engine No ██████. I do not know who took it. Please register an FIR for my stolen scooty. <a href="#">IPC 379.</a></p>	<p>I am Kapil ██████, son of ██████ ██████ from Nagina. I have a scooty number ██████ ██████ in white. I left my scooty on ██████ in a plot near ██████ University. I was giving exam from 2-5 o'clock when I came back, Scooty was not standing there... After that, I had gone to my hometown for some urgent work, and now I am submitting to police. I do not remember the Scooty's engine or chassis number, all papers were in Scooty itself. Please register an FIR for my theft. Phone No. ██████ ██████. <a href="#">IPC 379.</a></p>	<p>Scooter Theft</p>
<p>I am Vandana, wife of ██████ ██████ Caste Kamboj, resident of Village ██████ ██████. I am 30. Yesterday, my boy had gone to ██████ for tutoring. I was going to pick him up at 6.00 pm on my Aactiva, License No. ██████. While taking U-turn in front of ██████ ██████ Petrol Pump, a motorcycle driver from Yamunanagar crashed into me. I fell on the road, and my left leg was seriously injured... My brother noted the License number ██████ ██████...got admitted to ██████ ██████ Hospital ██████ ██████ for treatment. I am in full consciousness now. The motorcyclist ran away, but I can recognize him if he comes in front of me... <a href="#">IPC 279/337/338.</a></p>	<p>I am Harsha ██████, son of Pra ██████ ██████, Caste ██████...I study in B.T.Class. On date ██████ at around 9:40 PM, I was riding my cycle (License ██████) from Sector ██████ to ██████ Nagar. Behind me my friend Jag ██████ ██████, son of ██████ ██████, caste Jat, was sitting and I was driving. When we reached the telephone exchange, a car came from behind with great speed and carelessness, and hit me, from which I bounced off bike. My head went into the electric pole, and my friend fell on the road. The car no. was ██████ ██████18, a Honda 110...the driver's name is ██████...Strictest legal action should be taken against him. <a href="#">IPC 279/337.</a></p>	<p>Hit-and-Run</p>
<p>...I, Anuja ██████, daughter of late ██████...cruelty and violence which has completely left me traumatized and I am constantly living in fear for my own life...went to my parental house for Pag Phera and returned back at night to my matrimonial house, in ██████. In the evening all the leftover jewelry (which I was wearing) was taken by my sister-in-law on pretext that it is better to be kept safe with in-laws...After marriage I realised that my husband and my in-laws were downright greedy as they started making more illegal demands for dowry... They used to persistently taunt and harass me for not bringing sufficient amount of cash and gifts. My husband and his father ██████ also demanded that they have not been given a car according to their 'status,' and should be given a Mercedes or Pajero in dowry. Father-in-law ██████...is one of the leading lawyers in the town...his elder son is judge posted as Civil Judge Cum JMIC. My husband...taunting that my parents had not spent money...Since then health has started deteriorating, my mother-in-law and father-in-law became angry and beat me...<a href="#">IPC 323/406/498-A/506</a></p>	<p>Mr. Sir...█████ is my daughter who has studied up to M.Sc., B.Ed. and whose marriage we had with ██████ Mac ██████ from Delhi on ██████. We had an engagement ceremony which cost Rs.3,00,000 / and gave the boy a gold chain, a gold ring and Rs. 1,51,000 / cash. They then demanded a Scorpio. When we expressed our inability to deliver the Scorpio vehicle, he asked to meet after two days, and I met him on ██████, he said that we also want Rs.5,00,000 / - cash with the Scorpio. On our refusal, he refused to bring a procession. But we had completed the wedding preparations. Some relatives had arrived. We had booked confectioners, tents, banquet hall... we already spent Rs.10,00,000 /. Then I, and my boy ██████, my brother-in-law ██████, our neighbor met them. Sitting and talking, they refused to marry without Rs. 5,00,000 /...The culprits refused to marry my girl after being engaged in the greed of dowry, and I was humiliated and my Rs.20,00,000/ has been lost. Therefore, I pray that legal action should be taken against him and FIR should be lodged...my goods, cash should be returned...<a href="#">Dowry Prohibition Act, 1961;4/3.</a></p>	<p>Dowry Harassment</p>
<p>I am ██████. Late Shri ██████ married his girl ██████ Anuja to ██████, resident of ██████ on 21.04.20████. According to his status, everything was given, but after a few months, the accused started harassing the family and demanded a motorcycle. Her family members started beating her. In 20████, he tried to kill her by pouring kerosene on her, but she escaped. For this, ██████ and his father ██████ were caught and sent to jail, but later they started living together again. ██████ and ██████ brought ██████ Anuja to Delhi and started harassing her again, saying they want Rs. 1 lakh from her family to start business. The father and mother-in-law ██████ Devi...started to behave more wrongly till ██████ Anuja was hanged. Shrimanji is requested to investigate this and please get justice...information was received from Saf ██████ Hospital that ██████ Anuja has died...<a href="#">IPC 304-B.</a></p>	<p>I have come to complain that my sister Shilpa ██████ was wife of Sa ██████ ██████, resident of ██████ ██████ Ground. She was married to Sa ██████ 3 years ago at age 24. Today at 4 o'clock in the evening, we got the news that she and her son Rih ██████, aged 2 years, have both been found dead in the bathroom. We got a call from the hospital...Go to ██████ as soon as possible - we are sure that the death has been caused by dowry demands. We got a call from Shilpa on date ██████ ██████ from Po ██████, a resident of Delhi. Shilpa told her that she was being bullied for dowry - Rs. 10 lakh and a vehicle was being demanded...she was being beaten...Please fully investigate that Shilpa's husband Sa ██████ ██████ has definitely killed Shilpa and her son Rih ██████. We hope to take immediate action from you. <a href="#">IPC 304-B.</a></p>	<p>Dowry Murder</p>
<p>...Mr. Sir...I am Bimala, wife of ██████ ██████ from Sonipat. This morning my daughter, whose name is ██████, was abducted by Saga ██████ aka ██████ and family. She's been taken away. I am getting phone from No. ██████. Saga ██████ has threatened to kill her, and said that give 5 lakh rupees or else she will die. We do not know where she is, but the number is telling location Chandigarh. I pray to you that the police administration is involved and it is registered, please do not delay it. Phone no. ██████. <a href="#">IPC 365.</a></p>	<p>Mr. Sir...I am a resident of ██████ Road Punh ██████, Mewat, ██████ Khan. I am a man of peace who abides by the law. On the date ██████ at around 1 o'clock at night, Hak ██████ son of ██████ of ██████... asked me to open the door...there were two or three others. The men came in and put a <i>katta</i> [knife] on my neck and started saying that "if you make noise, we will kill you and your family." They took my girl ██████ bnam by force and cash of Rs. 32,000 / - and put my girl in a Scorpio. They said they will kill her if we go to police...when we went to Hak ██████ in the morning, he told us that he will not give her at any price...I request, Janab, to take legal action against the people and return my girl to a poor man. <a href="#">IPC 363/366-A.</a></p>	<p>Abduction</p>
<p>I am a Indira ██████ wife of Mr. ██████ ██████ from ██████ ██████ Colony, Hisar. I work as an assistant in ██████ ██████. In January 20████, I got a call from Sach ██████ ██████, JGS India Trading and Marketing PVT Ltd...a good scheme...where government employees have a big advantage... deposit two lakh twenty thousand rupees in the account of this company, you will get 8000 rupees per month for 12 months...He said that we have benefited thousands of people...Account ██████...IFSC Code ██████...Sach ██████ threatened me...stole Rs 2,20,000... <a href="#">IPC 406/420</a></p>	<p>Mr. Sir...I am Gulzar ██████ son Mr. Sa ██████ ██████, resident of ██████ ██████, Ambala city. I have known the suspects for 15-20 years. They said they would help me file to go to Canada in 2015...they told me that they work to send poor people abroad, and with down-payment of Rs 1,50,000 - to 2,00,000. / - one can easily earn more abroad...told me that you should give me all the documents...My shop is located in ██████ ██████ Mandi...They took my money and now saying they will kill me...retrieve my money which is Rs.6,50,000/... <a href="#">IPC 406/420.</a></p>	<p>Cheating</p>

**TABLE 8. Impact of Complainant Gender on Conviction/Acquittal of Suspect in Case after Text Matching**

	All complaints				Court docket only			
	Suspect conviction		Suspect acquittal		Suspect conviction		Suspect acquittal	
	1	2	3	4	5	6	7	8
Female	-0.008*** (0.001)	-0.008*** (0.001)	0.011*** (0.002)	0.015*** (0.002)	-0.014*** (0.002)	-0.012*** (0.002)	0.019*** (0.003)	0.021*** (0.003)
Constant	0.037*** (0.0003)	0.025*** (0.007)	0.123*** (0.001)	0.500*** (0.012)	0.064*** (0.001)	0.037*** (0.012)	0.213*** (0.001)	0.743*** (0.019)
No. of obs.	336,199	308,109	336,199	308,109	195,174	178,943	195,174	178,943
R <sup>2</sup>	0.0002	0.037	0.0001	0.095	0.0003	0.067	0.0002	0.139
Adj. R <sup>2</sup>	0.0002	0.036	0.0001	0.094	0.0003	0.065	0.0002	0.137
<i>Excluding VAW</i>								
	All complaints				Court docket only			
	Suspect conviction		Suspect acquittal		Suspect conviction		Suspect acquittal	
	1	2	3	4	5	6	7	8
Female	-0.011*** (0.001)	-0.011*** (0.001)	0.007*** (0.002)	0.013*** (0.002)	-0.019*** (0.002)	-0.019*** (0.002)	0.018*** (0.003)	0.025*** (0.003)
Constant	0.042*** (0.0004)	0.066*** (0.008)	0.110*** (0.001)	0.311*** (0.013)	0.077*** (0.001)	0.115*** (0.015)	0.202*** (0.001)	0.540*** (0.023)
No. of obs.	314,098	288,686	314,098	288,686	181,753	167,200	181,753	167,200
R <sup>2</sup>	0.0002	0.034	0.00004	0.081	0.0004	0.065	0.0001	0.119
Adj. R <sup>2</sup>	0.0002	0.033	0.00004	0.080	0.0004	0.063	0.0001	0.117
Controls	N	Y	N	Y	N	Y	N	Y
PS FE	N	Y	N	Y	N	Y	N	Y
Month-year FE	N	Y	N	Y	N	Y	N	Y

Note: Controls include a numeric variable for a crime's distance from a station and investigator rank. Columns 6 and 8 control for the rank of the judge. The top rows examine the effect of "female" post text matching, whereas the bottom rows exclude all cases with a VAW Penal Code so that the comparison is as far as possible restricted to generic or non-VAW cases. \*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01.

While the testimonies in Table 7 highlight other themes beyond this article's scope (e.g., the role of caste), the matched dataset facilitates an additional quantitative test for gender discrimination. In Table 8, the coefficients on female remain significant. Columns 1–4 in the top row show that the suspect implicated in a case registered by a woman has a significantly unlikely chance of being convicted or going to jail compared to those brought by men. Columns 5–8 restricts the sample to the court docket.

Matching cases of VAW between male and female complainants can be construed an odd comparison because, while typically similar, a man registering a case on behalf of a relative or friend might not have all the facts. Similarly, if a case of VAW is being lodged by a victim's male family member or friend, it could imply that the victim has significant support or belongs to a particular class. Moreover, victims of VAW may be under trauma when providing testimonies, and so cases could be systematically different between men and women complainants who report such crime. As seen in row 3 of Table 7, the male and female complainant

are both registering dowry cases; however, the man is emphasizing his economic losses whereas the woman is underscoring emotional and physical abuse.

For these reasons, one may want to make the comparison more parsimonious by matching on non-VAW (or generic) cases alone. In the bottom of Table 8, I re-run the algorithm to exclude VAW<sup>52</sup> such that the comparison is generally restricted to topics such as hit-and-runs, cheating, scooter theft, and burglary. The gender gap remains significant. Importantly, TIRM is likely an underestimate of discrimination. The approach understates the differential effort required by women to have reached the first stage or, say, norms about publicly coming forward. Moreover, women who register scooter theft (e.g., those that own such an asset or would even report it if stolen) may not be representative of women in Haryana society. The testimony upon which cases are matched might itself be gendered,

<sup>52</sup> See Supplementary Figure A45 for the balance check.

for example, based on a lifetime of discrimination that generates differences in speech and word-choice. And so, while matching rests on several assumptions (Feder et al. 2021), that discrepancies remain supports the preceding analyses in demonstrating that (gender) identity does have a bearing on criminal justice outcomes.

## DISCUSSION

Political science has had limited purchase, even basic descriptive evidence, as to whether the state treats groups seeking justice differently, especially in the Global South. With Indian crime records, combined with judicial files, I chart the full trajectory of citizen requests for help from their entrance into a police station until a court verdict. I establish a series of facts about how individuals navigate this system, and inductively illustrate a pattern of “multi-stage” discrimination in terms of a more onerous *process* and unequal *outcomes* for women at successive stages of seeking restitution. The study aims to re-direct discussions in criminal justice scholarship from *demand*-side factors (e.g., lack of trust in police or under-reporting by disadvantaged groups) to *supply*-side failures by institutions in providing help conditional on citizens turning to the state.

Specifically, I find that women are disadvantaged in terms of (1) police delays in registering cases, (2) fewer cases sent to court, (3) delays in investigations, (4) higher court dismissals, (5) delays in trials and verdict issuance, (6) higher accused acquittals, and (7) lower convictions of suspects. The effects hold when looking at each stage separately, or when analyzing outcomes as a function of all initial registrations. With structural topic modeling, I amplify victims’ voices and place their testimonies at the center of the research agenda. Then, with text matching, I utilize the first (police complaint) and final stage (judicial verdict), to provide credible evidence that the criminal justice administration discriminates based on the gender identity of the complainant.

Multi-stage discrimination can occur when groups approach institutions, including for grievance redressal; complaints, applications, and requests may be “squeezed” in terms of spending longer in-between stages or witness unsuccessful transitions. This funneling occurs at nodes in a system where either administrators have discretion or at inflection points wherein the routine implementation of rules indirectly disadvantage some over others. For example, mandating all citizens come forward in open court to describe their complaint might place undue burden on women more than men, thereby allowing for the formation of “gendered institutions” whereby disadvantages are maintained through official processes (Hawkesworth 2003). The findings underscore the importance of being attentive to the workings of criminal justice systems when complaints are being processed, long after initial gate-keeping by administrators in terms of registration; as we see, inequities in access to justice may reflect the

sum of episodic instances of discrimination that a majority of existing studies are likely overlooking. Discrimination that occurs at multiple stages may deter or dissuade disadvantaged groups from approaching the state for help altogether, and induce citizens to rely on alternate dispute resolution mechanisms.<sup>53</sup> Furthermore, policies aimed at mitigating inequity in any one institution (e.g., police) may be less effective unless successive administrators’ abilities to influence outcomes are accounted for.

The study expands discussions of VAW—which in political science largely focus on violence perpetrated during (or after) conflict—by highlighting the state’s response to day-to-day abuse. In India, dowry, for instance, is a complaint likely to be stalled; yet, topic modeling reveals that such crimes can involve heinous acts including marital rape. This is evocative of a double-bind: on the one hand, women may be faced with marital violence, and even (dowry) death, in an effort to extract resources from their natal homes; yet, delaying or avoiding marriage comes with its own costs (Carpena and Jensusius 2021; Corno, Hildebrandt, and Voena 2020). While studies on VAW in South Asia have focused on its relationship to property rights (Panda and Agarwal 2005), alcohol consumption (Luca, Owens, and Sharma 2015), and culture (Fernandez 1997), a question emerges as to whether perpetrators are aware of the inability, or unwillingness, of the state to provide punitive justice, and if this knowledge among abusers makes VAW more likely.

The cases capture—often in deeply poignant terms—the helplessness of victims, who sometimes express that they have turned to formal institutions as a last resort, despite uncertainty in a system’s ability to help when much seems lost or destroyed. Survey data show that women are likely to seek assistance from others when going to the police to file complaints (CSDS 2018); the findings imply that this is not an irrational decision. The study opens several avenues for future research. Do police discriminate because of supposed privilege that women exude by coming forward (e.g., without male support)? Are judges (who are generally well-educated) concerned that formal justice for women is a threat to a particular order, or easily dismissed to de-clog an overburdened docket? Do constraints for women such as lack of access to lawyers, political connections, or limited autonomy from the household intersect with administrators’ taste for discrimination? Might gender interact with caste, religion, or even age?<sup>54</sup> Is Haryana and the capital region surrounding Delhi representative of the subcontinent? How can we use administrative data to gain more precise estimates of the cumulative effect or magnitude of total and systemic discrimination (Bohren, Hull, and Imas 2022)? Can interventions that make the criminal justice

<sup>53</sup> In north India, such mechanisms may include informal assemblies of village elders called *khap panchayats*.

<sup>54</sup> Research shows gender disparities worsen for women at older ages in India (Dupas and Jain 2021).



administration more demographically representative (for women and minorities) affect the base-line statistics outlined herein?

While the notion that women face hardship in India may be unsurprising to many, others, including justices and policymakers, have maintained that female complainants send men to prison for frivolous offenses, that the Penal Code is stacked in women's favor, and that a burgeoning "men's rights movement" should be supported in deterring women's "legal terrorism" (Lodhia 2014; Naishadham 2018). The findings do not provide support for these assumptions. They make a theoretical argument for exploring the junctures at which linked institutions are connected, and the varying discretionary authority of bureaucrats across those bodies in order to understand layered, dynamic patterns of discrimination. Exploring whether discrimination repeats and evolves may promote theory-building and target reform<sup>55</sup> aimed at improving justice delivery and the quality of democracy.

## SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0003055423000916>.

## DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/F6UCQT>. Some limitations on data availability are discussed in the main text and readme file.

## ACKNOWLEDGMENTS

I thank Raminder Singh Jassal and Smita Tewari Jassal. I am grateful to Pascaline Dupas, Jessica Leino, and the King Center on Global Development at Stanford University. For feedback and helpful discussions, I thank Sam Asher, Paul Novosad, Aprajit Mahajan, Irfan Nooruddin, Alison Post, Abhijit Banerjee, Jeremy Bowles, Katherine Casey, Saumitra Jha, Eli Berman, Devesh Kapur, David Foster, Prashant Bhargadwaj, Brian Min, Adam Auerbach, Augustin Bergeron, Jake Grumbach, Susan Ostermann, Ryan Hübert, Andrew Little, Marcel Fafchamps, Renard Sexton, Hanif Qureshi, Manisha Choudhary, Margit Tavits, Susan Athey, Molly Roberts, Irine Morse, Alessandra Voena, Jayant Tripathi, Nayantara Roy, Radhika Jain, and Peter Hull. I am appreciative of Elliott Ash, Christoph Goessman, and the Development Data Lab. I thank colleagues at the Political

Economy and Gender Research working groups at Stanford, the Center for Statistics and Social Sciences at the University of Washington, the Empirical Studies of Conflict (ESOC) meeting at the University of California, Berkeley, the Berlin Meeting on the Political Economy of Development at the WZB Social Science Center, as well as the Comparative Politics seminars at the University of California, Santa Barbara, American University, and the University of Michigan. The study benefited from excellent research assistance by Emily Wu and Shirley Cheng.

## FUNDING STATEMENT

I acknowledge funding and support from Stanford University's King Center on Global Development.

## CONFLICT OF INTEREST

The author declares no ethical issues or conflicts of interest in this research.

## ETHICAL STANDARDS

The author affirms this research did not involve human subjects.

## REFERENCES

- Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race?" *Journal of Legal Studies* 41 (2): 347–83.
- Agerberg, Mattias, and Anne-Kathrin Kreft. 2020. "Gendered Conflict, Gendered Outcomes: The Politicization of Sexual Violence and Quota Adoption." *Journal of Conflict Resolution* 64 (2): 290–317.
- Agnes, Flavia, and Audrey D'Mello. 2015. "Protection of Women from Domestic Violence." *Economic and Political Weekly* 50 (44): 76–84.
- Alesina, Alberto, and Eliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing." *American Economic Review* 104 (11): 3397–433.
- Armstrong, Elizabeth A., Miriam Gleckman-Krut, and Lanora Johnson. 2018. "Silence, Power, and Inequality: An Intersectional Approach to Sexual Violence." *Annual Review of Sociology* 44: 99–122.
- Arnold, David, Will Dobbie, and Crystal S. Yang. 2018. "Racial Bias in Bail Decisions." *Quarterly Journal of Economics* 133 (4): 1885–932.
- Arrow, Kenneth J. 2015. "The Theory of Discrimination." In *Discrimination in Labor Markets*, eds. Orley Ashenfelter and Albert Rees, 1–33. Princeton, NJ: Princeton University Press.
- Ash, Elliott, Sam Asher, Aditi Bhowmick, Daniel L. Chen, Tanaya Devi, Christoph Goessmann, Paul Novosad, et al. 2021. "Measuring Gender and Religious Bias in the Indian Judiciary." Working Paper.
- Asher, Sam, and Paul Novosad. 2023. "Rent-Seeking and Criminal Politicians: Evidence from Mining Booms." *Review of Economics and Statistics* 105 (1): 20–39.
- Basu, Srimati. 2012. "Judges of Normality: Mediating Marriage in the Family Courts of Kolkata, India." *Signs: Journal of Women in Culture and Society* 37 (2): 469–92.
- Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science* 335 (6068): 582–6.

<sup>55</sup> While 30% of VAW cases are dismissed by law enforcement in Haryana, newspapers report prosecutors dropped 49% of sexual assault cases in New York City in 2019 (Ransom 2021).

- Becker, Gary S. 2010. *The Economics of Discrimination*. Chicago, IL: University of Chicago Press.
- Bertrand, Marianne, and Esther Duflo. 2017. "Field Experiments on Discrimination." In *Handbook of Economic Field Experiments*, eds. Abhijit Banerjee and Esther Duflo, 309–93. Amsterdam: North-Holland.
- Bhalotra, Sonia, Abhishek Chakravarty, and Selim Gulesci. 2020. "The Price of Gold: Dowry and Death in India." *Journal of Development Economics* 143: 102413.
- Bhatia, Rahul. 2018. "Opinion — India Loves Data but Fails to Protect It." *The New York Times*, April 3.
- Biderman, Albert D., and Albert J. Reiss. 1967. "On Exploring the "Dark Figure" of Crime." *Annals of the American Academy of Political and Social Science* 374: 1–15.
- Bindler, Anna, and Randi Hjalmarsson. 2020. "The Persistence of the Criminal Justice Gender Gap: Evidence from 200 Years of Judicial Decisions." *Journal of Law and Economics* 63 (2): 297–339.
- Blair, Graeme, Jeremy M. Weinstein, Fotini Christia, Eric Arias, Emile Badran, Robert A. Blair, Ali Cheema, et al. 2021. "Community Policing Does Not Build Citizen Trust in Police or Reduce Crime in the Global South." *Science* 374 (6571): eabd3446.
- Blair, Robert A., Sabrina M. Karim, and Benjamin S. Morse. 2019. "Establishing the Rule of Law in Weak and War-Torn States: Evidence from a Field Experiment with the Liberian National Police." *American Political Science Review* 113 (3): 641–57.
- Blank, Rebecca M. 2005. "Tracing the Economic Impact of Cumulative Discrimination." *American Economic Review* 95 (2): 99–103.
- Bohren, J. Aislinn, Peter Hull, and Alex Imas. 2022. *Systemic Discrimination: Theory and Measurement*. Cambridge, MA: National Bureau of Economic Research.
- Bohren, J. Aislinn, Alex Imas, and Michael Rosenberg. 2019. "The Dynamics of Discrimination: Theory and Evidence." *American Economic Review* 109 (10): 3395–436.
- Burulé, Rachel E. 2020. "Reform, Representation, and Resistance: The Politics of Property Rights' Enforcement." *Journal of Politics* 82 (4): 1390–405.
- Butler, Daniel M., and David E. Broockman. 2011. "Do Politicians Racially Discriminate Against Constituents? A Field Experiment on State Legislators." *American Journal of Political Science* 55 (3): 463–77.
- Carpna, Fenella, and Francesca R. Jentsch. 2021. "Age of Marriage and Women's Political Engagement: Evidence from India." *Journal of Politics* 83 (4): 1199–883.
- Chattopadhyay, Raghavendra, and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 72 (5): 1409–43.
- Chhibber, Pradeep. 2002. "Why are Some Women Politically Active? The Household, Public Space, and Political Participation in India." *International Journal of Comparative Sociology* 43 (3): 409–29.
- Chhibber, Pradeep, Francesca R. Jensenius, and Susan L. Ostermann. 2021. "Women's Education and Declining Child Sex Ratios in India." *Economic and Political Weekly* 56 (6).
- Cohen, Dara Kay. 2013. "Explaining Rape during Civil War: Cross-National Evidence (1980–2009)." *American Political Science Review* 107 (3): 461–77.
- Corno, Lucia, Nicole Hildebrandt, and Alessandra Voena. 2020. "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments." *Econometrica* 88 (3): 879–915.
- CSDS. 2018. "Status of Policing in India Report (SPIR): CSDS and Common Cause." Accessed May 12, 2018. <http://www.commoncause.in/page.php?id=85>.
- Dam, Abhirup. 2016. "For the First Time, Access First Information Reports (FIRs) Online." *The Quint*, September 8.
- DHS. 2017. "India National Family Health Survey NFHS-4 2015-16." *Demographic and Health Survey*.
- DiPrete, Thomas A., and Gregory M. Eirich. 2006. "Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical developments." *Annual Review of Sociology* 32: 271–97.
- Dupas, Pascaline, and Radhika Jain. 2021. *Women Left Behind: Gender Disparities in Utilization of Government Health Insurance in India*. Cambridge, MA: National Bureau of Economic Research.
- Emeriau, Mathilde. 2022. "Learning to be Unbiased: Evidence from the French Asylum Office." *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12720>.
- Feder, Amir, Katherine A. Keith, Emaad Manzoor, Reid Pryzant, Dhanya Sridhar, Zach Wood-Doughty, Jacob Eisenstein, et al. 2021. "Causal Inference in Natural Language Processing: Estimation, Prediction, Interpretation and Beyond." Preprint, <arXiv:2109.00725>.
- Fernandez, Marilyn. 1997. "Domestic Violence by Extended Family Members in India: Interplay of Gender and Generation." *Journal of Interpersonal Violence* 12 (3): 433–55.
- Goldsmith, Belinda, and Meka Beresford. 2018. "Poll Ranks India the World's Most Dangerous Country for Women." *The Guardian*, June 28.
- Green, Donald P., Anna M. Wilke, and Jasper Cooper. 2020. "Countering Violence against Women by Encouraging Disclosure: A Mass Media Experiment in Rural Uganda." *Comparative Political Studies* 53 (14): 2283–320.
- Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21 (3): 267–97.
- Grossman, Guy, Oren Gazal-Ayal, Samuel D. Pimentel, and Jeremy M. Weinstein. 2016. "Descriptive Representation and Judicial Outcomes in Multiethnic Societies." *American Journal of Political Science* 60 (1): 44–69.
- Hawkesworth, Mary. 2003. "Congressional Enactments of Race–Gender: Toward a Theory of Raced–Gendered Institutions." *American Political Science Review* 97 (4): 529–50.
- Holland, Alisha C. 2016. "Forbearance." *American Political Science Review* 110 (2): 232–46.
- Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–60.
- Htun, Mala, and S. Laurel Weldon. 2012. "The Civic Origins of Progressive Policy Change: Combating Violence against Women in Global Perspective, 1975–2005." *American Political Science Review* 106 (3): 548–69.
- Iyer, Lakshmi, Anandi Mani, Prachi Mishra, and Petia Topalova. 2012. "The Power of Political Voice: Women's Political Representation and Crime in India." *American Economic Journal: Applied Economics* 4 (4): 165–93.
- Jassal, Nirvikar. 2020. "Gender, Law Enforcement, and Access to Justice: Evidence from All-Women Police Stations in India." *American Political Science Review* 114 (4): 1035–54.
- Jassal, Nirvikar. 2021. "Segregation as Efficiency? Group-Specific Institutions in North India." *Journal of Asian Studies* 80 (3): 631–61.
- Jassal, Nirvikar. 2023. "Replication Data for: Does Victim Gender Matter for Justice Delivery? Police and Judicial Responses to Women's Cases in India." Harvard Dataverse. Dataset. <https://doi.org/10.7910/DVN/F6UCQT>.
- Jassal, Nirvikar, and Sharon Barnhardt. 2023. "Do Women Prefer In-Group Police Officers? Survey and Experimental Evidence from India." *Comparative Political Studies*. <https://doi.org/10.1177/00104140231194070>.
- Jayachandran, Seema. 2015. "The Roots of Gender Inequality in Developing Countries." *Annual Review of Economics* 7: 63–88.
- Jayachandran, Seema. 2017. "Fertility Decline and Missing Women." *American Economic Journal: Applied Economics* 9 (1): 118–39.
- Khan, Shamus, Joss Greene, Claude Ann Mellins, and Jennifer S. Hirsch. 2020. "The Social Organization of Sexual Assault." *Annual Review of Criminology* 3: 139–63.
- Knox, Dean, Will Lowe, and Jonathan Mummolo. 2020. "Administrative Records Mask Racially Biased Policing." *American Political Science Review* 114 (3): 619–37.
- Korver-Glenn, Elizabeth. 2018. "Compounding Inequalities: How Racial Stereotypes and Discrimination Accumulate across the Stages of Housing Exchange." *American Sociological Review* 83 (4): 627–56.
- Kothari, Jayna. 2005. "Criminal Law on Domestic Violence: Promises and Limits." *Economic and Political Weekly* 40 (46): 4843–9.
- Kumar, Avinash. 2017. "Bihar Police Upload FIRs on Web." *Hindustan Times*, January 12.

- Kurlychek, Megan C., and Brian D. Johnson. 2019. "Cumulative Disadvantage in the American Criminal Justice System." *Annual Review of Criminology* 2: 291–319.
- Kutateladze, Besiki L., Nancy R. Andiloro, Brian D. Johnson, and Cassia C. Spohn. 2014. "Cumulative Disadvantage: Examining Racial and Ethnic Disparity in Prosecution and Sentencing." *Criminology* 52 (3): 514–51.
- Lang, Kevin, and Ariella Kahn-Lang Spitzer. 2020. "Race Discrimination: An Economic Perspective." *Journal of Economic Perspectives* 34 (2): 68–89.
- Liebman, Benjamin L., Margaret E. Roberts, Rachel E. Stern, and Alice Z. Wang. 2020. "Mass Digitization of Chinese Court Decisions: How to Use Text as Data in the Field of Chinese Law." *Journal of Law and Courts* 8 (2): 177–201.
- Lodhia, Sharmila. 2014. "'Stop Importing Weapons of Family Destruction!': Cyberdiscourses, Patriarchal Anxieties, and the Men's Backlash Movement in India." *Violence Against Women* 20 (8): 905–36.
- Luca, Dara Lee, Emily Owens, and Gunjan Sharma. 2015. "Can Alcohol Prohibition Reduce Violence against Women?" *American Economic Review* 105 (5): 625–9.
- Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis* 23 (2): 254–77.
- McDougal, Lotus, Samuel Krumholz, Nandita Bhan, Prashant Bharadwaj, and Anita Raj. 2021. "Releasing the Tide: How has a Shock to the Acceptability of Gender-Based Sexual Violence Affected Rape Reporting to Police in India?" *Journal of Interpersonal Violence* 36 (11): NP5921–43.
- Naishadham, Suman. 2018. "Why India's Men's Rights Movement is Thriving." *Vice*, April 13.
- Olsen, Asmus Leth, Jonas Høgh Kyhse-Andersen, and Donald Moynihan. 2022. "The Unequal Distribution of Opportunity: A National Audit Study of Bureaucratic Discrimination in Primary School Access." *American Journal of Political Science* 66 (3): 587–603.
- Panda, Pradeep, and Bina Agarwal. 2005. "Marital Violence, Human Development and Women's Property Status in India." *World Development* 33 (5): 823–50.
- Parthasarathy, Ramya, Vijayendra Rao, and Nethra Palaniswamy. 2019. "Deliberative Democracy in an Unequal World: A Text-as-Data Study of South India's Village Assemblies." *American Political Science Review* 113 (3): 623–40.
- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62 (4): 659–61.
- Raghavan, R. K., and D. Sivanandhan. 2016. "A First Step to Wholesome Reform." *The Hindu*, September 14.
- Ransom, Jan. 2021. "'Nobody Believed Me': How Rape Cases Get Dropped." *The New York Times*, July 19.
- Rao, Vijayendra. 1997. "Wife-Beating in Rural South India: A Qualitative and Econometric Analysis." *Social Science & Medicine* 44 (8): 1169–80.
- Rehavi, M. Marit, and Sonja B. Starr. 2014. "Racial Disparity in Federal Criminal Sentences." *Journal of Political Economy* 122 (6): 1320–54.
- Reskin, Barbara. 2012. "The Race Discrimination System." *Annual Review of Sociology* 38: 17–35.
- Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airoidi. 2016. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association* 111 (515): 988–1003.
- Roberts, Margaret E., Brandon M. Stewart, and Richard A. Nielsen. 2020. "Adjusting for Confounding with Text Matching." *American Journal of Political Science* 64 (4): 887–903.
- Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2019. "Stm: An R Package for Structural Topic Models." *Journal of Statistical Software* 91 (1): 1–40.
- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, et al. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58 (4): 1064–82.
- Roychowdhury, Poulami. 2021. *Capable Women, Incapable States: Negotiating Violence and Rights in India*. New York: Oxford University Press.
- Sanders, James, Giulio Lisi, and Cheryl Schonhardt-Bailey. 2017. "Themes and Topics in Parliamentary Oversight Hearings: A New Direction in Textual Data Analysis." *Statistics, Politics and Policy* 8 (2): 153–94.
- Shayo, Moses, and Asaf Zussman. 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." *Quarterly Journal of Economics* 126 (3): 1447–84.
- Slough, Tara, and Christopher Fariss. 2021. "Misgovernance and Human Rights: The Case of Illegal Detention without Intent." *American Journal of Political Science* 65 (1): 148–65.
- Small, Mario L., and Devah Pager. 2020. "Sociological Perspectives on Racial Discrimination." *Journal of Economic Perspectives* 34 (2): 49–67.
- Soss, Joe, and Vesla Weaver. 2017. "Police Are Our Government: Politics, Political Science, and the Policing of Race-Class Subjugated Communities." *Annual Review of Political Science* 20: 565–91.
- Spohn, Cassia, and Katharine Tellis. 2019. "Sexual Assault Case Outcomes: Disentangling the Overlapping Decisions of Police and Prosecutors." *Justice Quarterly* 36 (3): 383–411.
- Srinivasan, Sharada, and Arjun S. Bedi. 2007. "Domestic Violence and Dowry: Evidence from a South Indian Village." *World Development* 35 (5): 857–80.
- Stolzenberg, Lisa, Stewart J. D'Alessio, and David Eitle. 2013. "Race and Cumulative Discrimination in the Prosecution of Criminal Defendants." *Race and Justice* 3 (4): 275–99.
- Sukhtankar, Sandip, Gabrielle Kruks-Wisner, and Akshay Mangla. 2022. "Policing in Patriarchy: An Experimental Evaluation of Reforms to Improve Police Responsiveness to Women in India." *Science* 377 (6602): 191–98.
- UN. 1994. "Declaration on the Elimination of Violence against Women: Resolution Adopted by the General Assembly [on the Report of the Third Committee (A/48/629)]." New York: UN General Assembly.
- Verma, Anshul. 2017. "Need to Amend Section 498A of Indian Penal Code and Section 41A of Code of Criminal Procedure." *Lok Sabha Debates*, August 3.
- Vries, Erik de, Martijn Schoonvelde, and Gijs Schumacher. 2018. "No Longer Lost in Translation: Evidence that Google Translate Works for Comparative Bag-of-Words Text Applications." *Political Analysis* 26 (4): 417–30.
- White, Ariel R., Noah L. Nathan, and Julie K. Faller. 2015. "What Do I Need to Vote? Bureaucratic Discretion and Discrimination by Local Election Officials." *American Political Science Review* 109 (1): 129–42.
- Zhao, Qingyuan, Luke J. Keele, Dylan S. Small, and Marshall M. Joffe. 2022. "A Note on Posttreatment Selection in Studying Racial Discrimination in Policing." *American Political Science Review* 116 (1): 337–50.