IIMB-WP N0. 700/2024

WORKING PAPER NO: 700

Fog to Pyre: The impact of Supreme Court Judgment Complexity on Dowry Deaths

Abhinav Anand
Assistant Professor
Finance & Accounting
Indian Institute of Management Bangalore
abhinav.anand@iimb.ac.in

Sanchit Jain
Doctoral Student
Indian Institute of Management Bangalore
sanchit.jain21@iimb.ac.in

Jalaj Pathak
Assistant Professor
Finance & Accounting
Indian Institute of Management Lucknow
jalaj.pathak@iiml.ac.in

Sheetal Sekhri
Associate Professor
University of Virginia
ssekhri@virginia.edu

Year of Publication – August 2024

Fog to Pyre: The impact of Supreme Court Judgment Complexity on Dowry Deaths*

Abhinav Anand[†] Sanchit Jain[‡] Jalaj Pathak[§] Sheetal Sekhri[¶]

Abstract

We use Natural Language Processing (NLP) techniques to examine if judgments' complexity in Supreme Court cases influences dowry deaths in India. Leveraging the quasi-random assignment of judges to cases, we find that a one-unit increase in the complexity of judgments captured by the Fog Index in a state year increases dowry deaths by 2% in the subsequent period. Based on estimates of the Value of Statistical Life (VSL), this increase in dowry-related homicides of women costs \$1.83 million in a state year or 51 million USD for the country annually. In otherwise similar judgments, very high complexity increases the incidence of future dowry deaths by 15%. However, having women justices on the panel of judges in such cases mitigates the effect of higher complexity.

Keywords: Dowry Deaths, Violence against women, Readability of judgments
JEL Codes: J16, J17, K41

*We are grateful to Elliot Ash, Sourabh Pal, and conference participants at the 7th Monash-Warwick-Zurich for their insightful suggestions. Author names appear in lexicographic ordering.

[†]Indian Institute of Management Bangalore, abhinav.anand@iimb.ac.in
[‡]Indian Institute of Management Bangalore, sanchit.jain21@iimb.ac.in
[§]Indian Institute of Management Lucknow, jalaj.pathak@iiml.ac.in
[¶]University of Virginia, ssekhri@virginia.edu

1 Introduction

Violent crimes against women are an acute manifestation of gender disparities often rooted in social norms and cultural practices. These crimes result in substantial costs for victims, including physical and emotional trauma as well as loss of productivity. They also impose significant costs on society due to intergenerational externalities (Carrell and Hoekstra, 2010; Carrell, Hoekstra and Kuka, 2018) and increased burdens on the healthcare and criminal justice systems.[1] These crimes are a leading cause of excess female mortality. In India, 40–50% of female homicides reported annually between 1999 and 2016 were due to dowry killings (United Nations Office on Drugs and Crime, 2019).[2] In 2021, violence against women cost 289 billion USD across Europe.

The criminal justice system plays a pivotal role in preventing these crimes. Recent evidence from the US (Miller and Segal, 2019) and India (Sukhtankar, Kruks-Wisner and Mangla, 2022; Amaral et al., 2023) shows that policing reforms can reduce gender-based violence. However, very little is known about the role of the judiciary and its rulings on crimes against women, especially in developing countries where data availability is limited. While judicial bias in developing countries has received attention (Ash et al., 2021), the complexity of rulings of judgments has evaded scrutiny. Complex judgments can impact the incentives of perpetrators and hence are very likely to impact crimes. In this paper, we take this *Beckerian* view and examine how the complexity of judicial rulings impacts crimes against women, particularly dowry deaths, a pre-meditated crime that requires planning. We do this in the context of India, which by some measures, has been identified as the most unsafe place for women.

We leverage *Natural Language Processing* (NLP) to capture and measure the com-

[1] In the United States, the cost of homicides is estimated to be 17.25 million USD.

[2] These deaths occur when brides are killed after being subjected to continuous harassment by the groom's family to extort dowry payments. Indian Penal Code (IPC) describes the death of a female as a dowry death if this occurs within 7 years of marriage and if the woman was subjected to harassment and abuse to demand more dowry.

plexity and similarity of written text.[3] Our data comprises the universe of Supreme Court judgments in India for the period 2002 to 2013. Supreme court judgments are vital in interpretation of law for lower courts and have far-reaching consequences.[4] These judgments in India are written by a bench of 2–3 judges assigned to a case. We use language models to compute two types of indices for the judgments in this universe. The first set captures the complexity of judgments—the *readability indices*—and the second set captures the similarity of judgments.

Readability is measured by the "Fog Index" Gunning (1952) which is a linear combination of two components: average words per sentence and the percentage of complex words (words with more than 2 syllabi). The index is always positive and it can be interpreted as the number of years of formal schooling required to understand the specific text in a first reading. For example, if the Fog Index for a particular text is 16 it implies that a person needs 16 years of formal education—a 4 year college degree—to comprehend the text. Hence, a higher Fog Index denotes harder-to-read, or more complex text.[5] We convert each judgment into a vector of words (tokens), and the angle between two such vectors (two judgments) is a metric for their similarity.

We also use NLP techniques to extract judge and state data for the respective judgments. This includes parsing the entire text of the judgment and extracting the names of the judges and the states for each judgment. The identification of text for crime against women is also done via NLP since manually computing complexity and similarity of the judgments has potential concerns regarding biases of the human evaluators and involves substantially higher time and cost (Ash et al., 2023). Since the corpus of judgments is quite large (11,663), our analysis is not "manual". Lin-

[3]NLP algorithms are gaining popularity in measuring similarity of texts.(see, for example Koffi, 2021)

[4]e-courts platform in India releases judicial data for lower courts since 2013 but it contains only metadata and not the entire judgments. Details can be found here: `https://www.devdatalab.org/judicial-data`. Hence, we are unable to replicate this for lower courts.

[5]We also analyze other formula-based indices of readability: the SMOG Index and the Flesch-Kincaid index.

guistically, 'readability analysis' emphasizes the syntactic aspect of the language. It focuses on word complexity as proxied via the polysyllabic nature of words and other related characteristics, such as the word length of the document and/or the average number of words per sentence. Insofar as these concepts capture judgments being hard to interpret, we measure their potential to create ambiguity and uncertainty regarding their connotation in the minds of their readers. Our approach circumvents the subjectivity inherent in having humans read and categorize the judgments as either simple or complex. We also examine several other features of judgments and conduct robustness tests, verifying that we are indeed assessing the impact of readability rather than other features such as sentiment or tone.

Equipped with our constructed indices for every judgment in our data, we construct state-by-year measures of median readability. This is the median readability index for cases that originated from a specific state in a given year related to women's issues. We then combine this data with the state-by-year crimes against women (CAW) data from the National Crimes Record Bureau. The complexity of judgments is influenced by the judges' interpretations and writing styles. For identification, we use the quasi-random assignment of judges to cases and assess the impact in a lagged two-way fixed effects model.

In our panel estimation of crimes against women, we find that an increase in the Fog index of judgments from a state in a given year increases the incidence of dowry deaths in the subsequent year. We include state and year-fixed effects in this lagged model and account for multiple hypothesis testing. A unit increase in the Fog index increases subsequent dowry deaths by 2 percent. In other words, an extra year of formal schooling needed to interpret the judgment corresponds to a 2% rise in next year's dowry death incidences. This corresponds to a loss of $2.3 million USD in terms of the value of statistical life (Viscusi and Masterman, 2017).[6] Dowry deaths are pre-meditated homicides requiring planning. Thus, complex judgments plausibly

[6]Rape and domestic violence also increase by 2.9 and 2.4 percent, respectively marginally significant at 10 percent. However, these other crimes are not statistically significant when accounting for multiple hypothesis testing.

affect the probability that perpetrators associate with being convicted and punished, influencing their willingness to commit crimes if the likelihood of punishment falls.[7] Increased judgment complexity can lower potential conviction probability owing to the uncertainty and ambiguity they engender in the minds of lawyers, judges, and potential criminals. Our estimates would be biased if readability were correlated with determinants of subsequent year dowry deaths. In our lagged model, current state conditions are not likely to affect the complexity of judgments in the year before. Nevertheless, we include a set of state-level controls: conviction rate of courts, government spending on new public goods projects, unemployment rate, police density, arrest rate, pending cases with police, police chargesheeting rate,[8] and pending cases with courts.

The Chief Justice (CJI) has discretionary powers to reassign benches and judges. If the CJI observed judges' judgment readability in year t relative to the period $t+1$ where the crime occurred and assigned them to cases in a state with high crimes against women, we could get biased estimates. To allay this concern and speak to limited interference by the CJI in the roster-based matching of judges and cases, we show that the crime incidents in the preceding year for which we include the median readability in our empirical model $(t-1)$ are not correlated with the median readability in the next year (t). Dowry deaths also do not exhibit serial correlation in states over time. To rule out home-state bias, we also demonstrate that judges are not assigned to cases from their home states. An alternative concern could be that the female victims are less motivated to report crimes and, consequently, dowry deaths in subsequent years if the readability index of women-related cases of a state increases. We find an increase in dowry deaths, likely not driven by attenuated reporting. This increase is net of effects on reporting behavior. Second, deaths are hard to hide. The findings are robust to using alternate indices for measuring

[7] Punishment can include life sentences or capital punishment.

[8] The 'chargesheet', in accordance with Section 174 of the Code of Criminal Procedure, 1973, is the document generated by the police officer subsequent to concluding a comprehensive investigation.

complexity, and small sample inference permutation tests.

Having established that an increase in the Fog index score of judgments in women's issues-related cases of a state in a given year leads to an increase in dowry deaths in the following year, we then examine if having women justices in the benches mitigates this effect. Here, we condition on the Gender Slant similarity of judgments. The Fog index median for women's issues-related cases in the preceding year continues to increase dowry deaths in the subsequent year. However, when there are more women justices on the panel, this effect is mitigated for otherwise similar judgments. Ash et al. (2021) examine judgments in India and show that there is no in-group bias among women judges. Sriram (2006) has also documented no elite bias in the assignment of cases to benches. Hence, this effect is not likely to emerge from an in-group bias of women judges towards female victims.

Our paper contributes to two strands of literature. First, we complement and extend a nascent but growing research that uses Natural Language Processing to identify gender disparities. Koffi (2021) uses NLP similarity to compare economics manuscripts of men and women economists to hold constant quality of papers in shedding light on citation omission bias. Adukia et al. (2023) examine widely read children's books and use text analysis to determine substantive gender underrepresentation.[9] Ash, Chen and Ornaghi (2024) propose a judge-specific measure of gender attitudes based on the usage of gender-stereotyped language in the judge's authored opinions and find that higher-slant judges vote more conservatively in gender-related cases in the US Circuit Courts, a finding that can explain underrepresentation of females in judiciary.[10] Our paper distills evidence that the complexity of judgments in cases related to violence against women increases the incidence of dowry death in the subsequent year.

There is a substantial body of work identifying causes that increase violence against women and, recently, what factors mitigate them. Poverty has been impli-

[9]Garg et al. (2018) use NLP to study the evolution of gender stereotyping in the US over 100 years.

[10]Other recent economics papers using NLP are reviewed in Ash and Hansen (2023)

cated in the killing of women (Rose, 1999; Miguel, 2005; Sekhri and Storeygard, 2014). An increase in female wages, on the other hand, has been shown to reduce domestic violence (Aizer, 2010; Bhalotra et al., 2021). Increasing women's representation in policing reduces CAW (Miller and Segal, 2019), and women's police stations, desks, and visible patrolling have an ameliorative effect as well (Jassal, 2020; Amaral et al., 2023). To the best of our knowledge, this is the first study that examines how judicial proceedings affect violence against women and establishes that the complexity of judgments matters for female homicides in the form of dowry deaths.

The rest of the paper is organized as follows: Sections 2 and 3 provide the contextual information on Supreme Court Judgments and dowry deaths in India, Section 4 describes the data used in our empirical work. Section 5 presents our empirical strategy. The main empirical results and robustness checks are presented in Sections 6 and 7. Section 8 provides concluding remarks.

2 Conceptual framework

Our primary motivation to focus on dowry deaths is because these are pre-meditated and require planning and coordination among several abetters, consistent with the classic framework of Becker (1998) in which the perpetrator derives utility from committing a crime and rationally weighs its costs and benefits.[11] In principle, the costs related to murdering for dowry can be high. According to the 1986 Dowry Prohibition (Amendment) Act, the guilty party receives a sentence of imprisonment for at least seven years. However, such seemingly high costs are substantially reduced owing to the prolonged court proceedings in India, which typically take several years to complete and are prone to suffer from poor conviction rates. On the other hand, there may be major benefits to committing a dowry-related murder: a remarriage brings in cashflows in the form of new dowry payments.

In India, Supreme Court cases set precedence and pertain to matters relating to

[11]Further, an extensive sociological and anthropological literature suggests that bride burning—a prominent form of dowry-related murders—requires detailed planning (Oldenburg, 2002).

interpreting the Indian constitution concerning specific crimes committed in different states over time (Bakar and Nambiar, 2023). Theoretically, all else equal, poor readability of a Supreme Court judgment creates more ambiguity and uncertainty in the minds of its interpreters, including High Court and trial court judges, lawyers, and the public. Such ambiguous, precedence-setting judgments can also be exploited by those willing to murder for the sake of dowry since the uncertainty induced by the unreadability of the judgment adds to the reduction in the expected costs of committing a crime by lowering the (already low) conviction probability. Hence, the rational Beckerian criminal estimates the average readability of the Supreme Court's judgments emanating from his state of residence at time t and forecasts future conviction probability before deciding to commit the crime at time $t+1$. For sufficiently poor levels of judgment readability, the estimated probability of future conviction can be low enough to justify a pre-meditated murder at time $t+1$.

3 Background

3.1 Dowry Killings

In India, dowry-related violence occurs after marriage when the dowry (paid at the time of the wedding) is in the hands of the husband and his family. Typically, the husband's family demands more payments after marriage, the refusal of which can lead to systematic domestic abuse.[12] In extreme cases, marital harassment escalates to murders, which frees the husband to remarry and potentially repeat the process all over again (Mullatti, 1995; Johnson and Johnson, 2001). Husbands and their families (which are usually complicit in the murder) use their kinship networks and leverage the extremely limited information available to future potential in-laws to distort facts or even conceal the occurrence of a prior marriage to remarry (Garg, 1990; Umar,

[12] Divorce and remarriage are considered taboo by the vast majority of Indians which precludes brides from walking out of abusive marriages and dowry demands.

1998; Musa, 2012).[13] In a related paper, Sekhri and Storeygard (2014) find evidence of an increase in dowry deaths in response to adverse weather shocks, suggesting that dowry killings act as an additional consumption smoothing mechanism. Despite legislative efforts and awareness campaigns, dowry-related violence and homicides persist, highlighting the complex social, economic, and cultural factors contributing to these heinous crimes.[14] Legally, dowry death became a crime in 1986 under the Dowry Prohibition (Amendment) Act of the Indian Penal Code (IPC). Any unnatural death of a woman within 7 years of marriage if harassed for dowry before the death is considered a dowry death. In case of suicide, section 306 of the IPC is also applicable.

Table 1 presents the summary statistics of dowry-related deaths across states in India during 2002–2013. The National Capital Territory (NCT) of Delhi (1.51), followed by Bihar (1.18), Madhya Pradesh (1.07) and Haryana (1.04) exhibit the highest mean dowry death rates in the country.

3.2 Reporting Bias

Dowry-related violence in India is often shrouded in a culture of silence, where victims hesitate to report the abuse until it reaches a breaking point. This phenomenon can be attributed to a complex interplay of socio-cultural, economic, and psychological factors.[15] Firstly, social stigma is pivotal in discouraging victims from speaking out. The prevailing patriarchal norms burden women immensely to maintain family honor and uphold societal expectations. Reporting dowry-related violence may be perceived as disgracing the family, leading victims to endure the abuse silently rather than jeopardizing their family's reputation. Furthermore, victims may fear social ostracism and isolation if they seek help or escape the abusive situation, leaving

[13] Typically, marriages tend to occur across villages separated over considerable distances which amplifies information asymmetry.

[14] See related news items: i) https://www.cnbctv18.com/india/19-women-were-killed-for-dowry-every-day-in-2020-ncrb-10758421.htm; and ii) https://www.bbc.com/news/world-asia-india-57677253

[15] https://economictimes.indiatimes.com/news/india/death-of-three-sisters-spotlights-india-dowry-violence/articleshow/92070742.cms?from=mdr

them with no support network and nowhere to turn for assistance.

Secondly, economic dependence can act as a barrier to reporting dowry-related violence. Many victims rely on their husbands or in-laws for financial support, making them vulnerable to exploitation and abuse.[16] The fear of economic insecurity, homelessness, or being burdened by the dowry repayment may deter victims from speaking up against their abusers. Additionally, the patriarchal control over women's finances in some households restricts their ability to seek legal aid or escape the abusive environment, further perpetuating their silence. As a result, victims often endure the violence until it reaches an extreme point, leading to dire consequences for their physical and mental well-being. The intersection of societal expectations and economic dependence creates a formidable barrier that prevents many victims from seeking help until it is too late.

Weak enforcement of existing laws and lax judicial processes contribute to the perception that such crimes can be committed with impunity. Furthermore, social stigma and the fear of disgrace prevent many victims and their families from seeking justice or reporting such crimes, enabling perpetrators to evade accountability.

3.3 Supreme Court Judgments

In this section, we provide a basic primer on how individual cases reach the Supreme Court of India and how judges in the Supreme Court get assigned to individual cases, both of which are important ingredients in informing the interpretation of our results.

3.3.1 How do Cases Reach the Supreme Court?

In India, the Supreme Court is set up by the Constitution of India and is the highest judicial authority and the final court of appeal.[17] A case reaches the Supreme Court

[16] https://www.livemint.com/money/personal-finance/59-working-women-do-not-make-their-own-financial-decisions-survey-11665555051026.html

[17] Article 124, Chapter IV of the Constitution. See link here: https://legislative.gov.in/constitution-of-india

through a hierarchical litigation process. The journey of a case to the Supreme Court typically begins at the trial court level, which can be a district court or a subordinate court. If any party is dissatisfied with the trial court's judgment, they can appeal to the appropriate High Court within the state. The High Court acts as the appellate court and reviews the decisions made by the lower court. If the High Court judgment is also contested, the aggrieved party can further escalate the matter to the Supreme Court of India.

To reach the Supreme Court, the aggrieved party files a Special Leave Petition (SLP) or a petition for appeal, stating why they believe the case warrants the Supreme Court's attention. The Supreme Court holds discretionary power to grant or reject the SLP based on the case's merits or if it involves a substantial question of law. Once the SLP is admitted, the case is presented before the Supreme Court, and the Justices of the Court review the arguments and evidence to arrive at a final, binding, and conclusive verdict.

3.3.2 Adjudication of Cases

Judgments by the Supreme Court are delivered in open court with the concurrence of most judges present at the case hearing. The judge's panel writes the judgment.[18] The Assistant Registrar maintains a record of proceedings. A copy of the judgment passed is maintained on the Supreme Court website and other private portals.[19] We obtain the judgment text from one such portal.

3.4 Assignment of Supreme Court Judges to Cases

No written procedure is followed to assign cases to judges. We interviewed two high court lawyers who have filed cases at the Supreme Court to ascertain the process. We also culled newspaper articles to determine the process and found the description

[18] Article 145 of the Constitution. Link: https://legislative.gov.in/constitution-of-india.

[19] See link: https://main.sci.gov.in/judgments

based on interviews with 4 Justices who serve on the Supreme Court. The process is as follows: cases are assigned based on a roster system.[20] When a case is filed, the Supreme Court registry officials scrutinize its details and subject matter and receive and process all documents. Cases are categorized based on subject matter. There are 47 broad categories such as, letter petitions, public interest matters, taxation, service matters, and criminal appeals. Each category has multiple sub-categories. The registry officials notify the roster for the benches to hear a case. This is done based on the subjects (or categories), and the Chief Justice of India approves it. More than one bench can be allocated the same subject matter. Details of a new case are entered into a computer, which automatically assigns the matter to the bench. The computer marks the matter to the bench as per the roster. If there are three benches dealing with a particular category, then the matters are marked sequentially to the three benches.[21]

The Chief Justice of India (CJI) prepares the roster, which is reviewed periodically. The Chief Justice of India also has the power to transfer cases from one bench to another, either to ensure that judges with the relevant expertise hear the case or that the workload is distributed equally among the judges. This power is exercised on a case-by-case basis and is intended to ensure that justice is served fairly and efficiently. Since this is in the public eye and scrutinized by the press, there are no known instances where the CJI assigned judges based on corrupt practices.

The assignment of judges to cases is quasi-random and not based on the origin of cases or the domicile of judges. This is important since in principle, it is possible that in penning judgments, judges act differently when presiding over cases that originate from their state. To demonstrate that the assignation of cases originating from a particular state is usually different from the state of domicile of the presiding judges, we compile Table A1, which presents the distribution of cases originating

[20]The roster of the Supreme Court is essentially a list of judges that specifies the areas of law in which each judge specializes.

[21]https://www.hindustantimes.com/india-news/sc-judges-vs-cji-dipak-misra-how-cases-are-allocated-in-india-s-top-court/story-fEQNyONGXpcoJY9GThAkdO.html

from each state as well as the count and percentage of judges with domicile from that state. Overall, for 57 out of the total of 1131 judgments, there is at least one presiding judge whose state of domicile matches the case's state of origination. In other words, about 5% of the cases feature judges from the same state as its state of origination. The highest proportion of such examples is from the state of Bihar (10 out 49 cases: $\sim 20\%$), followed by Gujarat (4 out of 35: $\sim 11\%$) and Uttar Pradesh (8 out of 77: $\sim 10\%$). In other words, the frequency of judges presiding over cases from their home state is relatively low, and there is no systematic bias in favor of assigning cases to judges from the same state.

Similarly, Table A2 illustrates the percentage of panels across states (from where cases originate) with at least one female judge. The numbers are uniformly low, with the highest number recorded for West Bengal (9.4%), followed by Uttar Pradesh (5.2%) and Punjab (5.1%). The overall mean proportion of panels with female presence is quite low, at 3.8%.

3.5 Readability of Judgments

A few recent studies have examined the nature of legal judgments from a readability perspective. Merchant and Pande (2018) present a new technique for summarising otherwise large and complex legal judgments into smaller and more comprehensible text based on Latent Semantic Analysis (LSA). On similar lines, Bansal, Sharma and Singh (2019) apply a fuzzy AHP-based automatic text summation technique for summarizing legal judgments based on their size and readability. Shridhar, Kayalvizhi and Thenmozhi (2021) specify how a pre-specified knowledge of "rhetoric roles", i.e., the sections in the thematic structure of the legal judgments e.g., "facts of the case", "issues being discussed" and "arguments given by parties" etc. improves the readability of the judgments.

While existing literature has predominantly concentrated on readability and topic classification within legal judgments, the objective numerical classification of readability of judgments and its potential influence on the incidence of crime remains an unexplored topic. We aim to fill this gap by examining this issue.

The level of judgment readability is variable for the same crimes. In Table A3, we present four judgments about dowry from our sample and demonstrate that they exhibit vastly different degrees of readability. Consequently, the level of readability cannot be solely attributed to the nature of the crime but also depends on additional factors, such as the writing style employed by the presiding judges.

4 Data and Measurement

We start the analysis with 11,663 judgments. Judges' and states' data are available for 5472 judgments. We analyze the text of these judgments to find those related to crime against women and arrive at 334 judgments. Among the 6,191 (11,663 - 5472) judgments dropped due to missing state and/or judge data, 433 judgments are related to crime against women. Table A8 presents the p-values for the difference in means of complexity (Fog Index) for the final sample (334 judgments) and the discarded full sample (6,191) as well as CAW judgments amongst the discarded sample (433). The mean difference is not significant for either group. Figure A1 presents the density plot of the Fog Index for the CAW judgments of the chosen sample (334) and the discarded sample (433). These distributions are similar. Kolmogorov-Smirnov test p-value is 0.14 for the chosen and the discarded sample. Hence, our sample is not systematically different than the population of judgments in terms of the Fog Index, implying this is not a likely source of bias.

4.1 Judgment Data

We utilize three primary sources of data in our analysis. Our first data source is 'Advocate Khoj'—an Indian portal for legal matters—from where we download Supreme Court judgments from 2002 to 2013.[22] Further, we extract relevant details

[22]Link for Supreme Court judgments can be accessed here: `https://www.advocatekhoj.com/library/judgments/index.php?go=2022/january/indexfiles/index1.php`

from each judgment, such as dates of judgments, names of presiding judges, and states from which cases originate.

The format of the judge's name in the judgment is inconsistent. For instance, the Hon'ble Justice "**First Name, Last Name**" is variously mentioned as '**Justice Last Name**' or '**Last Name, First Name**'. To overcome this problem, the list of judges' names from the Supreme Court of India website is used to identify each judge's name and then mapped to a standard format. The photograph attached to the judge's profile helps identify and map the gender of each judge. The extracted state names are used to map judgments to states. This dataset is then combined with the extracted date of judgment to create a state-year panel of crimes.

4.2 Judge Characteristics Data

We obtain the detailed profiles of the Supreme Court judges from the Supreme Court website.[23] We extract the following information from each profile: name of the judge, picture, date of birth, date of appointment as Supreme Court Justice, and date of appointment as CJI (where applicable). We analyze the picture of the judge to determine the gender. We compute the judge's age using the date of birth and the date of judgment. Similarly, we calculate the experience as a Supreme Court Justice by taking a difference between the date of appointment and the date of judgment. If the date of appointment as CJI is before the date of judgment, we assign the value of 1 to CJI and 0 otherwise.

4.3 Crimes Data

We use the Indian National Crime Records Bureau (NCRB) data on the number of crime incidences for the period 2002–2013. The local police encode crime data at the police station/district level, which then gets consolidated at the state level by the state police, which in turn gets aggregated at the national level by the NCRB.[24]

[23] https://main.sci.gov.in/

[24] NCRB uses the Principal Offence Rule for counting crimes (NCRB, 2020).

4.4 Other Data

We employ the Center for Monitoring Indian Economy's 'States of India' database to collect data on socio-economic controls, including that on unemployment, police density, arrest rate, pending cases with police, police chargesheeting rate,[25] pending cases with the court, conviction rate of the court and government spending on new projects.[26]

Finally, we use the World Bank's data on 'India States Briefs' to collect indicators related to the condition of women across different states of India.[27] We include data on the 'Child Sex Ratio' and 'Maternal Mortality Rate'.[28]

4.5 Measurement of Judgments' Readability

One of the most frequently used readability measures is the 'Fog Index'. First published in Gunning (1952), the Fog index is used to measure readability as a linear combination of the length of sentences and the proportion of complex words in each sentence. An informal interpretation of the index is the number of years of education needed to understand the text in the first reading. For example, a text with Fog index 16 means that the reader needs to have a college degree to interpret the text; another text with a value 18 corresponds to a reading ability associated with a post-graduate education and so on.

Formally, the Fog index is defined as follows:

$$\text{Fog} = 0.4 \times (\text{average words per sentence} + \text{percentage of complex words}) \qquad (1)$$

[25] In India, a 'chargesheet' refers to the set of criminal charges an individual is accused of. Once the chargesheet gets submitted to a court, the court pronounces orders regarding the framing of charges after which prosecution proceedings against the accused can begin.

[26] All variables are defined in detail in Table A7 in the appendix.

[27] The data is accessible at https://www.worldbank.org/en/news/feature/2016/05/26/india-states-briefs

[28] i) Child Sex Ratio is defined as the number of girls per 1,000 boys in the 0-6 age group; and ii) Maternal mortality ratio is computed as the ratio of maternal deaths per 1,00,000 live births reported.

The Fog index is an example of formula-based readability metric in which a formula links two critical components of a text: i) the average number of words in each sentence, and ii) the proportion of complex words in the text. These metrics use mathematical formulas that take into account the linguistic features of the text to provide a numerical readability score, typically associated with levels of formal schooling needed to interpret the text at a first reading.

To operationalize the definition of a sentence in a court judgment, we delineate it as the collection of words between i) two full stops; ii) a full stop and a question mark; and iii) two question marks.We compute the average words per sentence as the ratio of the total number of words in the text to the total number of sentences in the text; and define 'complex words' as polysyllabic words with at least three syllables.

Two alternative examples of formula-based readability metrics are the SMOG index (Laughlin, 1969) ('Simple Measure of Gobbledygook'); and the Flesch-Kincaid index (Kincaid et al., 1975), both of which are also functions of the proportion of complex (polysyllabic) words and the average number of words per sentence. Both these formula-based metrics are intended to be mapped to the years of formal schooling the reader ought to possess, to be able to interpret the text.[29] These readability scores provide insights into the text's complexity, structure, and readability; and offer quantitative indicators of its difficulty level which can be used to assess its accessibility to readers of differing education levels and abilities.

To demonstrate how text readability metrics can be used to judge the readability of court judgments, we use an extract from one of the case in our sample:[30]

> *"Taking all the facts and circumstances of the case into consideration in totality, it appears that the order to the extent of summoning the petitioners,* ***Mr. X*** *and* ***Mr. Y****, through non-bailable warrants does not appear justified and is liable to be quashed and set aside. However, the petitioners,* ***Mr. X*** *and* ***Mr. Y****, are directed to surrender before the learned*

[29]Exact formulas of the SMOG and Flesch-Kincaid indices are included in the appendix in Table A7.

[30]The actual names in the case are not specified for privacy concerns.

trial court and to move the application for their regular bail, which will be considered and allowed by that court on the same day on which it is moved."

The judgment fragment's readability, according to the Fog index is calculated as:

- Number of sentences: 2
- Average words per sentence: $89/2 = 44.5$
- Percentage of complex words: $16/89 = 0.180$
- Fog Index: $0.4 \times (44.5 + 100 * 0.180) = 24.99$.

Based on a Fog index value of about 25, one can conclude that the sample sentences have low readability and presumably require long years of formal education to parse and interpret them adequately.

4.6 Measuring Judgments' Similarities

We divide our judgment sample into state-year combinations and calculate their similarities using the cosine similarity measure for each judgment text. The cosine similarity measure quantifies the similarity of judgments by calculating the angle between text-vectors in a suitable linear space.

$$\text{Cosine Sim(A,B)} = \frac{\langle A, B \rangle}{\|A\| \|B\|} \tag{2}$$

where $\langle A, B \rangle$ denotes the inner product of vectors A and B; and $\|A\|$ and $\|B\|$ are the Euclidean norms (magnitudes) of text-vectors A and B, respectively (Ash and Hansen, 2023; Koffi, 2021). When the text-vectors are closely aligned, implying that corresponding terms have similar importance and distribution in the documents, the angle between the vectors is close to 0 and its similarity approaches 1. Conversely, when the vectors are dissimilar, the angle is close to being orthogonal ($= \pi/2$) the similarity measure equals 0, indicating dissimilarity.

We subdivide each group (state-year combination) into three sub-groups based on the similarity of the judgments using the similarity measure.[31] For each sub-group, we calculate the ratio of judgment panels containing a female judge to the total number of judgments. We use the average across the sub-groups to determine the share of judgments with female judges in the panels.[32]

4.7 Judgments Related to Crimes Against Women

To assess the impact, if any, of judgments' readability (or lack thereof) on dowry deaths, it is important to confine the analysis to only those judgments that pertain to the category of 'crimes against women' ('CAW' henceforth).

To this end, we first compile a list of keywords that capture special characteristics of text related to crimes against women. To identify these terms, we use the following procedure: i) elimination of stop words (semantically insignificant words), ii) identification of all verbs and nouns in the judgment according to the Merriam-Webster dictionary, iii) arrangement of the verbs and nouns in order of frequency of occurrence, iv) exclusion of all terms with frequency $\leq 5\%$, and v) identification of top terms usually used to characterize crimes against women. The final keyword list thus obtained features the following terms: "woman," "women," "female," "girls," "girl," "rape," "dowry," "eve-teasing,"[33] "stalking," "voyeurism," "molestation," "obscenity," "abduction," "kidnapping," "prostitution," and "witch-hunting."

Finally, to classify a judgment as related to crimes against women, we quantify the percentage of sentences containing at least one of the above-identified keywords. We define this as the judgment's "CAW percentage." Judgments with an above-median CAW percentage are classified as judgments regarding crimes against women.

[31] Our results are robust to other sub-group classifications ranging from 2-5.

[32] Our results are robust to the use of median values of female judge percentage

[33] This is a commonly employed Indian-English euphemism for street sexual harassment.

4.8 Descriptive Statistics

Table A4 presents the descriptive statistics of variables used in the analysis. These feature the incidences of crimes against women aggregated over the various types of CAW, including rape, domestic violence, dowry deaths, and molestation.[34] It also reports summary statistics for judgment readability metrics such as Fog, SMOG and Flesch-Kincaid indices; as well as that for socio-economic control variables such as unemployment rate, police density, conviction rate etc.

In particular, for all metrics of judgment readability, the values indicate uniformly high levels of word complexity and/or sentence length. Hence, judgments are not easily readable owing to long sentences and the presence of a high frequency of polysyllabic jargon. In other words, the judgments suffer from poor readability and the text requires expert knowledge for interpretation.

Further, table A5 presents the correlation of the Fog Index with all control variables used in this study. The Fog Index is not significantly correlated with any of the controls and the highest correlation is with population which is 0.14 (not significant).

Figure A2 presents the Lowess plot of the Fog Index and the gender slant (Ash, Chen and Ornaghi, 2024) of the judgments forming part of our sample. The blue line represents the association between the Fog Index and Gender Slant, with the gray shaded area indicating the confidence interval. There is no systematic pattern between the two. The overall correlation coefficient is 0.11 (not significant).

5 Estimation Strategy

We investigate the impact of Supreme Court judgments' readability (or lack thereof) on crime incidents against women using the following lagged panel specifications:

$$CAW^{j}_{i,t+1} = \beta \text{ Readability}_{i,t} + \Gamma Z_{i,t+1} + \lambda_s + \delta_t + \epsilon_{i,t} \tag{3}$$

$$CAW^{agg}_{i,t+1} = \gamma \text{ Readability}_{i,t} + \Omega Z_{i,t+1} + \lambda_s + \delta_t + \epsilon_{i,t} \tag{4}$$

[34] Detailed definitions are reported in the appendix in Table A7.

λ_s and δ_t represent state- and year-fixed effects respectively. We cluster the standard errors by state and year to account for potential heteroskedasticity and autocorrelation. The identifying assumption is that the lagged $Readability_{i,t}$ is orthogonal to the residual variation in contemporaneous crimes against women due to the quasi-random assignment of judges to cases conditional on state and year-fixed effects.

Given the multiple tests conducted in the study, it is essential to address the inherent risk of Type I errors, which occur when the null hypothesis is incorrectly rejected. To mitigate this risk, False Detection Rate (FDR) analysis is employed. We apply the widely used Benjamini-Hochberg (BH) procedure for the FDR analysis (Benjamini and Hochberg, 1995).

5.1 Correlates of Fog Index-Judge Characteristics

The complexity of a judgment is influenced by the writing styles of the judges comprising the panel. Hence, judge characteristics could potentially have an impact on the Fog Index of the judgments written by the panel of judges. We test this hypothesis using the following equation:

$$\text{Fog Index}_j = \text{Male}_{ij} + \text{SC experience}_{ij} + \text{Judge Age}_{ij} + \text{CJI}_{ij} + \epsilon_t \tag{5}$$

where, Fog Index_j is the Fog Index of the judgment j, Male_{ij} is a dummy variable indicating the gender of judge i, taking the value one for Male judge and 0 otherwise, SC experience_{ij} indicates the experience of judge i since appointment as a Supreme Court judge, as on the day of passing the judgment j, Judge Age_{ij} indicates the age of judge i as on the day of passing the judgment j, CJI_{ij} is a dummy variable, taking the value one where the judge was appointed as the Chief Justice of India as on the day of passing the judgment and 0 otherwise.

Our results are presented in Table A6. We find a positive and significant association between the judge's gender and the readability of the judgment penned by the judge panel. A higher number of male judges is associated with a higher Fog Index (lower readability) of the judgments. Having judges with more experience and those who have served as CJI on the panel also increases the complexity although this is imprecisely estimated.

5.2 Testing Serial Correlation

One concern can be that the CAW has serial correlation and CAW in previous years affects the lagged *Readability* measure we use. To address this, we first show that CAW, especially dowry deaths, are not serially correlated.

We conduct the Breusch-Godfrey test (Breusch, 1978; Godfrey, 1978) to examine the presence of serial correlation in the crimes against women across different states. The results show a p-value of 0.45, indicating an absence of serial correlation.

5.3 Reverse Causality

We further verify that the previous year's CAW does not affect the readability of the judgments in the subsequent year. The results are highlighted in Table 2. In the first column, we show that the coefficient on the first lag of CAW is small and statistically insignificant. Columns 2 and 3 show this for the second and third lag, respectively. Column 4 shows that the coefficient continues to be insignificant for a combination of all three lags.

6 Results

6.1 Main Findings

The results from the estimation of equations (3) and (4) are presented in Table 3. The first four columns include the specification in equation (3) for different types of crimes against women, namely rape, domestic violence, dowry deaths, and molestation. The final column presents the result for the specification in equation (4) in which the dependent variable is the aggregate of all types of crimes against women.

For the case of dowry deaths, we find the coefficient to be positive and statistically highly significant (BH q-value 0.001), implying that larger Fog indices (poorer readability) of judgments in year t are associated with significantly more dowry deaths in year $t+1$. In particular, a unit increase in the judgments' Fog index—an additional year of formal schooling required to interpret the text—corresponds to a 2.02% rise in the dowry deaths next year. We also observe an increase in rape and domestic violence, but these are less precisely measured. Permutation tests for robustness of the small sample inference are discussed in Section 7. One concern might be that readability is positively correlated with the judge being male, and the gender of the judges has an effect on crimes in the following year. To allay this concern, we control for the percentage of male judges in panels in a state year in our specification and present the results in Table 4. Dowry death is unencumbered, indicating that it is impacted by readability.

6.2 Economic Cost

To determine the economic cost associated with dowry deaths, we convert the percentage change into numbers by considering mean dowry deaths. The mean for dowry deaths is 328 women in our sample. We multiply this average by 2.02% to calculate the increase in dowry deaths associated with a unit rise in the (median) judgment Fog Index, i.e., $328.66 \times 2.02/100 = 6.64$. This is then multiplied by the value of statistical life in India as per Viscusi and Masterman (2017), which is 275,000 USD. In other words, a unit rise in the formal schooling level needed to interpret a Supreme Court judgment corresponds to a loss of $275,000 \times 6.64 = 1.83$ million USD.

6.3 Do Female Judges Matter?

We previously provided evidence that poorer readability of judgments in CAW cases corresponds to increases in subsequent year's incidences of dowry deaths. We now examine whether the gender composition of judge panels amplifies or ameliorates this effect.

We construct a dummy variable 'Fog_high', which assumes the value of 1 when judgments' readability is greater than the 20th percentile of Fog Index and 0 otherwise.[35] The following regression specification is employed:

$$\begin{aligned} Dowry\_deaths_{i,t+1} = \beta_1 \; Fem\_judge\_\%_{i,t} + \beta_2 \; Fog\_high_{i,t} \\ + \beta_3 \; Fem\_judge\_\%_{i,t} \times Fog\_high_{i,t} + \Gamma Z_{i,t} + \lambda_s + \delta_t + \epsilon_{i,t} \quad (6) \end{aligned}$$

where, $Dowry\_death_{i,t+1}$ is the logarithm of dowry deaths in the state i in year $t+1$, $Fem\_Judge\_\%$ is the percentage of judgments where a female judge is a part of the panel that presides over the case, Fog_high is an indicator variable taking the value of one when Fog Index exceeds the 20^{th} percentile benchmark and zero otherwise. Similarly, $Smog\_high$ and Fk_high are indicator variables that take the value of one when Smog Index and Flesch Kincaid Index, respectively, exceed the 20^{th} percentile benchmark, and zero otherwise.

Table 5 presents the results for equation (6). To do this analysis, we condition on natural language processing-based similarity of judgments. This allays the concern that women might be in benches that are assigned different types of cases. In this pool of similar judgments, poorer readability of judgments corresponds to significantly increased dowry deaths in the subsequent year, even after controlling for judgments with female judges. The coefficient on the dummy Fog_high indicates a 14.9% increase in subsequent dowry deaths. This implies that a unit increase in Fog index when it exceeds the 20th percentile leads to an especially high female mortality on

[35] Appendix Figure A1 shows the density of Fog Index. Results are similar if we use bottom 10% and bottom 1% of readability to define *Fog high*.

account of dowry deaths. However, the interaction effect is negative implying having female judges in panels where readability is poor, ameliorates the impact by 1.3%. Columns (2) and (3) present similar results for readability measures Smog Index and Flesch kincaid.

6.4 Heterogeneity Analysis: Does the Progressive Status of Women Matter?

We verify if the results vary according to women's status in different states. To test this, we use the Child Sex ratio obtained from World Bank for 1995. In states with progressive attitudes towards women, we expect that the child sex ratio would be less skewed.[36] We divide the states in our sample into two categories based on the mean value of the indicator. Our results are reported in Table 6. In column 1, we present our main specification and interaction with the Child Sex Ratio dummy. Columns 2 and 3 present results for alternate readability measures: SMOG and Flesch-Kincaid. While the lagged readability measure remains significant, the interaction is negligible and statistically insignificant. The results imply the impact of lagged judgment readability on subsequent periods' dowry deaths is not significantly different in states with better status of women.

7 Robustness

7.1 Alternate Readability Metrics

Table 7 presents the impact of the readability of the judgments on the incidences of crimes against women using two alternative formula-based metrics for readability: the SMOG Index and the Flesch-Kincaid Index. The results are similar to benchmarks presented in Table 3 with the readability metrics showing significant positive

[36]Previous research documents gender disparities in regions of India in sex ratios. Despite biological advantage at birth, there are fewer girls per 1000 boys. As per the population census 2011, there were 918 girls per 1000 boys in ages 0-6 in India.

associations between judgments' readability and future dowry deaths. In particular, a unit increase in the SMOG index is associated with 4.1% rise in next year's dowry deaths, while a unit rise in the FK index corresponds to an increase of 1.8% in future dowry deaths.

7.2 Permutation Test

We have 111 state-year combinations for our analysis. We perform permutation tests using 1000 iterations to ensure the robustness of our inference to small sample bias. Our results are presented in Table 8. The results for dowry deaths continue to be significant, in line with Table 3. We further extend our sample period described in a later subsection and do not observe a change in our benchmark results.

7.3 Reporting Bias

Can poor readability lead to an increase in reporting and could it be that we are picking up only increased reporting instead of increased incidence? Low judgment readability, if anything, would delay judicial resolutions and amplify the range of valid interpretations of cases for lower courts.[37] This should deter reporting by women, not increase it. Moreover, homicides due to dowry are not easy to hide. Our heterogeneity analysis also casts doubt on the plausibility of the hypothesis that results are explained by changes in reporting. If this were the case, and an increase in dowry deaths was due to increased reporting, we would observe higher reporting effects in more states with a more progressive status of women. However, as we demonstrate in Table 6, the results are agnostic to the status of women in the state. Hence we think it is highly unlikely that this factor is driving our findings.

[37] To the extent there is under-reporting overall in CAW, we estimate lower bounds of the effects.

7.4 Excluding Data for Judgments related to the State of Bihar (Highest State Domicile Judges Assigned to CAW Cases)

We noted in Table A1 that around 20.41% of the cases originating from Bihar have Supreme Court judges whose domicile is Bihar. To show that our results are not driven by home state bias, we exclude Bihar as a robustness exercise. Table 9 presents the results from the estimation of the equations (3) and (4) excluding data for judgments related to Bihar. The results are similar to benchmarks presented in Tables 3 and 7 with the readability metrics showing significant positive associations between judgments' readability and future rates of dowry deaths.

7.5 Enhanced Sample Analysis

To demonstrate that the choice of the sample period from 2002 to 2013 is not generating results, we extend our analysis to a longer duration from 1996 to 2013. Table 10 presents results of estimating equation (3) on the larger sample. The advantage of this exercise is that we have a larger sample, but the disadvantage is that we do not have all the controls in the specifications. Reassuringly, however, we find that results are positive and significant for dowry deaths, i.e., a fall in the readability of judgments is associated with an increase in future dowry death incidents. As before, the magnitude of rape and domestic violence is large and positive, but the estimates are imprecise for a conclusive empirical assertion.

7.6 Additional Text Measures

We ensure that the impact of judgment complexity on dowry deaths is not due to other text-based dimensions of the judgments. Thus, we quantify three more text-based measures based on the following dimensions – sentiment, semantic complexity, and Gender Slant (Ash, Chen and Ornaghi, 2024). The methodology is specified in section 8.

Tables 11 and 12 present the results for the impact of judgment readability on dowry deaths for next year with combinations of the valence shifters, sentiment and gender slant as additional controls and the effect of fog index is still positive and significant on next year dowry deaths.

7.7 Impact of Trial Duration

Another concern is that complex judgments stem from trials that last a long time, and we are discerning the effect of duration rather than complexity. We allay this concern by computing the duration of the trials and controlling for it. We determined the duration of each trial by calculating the difference between the year in which the appeal was filed with the Supreme Court and the year in which the judgment was rendered. Table 13 presents the results of equations (3) and (4), with trial duration included as an additional control variable. These results remain consistent with the main findings in Table 3.

7.8 Impact of Media Attention

Relatedly, complex cases could garner media attention and influence how perpetrators behave. We create a measure of media salience/attention and control that in our analysis. Using the Factiva database,[38] we analyze each of the 334 judgments in our final sample and obtain the number of news articles featuring the case before judgment is announced. We find that the media has covered very few cases before judgment. We use an indicator variable *Media Attention*, which takes the value one if the judgment had prior news coverage, else zero. Our results are presented in Table 14. The results are consistent with the main finding presented in Table 3.[39]

[38]Factiva, a business intelligence platform operated by Dow Jones, aggregates content from 33,000 news, data, and information sources spanning 200 countries and 32 languages. It is widely utilized for conducting comprehensive analyses of media coverage.

[39]The results remain consistent if we use a continuous variable (number of news articles) instead of the indicator variable.

8 Conclusion

We use Natural Language Processing to determine the complexity of Supreme Court rulings/ judgments in India. A higher complexity of judgments in a state year leads to a statistically significant increase in dowry deaths in the subsequent period. Female homicides by way of dowry deaths are indicative of disparities against females that disempower them. We highlight that judicial judgment complexity in apex courts exacerbates this. The institutions meant to protect citizens have contrary effects. Training of justices to address complexity can play a role in reducing this effect. Our results also underscore the potential influence of gender representation within the judiciary—specifically the inclusion of female judges in Supreme Court panels—in addressing and mitigating the occurrence of dowry deaths. The findings suggest that the unique perspectives, experiences, and insights brought by female judges may contribute to a more nuanced and empathetic approach toward cases involving dowry-related crimes, resulting in improved outcomes and reductions in dowry deaths. This evidence highlights the importance of diversifying judicial representation to promote gender equality and enhance the legal system's effectiveness in combating violence against women.

Tables

Table 1: Dowry death rates across states 2002–2013

State	Min	Mean	Median	SD	IQR	Max
Assam	0.20	0.33	0.31	0.11	0.13	0.51
Bihar	1.01	1.18	1.23	0.11	0.15	1.32
Chhattisgarh	0.26	0.38	0.39	0.06	0.07	0.50
Gujarat	0.03	0.07	0.07	0.03	0.06	0.13
Haryana	0.87	1.04	1.03	0.12	0.10	1.33
Himachal Pradesh	0.01	0.07	0.05	0.05	0.07	0.16
Jharkhand	0.75	0.86	0.86	0.06	0.07	0.96
Karnataka	0.31	0.40	0.41	0.04	0.05	0.46
Kerala	0.04	0.07	0.07	0.02	0.03	0.10
Madhya Pradesh	0.96	1.07	1.09	0.06	0.10	1.17
Maharashtra	0.25	0.33	0.31	0.04	0.06	0.41
NCT of Delhi	1.26	1.51	1.46	0.20	0.28	1.88
Punjab	0.27	0.39	0.40	0.07	0.08	0.53
Rajasthan	0.59	0.63	0.63	0.03	0.03	0.68
Tamil Nadu	0.15	0.26	0.28	0.06	0.09	0.34
Uttar Pradesh	0.75	1.04	1.07	0.14	0.14	1.32
Uttarakhand	0.40	0.78	0.76	0.16	0.16	1.05
West Bengal	0.32	0.47	0.50	0.08	0.10	0.59

Note: This table presents summary statistics of the dowry death rate for different states across the sample period. 'SD' and 'IQR' refer to standard deviation and interquartile range, respectively. We compute states' yearly dowry death rates by dividing the total dowry deaths in that state by the state's population in that year; and report summary statistics for dowry death rates for each state over the whole sample period 2002–2013, by computing means/medians etc. for the yearly death rates.

Table 2: Impact of crime incidents on judgment readability in subsequent year

	Dependent variable:			
	Fog Index (t)			
	(1)	(2)	(3)	(4)
CAW Incidents (t-1)	2.401			1.238
	(1.667)			(1.871)
CAW Incidents (t-2)		2.675		−0.431
		(2.123)		(2.897)
CAW Incidents (t-3)			2.435	2.159
			(2.125)	(1.554)
R-squared	0.047	0.055	0.041	0.046
Observations	111	98	84	84

Note: This table reports the results from the regression of the readability index (Fog) of the judgments related to crimes against women on lagged values of crime incidents. ***, ** and * indicate that the coefficient estimates are significantly different from zero at the 1 percent, 5 percent, and 10 percent levels, respectively.

Table 3: Impact of judgment readability on the incidents of crimes against women

	Crime Incidents (t+1)				
Readability Index	Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Fog Index (t)	0.029^{+}	0.024^{+}	0.020^{***}	0.001	0.018^{+}
	(0.017)	(0.014)	(0.004)	(0.007)	(0.010)
BH q-value	0.439	0.439	0.001	0.812	0.439
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.532	0.631	0.210	0.214	0.633
Observations	111	102	111	102	102

Note: This table reports the results of equations (3) and (4) from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year. BH q-value indicates the cut-off above which we fail to reject the null hypothesis for the regression equation. '+' denotes specifications which are significant at 10% for traditional standard errors, but BH q value is not significant.

Table 4: Impact of judgment readability on the incidents of crimes against women, controlling for percentage of male judges

	Crime Incidents (t+1)				
Readability Index	Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Fog Index (t)	0.029*	0.024*	0.021***	0.002	0.018*
	(0.016)	(0.014)	(0.004)	(0.005)	(0.010)
BH q-value	0.109	0.109	0.001	0.747	0.109
Control- Male Judge	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.533	0.632	0.212	0.228	0.635
Observations	111	102	111	102	102

Note: This table reports the results of equations (3) and (4) from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women and control variables including share of male judges, unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year. BH q-value indicates the cut-off above which we fail to reject the null hypothesis for the regression equation.

Table 5: Impact of female judge presence on dowry deaths

	Dowry Deaths (t+1)		
Readability Index	(1)	(2)	(3)
Female Judge % (t)	0.012**	0.012**	0.003
	(0.005)	(0.006)	(0.005)
Fog High	0.149***		
	(0.048)		
Female Judge % (t)* Fog High	−0.013***		
	(0.005)		
Smog High		0.125**	
		(0.050)	
Female Judge % (t)* Smog High		−0.013**	
		(0.005)	
Fk High			0.155***
			(0.042)
Female Judge % (t)* Fk High			−0.002
			(0.006)
Controls	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes
R-squared	0.233	0.217	0.239
Observations	111	111	111

Note: This table reports the results of Equation (6) from the regression of dowry deaths on percentage of female judges in the judicial panel, readability measures of the judgments related to crimes against women, and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases and new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year. ***, ** and * indicate that the coefficient estimates are significantly different from zero at the 1 percent, 5 percent, and 10 percent levels, respectively.

Table 6: Heterogeneity Analysis

	Dowry Death Incidents (t+1)		
Readability Index	(1)	(2)	(3)
Fog Index (t)	0.026***		
	(0.010)		
Smog Index (t)		0.048***	
		(0.015)	
Flesch-kincaid (t)			0.025**
			(0.011)
Child Sex Ratio*Fog Index (t)	−0.009		
	(0.016)		
Child Sex Ratio*Smog Index (t)		−0.010	
		(0.023)	
Child Sex Ratio*Flesch-kincaid (t)			−0.012
			(0.020)
Controls	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes
R-squared	0.284	0.305	0.273
Observations	103	103	103

Note: This table reports the results from the regression of dowry death incidents on the interaction of readability measure (fog index, smog index and flesch-kincaid index respectively) of the judgments related to crimes against women and various indicators of state of women in India as per World Bank Survey, with control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year. ***, ** and * indicate that the coefficient estimates are significantly different from zero at the 1 percent, 5 percent, and 10 percent levels, respectively.

Table 7: Impact of judgment readability—SMOG and FK indices—on the incidents of crime against women

Panel A: SMOG index					
	Crime Incidents (t+1)				
	Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Smog Index (t)	0.041	0.040^{+}	0.041^{***}	0.003	0.030^{+}
	(0.028)	(0.023)	(0.010)	(0.014)	(0.017)
BH q-value	0.397	0.397	0.001	0.81	0.397
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.523	0.635	0.230	0.214	0.637
Observations	111	102	111	102	102
Panel B: FK index					
Flesch-Kincaid Index (t)	0.032^{+}	0.022	0.018^{***}	-0.0004	0.016
	(0.017)	(0.014)	(0.004)	(0.006)	(0.010)
BH q-value	0.543	0.544	0.001	0.967	0.544
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.535	0.624	0.198	0.214	0.626
Observations	111	102	111	102	102

Note: This table reports the results from the regression of crime rate on readability measures (smog index and flesch-kincaid index) of the judgments related to crimes against women and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year. BH q-value indicates the cut-off above which we fail to reject the null hypothesis for the regression equation.

Table 8: Results of Permutation Test

	Crime Incidents (t+1)				
	Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Permutation Test p-value	0.267	0.341	0.001	0.783	0.302

Note: This table reports the results (p-value) from permutation tests with 1000 iterations from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects. .

Table 9: Impact of judgment readability on the crimes against women (except Bihar)

	Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Panel A: Fog index					
	Crime Incidents (t+1)				
	Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Fog Index (t)	0.021	0.020	0.026***	0.010	0.016
	(0.014)	(0.014)	(0.004)	(0.007)	(0.011)
BH q-value	0.194	0.194	0.001	0.194	0.194
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.670	0.635	0.226	0.383	0.660
Observations	101	93	101	93	93
Panel B: SMOG index					
	Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Smog Index (t)	0.027	0.036	0.052***	0.016	0.028
	(0.022)	(0.024)	(0.008)	(0.015)	(0.018)
BH q-value	0.279	0.234	0.001	0.295	0.234
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.664	0.641	0.258	0.385	0.666
Observations	101	93	101	93	93
Panel C: Flesch-Kincaid index					
Flesch-Kincaid Index (t)	0.022	0.017	0.024***	0.008	0.014
	(0.014)	(0.014)	(0.004)	(0.007)	(0.011)
BH q-value	0.282	0.282	0.001	0.282	0.282
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.67	0.628	0.211	0.377	0.653
Observations	101	93	101	93	93

Note: This table reports the results from the regression of crime incidents on readability measures (fog index, smog index and flesch-kincaid index) of the judgments related to crimes against women (except those mapped to the state of Bihar) and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year. BH q-value indicates the cut-off above which we fail to reject the null hypothesis for the regression equation.

Table 10: Impact of judgment readability on the incidents of crimes against women (Sample 1996-2013)

	Crime Incidents (t+1)		
Readability Index	Rape	Domestic Violence	Dowry Deaths
Fog Index (t)	0.011	0.011	0.009***
	(0.014)	(0.014)	(0.003)
BH q-value	0.441	0.441	0.018
Year Fixed Effects	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes
R-squared	0.009	0.005	0.008
Observations	210	210	210

Note: This table reports the results of equations (3) and (4) from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women for the period 1996 to 2013. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year. BH q-value indicates the cut-off above which we fail to reject the null hypothesis for the regression equation.

Table 11: Impact of judgment readability on the incidents of crimes against women (Additional Controls)

	Dowry Deaths (t+1)			
	(1)	(2)	(3)	(4)
Fog Index (t)	0.018**	0.015**	0.020***	0.012***
	(0.008)	(0.006)	(0.003)	(0.005)
Sentiment (t)	Yes			Yes
VS (t)		Yes		Yes
Gender Slant			Yes	Yes
Controls	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.213	0.223	0.228	0.245
Observations	111	111	111	111

Note: This table reports the results of equation (3) from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects along with additional text based controls (sentiment and % sentences with valence shifters (VS) and gender slant). The standard errors (reported in parentheses) are clustered by state and year.

Table 12: Impact of judgment readability on the incidents of crimes against women (VS Types)

	Dowry Deaths (t+1)			
	(1)	(2)	(3)	(4)
Fog Index (t)	0.019***	0.016***	0.020***	0.019***
	(0.003)	(0.006)	(0.002)	(0.006)
VS Negators (t)	Yes			
VS Amplifier (t)		Yes		
VS De Amplifier (t)			Yes	
VS Adversative Conjunction (t)				Yes
Controls	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes
R-squared	0.211	0.225	0.210	0.210
Observations	111	111	111	111

Note: This table reports the results of equation (3) from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value and state-year median population. The regression includes State fixed effects and year fixed-effects along with additional text based controls (four types of Valence Shifters - VS). The standard errors (reported in parentheses) are clustered by state and year.

Table 13: Impact of trial duration

Readability Index	Crime Incidents (t+1) Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Fog Index (t)	0.025*	0.024*	0.021***	0.001	0.018*
	(0.014)	(0.013)	(0.004)	(0.007)	(0.010)
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.582	0.649	0.21	0.219	0.652
Observations	111	102	111	102	102

Note: This table reports the results of equations (3) and (4) from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value, state-year median population and trial duration. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year.

Table 14: Impact of media attention

Readability Index	Crime Incidents (t+1) Rape	Domestic Violence	Dowry Deaths	Molestation	CAW
Fog Index (t)	0.029*	0.024*	0.020***	-0.0005	0.018*
	(0.016)	(0.014)	(0.004)	(0.008)	(0.010)
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
R-squared	0.536	0.637	0.214	0.235	0.638
Observations	111	102	111	102	102

Note: This table reports the results of equations (3) and (4) from the regression of crime incidents on readability measure (fog index) of the judgments related to crimes against women and control variables including unemployment rate, police charge-sheeting rates, court conviction rates, arrest rates, pending police cases, police density, pending court cases, new projects value, state-year median population and media attention. The regression includes State fixed effects and year fixed-effects. The standard errors (reported in parentheses) are clustered by state and year.

Appendices

Table A1: Cases' state of origin and judges' state of domicile.

Cases' state of origination	#cases with judge domicile state the same as case origination state	Total #cases originating from state	%
Assam	2	20	10.00
Bihar	10	49	20.41
Chhattisgarh	0	15	0.00
Gujarat	4	35	11.43
Haryana	0	154	0.00
Himachal Pradesh	0	10	0.00
Jharkhand	0	24	0.00
Karnataka	5	58	8.62
Kerala	2	21	9.52
Madhya Pradesh	3	141	2.13
Maharashtra	3	102	2.94
NCT of Delhi	2	43	4.65
Punjab	8	176	4.55
Rajasthan	3	80	3.75
Tamil Nadu	4	68	5.88
Uttar Pradesh	8	77	10.39
Uttarakhand	0	5	0.00
West Bengal	3	53	5.66
Total	57	1131	5.04%

Table A2: Gender representation of judges during the sample period

Cases' state of origination	#Judgments with at least one female judge	#Judgments with all male judges	Total # judgments	% judgments with at least one female judge
Assam	2	18	20	10.00
Bihar	1	48	49	2.04
Chhattisgarh	0	15	15	0.00
Gujarat	1	34	35	2.86
Haryana	6	148	154	3.90
Himachal Pradesh	0	10	10	0.00
Jharkhand	1	23	24	4.17
Karnataka	2	56	58	3.45
Kerala	0	21	21	0.00
Madhya Pradesh	4	137	141	2.84
Maharashtra	3	99	102	2.94
NCT of Delhi	2	41	43	4.65
Punjab	9	167	176	5.11
Rajasthan	1	79	80	1.25
Tamil Nadu	1	67	68	1.47
Uttar Pradesh	4	73	77	5.19
Uttarakhand	1	4	5	20.00
West Bengal	5	48	53	9.43
Total	43	1088	1131	3.80

Table A3: Illustration of dowry-related judgments' Fog index values.

Case	Judgment Date	Judgment State	Judge Panel	Fog Index
Case 1	XX-Jul-2010	Madhya Pradesh & Rajasthan	Judge A & Judge B	14.37
Case 2	XX-Aug-2010	Punjab	Judge C & Judge D	16.28
Case 3	XX-Feb-2010	Karnataka & Haryana	Judge E & Judge F	16.37
Case 4	XX-May-2010	Madhya Pradesh	Judge G & Judge H	22.99

Note: This table reports four sample judgments pertaining to dowry-related cases from 2010 and their respective Fog index values.

Table A4: Descriptive statistics

Variable	Mean	Median	SD	IQR
Dependent variables:				
Rape Incidents	993.30	706.00	827.26	893.25
Domestic violence Incidents	3547.74	3250.00	2598.14	3509.00
Dowry death Incidents	328.66	247.00	320.71	289.00
Molestation Incidents	1686.38	1036.00	1705.53	1947.50
Crimes against women	6490.57	5250.00	4409.32	6220.00
Independent Variables:				
Fog index	17.96	17.59	2.22	2.73
Smog Index	12.25	12.00	1.35	1.57
Flesch Kincaid index	29.40	28.95	2.11	2.52
Controls:				
Unemployment rate	3.22	2.56	2.28	2.34
Government spending on new projects	672.72	350.00	896.05	728.92
Police density	3.00	0.41	9.52	0.63
Arrest rate	281.37	270.94	88.63	66.10
Pending cases with police (%)	24.70	22.10	16.49	27.50
Police chargesheeting rate	76.52	78.40	8.90	9.53
Pending cases with court (%)	83.12	81.95	6.93	11.12
Conviction rate of court	35.87	36.50	15.12	25.95

Note: This table presents summary statistics for the sample variables based on a state-year classification. The 'Crimes against women' category is an aggregation of individual crime-types such as rape, domestic violence, dowry deaths etc. Detailed variable definitions are reported in the appendix in Table A7.

Table A5: Correlation Table

	Unemployment Rate	Police Chargesheeting Rate	Court Conviction Rate	Arrest Rate	Pending Cases with Police	Police Density	Pending Cases with Court	New Projects Spending	Population
Fog Index	0.07	-0.11	0.03	0.03	0.03	0.04	0.08	0.10	0.14

Note: This table presents the correlation of Fog Index with all control variables used in this study. Detailed variable definitions are reported in the appendix in Table A7.

Table A6: Impact of judge characteristics on Fog Index

	Dependent variable:
	Fog Index (t)
Male	1.189***
	(0.366)
SC Experience	0.033
	(0.101)
Judge Age	−0.071
	(0.115)
CJI	0.229
	(0.381)
Year Fixed Effects	Yes
Observations	649
R^2	0.054

Note: This table reports the results of equation (5) from the regression of fog index on the characteristics of the judges that form the panel presiding over the case and penning the judgment. The regression includes year-fixed effects. The standard errors (reported in parentheses) are clustered by year. ***, ** and * indicate that the coefficient estimates are significantly different from zero at the 1 percent, 5 percent, and 10 percent levels, respectively.

Table A7: Variable definitions

Variable	**Definition**
Readability measures	
Number of sentences	We classify sentences as a collection of words between i) two full stops, ii) a full stop and a question mark, and iii) two question marks.
Average words per sentence	The number of words in the judgment text divided by the total number of sentences.
% complex words	The percentage of words with more than two syllables.
Fog index	0.4×(*average words per sentence* + *% complex words*). High values of the *Fog index* imply less readable text.
Smog Index	$1.043 \times \sqrt{\% \, complex \, words \times \frac{30}{\# \, of \, sentences}}$
Flesch-Kincaid index	$0.39 \times \text{average words per sentence} + 11.8 \times \frac{\text{total syllables}}{\text{total words}}$
Crimes against Women:	

Continued on next page

Table A7: continued from previous page

Variable	Definition
Rape	Sexual intercourse with a woman against her will, without her consent, by coercion, misrepresentation or fraud or at a time when she has been intoxicated or duped, or is of unsound mental health and in any case if she is under 18 years of age.
Molestation	Molestation is defined under sec 354 of IPC as: "Whoever assaults or uses criminal force to any woman, intending to outrage or knowing it to be likely that he will thereby outrage her modesty, shall be punished with imprisonment of either description for a term which shall not be less than one year but which may extend to five years, and shall also be liable to fine."
Dowry Death	Where the death of a woman is caused by any burns or bodily injury or occurs otherwise than under normal circumstances within seven years of her marriage and it is shown that soon before her death she was subjected to cruelty or harassment by her husband or any relative of her husband for, or in connection with, any demand for dowry, such death shall be called *dowry death*.

Continued on next page

Table A7: continued from previous page

Variable	Definition
Domestic Violence	(a) Any wilful conduct which is of such a nature as is likely to drive the woman to commit suicide or to cause grave injury or danger to life, limb or health (whether mental or physical) of the woman; or (b) harassment of the woman where such harassment is with a view to coercing her or any person related to her to meet any unlawful demand for any property or valuable security or is on account of failure by her or any person related to her to meet such demand.
Fog High	An indicator variable taking the value of one when Fog Index exceeds the 20th percentile benchmark and zero otherwise. Thus the complex judgments are assigned a value of one.
Female judge %	Percentage of judgments with at least one female judge on the panel, calculated at state-year level.

Continued on next page

Table A7: continued from previous page

Variable	**Definition**
Smog High	An indicator variable taking the value of one when Smog Index exceeds the 20th percentile benchmark and zero otherwise. Thus the complex judgments are assigned a value of one.
Fk High	An indicator variable taking the value of one when Flesch Kincaid Index exceeds the 20th percentile benchmark and zero otherwise. Thus the complex judgments are assigned a value of one.
Control variables:	
Conviction rate of the court	Calculated as the ratio of the number of convictions by the court to the cases where trials were completed during the year (under IPC).
Government spending on new projects	Value of new investment projects announced (in millions of Rupees *(INR)*), obtained from CMIE States of India
Unemployment rate	Calculated as an average of urban and rural unemployment rate (in percentage terms).

Continued on next page

Table A7: continued from previous page

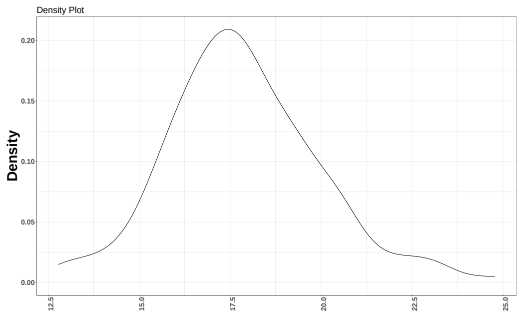
Variable	Definition
Police density	Number of policemen per 1,000 sq km, calculated as a ratio of Civil Police to Geographical Area of the State according to land use classification (both obtained from CMIE States of India).
Arrest rate	Calculated as a ratio of the number of persons arrested under IPC to population (in hundred thousand)
Pending cases with police	Percentage of pending cases by police: Indian Penal Code, obtained from CMIE States of India
Police chargesheeting rate	The ratio of persons chargesheeted by the police to the total cases disposed off by the police during the year, obtained from CMIE States of India
Pending cases with court	Percentage of pending cases by courts: Indian Penal Code, obtained from CMIE States of India

Table A8: Sample

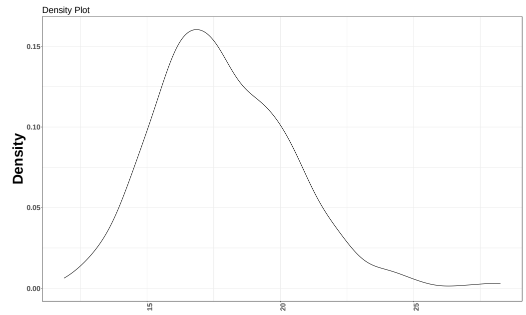
	Comparing Selected Sample (334) with Full Discarded Sample (6,191)			
	Fog Index Final Sample	Fog Index Discarded Sample	Difference	P Value
Mean Fog Index	17.96	18.21	0.25	0.18

	Comparing Selected Sample (334) with CAW Discarded Sample (433)			
	Fog Index Final Sample (334)	Fog Index CAW Discarded Sample	Difference	P Value
Mean Fog Index	17.96	17.85	0.11	0.56

Note: This table reports the mean value of Fog Index for the final sample (334 judgments) and the discarded sample (6,191 overall and 433 CAW judgments) due to missing judge and/or state data along with the respective p-value for the difference in mean values.



(a) Final CAW Judgments



(b) Discarded CAW Judgments

Figure A1: The figures present the Fog Density Index for the final sample of CAW judgements (panel A, 334) and discarded sample of CAW judgments (panel B, 433). The Kolmogorov-Smirnov test p-value for the two distributions is 0.14.

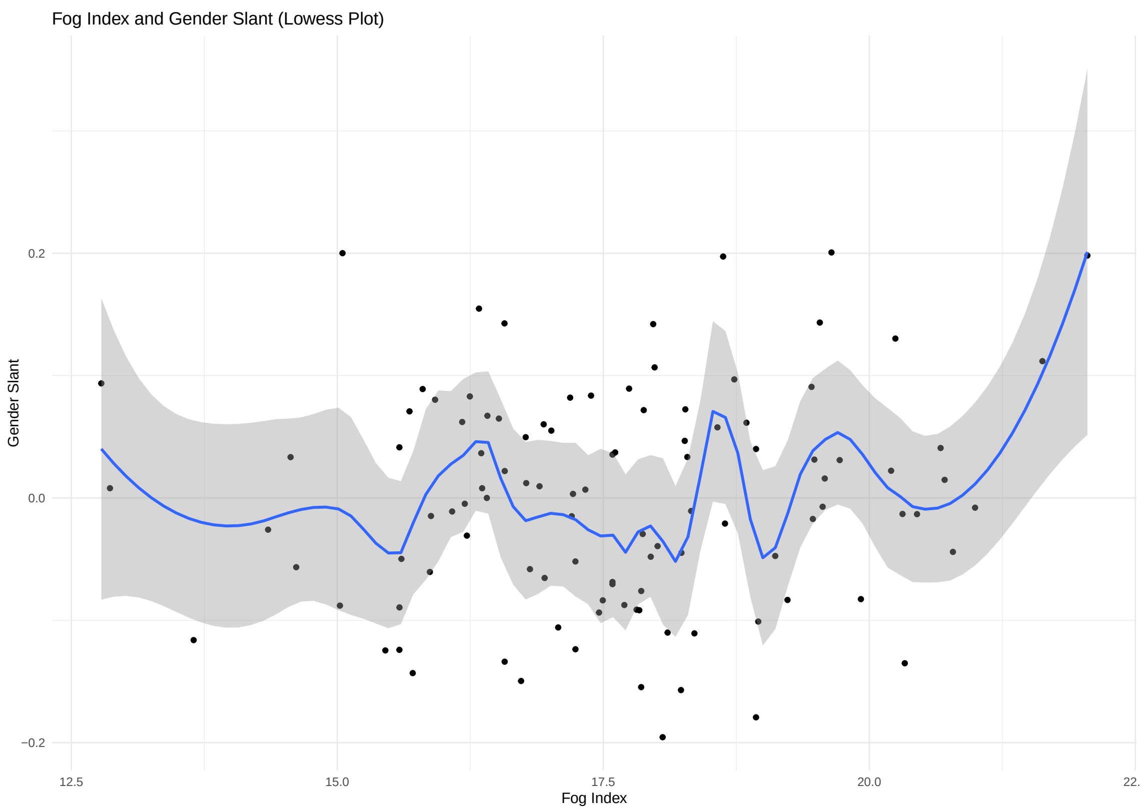


Figure A2: The figures presents the Lowess plot of Fog Index and Gender Slant

Alternative metrics

We include the size-based 'log(total words)' as a measure of the judgment length. All else equal, a shorter world length has lower text complexity and hence judgments with larger aggregate log(total words) should be deemed more complex than shorter counterparts (Loughran and McDonald, 2014).

To expand our coverage of readability metrics, we include an alternate measure that is based on 'valence shifters', which are text modifiers that alter the connotation of sentences. These could be of four different types: amplifiers ("absolutely", "acutely", "very"), de-amplifiers ("barely", "faintly", "few"), negators ("not", "cannot") and adversative conjunction ("despite", "but"). For example, contrast the sentence 'The evidence is serious', with 'The evidence is *very* serious", with 'very' amplifying the import of 'serious'.[40] For a judgment text, we define its 'semantic complexity' as the proportion of sentences which contain at least one valence shifter.

Next, we calculate the sentiment for each judgment using the 'word2vec algorithm' (Mikolov et al., 2013). For the calculation of judgment sentiment, we first create a vector space representation of words ("word embeddings") and then use semantically related words to generate a dictionary. This approach has been shown to perform better than both the existing dictionaries as well as the machine learning-based approaches (Cochrane et al., 2021). The word2vec algorithm uses a shallow neural net layer to predict the occurrence of words based on surrounding words. The coefficients from this model are defined as "word embeddings", based on the property that commonly used words in the same context are close to each other. Closeness in this context is captured by their cosine similarity.

Lastly, we calculate the 'gender slant', as defined in Ash, Chen and Ornaghi (2024) for all judgments at the state-year level. We also use 'Global Vectors for Word Representation' (GloVe) (Pennington, Socher and Manning, 2014), a weighted least square model that trains word vectors on global co-occurrence counts. GloVe

[40]The idea of valence shifters and their list is based on Kennedy and Inkpen (2006); Polanyi and Zaenen (2006) and Schulder et al. (2018).

first computes a global co-occurence matrix, which reports the count of incidences in which two words have appeared together in a given context (Ash, Chen and Ornaghi, 2024). It then obtains word vectors $w_i \in w$ to minimize the objective function below:

$$\min J(w) = \min \sum_{i,j} f(X_{ij}) \left(w_i^T w_j - \log(X_{ij}) \right)^2 \tag{7}$$

where X_{ij} is the co-occurrence matrix for count between words i and j and $f(\cdot)$ is a weighing function that down-weighs frequent words. Hence, the objective function $J(\cdot)$ minimizes the squared distance between the inner product of the vectors representing two words and their co-occurrence corpus.

In line with Pennington, Socher and Manning (2014) we keep the dimensionality to be 300 and window size of 10. We define 'Gender Slant' as the Gender Slant between the $(\overrightarrow{\text{male}} - \overrightarrow{\text{female}})$ vector and the $(\overrightarrow{\text{career}} - \overrightarrow{\text{family}})$ vector. In line with Ash, Chen and Ornaghi (2024), the $(\overrightarrow{\text{male}})$ vector is calculated as the normalized average of male-centric words such as "his", "he", "him", "mr", "himself", "man", "men", "male". Similarly we calculate the normalized vector for $\overrightarrow{\text{female}}$, $\overrightarrow{\text{career}}$, and $\overrightarrow{\text{family}}$ and finally we calculate the Gender Slant between the $\overrightarrow{\text{male}} - \overrightarrow{\text{female}}$ vector and the $\overrightarrow{\text{career}} - \overrightarrow{\text{family}}$ vector as:[41]

$$\text{Gender Slant Sim(A,B)} = \frac{\langle A, B \rangle}{\|A\| \|B\|} \tag{8}$$

The correlation of the Fog Index with Sentiment, valence shifter, and gender slant is $+0.37$, $+0.45$, and -0.02 respectively.

[41] The female, career, and family specific words are — ("her", "she", "ms", "women", "woman", "female", "herself", "girl"), ("family", "wife", "husband", "mother", "father", "parents", "son", "brother"), and ("company", "work", "business", "service", "pay", "corp", "employee", "employment"). These words are identified using LIWC (Linguistic Inquity and Word Count) technique and are similar to the words used in Ash, Chen and Ornaghi (2024).

References

Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz. 2023. "What we teach about Race and Gender: Representation in Images and Text of Children's Books." *Quarterly Journal Of Economics*, 138: 2225–2285.

Aizer, Anna. 2010. "The gender wage gap and domestic violence." *American Economic Review*, 100(4): 1847–1859.

Amaral, Sofia, Girija Borker, Nathan Fiala, Anjani Kumar, Nishith Prakash, and Maria Micaela Sviatschi. 2023. "Sexual Harassment in Public Spaces and Police Patrols: Experimental Evidence from Urban India." Working Paper.

Ash, Elliot, Sam Asher, Aditi Bhowmick, Sandeep Bhupatiraju, Daniel Chen, Tanaya Devi, Christoph Goessmann, Paul Novosad, and Bilal Siddiqi. 2021. "In-group bias in the Indian judiciary: Evidence from 5 million criminal cases." Working Paper.

Ash, Elliott, and Stephen Hansen. 2023. "Text Algorithms in Economics." *Annual Review of Economics*, 15: 659–688.

Ash, Elliott, Daniel L. Chen, and Arianna Ornaghi. 2024. "Gender Attitudes in the Judiciary: Evidence from US Circuit Courts." *American Economic Journal: Applied Economics*, 16: 314–50.

Ash, Elliott, Sam Asher, Aditi Bhowmick, Sandeep Bhupatiraju, Daniel Chen, Tanaya Devi, Christoph Goessmann, Paul Novosad, and Bilal Siddiqi. 2023. "In-Group Bias in the Indian Judiciary: EVIDENCE FROM 5 MILLION CRIMINAL CASES."

Bakar, Abu, and Varun Vinod Nambiar. 2023. "Precedent as a source of Law." *Indian Journal of Law and Legal Research*, 5: 1–12.

Bansal, Neha, Arun Sharma, and RK Singh. 2019. "Fuzzy AHP approach for legal judgement summarization." *Journal of Management Analytics*, 6: 323–340.

Becker, Gary S. 1998. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76: 169–217.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 57: 289–300.

Bhalotra, Sonia, Diogo GC Britto, Paolo Pinotti, and Breno Sampaio. 2021. "Job displacement, unemployment benefits and domestic violence."

Breusch, T. S. 1978. "Testing for Autocorrelation in Dynamic Linear Models." *Australian Economic Papers*, 17: 334–355.

Carrell, Scott, and Mark Hoekstra. 2010. "Externalities in the Classroom: How children exposed to Domestic Violence affect everyone's kids." *American Economic Journal: Applied Economics*, 2: 211–228.

Carrell, Scott E., Mark Hoekstra, and Elira Kuka. 2018. "The Long-Run Effects of Disruptive Peers." *American Economic Review*, 108: 3377–3415.

Cochrane, Christopher, Ludovic Rheault, Jean-François Godbout, Tanya Whyte, Michael W-C Wong, and Sophie Borwein. 2021. "The Automatic Analysis of Emotion in Political Speech Based on Transcripts." *Political Communication*, 1–24.

Garg, A. S. 1990. *Bride burning : Crime against Women.* Sandeep Publication.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. "Word embeddings quantify 100 years of gender and ethnic stereotypes." *Proceedings of the National Academy of Sciences*, 115(16): E3635–E3644.

Godfrey, L. G. 1978. “Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables.” *Econometrica*, 46: 1293–1301.

Gunning, Robert. 1952. *The Technique of Clear Writing.* McGraw-Hill, New York.

Jassal, Nirvikar. 2020. “Gender, law enforcement, and access to justice: Evidence from all-women police stations in India.” *American Political Science Review*, 114(4): 1035–1054.

Johnson, Pamela S., and Jennifer A. Johnson. 2001. “The oppression of women in India.” *Violence Against Woman*, 7: 1051–1068.

Kennedy, Alistair, and Diana Inkpen. 2006. “Sentiment classification of movie reviews using contextual valence shifters.” *Computational Intelligence*, 22: 110–125.

Kincaid, J. P. JP, RP Fishburne Jr, RL L. Rogers, B. S BS Chissom, R. P. Fishburne Jr, RL L. Rogers, and B. S BS Chissom. 1975. “Derivation of new readability formulas (automated readability index, Fog count and Flesch reading ease formula) for Navy enlisted personnel (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch.”

Koffi, Marlène. 2021. “Innovative ideas and gender inequality.” Working Paper Series.

Laughlin, G. Harry Mc. 1969. “SMOG Grading-a New Readability Formula.” *Journal of Reading*, 12: 639–646.

Loughran, Tim, and Bill McDonald. 2014. “Measuring readability in financial disclosures.” *The Journal of Finance*, 69: 1643–1671.

Merchant, Kaiz, and Yash Pande. 2018. “NLP based Latent Semantic Analysis for legal text summarization.” 1803–1807, Ieee.

Miguel, E. 2005. "Poverty and witch killing." *Review of Economic Studies*, 72: 1153–1172.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. "Distributed representations of words and phrases and their compositionality." 3111–3119.

Miller, Amalia R., and Carmit Segal. 2019. "Do Female Officers improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence." *The Review of Economic Studies*, 86: 2220–2247.

Mullatti, Leela. 1995. "Families in India : Beliefs and Realities." *Journal of Comparative Family Studies*, 26: 11–25.

Musa, Sainabou. 2012. "Dowry-murders in India: The law & its role in the continuance of the wife burning phenomenon." *Northwestern Interdisciplinary Law Review*, V: 227–245.

NCRB. 2020. "Methodology for Data Collection & Publication."

Oldenburg, Veena Talwar. 2002. *Dowry Murder: The Imperial Origins of a Cultural Crime.* Oxford University Press.

Pennington, Jeffrey, Richard Socher, and Christopher D Manning. 2014. "Glove: Global vectors for word representation." 1532–1543.

Polanyi, Livia, and Annie Zaenen. 2006. "Contextual valence shifters." In *Computing Attitude and Affect in Text: Theory and Applications.* 1–10. Springer.

Rose, Elaina. 1999. "Consumption Smoothing and Excess Female Mortality in Rural India." *The Review of Economics And Statistics*, 81: 41–49.

Schulder, Marc, Michael Wiegand, Josef Ruppenhofer, and Stephanie Köser. 2018. "Introducing a lexicon of verbal polarity shifters for English."

Sekhri, Sheetal, and Adam Storeygard. 2014. "Dowry deaths: Response to weather variability in India." *Journal of Development Economics*, 111: 212–223. Special Issue: Imbalances in Economic Development.

Shridhar, B Sai, S Kayalvizhi, and D Thenmozhi. 2021. "Simple Transformers in Rhetoric Role Labelling for Legal Judgements."

Sriram, Shyam. 2006. "Caste and the Court: Examing Judicial Selection Bias on Bench Assignments in the Indian Supreme Court."

Sukhtankar, Sandip, Gabriele Kruks-Wisner, and Akshay Mangla. 2022. "Policing in patriarchy: An Experimental Evaluation of reforms to improve police responsiveness to Women in India." *Science*, 377: 191–198.

Umar, Mohd. 1998. *Bride Burning in India: A Socio Legal Study.* A.P.H. Publishing Corporation.

United Nations Office on Drugs and Crime. 2019. "Global Study on Homicide."

Viscusi, W Kip, and Clayton J Masterman. 2017. "Income elasticities and global values of a statistical life." *Journal of Benefit-Cost Analysis*, 8(2): 226–250.