DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY

Does Rape Culture Predict Rape? Evidence From U.S. Newspapers, 2000–2013

Citation

Baum, Matthew, Dara Cohen, and Yuri Zhukov. "Does Rape Culture Predict Rape? Evidence from U.S. Newspapers, 2000–2013." Quarterly Journal of Political Science 13, no. 3 (2018): 263-289.

Published version

https://doi.org/10.1561/100.00016124

Link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:38435482

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles (OAP), as set forth at
https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

Quarterly Journal of Political Science
13:3: 263-89 (2018)

Does Rape Culture Predict Rape Evidence from U.S. Newspapers, 2000-2013

Matthew A. Baum
John F. Kennedy School of Government
Mailbox 113
79 JFK Street
Cambridge, MA 02138
Matthew_Baum@hks.harvard.edu

Dara Kay Cohen
John F. Kennedy School of Government
Mailbox 74
79 JFK Street
Cambridge, MA 02138
dara_cohen@hks.harvard.edu

Yuri M. Zhukov
Department of Political Science
University of Michigan
5700 Haven Hall
Ann Arbor, MI 48109
zhukov@umich.edu

Abstract: We offer the first quantitative analysis of *rape culture* in the United States. Observers have long worried that biased news coverage of rape - which blames victims, empathizes with perpetrators, implies consent, and questions victims' credibility - may deter victims from coming forward, and ultimately increase the incidence of rape. We present a theory of how rape culture might shape the preferences and choices of perpetrators, victims and law enforcement, and test this theory with data on news stories about rape published in U.S. newspapers between 2000 and 2013. We find that rape culture in the media predicts both the frequency of rape and its pursuit through the local criminal justice system. In jurisdictions where rape culture was more prevalent, there were more documented rape cases, but authorities were less vigilant in pursuing them.

Acknowledgements: The authors thank the Folke Bernadotte Academy (FBA) (Sweden), the Women and Public Policy Program and the Dean's Research Fund at the Harvard Kennedy School and the Behavioral Laboratory in the Social Sciences at Harvard for their generous financial support of this project. We also thank Susanne Schwarz and our team of 11 undergraduates (Natalie Chang, Vicente de la Torre, Renzo Falla, Jessica Fournier, Nora Garry, Alice Han, Ren Jie Teoh, Kelsey Jost-Creegan, Lauren Kang, Laura Riccardione, Andrew Wyner) for excellent research assistance. We received helpful comments from Dan Hopkins and other participants at the 2014 APSA meeting, and the UNSCR 1325 Research Working Group of the FBA. The authors are grateful to the many subject matter experts we consulted in the course of this project, especially Soraya Chemaly.

In recent years, the United States has seen growing public debate around media bias in news reporting of sexual violence. News coverage of cases like the 2012 Steubenville, Ohio gang rape raised concerns about media empathizing with the accused and neglecting the victim's perspective. Much of this discussion centers on the idea of *rape culture*, which scholars and activists define as "a set of values and beliefs that provide an environment conducive to rape" (Boswell and Spade 1996), where "rape is often not acknowledged as a crime and its victims are frequently blamed ... for their own violation" (Vogelman 1990). If news reporting reflects the norms and policy preferences of journalists and their audiences (Hamilton 2004), a closer look at this media coverage may help us understand why some local authorities are less assertive in investigating rape allegations, some victims are less likely to report assaults, and some perpetrators are more likely to commit rape.

Rape culture has important implications. Social perceptions of sexual violence -- for example, whether rape is even possible within marriage -- bound women's political, social and economic rights. Feminist scholars have long viewed social norms surrounding rape as defining features of gender equality (Brownmiller 1975), with "rape-prone" societies characterized by high levels of sex segregation, devaluation of women, and interpersonal violence (Sanday 1981). Despite growing public interest (Madden 2014), social science research on rape culture remains limited (Wolf 2013).

Our study brings together literatures on political communication, sexual violence, and economic models of crime, providing a theoretical logic by which local norms about rape might affect the choices of victims, perpetrators and police. It is also the first to move beyond small group experiments and qualitative case studies to link quantitatively these norms to the local prevalence of rape. We employ local newspaper coverage as a measure of local norms, and

develop a text classification model to detect several aspects of rape culture in the media, including victim-blaming, empathy for the accused, implications of consent and incredulity toward victims. We classify over 300,000 news articles about rape across 279 mostly local U.S. newspapers between 2000 and 2013. We then analyze whether rape culture in the press helps predict local variation in reports and arrests for rape, based on data from F.B.I. Uniform Crime Reports. In doing so, we account for a host of political, social and economic confounding factors, and exploit exogenous variation in news content due to shocks to the local media market.

Does rape culture predict rape? In a word, yes. We find that where there is more rape culture in the press, there is more rape. In areas with more prevalent rape culture in the press, police receive more frequent reports of rape, but make *fewer* arrests in response. Rape culture in the press, in turn, is most prevalent during the arrest and prosecution phases of the criminal justice process. Because lower police vigilance or courtroom mistreatment may deter future victims from reporting, while raising potential perpetrators' senses of impunity, the association between rape culture and crime likely reflects an increased *incidence* of rape, rather than increased *reporting* by victims. We do not find similar patterns for other violent or non-violent crimes. To the extent that rape culture, as reflected in news coverage, can help explain the choices of perpetrators, victims and police, this finding highlights a key missing element from existing theories of crime, advancing our understanding of rape and conditions that enable it.

Causes of rape and sexual violence

Although policy discourse often assumes that rape culture is socially harmful, scholars have yet to link empirically gendered media biases to the prevalence of rape. Economic theories of crime generally assume that a person commits an offense if the expected utility gained from it exceeds that of investing time and resources into other activities (Becker 1968, 176). As the

expected costs of committing a crime increase -- as discovery becomes more likely or punishment more severe -- fewer people will commit it. Such theories have greatly influenced policy by offering a simple, generalizable logic of deterrence. Yet they overlook normative drivers of criminal behavior, its cultural context, and the effects of age, race and gender (Eide et al. 2006).

Applications of economic models of crime to sexual violence remain nascent (Beauregard et al. 2007; Benson and Zimmerman 2007), and their empirical validity uncertain. Although police resources (i.e. staffing, training, equipment) should affect perpetrators' expectations of being caught and punished, empirical research finds that such resources have a stronger deterrent effect on property crime (Levitt 1997) than on violent crimes like rape (Evans and Owens 2008). The same is true for other political-economic factors, like women's demographic presence (Iyer et al. 2011), or local attitudes on criminal justice (Jost et al. 2003).

Gendered biases in media coverage of rape. Existing research has not yet addressed a potentially important source of variation in sexual crime: community norms on gender and sexual violence, as reflected in mass media.

Social scientists have examined select forms of gendered bias in the media, like exposure to violent pornography (Malamuth et al. 1986; Russell 1988), false beliefs or "myths" about rape (Benedict 1992; Soothill and Walby 1991), and the conduciveness to rape of specific social settings (Armstrong, Hamilton and Sweeney 2006; Boswell and Spade 1996). Most of this research holds that news coverage shapes, rather than reflects, prevailing norms (Benedict 1992). Yet unless potential victims, perpetrators and police all read their local newspapers, it is unclear how newspaper reporting would affect their behavior.

Political communication scholars (Entman 2004; Strömbäck and Dimitrova 2006) argue

that, with rare exceptions, news reporting *reflects* rather than *challenges* the normative context within which it is produced. Reporters' biases mirror the norms of the communities in which they live and work. Their audiences, meanwhile, resist or avoid frames (that is, emphases or perspectives on a story) that conflict with their preexisting norms or values (Goldman and Mutz 2011; Iyengar and Hahn 2009). News organizations that neglect these norms risk losing their audiences, particularly in diverse media markets where consumers have numerous options (Mutz and Martin 2001).

A new model of rape culture

We introduce a stylized model to assess the relationship between rape culture and rape. Following Donohue & Levitt (2001) and Iyer et al. (2011), we assume that rape results from an interaction between three actors: perpetrators, victims and police. Initially, a perpetrator decides whether to commit rape. If they do so, the victim decides whether to contact the police. If the victim reports, the police decide whether to arrest the perpetrator.

This interaction produces four possible outcomes: (1) perpetrator does not commit rape, (2) perpetrator commits rape, but victim doesn't report, (3) perpetrator commits rape, victim reports, but the police make no arrest, and (4) perpetrator commits rape, victim reports, police make an arrest. The prevalence of rape culture in society affects actors' preferences over these outcomes and, by extension, the probability that each occurs.

Police. By backward induction, the perceived likelihood that law enforcement will pursue a rape allegation affects decisions both to commit and to report the crime. The police have limited resources and are more likely to make an arrest if they view the allegation as credible. Where rape culture is prevalent, the threshold for justifying an investigation will be high. If local norms favor victim-blaming and empathy for the accused, police embedded in those norms may

be more likely to scrutinize the victim's account, more wary of making a false arrest, and more likely to drop cases early in the criminal justice process (Jordan 2006; Schuller and Stewart 2000). Where rape culture is less pervasive, police may be more likely to believe the victim, to conduct a full investigation, and to make arrests. Higher levels of rape culture should be associated with decreased police vigilance in rape cases.

Victims. Rape survivors' decisions to report violations depend on various factors, including trauma, fear of reprisal, their relationship to the perpetrator, and trust in law enforcement. Assuming that victims seek increased security or justice, they may prefer that police make an arrest following a report over either no arrest or no report.[1] The greater a victim's expectation of an arrest, the stronger their incentive to report the crime.

If rape culture is pervasive, victims will expect the investigation threshold to be high, and may not expect police to bring alleged rapists to justice. Reporting is also likely to be costlier -- proceedings take longer, with victims subject to greater public scrutiny and risk of retaliation. A victim in a high rape culture context may conclude that it is least costly -- in terms of damage to dignity and reputation -- to forgo reporting the crime. Higher levels of rape culture should therefore be associated with decreased reporting of rape by victims.

Perpetrators. Decisions to engage in criminal activities depend, at least in part, on the perceived probability of being caught, and the severity and immediacy of punishment. Perpetrators prefer committing rape when victims don't report and police don't arrest, and prefer not offending over doing so and being punished. The costs of rape therefore depend on how victims and police respond.

Where rape culture is limited, potential perpetrators may expect a higher probability of

1 Empirically, rape victims are more likely to report if they experienced high levels of violence and injuries, and less likely if the perpetrator was a relative or acquaintance (Pino and Meier 1999).

detection and arrest, because police are more likely to pursue victims' allegations vigilantly, and victims are more likely to come forward. By reducing the probability of arrest and severity of punishment, rape culture lowers the expected costs of rape, and raises the likelihood that perpetrators commit it.[2]

Empirical implications. The previous discussion implies that higher levels of rape culture should be associated with a higher incidence of rape. Yet an increase in documented cases could result from either a higher incidence of rape, or from more reporting by victims.

To distinguish empirically between these possibilities, we can examine the number of rape reports that result in arrests. We predict that where rape culture is high, police are less likely to investigate rape reports, and are less likely to make arrests. Victims are thus less likely to report the crime, and perpetrators more likely to offend. Consequently, if we observe lower police vigilance in high-rape culture contexts, we can have greater confidence that increases in *crime*, rather than victims' propensities to report, account for increases in documented rape cases.

News content as an indicator of local rape culture

If the threshold for investigating rape cases indeed depends on the local normative context, then an empirical measure of these norms may help predict the local prevalence of rape. One such indicator is public information about rape that local communities produce and consume. In the context of print media, this information is of two types: *news content* (i.e. *what* journalists write about rape) and *news volume* (i.e. *frequency* of stories about rape). We take the first of these to be more informative of journalistic ethics and community norms.

Following the emerging consensus in political communication, we assume that

2 Perpetrators are sensitive to these expected costs, and use strategies to avoid capture -- targeting acquaintances and vulnerable groups who are less likely to report, and drugging victims before the assault (Lisak and Miller 2002; Tjaden and Thoennes 2006).

commercially-oriented media organizations are reticent to risk alienating readers by publishing content that falls far outside the normative mainstream of its target audience (Hamilton 2004). If local news coverage of rape systematically features victim-blaming language, empathy for the accused, implications of consent and incredulity toward victims, we can reasonably interpret such content as a noisy indicator of attitudes that local news consumers and journalists find normatively acceptable and commercially viable. We should therefore expect lower police vigilance and a higher incidence of rape where rape culture is readily observable in news content.

What explains variation in news coverage about rape? At the community level, consumer preferences depend both on observable factors like local demographics, wealth, education, religiosity and politics, and on more static, but difficult-to-measure local norms about gender and sexual violence. Yet two newspapers within a community may cover the same story differently, depending on the nature and scope of their audiences, differences in ethical standards, or normative or ideological considerations across editorial offices and newsrooms (Atwater 1984; Carroll 1985; Ho and Quinn 2009; Zaller 1999). Even the same newspaper may cover the same story differently, depending on an article's timing in the news cycle – for instance, a high volume of coverage may incentivize differentiation across a range of dimensions, such as the extent of provocative and contrarian perspectives (Baum and Zhukov n.d.; Carroll 1985). More consequentially, a newspaper's reporting may depend on an article's timing in the life cycle of a criminal case – for instance, coverage of courtroom testimony and cross-examination, due to their inherently adversarial nature, may invite greater scrutiny of a victim's account than coverage of sentencing after a suspect's conviction.

News coverage of the criminal justice process is particularly salient to the choices of potential perpetrators, victims and police – not because it drives readers' behavior, but because it

reveals when and where local norms are most likely to surface and to be publicly reinforced. If victims know that heightened public attention to rape cases invites increased scrutiny of victims' accounts -- especially when (and if) these cases come to trial -- they may be less likely to come forward. If perpetrators expect public sympathies to shift in favor of the accused during court proceedings, they may see the likelihood of arrest and prosecution as relatively low, especially if these same forces also deter a victim from reporting the crime. Police, victims and perpetrators do not necessarily need to read the news to reach these conclusions -- chances are, if a newspaper prints a victim-blaming story, such sentiments already exist in the community. News coverage merely amplifies these sentiments, and makes them more visible.

Measuring rape culture

To evaluate the empirical relationship between rape culture and rape, we collected original data on news coverage of sexual violence: whether a newspaper published a story about rape on a given day, and whether the content of that story (or stories) demonstrates evidence of rape culture, as defined below. By connecting newspapers to communities that either produce or consume their coverage, we can estimate the local extent of rape culture.

Data collection. We collected every article mentioning the keywords "rape" or "sexual assault" published in all daily and weekly U.S. newspapers listed in Lexis-Nexis between 2000 and 2013. We also collected information on the physical address of newspapers' main bureaus and average daily circulations by county. We focus on newspapers due to their prevalence as primary sources of local information on political, economic and social events, and our ability to collect a consistent and representative data sample across the largest set of geographic units.

Our corpus includes 310,938 articles published in 279 newspapers (Appendix A.1). The

median newspaper published 52 articles about rape, including both news and opinion-editorials.[3] The *Washington Post* and *New York Times* featured the most coverage, with over 20,000 articles each, followed by the *St. Louis Post Dispatch*, *Washington Times* and *New York Daily News.*[4] One hundred and forty-three newspapers published fewer than 10 stories, and 49 published just one.

Measurement. Rape culture is difficult to quantify because most existing definitions are imprecise. Our review of previous work revealed convergence around four main categories: (1) victim-blaming language, (2) empathy for perpetrators, (3) implied victim consent and (4) questioning of victims' credibility. These categories feature prominently in toolkits and guidelines for journalists (Dart Center for Journalism and Trauma 2011; Garcia-Rojas 2012), and have dominated recent public discourse.

To develop a coding instrument around these categories, we solicited feedback from two dozen experts, including academic researchers, journalists and activists. Through these consultations, we disaggregated our four main categories into 76 components, and developed the coding instrument as an online survey form (summarized in Table 1; detailed in Appendix A.2).

[Table 1]

We used supervised machine learning to classify each news story into these categories, based on a training set of randomly-selected reference articles manually classified by research assistants (RAs) (Appendix A.2). Our team of 10 RAs created a combined training set of 21,911 manually coded newspaper articles. Intercoder reliability statistics, based on 341 overlapping articles, meet or exceed conventional standards of agreement (Appendix A.3).

---

3 Even national and international stories can reveal local cultural attitudes about rape, because editors retain discretion over which wire service stories they print.

4 Even national papers, like the *New York Times* have strong local components: they cover local and regional issues and events, and local residents and businesses account for most print subscriptions.

With these training data, we used a Support Vector Machine (SVM) classifier to assign each document to the categories in Table 1 (Appendix A.4).[5] We used SVM because it is well-suited to sparse, high-dimensional data, is highly robust, and can handle a low training-to-test data ratio. The classifier's out-of-sample prediction accuracy was high, ranging across categories from 80 to 90 percent (Appendix A.5).

Overview of rape culture in the press. Along with the four main categories of rape culture, we created a combined variable, coded 1 if an article contained *any* of the four types of coverage. Overall, rape culture is relatively rare in the news. The SVM algorithm classified about 3% of rape-related stories as containing any of the four components of rape culture (Appendix A.6). The most common sub-category was victim blaming (1.3 percent), and least common was incredulity toward victims (0.5 percent). The average U.S. county saw 417 newspaper articles about rape per year, 12 of which featured some element of rape culture.

Figure 1 reports the relative probability that SVM-classified articles on rape mention a specific term.[6] Solid circles indicate that a term is more likely to appear in articles that belong to each rape culture category, compared to articles not in that category. These probabilities reveal stark separation between articles with and without rape culture, and across the main categories. For instance, articles with *no rape culture* focus more on investigations and judicial proceedings, commonly mentioning terms like "suspect," "convict" and "sentence," but only rarely contextual language like "drink" or "night." They also tend to focus on more violent crimes ("kill", "murder").

---

5 SVM classifies documents by fitting a maximally-separating hyperplane to a feature space (i.e. relative frequency of words across documents), examining combinations of features that best yield separable categories. We trained this algorithm on human-coded data for each variable, using median values (majority vote) from training articles that overlapped between coders.

6 For parsimony, Figure 1 lists only terms in the top 95th percentile by td/idf weights. See Appendix A.6 for additional classification results.

By contrast, articles containing at least one of the four categories (*any rape culture*) focused more on the individuals at the center of the case ("student," "player," "team"), and less on the severity or criminal nature of an incident ("abus[e]", "crime"). Breaking these probabilities down by category, *victim-blaming* articles focus on the circumstances of the incident, particularly those that might cast doubt on the victim's physiological state ("drink," "parti[es]"). *Empathy for the accused* features terms associated with athletic institutions ("player," "team," "coach," "football") and standards of evidence ("evid[ence]," "innoc[ent]," "test"). This category is also unlikely to refer to the accused as "suspect[s]." Articles that imply consent are more likely to mention a "sexual" "relationship" between victim and accused, particularly in an educational setting ("student," "teacher," "school"). Finally, articles that question a victim's credibility emphasize the victim's account of events ("accus[e]", "alleg[e]") during adversarial court proceedings ("defens[e]," "prosecutor," "attorney").

[Figure 1]

To analyze the relationship between rape culture and crime, we aggregated these article-level indicators to county-years, as local proportions of newspaper stories containing each category of rape culture. We matched newspapers to counties in two ways (Appendix A.7). First, we examined communities that produce the news (*producers*), weighing newspaper articles by the geographic proximity of each newspaper's main bureau to the county center.[7] Second, we examined communities that consume the news (*consumers*), with weights based on each newspaper's market share in the county, according to circulation data from the Alliance for

7 The weights for newspaper k in county i are w_{ik}= 1 if $d(i,k) \leq d_{(r)}(i,k)$ and w_{ik}=0 otherwise, where $d(i,k)$ is the distance between country i and main bureau of k, and $d_{(r)}(i,k)$ is the r-th nearest bureau to i (details in Appendix A.7). The nearest-neighbor weights ensure that all counties, including ones without nearby bureaus, have the same number of newspapers. Below, we used r=5, but Appendix C.1 reports sensitivity analyses over r \in {1,20}.

Audited Media.[8]

These measures carry different theoretical interpretations. The *producers* measure assumes that news content reflects local norms in communities where journalists and editors live. The *consumers* measure assumes that it reflects the norms of communities where potential readers live. The largest differences are in how the measures treat national newspapers like the *Washington Post* and *New York Times.* The first measure gives greater weight to local, small circulation papers, which may claim to more closely represent the voice of a local community, even if they have a smaller local market share than their national counterparts. The second measure places more weight on large-circulation papers, and assumes that journalists aim to reflect the norms of a broader audience than resides in the immediate vicinity of their home bureau. If the *New York Times* has more local subscribers than the small-town newspaper, this measure will consider coverage in the *Times* to be more informative of local norms. The second measure is also more dynamic than the first: *producer* weights change only as nearby bureaus open and close, although *consumer* weights fluctuate with local subscribership.

Figures 2a and 2b show the geographic distribution of rape culture in newspaper articles according to these two measures, averaged over 2000-2013. The percent of local news stories about rape containing rape culture language ranges from 0 (dark blue) to 5 (bright yellow).

The two maps convey different distributions of news content about rape. According to the *producers* measure, areas with the highest prevalence of rape culture include the Mountain States, parts of central California and the Upper Midwest. Of 100 counties with the highest levels of rape culture, 52 were in Minnesota and Iowa, 13 were in North Carolina and 10 were in California. Conversely, the Midwestern states of Indiana, Ohio and Michigan had relatively little

8 The weights are $w_{ik} = \text{circulation}_{ik} / \Sigma_k \text{circulation}_{ik}$, or newspaper *k*'s local circulation in county *i* as a share of all newspapers' local circulation (Appendix A.7).

rape culture; these states include 75 of the 100 counties that scored lowest on the index.

The county map based on circulation (Figure 2b, *consumers*) offers a more conservative estimate of local rape culture in the press. Although some areas of high rape culture overlap with those in Figure 2a, like the Upper Midwest, Mid-Atlantic and parts of California, there are also significant disparities, as in the Mountain States and Florida. Much of this discrepancy is due to local media consumption favoring national or regional newspapers, which offsets the influence of locally-based, low-readership media. For instance, areas that appear orange in Figure 2a, but blue in 2b, indicate that local journalists produce more rape culture in their content than local readers tend to consume. To ensure that our results are not artifacts of geographic aggregation, we conduct all analyses separately for these two measures.

[Figure 2]

Figure 2c shows the distribution of reported rapes per 1,000 county residents, from the FBI's Uniform Crime Reports statistics, averaged over 2000-2013.[9] Figure 2d shows the difference between arrest rates and rape reports, with higher numbers (yellow) indicating higher police vigilance, and lower, negative numbers indicating lower vigilance.

Descriptive statistics support our expectation of a positive relationship between rape culture and the reported incidence of rape. According to the *producers* measure, county-years with above-average rape culture in local newspapers (greater than 3 percent) saw 93% more reported rapes than county-years with below-average rape culture (0.93 vs. 0.50 rapes per 1,000 residents, respectively, $p<.001$). The difference for the more conservative *consumers* measure is smaller (1 percent, $p<.001$), but also statistically significant.

9 The FBI defines rape as "penetration... of the vagina or anus with any body part or object, or oral penetration by a sex organ of another person, without the consent of the victim." Before 2013, the FBI used a narrower interpretation of "forcible rape," or "the carnal knowledge of a female forcibly and against her will" (FBI, 2014).

An even stronger relationship appears between rape culture and local police vigilance. In counties with higher-than average rape culture, between 1 and 2 additional reported rapes per year did not result in arrest, according to the *producers* measure. According to the *consumers* measure, there were up to eight additional reports without arrests per year. These patterns suggest a disturbing pattern: where rape culture is prevalent, there are more reported rapes, but fewer of the reports result in arrests.[10]

The strength of this relationship, however, is difficult to discern from summary statistics. Like other violent crime, rape results from multiple economic, demographic and political risk factors. Any correlation between rape culture and rape might disappear once we account for these confounders. Alternatively, it may reflect heightened media attention in locations with high preexisting rates of rape. If so, a third, unobserved factor may be driving variation in both news and crime. We thus perform a series of more rigorous empirical tests.

Regression Analysis

To examine the relationship between sexual violence and rape culture in local print media, we use the following core model specification:

$$y_{it} = \rho' C_{it-1} + \beta' X_{it} + \alpha_i + u_t + \varepsilon_{it} \quad (1)$$

where y_{it} is the number of reported rapes per 1,000 residents in county *i* in year *t*. The covariate of central interest, C_{it-1}, is the proportion of local newspaper articles on rape published in the previous year containing one or more of the rape culture categories listed in Table 1. The coefficient ρ captures the impact of such coverage on local crime.

We also control for several covariates, X_{it}, which correspond to the previously discussed alternative explanations of crime. These include median personal income (as a proxy for police

10 Kolmogorov-Smirnov test statistics for the *producers* and *consumers* measures are .04 ($p<.001$) and .03 ($p<.001$) for rape reports and.03 ($p<.001$) and .04 ($p<.001$) for police vigilance.

resources), unemployment, female share of the population, religious affiliation and political party vote during the last presidential election (summary statistics in Appendix A.8). Admittedly, these controls include only some of the potential determinants of rape, and omit important yet endogenous factors -- like local resources for survivors and sex crime units -- which may correlate with both local crime and news coverage. To account for these and other differences across space and time, we include county and year fixed effects, α_i and u_t. , and robust standard errors clustered on media markets.

Because news coverage is event-driven, local crime trends likely influence the volume and content of news stories about rape. We address this issue in two ways. First, we lag our rape culture variables by one year to avoid contemporaneous feedback from crime to news. Second, we consider a separate specification employing changes in local newspaper circulation as an instrument for rape culture. Assuming that shocks to the local print media market are unlikely to affect crime rates, except through their impact on news content, this approach, detailed below (and Appendix B.1), allows us to exploit exogenous variation in rape culture.

Is there more rape where there is more rape culture? The first two columns of Table 2 report the results of the model in equation (1), for reports of rape at the county-year level. Model 1 presents estimates based on the *producers* measure, and Model 2 presents estimates based on the *consumers* measure. The standardized coefficients represent the impact of a standard deviation increase in each variable on a standard deviation change in the outcome.

[Table 2]

The results confirm our core hypothesis: rape is significantly more prevalent where there is more rape culture in the press. The size of this impact varies across the two measures, but its direction is the same. A standard deviation increase above the mean in *any rape culture* yields,

respectively, 11 (95% CI: 6.1, 15.8) and 5.8 (95% CI: 3.0, 8.6) percent increases in reported rapes.[11]

What do these numbers mean in practice? Because we employ news content as a proxy for local norms, the precise magnitude of rape culture's impact on crime is difficult to establish. In the current context, a standard deviation increase in rape culture is equivalent to 8 additional newspaper articles per county-year (recall that an average county saw 12 articles with rape culture, out of 417 articles per year). Whether or not such an increase in rape culture – from 2.9 to 4.8 percent of local articles – represents a substantial shift in norms, its estimated impact on crime is modest: on average, one additional rape case per thousand local residents per year.

Rape culture in the press is a stronger predictor of reported rapes than several widely-cited variables, like local unemployment and the sex ratio. Still, even the largest standardized coefficient for rape culture is one-tenth of that for income or religiosity. Although it is a consistent and meaningful contributing factor, rape culture cannot by itself explain fundamental differences in the incidence of rape across counties.

Exogenous variation in rape culture. That rape culture correlates with increases in documented rape cases reveals little about the direction of the relationship. Journalists may simply be less sensitive where rape is more common, or some other, unobserved factor may drive both local news content and sexual violence. We thus consider another specification of the model in equation (1), employing changes in the average daily circulation of local newspapers as an instrument for rape culture in the media (Appendix B.1).[12]

Circulation shocks are a valid instrument because newspapers adapt the content of their

11 The SD shift is from .03 to .05 for *producers* and from .007 to .023 for *consumers*.

12 Although our *consumer* measure uses a subscription-based weighted average, shocks to overall local newspaper circulation do not mechanistically bias this weighted average up or down (see simulation in Appendix B.2).

coverage in response to market pressures, but market pressures are unlikely to affect crime rates, except through their effect on news coverage.[13] The relationship between market incentives and media bias is a subject of ongoing debate in political science and economics (Gentzkow and Shapiro 2010; Hamilton 2004). This research assumes that news organizations modify coverage to optimize readership, but disagrees on whether market pressures create cutbacks in reporting and editorial quality (Zaller 1999), or increase the diversity of published opinions (Gentzkow and Shapiro 2008). Regardless, market pressures create incentives against news coverage that alienates marginal consumers -- the readers *most* likely to cease consuming if displeased with content, and whom publishers can *least* afford to lose (Hamilton 2004).

Who are these marginal consumers today? In the last decade, the print media market has steadily declined across the United States. From 2004 to 2012, U.S. weekday newspaper circulation fell by 20%, from 55 to 44 million. According to the Pew Center (2012), one of the few groups bucking this trend is unmarried women, with those who report reading a local or community newspaper "sometimes" or "regularly" rising from 56 to 61 percent over the same eight years. The corresponding trend among single men is an eight-point *decline*, from 50 to 42 percent. If unmarried females are more attentive to press coverage of gender issues than are consumers less essential to newspapers' bottom lines, outlets in declining local markets should be less likely to feature coverage with gendered biases about rape.

We find that rape culture is indeed more common where market pressures are less severe. The proportion of newspapers with *any rape culture* is significantly higher in counties where newspaper readership increased or stayed constant from year-to-year, and lower where

13 Although we cannot directly test the exclusion restriction, the pairwise correlation between circulation shocks and reported rape cases is quite small (0.02), and is even smaller for police vigilance (0.004).

readership declined.[14] Rape culture is also lower where single women are more important to newspapers' bottom lines. Where a higher-than-average proportion of readers were single women, there was far less rape culture in local newspapers than would be expected by chance.[15]

The rape-culture-inhibiting influence of unmarried female newspaper readers dramatically strengthens in more competitive markets, with more than one newspaper based in the county (Appendix B.1). If declining markets and heightened competition reduce gendered biases in the media, we can exploit this variation to find a consistent estimate of the relationship between rape culture and crime.

Models 3 and 4 in Table 2 report the results of the instrumental variable regressions, estimated by two-stage least squares (2SLS) with county and year fixed effects. Test statistics for underidentification, overidentification and weak instruments all fall within conventional bounds of significance. For both measures of rape culture, the 2SLS coefficients are substantially larger than the reduced form estimates in Models 1 and 2. Although a one standard deviation increase above the mean in *any rape culture* previously produced a 3.1 or 2 percent rise, respectively, for the *producers* and *consumers* measures, the same counterfactuals yield increases of 81 (95% CI: 55, 108) and 202 (95% CI: 89, 316) percent in the 2SLS specification.

Figure 3a breaks this result down for each component of rape culture, the colored areas representing the density of the distribution of each standardized coefficient and its 95% confidence interval. White vertical dashes represent point estimates. The impact of rape culture

---

14 We define "decline in readership" as either a negative change in average self-reported frequency of newspaper readership (i.e. from "Regularly" to "Sometimes," "Sometimes" to "Hardly ever"), or as a decline in the proportion of local Pew survey respondents who reported reading a newspaper "yesterday." According to the two measures, there were 141 (expected: 153) and 103 (116) counties with both a decline in readership and higher-than-average rape culture, with Chi-squared tests significant at $p<.10$.

15 A total of 102 (expected: 121) counties had above-average proportions of single, female newspaper readers and above-average rape culture, compared to 179 (160) with below-average rape culture. Chi-squared statistics were significant at $p<.01$.

on rape is more robust in communities that produce biased news coverage than in communities that consume this coverage, but general patterns are the same. For the *producers* measure, every category of rape culture yields a positive and significant association with rape, with some minor differences in magnitude. For the *consumers* measure, *empathy for accused* was the only subcategory with a strong, positive relationship to reported rapes.

[Figure 3]

Instrumental variable results suggest that the direction of bias due to endogeneity is the opposite of what one would expect if news outlets in high-rape areas were systematically more skeptical of victims' accounts and more supportive of the accused. Instead, we see heightened sensitivity and caution among news outlets in such markets. Where there is more rape, journalists and editors produce *less* news coverage that might increase it. Therefore, reduced form estimates understate the true magnitude of rape culture's potential impact. When we purge these estimates of their correlation with the error term, the size of the estimate spikes: a standard deviation increase in rape culture is associated with two to five additional reported rapes per thousand residents in a county year.

To be clear, our results do not suggest that the publication of several newspaper articles produces a major impact on crime. We employ rape culture in the news as a proxy for local community norms, and we make no claims about the precise exchange rate between the number of articles with rape culture and the extent of underlying rape culture in communities that produce or consume them. It may take a small or huge increase in rape culture to generate an additional article containing it, and this relationship likely varies across time and space. Due to the noisiness of our proxy, the magnitude of the impact of one additional newspaper article is, for our purposes, secondary to the direction of the relationship. Our analyses clearly show this

direction is consistently positive.

Does rape culture increase the incidence of rape, or victims' reporting of rape? An increase in documented rape cases may reflect either a true increase in incidence, or an increase in victims' reporting. Only the former is consistent with our theory, which holds that rape culture in the media reflects local norms that fail to deter potential perpetrators, and that this increased incidence outweighs decreases in reporting by victims. If the latter is true, then rape culture may actually motivate outraged victims to come forward.

To determine which of these two stories is more plausible, we reexamine the incentives facing perpetrators, victims and police. Our theory implies that rape culture should produce the following interaction (working backwards): police are less likely to make arrests, which makes victims less likely to report, which, in turn, emboldens perpetrators to offend. Because expectations of police behavior drive others' choices, the empirical challenge is to establish how police react to victims' reports.

An analysis of per capita arrest rates does not necessarily capture the vigilance of police -- there may simply be more arrests where there are more reports. A more revealing measure is the *difference* between rape-related arrests and reports. Where this number is high and positive, police are highly vigilant, arresting more suspects than there are reports. Where it is negative, there are fewer arrests than reports, and police vigilance is low.

We model the determinants of police vigilance using the same general specification as in equation (1), substituting county-level arrest-report differences as the new dependent variable. The covariates, instruments and fixed effects are the same as before.

The last four columns of Table 2 report the determinants of police vigilance in rape cases, in both reduced form (Models 5 and 6) and instrumented (Models 7 and 8). Figure 3b further

reports instrumented standardized coefficient estimates for each component of rape culture.

According to the *producers* measure, there is a negative and significant relationship with police vigilance for every category of rape culture except victim blaming. A standard deviation increase in *empathy for the accused*, for instance, yields a .18 standard deviation decrease in police vigilance, equivalent to 9 fewer rape cases with perpetrators arrested (95% CI: -15, -1.6) per county-year. The same increase on the *consumers* measure produces 19 fewer reported cases with arrests (95% CI: -31.8, -6.9).

The difference between arrests and reports may also reflect police resources, electoral politics or other confounding factors. As Table 2 shows, police vigilance in rape cases is significantly lower in Republican-leaning counties, and -- depending on specification -- in counties where unemployment and religiosity are high. Yet the impact of rape culture on police vigilance still holds even after accounting for these and other unobservable characteristics captured by county and year fixed effects.

One additional concern is that lower police vigilance may not be unique to rape, and that police in communities with high rape culture may generally struggle to keep up with crime. To investigate this possibility, we replicated Models 7 and 8 for other categories of violent and property crime. Figure 3c shows that the negative association between rape culture and police vigilance does not emerge for murder, where under-reporting is less of a concern, or robbery, a non-sexual, opportunistic crime. In neither case does news content have a discernible effect, for either the general measure of rape culture, or its subcategories. Given that police vigilance is generally higher for murder and robbery than for rape -- 18 and 12 percent of county-years, respectively, had more arrests than reports, compared to 7 percent for rape -- the negative impact of rape culture is even more compelling. Unless victims are most likely to seek justice in places

where they know it is most elusive, variation in documented rape cases more likely reflects a true increase in rape rather than in victims' reporting.

News content or news volume? A key assumption behind our empirical approach is that news content (i.e. proportion of stories that contain rape culture) is more informative of local norms than news volume (i.e. number of stories about rape). Yet if the volume of coverage is indeed decreasing in local rape culture, then our proportional measure is capturing only one dimension of this phenomenon.

A deeper look at the data reveals that news volume has a very different relationship to crime than news content. To investigate this relationship, we replicated our empirical models, substituting the average number of articles on rape (per newspaper) per county-year in place of our proportional measures of rape culture. The results (Appendix B.7) show a positive and significant link between the volume of rape coverage and both reported rape cases and police vigilance. A standard deviation rise in the local number of stories about rape -- from 81 to 126 articles per newspaper, according to the *producers* measure -- is associated with 2 additional reported rape cases per 1,000 residents in a county-year (95% CI: 1.1, 2.6). However, ten more of these rape reports are likely to result in arrests (95% CI: 3.9, 16.2).

Contrary to our prior findings, police are *more likely* to follow through on victims' reports and make arrests where newspapers publish more stories about rape. Theoretically, this should make victims more willing to come forward. Hence, the increased incidence of rape in these locations likely results from increased reporting by victims rather than a genuine rise in sexual crime. As we have seen, however, much depends on how news organizations actually cover these stories -- and what their tone and content reveal about the local normative context.

Sources of Rape Culture in the News. The possibility that communities with more rape

culture in the news see a greater incidence of rape -- and less reporting by victims -- begs a question: what drives rape culture in the news? Where does it enter into a newspaper's editorial process and what accounts for variation across newspapers?

To address these questions, we undertake an additional set of article-level analyses (presented in detail in Appendix B.8). We model individual articles' propensity to include *any rape culture* -- or its four subcategories -- as a function of county-level attributes (e.g. sex ratio, wealth, education, religiosity, politics) and newspaper-level covariates (e.g. national vs. local distribution, general audience vs. trade), along with common temporal shocks, and unobserved cultural and normative factors that remain relatively static within counties and newspapers. We also account for article-level factors, such as the volume of a newspaper's coverage about rape in the past week, and the article's topic: whether it mentions a high-profile case, like the 2012 Steubenville, Ohio gang rape, or mentions a specific stage of the criminal justice process (i.e. arrest, prosecution, corrections).

Our results reinforce the centrality of the criminal justice process (and perceptions thereof) to victims' and perpetrators' decision-making. All else equal, rape culture is most likely to appear in the press when a criminal case reaches the point in its life cycle that is most consequential for victims' pursuit of justice: when the level of public attention is high (that is, when news coverage is intense, or focused on a high-profile case), when law enforcement is in the process of investigating and arresting perpetrators, and especially when those perpetrators are on trial. Articles that mention the arrest of suspects are 35 percent (95% CI: 18, 54) more likely to feature rape culture than ones that do not, although articles that mention court proceedings are even more likely (156 percent, 95% CI: 111, 211) to do so. These patterns are concerning if, as we anticipate, police vigilance depends on expectations of successful prosecution, or if victims'

expectations of justice influence their propensity to seek it. If newspapers are most likely to blame victims or empathize with perpetrators during these vital stages of the criminal justice process -- due to perceived consumer demand or editorial discretion -- it is hardly surprising that communities with more rape culture in the press also experience more rape.

Robustness checks. To account for potential sources of error in our analysis, we conducted a set of robustness checks and placebo tests, which we report fully in Appendix C. Some of these checks pertain to measurement: our geographic aggregation method, and newspapers included in our sample. First, we replicated our models with alternate *producers* measures (Appendix C.1), using search radii from 1 to 20 nearest newspaper bureaus to each county (the main analyses used only the 5 closest bureaus). Second, we investigated the potentially disproportionate influence of major newspapers like the *Washington Post* and *New York Times* -- which together account for 17 percent of the articles in our dataset -- by iteratively excluding from our analysis each of the top ten newspapers by volume (Appendix C.2).

The remaining robustness checks addressed our estimation strategy. First, because the fixed effects models in equation (1) are static, omitting a lagged dependent variable due to a high ratio of counties to years, we fit a series of dynamic panel data estimators to account for temporal autocorrelation while avoiding the problem of Nickell Bias (Appendix C.3). Second, to account for shocks common to media markets, we fit models with media market-level fixed effects (Appendix C.4).[16] Third, we conducted a series of placebo tests with leads of the rape culture variable, to address the possibility that our coefficient estimates are capturing anticipatory effects -- where crime increases not because local norms have recently changed, but because perpetrators expect them to change, based on pre-existing trends (Appendix C.5). Fourth, we added media market- and state-specific linear time trends to our models, to account for

16 We used Nielsen's Designated Market Areas to denote media market boundaries.

potentially heterogeneous regional trends -- in demographic and economic growth, police surveillance and forensics, or crime reporting -- any one of which could conceivably alter incentives for sexual crime for reasons that have little to do with underlying community norms (Appendix C.6). Finally, because our first dependent variable (rape reports per 1,000 people) is highly skewed, we re-estimated our models with logarithmic transformations (Appendix C.7).

The results (at Appendix C), are consistent with those in Table 2 and Figure 3. Coefficients on the rape culture variable remain generally positive for rape reports and negative for police vigilance.

Caveats. Our study has several limitations, each of which suggests potential areas for future research. First, despite our extensive consultation with numerous experts, many of the specific elements of our rape culture measure are not without controversy. The significance (Luciano 2015), and even existence (Kitchens 2014) of rape culture are both contested, and news coverage is only one potential indicator. Future studies might gather data on other local measures of culture, such as surveys about attitudes toward rape or local school curricula on preventing sexual assault.

Second, future research could further validate our instrument for rape culture. For instance, longitudinal data on individual newspaper demographics may help determine the nature of market incentives facing media owners.

Third, official crime statistics may not accurately measure the true frequency of rape. Future research may reveal alternative means of triangulating, if not necessarily directly measuring, the frequency of such crimes.

Finally, future research can assess, ideally through controlled experiments, whether and how rape culture actually produces individual-level effects (i.e. attitudes, behavior) that

correspond to our observational, aggregate-level findings.

Conclusion

Where there is more rape culture, there is more rape, and also less vigilance among local police. This finding, and our broader attempt to assess rape culture empirically, represents an important step forward for social science research on rape, as well as for the public debate. Our research can potentially help journalists and editors uncover implicit biases in their work, allow policymakers to gauge police responsiveness, activists to devise methods to reduce or mitigate sexual crime, and scholars to systematically investigate the causes and consequences of rape. It also validates a long-standing concern that some social norms can enable, or at least fail to deter, sexual violence.

Because, even in the best of circumstances, rape culture in the press is a noisy and indirect indicator of local norms, some caution is warranted in over-interpreting these results. Our empirical strategy shows that rape culture in the media is a reliable local predictor of sexual crime, but these estimates do not represent a causal effect. That said, a variety of tests suggest that rape culture is more likely to predict rape reports than to be predicted by them. These patterns hold after we account for a variety of confounding factors, and after we exploit fluctuations in newspaper circulation as an exogenous source of variation. Our indirect tests of alternative causal pathways further suggest that rape culture drives the actual frequency of rape, rather than the propensity of victims to report it. We did not find a similar relationship for other violent and non-violent crimes, like murder or robbery. Local community norms and biases, as reflected in this news reporting, appear to have a non-trivial impact on sexual crime.

Even if we stipulate that rape culture is both real and significant, there likely exists no

unimpeachable and entirely exogenous empirical measure of it. Our approach has been to rely on the collective judgments of experts, to maximize the transparency of measurement and analyses, to test hypotheses against both aggregated and disaggregated measures of rape culture, and to conduct parallel tests of leading alternative explanations.

This study contributes to several important strands of social science research. First, we demonstrate that economic models of crime have largely overlooked the role of culture, a key element in understanding the causes and consequences of rape. Second, we advance the literature on gendered violence, by bringing together insights from criminology, gender studies and political communication to present a novel argument about the choices of potential perpetrators, victims and police. Finally, this work joins other recent studies that analyze large amounts of text data (e.g., Shor et al. 2015), to offer an initial test of empirical linkages between community norms about rape and the prevalence of rape in society.

Although our results confirm the worst fears of many observers, they also have the potential to shift the study of rape culture away from the obscurity of ivory tower debates and internecine discussions among activists, and firmly into the domain of empirical social science and evidence-based policymaking. Rape culture has devastating consequences, and a better understanding of this phenomenon is the first step toward change.

References

Armstrong, Elizabeth, Laura Hamilton, and Brian Sweeney. 2006. "Sexual Assault on Campus: A Multilevel, Integrative Approach to Party Rape." *Social Problems* 53(4): 483–99.

Atwater, Tony. 1984. "Product Differentiation in Local TV News." *Journalism Quarterly* 61(4), pp.757-762.

Baum, Matthew A. and Yuri M. Zhukov. N.d. "Media Ownership and News Coverage of International Conflict." Typescript.

Benedict, Helen. 1992. *Virgin or Vamp: How the Press Covers Sex Crimes*. New York: Oxford.

Beauregard, Eric, Jean Proulx, and Kim Rossmo. 2007. "A Descriptive Model of the Hunting Process of Serial Sex Offenders." *Journal on Family Violence* 22(6): 449–63.

Benson, Bruce and Paul Zimmerman. 2007. "Alcohol and Rape: An Economics-of-Crime Perspective." *International Review of Law and Economics* 27(4): 442-73.

Boswell, Ayres, and Joan Spade. 1996. "Fraternities and Collegiate Rape Culture: Why Are Some Fraternities More Dangerous Places for Women?" *Gender & Society* 10(2): 133-47.

Brownmiller, Susan. 1975. *Against Our Will. Men, Women, and Rape*. New York: Ballantine.

Carroll, Glenn. 1985. Concentration and Specialization: Dynamics of Niche Width in Populations of Organizations." *American Journal of Sociology* 90(6): 1262-83.

Dart Center for Journalism and Trauma. 2011. "Tip Sheet: Reporting on Sexual Violence," July 15. Source: http://dartcenter.org/content/reporting-on-sexual-violence#.U_XpV7xdVy8

Demers, David. 1996. "Corporate Newspaper Structure, Editorial Page Vigor, and Social Change." *Journalism & Mass Communication Quarterly* 73(4): 857-77.

Donohue, John, and Steven Levitt. 2001. "The Impact of Race on Policing and Arrests." *Journal of Law and Economics* 44(2): 367-94.

Entman, Robert. 2004. *Projections of Power.* Chicago: University of Chicago Press.

Eide, Erling, Paul Rubin, and Joanna Sheperd. 2006. "Economics of Crime." *Foundations and Trends in Microeconomics* 2(3): 205-79.

Evans, William N., and Emily Owens. 2007. "COPS and Crime." *Journal of Public Economics* 91(1-2): 181–201.

Garcia-Rojas, Claudia, ed., and the Chicago Taskforce on Violence Against Girls & Young Women. 2012. "Reporting on Rape and Sexual Violence: A Media Toolkit for Local and National Journalists to Better Media Coverage." Source: http://www.chitaskforce.org/wp/wp-content/uploads/2012/10/Chicago-Taskforce-Media-Toolkit.pdf

Gentzkow, Matthew, and Jesse Shapiro. 2008. "Competition and Truth in the Market for News." *Journal of Economic Perspectives* 22(2): 133-54.

Gentzkow, Matthew, and Jesse Shapiro. 2010. "What Drives Media Slant?" *Econometrica* 78(1): 35-71.

Goldman, Seth, and Diana Mutz. 2011. "The Friendly Media Phenomenon: A Cross-national Analysis of Cross-Cutting Exposure." *Political Communication* 28(1): 42-66.

Hamilton, James. 2004. *All the News That's Fit to Sell.* Princeton: Princeton University Press.

Ho, Daniel and Kevin Quinn. 2008. "Viewpoint Diversity and Media Consolidation: An Empirical Study. *Stanford Law Review* 61(4): 781-868.

Iyengar, Shanto, and Kyu Hahn. 2009. "Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use." *Journal of Communication* 59(1): 19–39.

Iyer, Lakshmi, Anandi Mani, Prachi Mishra, and Petia Topalova. 2012. "The Power of Political Voice." *American Economic Journal: Applied Economics* 4(4): 165-93.

Jordan, Jan. 2006. "Beyond Belief? Police, Rape and Women's Credibility." *Criminal Justice* 4(1): 29-59.

Jost, John, Jack Glaser, Arie Kruglanski, and Frank Sulloway. 2003. "Political Conservatism as Motivated Social Cognition." *Psychological Bulletin* 129(3): 339-75.

Kitchens, Caroline. 2014. "It's Time to End 'Rape Culture' Hysteria," *Time* (March 20).

Levitt, Steven. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review* 87(3): 270–90.

Lisak, David, and Paul Miller. 2002. "Repeat Rape and Multiple Offending Among Undetected Rapists." *Violence and Victims* 17(1): 71-84.

Luciano, Michael. 2015. "'Rape Culture' Exists, but Not in the United States" *The Daily Banter* (4/8). Source: http://thedailybanter.com/2015/04/rape-culture-exists-but-not-in-the-united-states/

Madden, Katie. 2014. "Rape Culture: The Media's Role in Normalizing Assault." *Campus Times* (May 16). Source: http://laverne.edu/campus-times/2014/05/rape-culture-the-medias-role-in-normalizing-assault/

Malamuth, Neil, and John Briere. 1986. "Sexual Violence in the Media: Indirect Effects on Aggression against Women." *Journal of Social Issues* 42(3): 75-92.

Mutz, Diana, and Paul Martin. 2001. "Facilitating Communication Across Lines of Political Difference: The Role of Mass Media." *American Political Science Review* 95 (1): 97-114.

Pew Research Center for the People and the Press. 2012 Biennial Media Consumption Survey. Available at: http://www.people-press.org/2012/06/03/2012-media-consumption-survey/ (accessed 1/27/19).

Pino, Nathan, and Robert Meier. 1999. "Gender Differences in Rape Reporting." *Sex Roles* 40(11): 979-90.

Russell, Diana. 1988. "Pornography and Rape." *Political Psycholog*y 9(1): 41-73.

Salton, Gerard, and Christopher Buckley. 1988. "Term-weighting Approaches in Automatic Text

Retrieval." *Information Processing and Management* 24(5): 513-23.

Sanday, Peggy Reeves. 1981. "The Socio-Cultural Context of Rape: A Cross-Cultural Study." *Journal of Social Issues* 37(4): 5–27.

Schuller, Regina and Anna Stewart. 2000. "Police Responses to Sexual Assault Complaints: The Role of Perpetrator/Complainant Intoxication." *Law and Human Behavior* 24(5): 535-51.

Shor, Eran, Arnout van de Rijt, Alex Miltsov, Vivek Kulkarni, and Steven Skiena. 2015. "A Paper Ceiling: Explaining the Persistent Underrepresentation of Women in Printed News," *American Sociological Review* 80(5): 960–84.

Soothill, Keith, and Sylvia Walby. 1991. *Sex Crime in the News*. London: Routledge.

Strömbäck, Jesper and Daniela Dimitrova. 2006. "Political and Media Systems Matter." *Harvard International Journal of Press/Politics* 11(4): 131-47.

Tjaden, Patricia, and Nancy Thoennes. 2006. *Extent, Nature, and Consequences of Rape Victimization*. US National Institutes of Justice. Source: https://www.ncjrs.gov/pdffiles1/nij/210346.pdf

Vogelman, Lloyd. 1990. *The Sexual Face of Violence*. Johannesburg: Ravan Press.

Wolf, Naomi. 2013. "The US Military's Rape Culture," *Project Syndicate* (June 30).

Zaller, John. 1999. "A Theory of Media Politics." Unpublished Manuscript. Source: http://www.uky.edu/AS/PoliSci/Peffley/pdf/ZallerTheoryofMediaPolitics(10-99).pdf

FIGURE 1. Relative probabilities of articles' mentioning terms. Point size proportional to difference in probability of term being mentioned in articles in vs. not in each category. Solid points indicate that term is more likely to appear in articles classified by SVM as containing rape culture. Terms listed are from corpus' 95-th percentile by mean tf-idf weights.

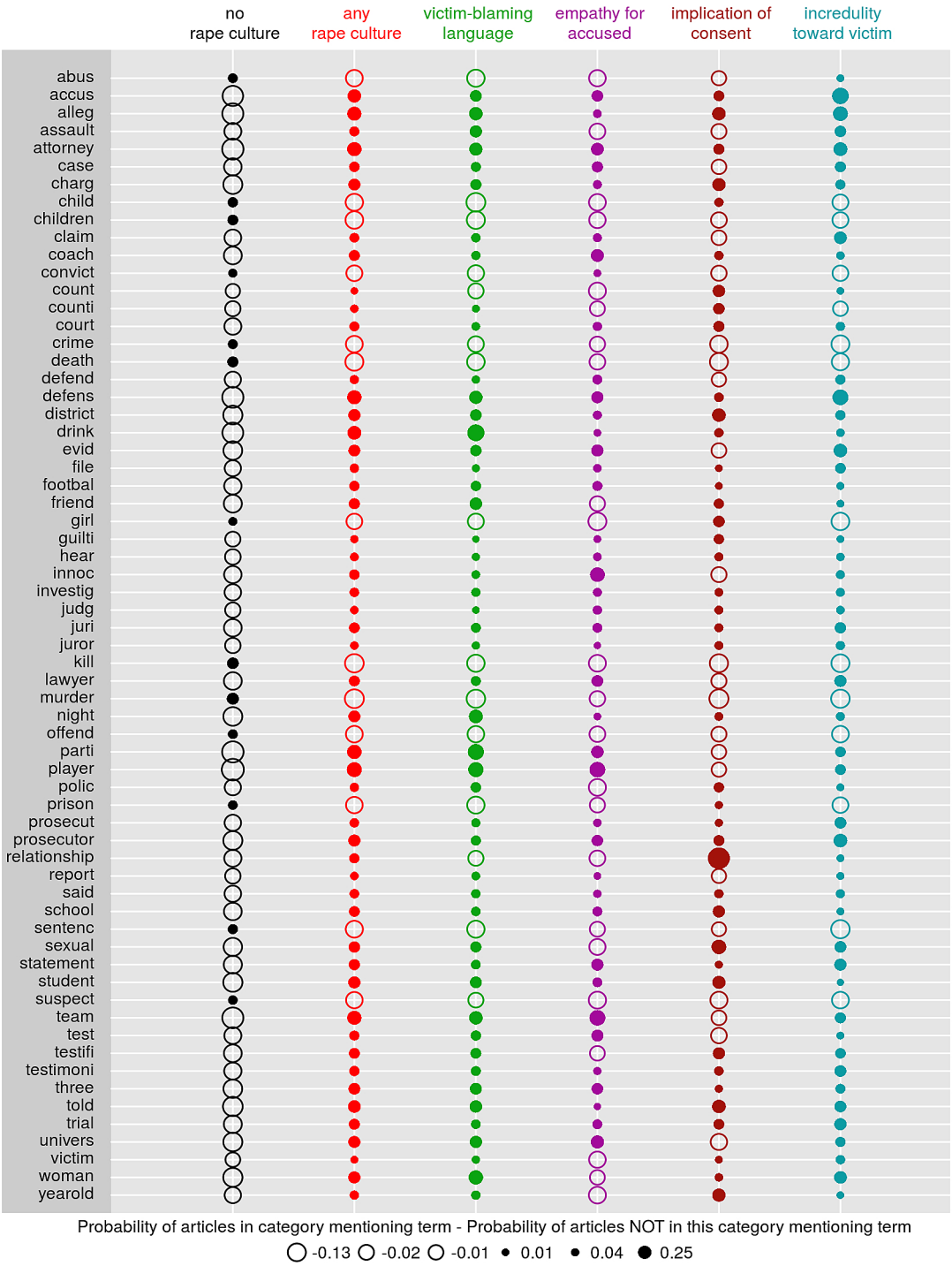


FIGURE 2. Geographic distribution of rape culture (a and b), reported rape (c) and police vigilance in rape cases (d).

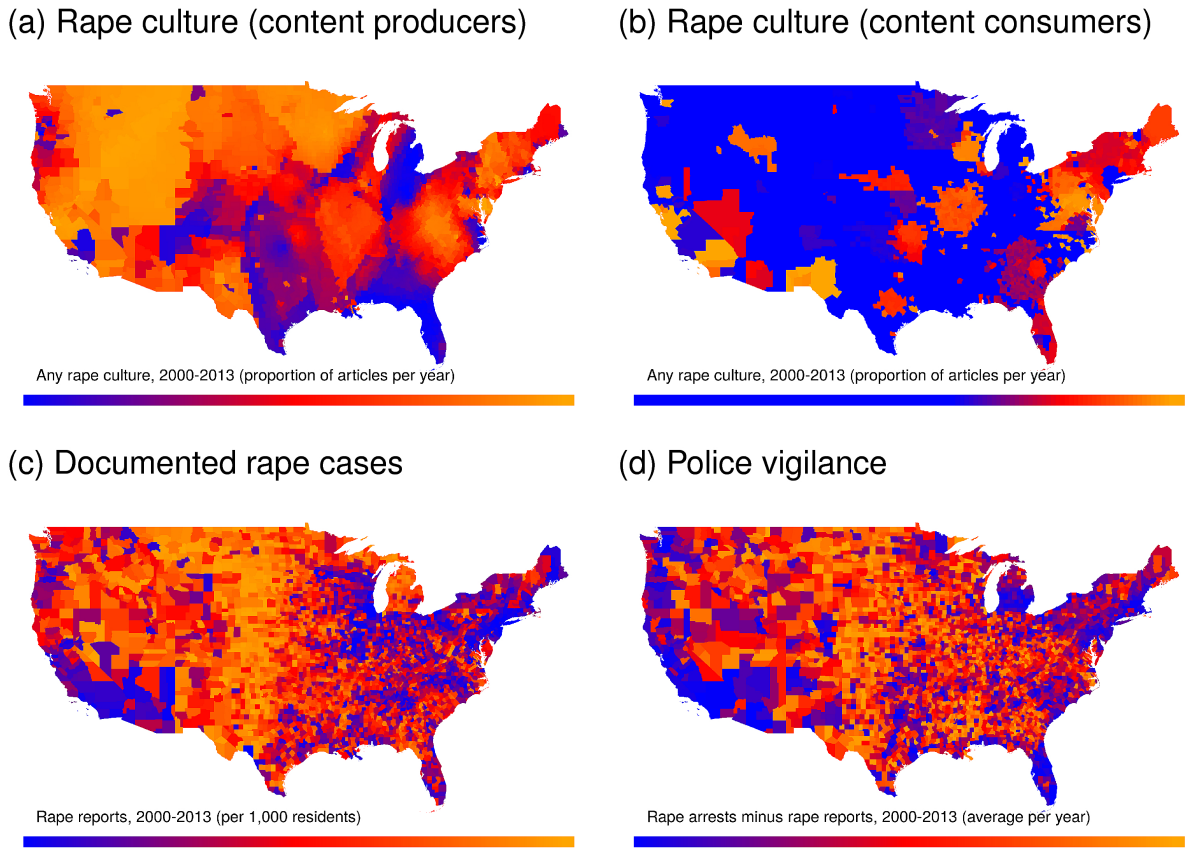
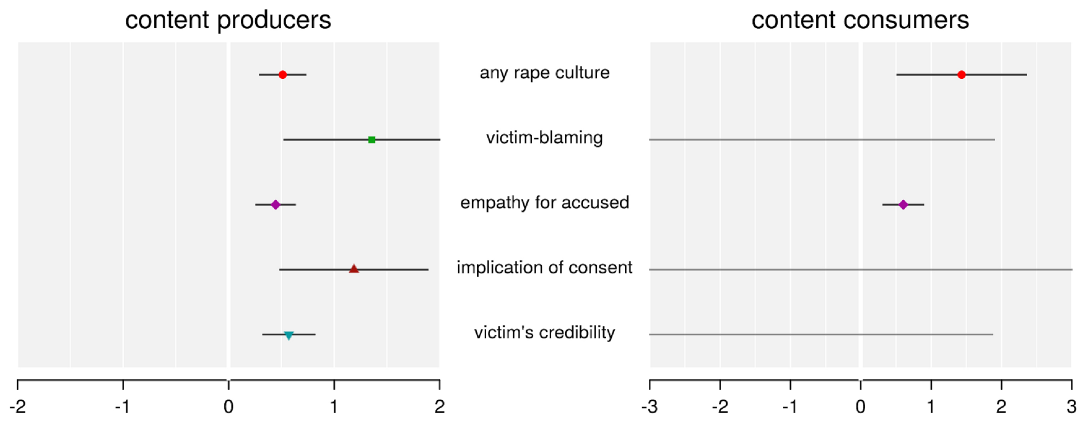
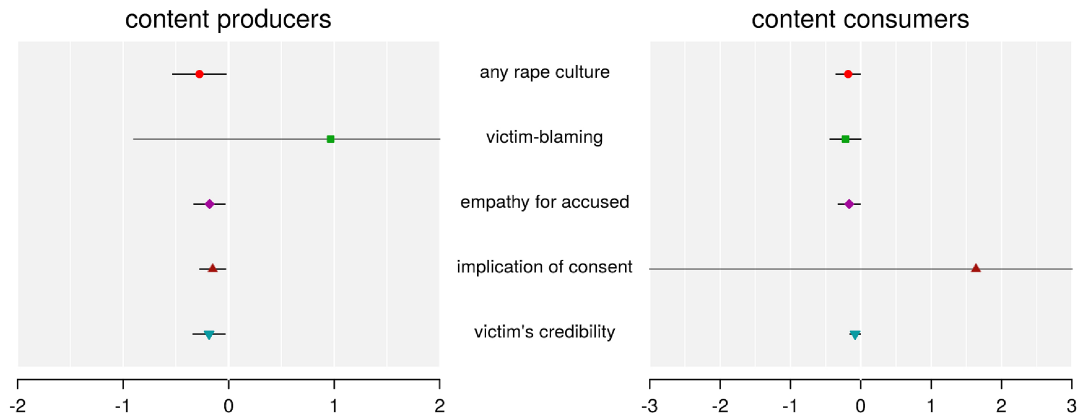


FIGURE 3. Impact of rape culture on crime. Quantities reported are standardized two-stage least squares coefficient estimates and 95% confidence intervals, interpreted as the impact of a standard deviation increase in each category of rape culture on standard deviation changes in each outcome measure.

(a) Rape culture and documented rape cases



(b) Rape culture and police vigilance in rape cases



(c) Police vigilance for other categories of crime

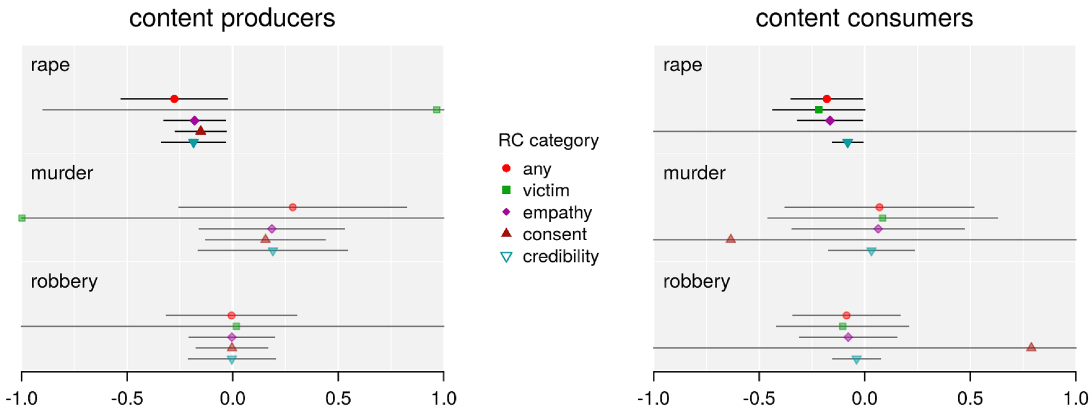


TABLE 1. Coding instrument (abridged, full list in online appendix).

Category of bias	*Individual components*
1. Victim-blaming language	Clothing, makeup of victim
	Victim's physiological state at time of incident (e.g. drunk, high, had consumed alcohol)
	Victim's former/current job as a sex worker or prostitute
	Victim's sexual history or promiscuity
	Victim's upbringing as explanation for current behavior (e.g. absentee parents, socioeconomic status)
	Locations that suggest victim culpability (e.g. victim had invited accused into own home)
	Use of loaded terms to describe rape self-reporting (e.g. the victim "complained", "admitted", "confessed")
2. Empathy for perpetrators	Mitigating factors and circumstances for accused (e.g. rape is "predictable outcome" of war, sports, substance abuse, age)
	Focus on suffering of the community as opposed to the victim
	Accused's promising future is now at risk (e.g. hopes dashed for honor-roll student, star athlete)
	Accused has high credibility or stature in the community
	Accused was the target of an unfair trial or overzealous prosecution
3. Implication of consent	Absence of physical resistance on part of victim
	Description of long-term abuse as an "affair" or a "sex scandal"
	Description of victim's past romantic relationship with the perpetrator
	Rape is referred to as "sex", "intercourse" or non-specific terms that diminish the force of rape
4. Questioning victim's credibility	Victim's past criminal record
	Inconsistencies in victim's account
	Victim's past or current history of substance abuse
	Victim's mental health
	Time elapsed between rape and report of rape; failure to report previous instances of abuse
	Ulterior motives on the victim's part (e.g. divorce proceedings)

TABLE 2. Determinants of rape reporting (1-4) and police vigilance (5-8).

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	Reported rapes per 1,000 people				Arrest to report difference per 1,000 people			
Measure:	prod.	cons.	prod.	cons.	prod.	cons.	prod.	cons.
Any rape culture (t-1)	0.019	0.011	0.51	1.43	-0.0071	-0.0037	-0.28	-0.31
	(0.0043)	(0.0028)	(0.11)	(0.47)	(0.0020)	(0.0027)	(0.13)	(0.10)
Median personal	0.55	0.55	0.60	0.56	-0.015	-0.015	-0.029	-0.021
	(0.063)	(0.063)	(0.021)	(0.032)	(0.0089)	(0.0090)	(0.011)	(0.0087)
Percent female	-0.047	-0.047	-0.062	-0.062	0.00039	0.00023	-0.0059	0.0057
population	(0.017)	(0.017)	(0.012)	(0.017)	(0.0037)	(0.0037)	(0.0058)	(0.0053)
Percentage of workers	0.034	0.030	0.16	0.086	-0.0076	-0.0061	-0.032	-0.030
unemployed	(0.015)	(0.015)	(0.029)	(0.025)	(0.0096)	(0.0096)	(0.014)	(0.010)
Percent population in	0.17	0.17	0.0039	-0.041	-0.0019	-0.0029	0.091	0.035
religious congregation	(0.078)	(0.078)	(0.063)	(0.095)	(0.024)	(0.024)	(0.057)	(0.030)
Percent presidential	0.087	0.083	0.27	0.18	-0.061	-0.059	-0.16	-0.083
vote for Republican	(0.022)	(0.022)	(0.044)	(0.044)	(0.015)	(0.015)	(0.045)	(0.012)
Constant	1.19	1.19			-0.58	-0.58		
	(0.045)	(0.045)			(0.018)	(0.018)		
Observations	40,349	40,358	37,244	37,253	40,349	40,358	37,244	37,253
County & Year FE	YES	YES	YES	YES	YES	YES	YES	YES
LL	-51506	-51519	-52673	-64384	-14035	-14043	-21089	-19992
Craig-Donald F			105	16.2			13.6	30.1
Kleiberger-Popp F			105	16.2			13.6	30.1
Anderson-Rubin F			26.2	21.6			7.01	12.8
Kleiberger-Popp LM			105	16.2			13.6	30.1

Notes: Fixed effect regression (1, 2, 5, 6), fixed effects two-stage least squares (3, 4, 7, 8), county-year level data. Standardized coefficients reported. Robust standard errors in parentheses.

Online Appendix:
Rape Culture and Sexual Crime

Matthew A. Baum, Dara Kay Cohen and Yuri M. Zhukov

January 20, 2018

Contents

A Data **A1**
A.1 Newspapers included in data . . . A1
A.2 Coding instrument . . . A5
A.3 Intercoder reliability . . . A8
A.4 Classification algorithm . . . A15
A.5 Cross-validation . . . A16
A.6 Classification results . . . A20
A.7 Aggregation to county-years . . . A24
A.8 Summary statistics at county-year level . . . A25

B Analysis **A26**
B.1 Instrumental variable design and newspaper readership . . . A26
B.2 Monte Carlo study: circulation shocks and newspaper weights . . . A33
B.3 Main results . . . A35
B.4 Rape culture and incidence of rape (Models in Figure 3.a) . . . A36
B.5 Rape culture and police vigilance (Models in Figure 3.b) . . . A38
B.6 Other types of crime (Models in Figure 3.c) . . . A40
B.7 Impact of news volume . . . A44
B.8 Article-level analysis . . . A45

C Robustness checks **A51**
C.1 Alternate "content producers" weights . . . A51
C.2 High-volume newspapers . . . A52
C.3 Dynamic panel data analysis . . . A53
C.4 Media market effects . . . A55
C.5 Placebo tests with temporal leads . . . A56
C.6 Heterogeneous time trends . . . A57
C.7 Logarithmic transformations . . . A59

A Data

A.1 Newspapers included in data

The following list enumerates, by location of main bureau, the 279 newspapers included in our database. Note that these newspapers are ones that published at least one article about rape or sexual assault in the study period, rather than an exhaustive list of all U.S. newspapers in the Lexis-Nexis database. Overall, Lexis-Nexis listed 333 unique U.S. newspapers in 2013 (the year of our data collection), excluding evening editions, weekend supplements and other duplicate entries. The same year, a total of 593 U.S. newspapers reported circulation statistics to the Alliance for Audited Media (AAM). If we take AAM's sample to be universal, then Lexis-Nexis includes about 56 percent of all U.S. newspapers.

Most of the missing entries are smaller, local newspapers and trade publications. Omissions among major newspapers are relatively rare. For example, our data include 21 of the 25 highest-circulation newspapers in the United States, and all of the top 10, according to the AMM.[1] High-circulation newspapers excluded from our sample include Chicago Tribune, The Arizona Republic, Honolulu Star-Advertiser, and The Boston Globe – all of which were absent from Lexis-Nexis at the time of data collection.

Alaska
Fairbanks Daily News Miner Alaska
Arizona
Arizona Capitol Times
The Arizona Capitol Times
California
Alameda Journal California
Alameda Times Star Alameda CA
Berkeley Voice California
Brentwood News California
Cathryn Ware and Robert Ware
Chico Enterprise Record California
Concord Transcript California
Contra Costa Sun California
Contra Costa Times California
Craig Riordan and Nelya Riordan
Daily Variety
El Cerrito Journal California
Enterprise Record Chico California
Eureka Times Standard California
Inland Valley Daily Bulletin
Inside Bay Area California
Investors Business Daily
La Opinion
Ledger Dispatch California
Long Beach Press Telegram
Los Angeles Times
Marin Independent Journal California
Martinez Record California
Metropolitan News Enterprise
Montclarion California
Monterey County Herald California
Nichole Moritz and Anthony Moritz
Oroville Mercury Register California
Pasadena Star News Pasadena CA
Piedmonter California
Pleasant Hill Record California
Ronald Straight and Pamella Straight
San Bernardino Sun California
San Gabriel Valley Tribune California
San Jose Mercury News California
San Mateo County Times San Mateo CA
San Ramon Valley Times California
SqueezeOC Orange County California
The Argus Fremont CA
The Daily News of Los Angeles
The Daily Review Hayward CA
The Oakland Tribune Oakland CA

[1] Neal Lulofs, "Top 25 U.S. Newspapers for March 2013," *Alliance for Audited Media*, April 30, 2013.

The Orange County Register California
The San Diego Union-Tribune
Tri Valley Herald Pleasanton CA
Vallejo Times Herald California
Valley Times California
Variety
Walnut Creek Journal California
West County Weekly California
Whittier Daily News California

Colorado

Colorado Springs Business Journal
Daily Camera Boulder Colorado
The Denver Post

Connecticut

Hartford Courant Connecticut
American Spectator

District of Columbia

CongressNow
Knight Ridder Washington Bureau
McClatchy Washington Bureau
Roll Call
States News Service
The Chronicle of Higher Education
The Chronicle of Philanthropy
The Hill
The Washington Post
The Washington Times
Washingtonian

Florida

Florida Times Union Jacksonville FL
Fox Huckabee
Ledger Lakeland Florida
Palm Beach Post Florida
Sarasota Herald Tribune Florida
St Petersburg Times Florida
Tampa Bay Times
Tampa Tribune Florida
The Ledger Lakeland FL
The Tampa Tribune Florida

Georgia

The Atlanta Journal Constitution
AJC Community News
The Augusta Chronicle Georgia

Idaho

Idaho Falls Post Register Idaho
Latah Eagle Moscow Idaho
Lewiston Morning Tribune Idaho
The Idaho Business Review Boise ID

Illinois

Chicago Daily Herald
Chicago Sun-Times
The Pantagraph Bloomington IL
The Pantagraph Bloomington Illinois
The State Journal Register Springfield

Indiana

Associated Press Online
Indianapolis Business Journal
South Bend Tribune Indiana
The Indianapolis Business Journal
The News Sentinel

Iowa

Telegraph Herald Dubuque IA

Kansas

Topeka Capital Journal Kansas

Lousiana

CityBusiness North Shore Report
New Orleans City Business
The Journal of Jefferson Parish

Maine

Bangor Daily News Maine
Portland Press Herald Maine

Maryland

The Baltimore Sun
The Capital Annapolis MD
The Daily Record Baltimore MD
The Maryland Gazette

Massachusetts

Lowell Sun Massachusetts
Massachusetts Lawyers Weekly
Montachusett Telgram Gazette
Rhode Island Lawyers Weekly
Sentinel Enterprise Fitchburg
Sunday Telegram Massachusetts
Telegram Gazette Massachusetts
The Berkshire Eagle Pittsfield
The Christian Science Monitor
The Patriot Ledger Quincy MA
The Union Leader Manchester NH

Michigan
Crains Detroit Business
Michigan Lawyers Weekly
The Detroit News Michigan
Minnesota
Finance Commerce Minneapolis MN
Lawyers Weekly USA
Saint Paul Pioneer Press Minnesota
St Paul Pioneer Press Minnesota
Star Tribune Minneapolis MN
The Legal Ledger St Paul MN
The Minnesota Lawyer Minneapolis MN
Mississippi
Mississippi Business Journal
Missouri
Kansas City Daily Record
Missouri Lawyers Weekly
St Charles County Business Record
St Louis Daily RecordSt Louis Countian
St Louis Post Dispatch Missouri
Nebraska
Lincoln Journal Star Nebraska
Omaha World Herald Nebraska
Nevada
Las Vegas Review Journal Nevada
New Hampshire
New Hampshire Sunday News
The Cabinet Milford New Hampshire
New Jersey
AIM Jefferson Morris
AIM Vernon Sussex
AIM West Milford Passaic
American Banker
Argus Cumberland
Belleville Times Essex
Bloomfield Life Essex
Bogota Bulletin
Cliffside Park Citizen
Clifton Journal Passaic
Community News
Daily News New York
Edgewater View
Fort Lee Suburbanite
Franklin Lakes Oakland Suburban News
Glen Ridge Voice Essex
Glen Rock Gazette
Hackensack Chronicle
Herald News Passaic County NJ
Leonia Life
Little Ferry Local
Mahwah Suburban News
Metropolitan Corporate Counsel
Midland Park Suburban News
Montclair Times Essex
Northern Valley Suburbanite
Nutley Sun Essex
Parsippany Life Morris
Pascack Valley Community Life
Passaic Valley Today
Ridgefield Park Patriot
South Bergenite
Suburban Trends Morris
Teaneck Suburbanite
Teaneck Suburbanite
The Forward
The Gazette Fairlawn
The Metropolitan Corporate Counsel
The Neighbor News Morris
The New York Sun
The Record Bergen County NJ
The Ridgewood News
The Star-Ledger
Town Journal
Town News
Twin Boro News
Verona Cedar Grove Times Essex
Wayne Today Passaic
New Mexico
Alamogordo Daily News New Mexico
Carlsbad Current Argus New Mexico
Deming Headlight New Mexico
Farmington Daily Times New Mexico
Las Cruces Sun News New Mexico
Richard Bonney 59 Santa Fe
Ruidoso News New Mexico
Santa Fe New Mexican New Mexico
Silver City Sun News New Mexico
The Alamogordo Daily News New Mexico

The Santa Fe New Mexican New Mexico
The Taos News New Mexico

New York

Advertising Age
Buffalo News New York
Long Island Business News
New York Observer
Newsday New York
Record The
The Daily Record of Rochester
The Gun Report
The New York Post
The New York Times
Wall Street Journal Abstracts

North Carolina

Chapel Hill Herald Durham NC
Morning Star Wilmington NC
North Carolina Lawyers Weekly
Star News Wilmington NC
Sunday Star News Wilmington NC
The Herald Sun Durham NC
The Mecklenburg Times Charlotte NC
Wilmington Star Wilmington NC

North Dakota

Bismarck Tribune North Dakota
The Bismarck Tribune

Ohio

Crains Cleveland Business
Dayton Daily News Ohio
The Plain Dealer

Oklahoma

Daily Oklahoman Oklahoma City OK
Journal Record Legislative Report
The Daily Oklahoman Oklahoma City OK
The Journal Record Oklahoma City OK
The Oklahoman
The Sunday Oklahoman Oklahoma City OK
Tulsa World Oklahoma

Oregon

Daily Journal of Commerce Portland OR

Pennsylvania

Chambersburg Public Opinion
Intelligencer Journal Lancaster
Lancaster New Era Pennsylvania
Morning Call Allentown Pennsylvania
Pittsburgh Post Gazette Pennsylvania
Pittsburgh Tribune Review
Public Opinion Chambersburg PA
Sunday News Lancaster Pennsylvania
The Evening Sun Hanover Pennsylvania
The Lebanon Daily News Pennsylvania
The Philadelphia Daily News
The Philadelphia Inquirer
The York Dispatch Pennsylvania
Tribune Review Greensburg PA
York Sunday News York PA

Rhode Island

Providence Journal Bulletin
The Providence Journal

South Carolina

South Carolina Lawyers Weekly
The Myrtle Beach Sun News

Texas

Austin American Statesman Texas
El Paso Times Texas
Houston Chronicle
The Dallas Morning News

Utah

Deseret Morning News Salt Lake City
Deseret News Salt Lake City Utah
Salt Lake Tribune Utah
The Salt Lake Tribune

Vermont

Brattleboro Reformer Vermont

Virginia

Dolans Virginia Business Observer
Richmond Times Dispatch Virginia
The Roanoke Times Virginia
The Virginian Pilot Norfolk VA
USA TODAY
Virginia Lawyers Weekly

Washington

Spokesman Review Spokane WA
The Columbian Vancouver WA
The Columbian Vancouver Washington

West Virginia

Charleston Daily Mail West Virginia
Charleston Gazette West Virginia

Wisconsin
Capital Times Madison WI
The Capital Times Madison Wisconsin
The Daily Reporter Milwaukee WI
Wisconsin Law Journal Milwaukee WI
Wisconsin State Journal

Wyoming
Wyoming Tribune Eagle Cheyenne WY

Puerto Rico
El Nuevo Dia Puerto Rico
Business Peru Peru
Business Wire Latin America
El Reporte Delta Latin America
Expansion Mexico
IPS Latin America

A.2 Coding instrument

The following is an unabridged version of the coding instrument summarized in Table 1 in the main text. Coders completed the survey below as a Google Form, after reading each article in their training set. The symbol "[x]" indicates a checkbox, "[...]" indicates an open text entry field, "[list]" indicates a drop-down list menu.

```
* Required field

Document ID * Copy and paste from first column of spreadsheet
[...]
Coder *
[list]

Victim-blaming language
Does the document make reference to any of the following? Check as many as apply.
[x] Clothing, makeup of victim
[x] Victim's physiological state at time of incident (e.g. drunk, high, had consumed
     alcohol)
[x] Victim's former/current job as a sex worker or prostitute
[x] Victim's sexual history or promiscuity
[x] Victim's upbringing as explanation for current behavior (e.g. absentee parents,
    socioeconomic status)
[x] Locations that suggest victim culpability (e.g. victim had invited accused into
    own home)
[x] Use of loaded terms to describe rape self-reporting (e.g. the victim "complained
    ", "admitted", "confessed")
[x] Other: [...]

Empathy for perpetrators
Does the document make reference to any of the following? Check as many as apply.
[x] Mitigating factors and circumstances for accused (e.g. rape is "predictable
    outcome" of war, sports, substance abuse, age)
[x] Focus on suffering of the community as opposed to the victim
[x] Accused's promising future is now at risk (e.g. hopes dashed for honor-roll
    student, star athlete)
[x] Accused has high credibility or stature in the community
[x] Accused was the target of an unfair trial or overzealous prosecution
```

```
[x] Other: [...]

Implication of consent
Does the document make reference to any of the following? Check as many as apply.
[x] Absence of physical resistance on part of victim
[x] Description of long-term abuse as an "affair" or a "sex scandal"
[x] Description of victim's past romantic relationship with the perpetrator
[x] Rape is referred to as "sex", "intercourse" or non-specific terms that diminish
    the force of rape
[x] Other: [...]

Questioning victim's credibility
Does the document make reference to any of the following? Check as many as apply.
[x] Victim's past criminal record
[x] Inconsistencies in victim's account
[x] Victim's past or current history of substance abuse
[x] Victim's mental health
[x] Time elapsed between rape and report of rape; failure to report previous
    instances of abuse
[x] Ulterior motives on the victim's part (e.g. divorce proceedings)
[x] Other: [...]

Rape jokes
Does the document make reference to any of the following? Check as many as apply.
[x] Puns and word play (e.g. "rapetastic," "rapetard")
[x] Disparaging remarks with reference to sexual violence (e.g. "who would want to
    rape her?"; "I hope he gets raped in jail")
[x] "Rape" as synonym for assertiveness/empowerment (e.g. "I'm going to rape that
    physics exam")
[x] "Rape" as synonym for weakness/domination by others (e.g. "Our team got raped
    last night")
[x] Threats of rape, including in jest (e.g. "Shut up or I'll rape you")
[x] Other language that trivializes or makes light of rape (including #hashtags)
[x] Other: [...]

Privacy protection for victim
Does the document make reference to any of the following? Check as many as apply.
[x] Name of victim
[x] Age of victim
[x] Ethnicity/race of victim
[x] Sexual orientation of victim
[x] Religion of victim
[x] Profession of victim
[x] Victim is a public figure or celebrity
[x] Use of term "incest" or other indication of blood relations between victim and
    accused
[x] Other: [...]
```

```
Privacy protection for accused
Does the document make reference to any of the following? Check as many as apply.
[x] Name of accused
[x] Age of accused
[x] Ethnicity/race of accused
[x] Sexual orientation of accused
[x] Religion of accused
[x] Profession of accused
[x] Accused is a public figure or celebrity
[x] Other: [...]

Other
Check as many as apply.
[x] Rapist as "monster" or exceptional case
[x] Rape as symptom of systemic failure (e.g. education, culture, early detection,
    mental health treatment)
[x] Dismissal of statements defending the accused (e.g. "victim's credibility is
    beside the point")
[x] Criticism of "rape culture"
[x] Incident of domestic violence
[x] Rape discussed in the context of abortion (pro-life/pro-choice debate)
[x] Graphic description of the act of violence
[x] Graphic description of the rape event (sexually explicit)
[x] Other: [...]

Type of coverage
Check box if discussion of rape is in context of one or both of the following
[x] News reporting (e.g. court proceedings, witness testimony)
[x] Opinion, editorial, letter, interview
[x] Announcement
[x] Entertainment or Fiction

Good example?
Check box if text is a particularly clear, unambiguous example of "rape culture."
[x] GOOD EXAMPLE

Tough Call?
Check box if text is ambiguous, or if you were otherwise uncertain about how to code
     it.
[x] TOUGH CALL

Errors
[x] Gibberish / Incomprehensible / Mistranslated / Missing text
[x] Article is not about an incident of rape / sexual assault
[x] Topic is not rape or sexual assault
[x] Multiple stories in one article
```

A.3 Intercoder reliability

Following initial training, each coder received a common set of 50-100 random articles, which they classified according to the coding instrument described above. We analyzed their data and met with coders to identify potential points of confusion and provide early corrective feedback. Next, the the coders received multiple larger training sets of 500 articles each, including a subset of overlapping articles to facilitate inter-coder reliability diagnostics. The team met bi-weekly, to reach a common conceptual understanding and identify difficult cases.

Table A.1 summarizes the human-coded training set data. By coder, the size of each training set ranges from $N_9 = 695$ (coder 9) to $N_{10} = 3034$ (coder 10). The total size of the training set (pooled across the 10 coders) is $N = 21,911$. $N_{\text{overlap}} = 341$ of these articles are common across all coders. Although our coding instrument identifies 76 discrete dimensions of rape culture, we focus on combined measures of the four main categories of rape culture. Each measure received a value of 1 if coders identified any of the sub-categories as appearing in a story, and 0 otherwise.

Table A.2 reports several measures of inter-coder reliability for the 341 overlapping articles in the training set. These statistics include:

1. *All agree*: percent of articles for which all coders assigned same value.

2. *Fleiss' Kappa*: reliability of agreement between multiple coders assigning categorical ratings. $\kappa = 1$ if the coders are in complete agreement and $\kappa \leq 0$ if there is no agreement beyond what would be expected by chance.

3. *Kendall's W*: non-parametric statistic of concordance. $w = 0$ indicates no agreement between coders and 1 indicated unanimous agreement.

4. *Krippendorff's Alpha*: coefficient of inter-coder reliability, with bootstrapped 95% confidence intervals. $\alpha = 1$ indicates perfect agreement, $\alpha = 0$ indicates random assignment of values, and $\alpha < 0$ indicates more disagreement than what can be expected by chance.

All intercoder reliability statistics in Table A.2 are positive and statistically significant, indicating more agreement that we would expect by chance. However, test statistics for Krippendorff's Alpha are consistently below .67, indicating moderate rather than high inter-coder reliability on that particular measure.

To take a closer look at these statistics, Figures A.1-A.4 report pairwise versions of these statistics, calculated separately for each individual pair of coders. The figures show that pairwise reliability statics vary considerably, by both coders and variables. While some individual statistics approach perfect agreement, no one pair consistently sees "eye to eye." For example while coders 8 and 9 received a Krippendorff's Alpha statistic of .75 for their work on the "victim-blaming" variable, their pairwise Alpha statistic dropped to .34 for "empathy for the accused." For this reason, rather than to cull a subset of training data from individual pairs of "kindred spirit" coders, we trained the classification algorithm of a pooled training set from all 10 coders, using median values (equivalent to

a majority vote) for overlapping articles. This approach effectively optimized training set size, at the partial expense of inter-coder reliability.

How does training set pooling affect subsequent analysis? Table A.1 shows that rape culture appears in between 3 and 27 percent of training documents, depending on coder and category. These baseline differences are due not only to differences in coder judgment – which Table A.2 shows to be generally acceptable – but also due to the random selection of articles into each training set. A classification algorithm that relies on one training set may therefore see rape culture as more (or less) frequent than an algorithm trained on another coder's data. Does pooling the training sets aggravate or reduce the potential risk of false negatives and false positives in the data?

Figure A.5 reports the proportion of test set documents classified into each category by SVM (see below for details on the algorithm), according to each individual training set (open circles) and the pooled training set (closed circles). In most cases, classifications generated with pooled training data are less sparse than ones generated with individual training sets. This is not surprising, considering that pooled training data contain more examples of documents that potentially belong to each category – and may thus pick up on nuances that individual training sets would miss. For some categories, however (e.g. empathy for accused, incredulity toward victims), pooled training data yield lower proportions than some of the individual training sets. For example, while one of the training sets produced a final dataset with rape culture in 20 percent of documents, the proportion with pooled training data is closer to 5 percent. In this sense, pooling potentially reduces the risk of false negatives due to overly-sparse training data, and offsets the influence of extreme outliers in the opposite direction. As we show in the section on cross-validation below, pooling the training data also significantly improves out-of-sample predictive accuracy (Figure A.7).

Table A.1: Training set summary statistics. Values shown are number of articles (percent of articles) that coders assigned to each category. *N* represents the total number of training set articles assigned to each coder.

	Victim-blaming	Empathy for accused	Implication of consent	Incredulity toward victim	Wrong topic	No error	N
Coder 1	59 (5%)	128 (12%)	88 (8%)	52 (5%)	385 (36%)	498 (46%)	1075 (100%)
Coder 2	153 (14%)	217 (20%)	88 (8%)	124 (11%)	442 (40%)	490 (45%)	1101 (100%)
Coder 3	223 (7%)	391 (13%)	397 (13%)	165 (5%)	873 (29%)	2069 (69%)	3013 (100%)
Coder 4	160 (13%)	344 (27%)	161 (13%)	123 (10%)	500 (40%)	558 (44%)	1259 (100%)
Coder 5	38 (4%)	101 (10%)	45 (4%)	33 (3%)	182 (18%)	777 (77%)	1005 (100%)
Coder 6	214 (7%)	349 (12%)	248 (8%)	192 (6%)	596 (20%)	1837 (61%)	3029 (100%)
Coder 7	258 (10%)	390 (15%)	357 (14%)	175 (7%)	1122 (43%)	1111 (42%)	2627 (100%)
Coder 8	107 (5%)	251 (12%)	87 (4%)	118 (6%)	811 (39%)	923 (45%)	2067 (100%)
Coder 9	93 (13%)	164 (24%)	144 (21%)	73 (11%)	219 (32%)	358 (52%)	695 (100%)
Coder 10	154 (5%)	174 (6%)	241 (8%)	105 (3%)	1111 (37%)	1381 (46%)	3034 (100%)

Table A.2: Intercoder reliability statistics (over all 10 coders).

	All Agree	*Fleiss' Kappa*	*Kendall's W*	*Krippendorff's Alpha* (95% confidence interval)	*N*
Victim-blaming	85.63	0.57***	0.66***	0.56 (0.36,0.72)	341
Empathy for accused	71.26	0.3***	0.45***	0.29 (0.12,0.46)	341
Implication of consent	76.25	0.42***	0.55***	0.41 (0.25,0.56)	341
Incredulity toward victim	87.68	0.41***	0.53***	0.38 (0.09,0.62)	341
Wrong topic	62.76	0.59***	0.69***	0.59 (0.48,0.68)	341
No error	50.44	0.54***	0.67***	0.54 (0.45,0.63)	341
Errors and wrong topics removed					
Victim-blaming	75.23	0.6***	0.69***	0.6 (0.49,0.71)	109
Empathy for accused	58.72	0.37***	0.51***	0.36 (0.23,0.5)	109
Implication of consent	53.21	0.42***	0.54***	0.42 (0.32,0.52)	109
Incredulity toward victim	75.23	0.43***	0.54***	0.41 (0.23,0.58)	109

Figure A.1: Intercoder reliability statistics: proportion of training set articles with identical labels ('All Agree'). Box shadings by quintile, light to dark.

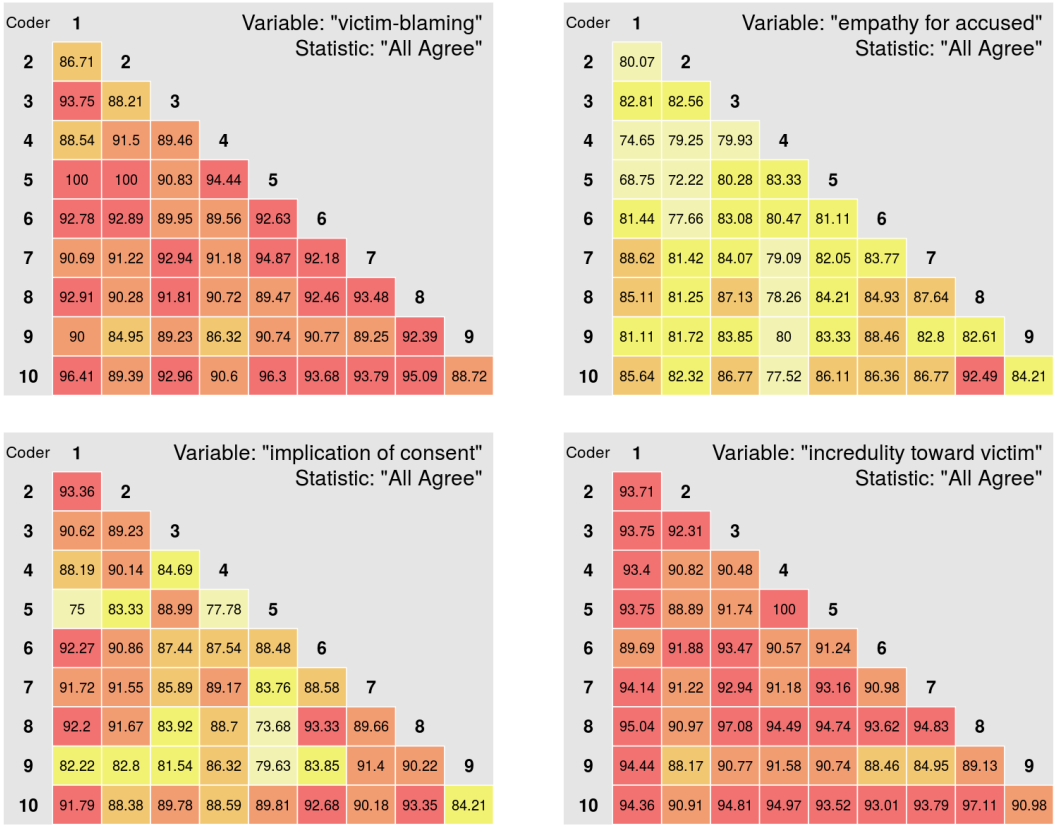


Figure A.2: Intercoder reliability statistics: pairwise Fleiss' Kappa statistics. Shadings by quintile, light to dark.

Variable: "victim-blaming"
Statistic: "Fleiss' Kappa"

Coder	1	2	3	4	5	6	7	8	9
2	0.22								
3	0.37	0.37							
4	0.24	0.62	0.4						
5	NaN	NaN	0.36	-0.03					
6	0.42	0.67	0.38	0.46	0.39				
7	0.29	0.58	0.54	0.54	0.47	0.55			
8	0.41	0.51	0.47	0.46	-0.06	0.56	0.61		
9	0.51	0.44	0.53	0.44	0.39	0.61	0.55	0.73	
10	0.57	0.4	0.46	0.41	0.65	0.54	0.57	0.63	0.45

Variable: "empathy for accused"
Statistic: "Fleiss' Kappa"

Coder	1	2	3	4	5	6	7	8	9
2	0.28								
3	0.32	0.35							
4	0.11	0.36	0.38						
5	0.09	0.26	0.24	-0.09					
6	0.29	0.2	0.38	0.41	0.32				
7	0.4	0.21	0.36	0.26	0.33	0.37			
8	0.3	0.29	0.4	0.25	-0.09	0.3	0.35		
9	0.42	0.45	0.47	0.18	0.5	0.61	0.33	0.36	
10	0.14	0.05	0.31	0.13	0.3	0.34	0.25	0.39	0.37

Variable: "implication of consent"
Statistic: "Fleiss' Kappa"

Coder	1	2	3	4	5	6	7	8	9
2	0.52								
3	0.52	0.4							
4	0.28	0.38	0.37						
5	-0.14	-0.09	0.27	-0.13					
6	0.5	0.39	0.35	0.34	0.26				
7	0.47	0.44	0.41	0.49	0.08	0.44			
8	0.46	0.41	0.21	0.37	-0.15	0.53	0.38		
9	0.28	0.18	0.23	0.4	0.23	0.34	0.64	0.63	
10	0.55	0.35	0.47	0.46	0.26	0.54	0.51	0.57	0.39

Variable: "incredulity toward victim"
Statistic: "Fleiss' Kappa"

Coder	1	2	3	4	5	6	7	8	9
2	0.56								
3	0.47	0.44							
4	0.45	0.4	0.28						
5	-0.03	-0.06	0.21	NaN					
6	0.18	0.46	0.43	0.34	0.19				
7	0.53	0.45	0.39	0.41	0.16	0.3			
8	0.58	0.41	0.67	0.57	-0.03	0.32	0.57		
9	0.68	0.29	0.4	0.38	0.24	0.22	0.22	0.44	
10	0.45	0.26	0.43	0.59	0.09	0.29	0.42	0.6	0.2

Figure A.3: Intercoder reliability statistics: pairwise Kendall's W statistics. Shadings by quintile, light to dark.

Variable: "victim-blaming"
Statistic: "Kendall's W"

Coder	1	2	3	4	5	6	7	8	9
2	0.63								
3	0.69	0.7							
4	0.63	0.81	0.71						
5	NaN	NaN	0.69	0.5					
6	0.72	0.84	0.69	0.73	0.7				
7	0.65	0.79	0.77	0.77	0.74	0.78			
8	0.71	0.76	0.74	0.73	0.5	0.78	0.81		
9	0.78	0.72	0.77	0.72	0.7	0.81	0.78	0.86	
10	0.79	0.72	0.73	0.72	0.82	0.78	0.79	0.82	0.74

Variable: "empathy for accused"
Statistic: "Kendall's W"

Coder	1	2	3	4	5	6	7	8	9
2	0.65								
3	0.66	0.68							
4	0.56	0.68	0.7						
5	0.63	0.7	0.62	0.46					
6	0.65	0.6	0.69	0.71	0.67				
7	0.7	0.63	0.68	0.65	0.67	0.68			
8	0.65	0.65	0.71	0.64	0.46	0.65	0.68		
9	0.71	0.72	0.74	0.61	0.77	0.82	0.68	0.7	
10	0.6	0.56	0.67	0.61	0.66	0.69	0.64	0.71	0.72

Variable: "implication of consent"
Statistic: "Kendall's W"

Coder	1	2	3	4	5	6	7	8	9
2	0.76								
3	0.76	0.7							
4	0.64	0.69	0.68						
5	0.44	0.5	0.64	0.45					
6	0.75	0.69	0.68	0.68	0.63				
7	0.74	0.72	0.71	0.74	0.55	0.72			
8	0.73	0.7	0.62	0.69	0.43	0.76	0.7		
9	0.65	0.6	0.63	0.71	0.63	0.68	0.83	0.82	
10	0.78	0.67	0.74	0.73	0.63	0.77	0.76	0.79	0.7

Variable: "incredulity toward victim"
Statistic: "Kendall's W"

Coder	1	2	3	4	5	6	7	8	9
2	0.78								
3	0.73	0.72							
4	0.73	0.7	0.64						
5	0.5	0.5	0.61	NaN					
6	0.59	0.73	0.72	0.67	0.6				
7	0.77	0.73	0.69	0.71	0.6	0.65			
8	0.79	0.71	0.84	0.78	0.5	0.66	0.79		
9	0.84	0.65	0.7	0.71	0.62	0.61	0.61	0.72	
10	0.73	0.64	0.72	0.8	0.55	0.65	0.71	0.8	0.62

Figure A.4: Intercoder reliability statistics: pairwise Krippendorff's Alpha statistics. Shadings by quintile, light to dark.

Variable: "victim-blaming"
Statistic: "Krippendorff's Alpha"

Coder	1	2	3	4	5	6	7	8	9
2	0.29								
3	0.34	0.39							
4	0.22	0.51	0.45						
5	1	1	0.35	-0.03					
6	0.36	0.72	0.37	0.47	0.28				
7	0.22	0.48	0.54	0.59	0.45	0.55			
8	0.44	0.51	0.54	0.48	-0.06	0.57	0.72		
9	0.5	0.39	0.43	0.41	0.35	0.62	0.61	0.75	
10	0.44	0.39	0.55	0.29	0.5	0.51	0.55	0.62	0.45

Variable: "empathy for accused"
Statistic: "Krippendorff's Alpha"

Coder	1	2	3	4	5	6	7	8	9
2	0.26								
3	0.36	0.3							
4	0.13	0.46	0.32						
5	0.07	0.29	0.21	-0.09					
6	0.31	0.2	0.41	0.39	0.32				
7	0.35	0.28	0.36	0.24	0.21	0.35			
8	0.25	0.27	0.4	0.24	-0.08	0.17	0.32		
9	0.42	0.48	0.52	0.14	0.43	0.63	0.35	0.34	
10	0.14	0.04	0.27	0.19	0.32	0.33	0.25	0.42	0.36

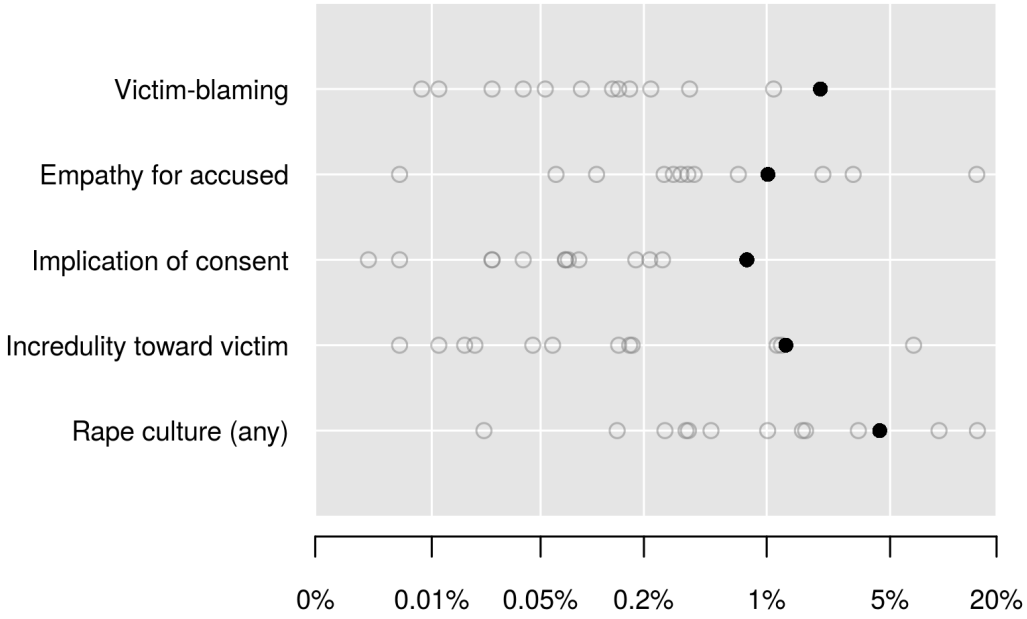
Variable: "implication of consent"
Statistic: "Krippendorff's Alpha"

Coder	1	2	3	4	5	6	7	8	9
2	0.53								
3	0.45	0.36							
4	0.31	0.38	0.39						
5	-0.15	-0.08	0.31	-0.13					
6	0.51	0.45	0.39	0.28	0.32				
7	0.53	0.45	0.41	0.43	0.1	0.42			
8	0.5	0.36	0.22	0.41	-0.14	0.5	0.41		
9	0.26	0.2	0.28	0.4	0.27	0.28	0.69	0.67	
10	0.48	0.27	0.42	0.43	0.29	0.45	0.55	0.54	0.35

Variable: "incredulity toward victim"
Statistic: "Krippendorff's Alpha"

Coder	1	2	3	4	5	6	7	8	9
2	0.5								
3	0.43	0.39							
4	0.57	0.21	0.27						
5	-0.03	-0.05	0.25	1					
6	0.23	0.45	0.34	0.39	0.27				
7	0.51	0.4	0.41	0.45	0.13	0.36			
8	0.54	0.28	0.65	0.51	-0.02	0.35	0.65		
9	0.69	0.33	0.43	0.41	0.25	0.19	0.27	0.34	
10	0.42	0.12	0.39	0.59	0.11	0.3	0.35	0.45	0.25

Figure A.5: Effect of training set pooling on SVM classification. Open circles represent proportions of test set documents classified as belonging to each category, using individual training sets. Closed circles represent proportions of test set documents classified as belonging to each category, using pooled training set. X-axis on logarithmic scale.



A.4 Classification algorithm

We used the randomly-selected reference texts in each training set to train a Support Vector Machine (SVM) classifier to predict the categories for all previously unseen corpus texts. The advantage of the SVM is that it is well-suited to sparse, high-dimensional data, is highly robust, and can handle a low training-to-test data ratio. Let $d \in \{1, \ldots, D\}$ index documents in a text corpus of size D. In our case, each d is a newspaper article on the topic of rape, published in the United States between 2000 and 2013. Let c denote a category (or label) to which document d may or may not belong. In our case, c is any one of the categories of rape culture described above.

The SVM classifies documents by fitting a maximally-separating hyperplane to a feature space, examining combinations of features that best yield separable categories. Formally, the SVM separates D data points from each other according to their labels ($y_d^c \in \{-1, 1\}$), and finds maximum marginal distance Δ between the points labeled $y_d^c = 1$ and $y_d^c = -1$, solving the optimization problem

$$\underset{\Delta,\alpha,\phi}{\arg\max}\, \Delta \text{ s.t. } y_d^c(\alpha + \phi(X_d)) > \Delta$$

where $y_d^c(\alpha + \phi(X_d)\beta)$ is a functional margin, $\phi()$ is a function that maps the training data X to a high-dimensional space, and $\mathbf{K}(x_d, x_{-d}) = \exp(-\gamma|\phi(x_d) - \phi(x_{-d})|^2)$ is a radial basis kernel function.

We created a document-term matrix for our corpus of newspaper articles, where the rows are documents $d \in \{1, \ldots, D\}$, columns are terms $t \in \{1, \ldots, T\}$, cell entries are weighted term frequencies, and each row vector $\mathbf{y}_d \in \mathbb{R}^T$ represents document d in a T-dimensional feature space. We weighted the features in the document-term matrix by term frequency/inverse document frequency (tf/idf) and normalized word counts,

$$tf.idf_{dt} = tf_{dt} \log\left(\frac{D}{df_t}\right)$$

where tf_{dt} is term frequency (number of times term appears in d), and df_t is document frequency (# documents with term t). A high $tf.idf_{dt}$ weight indicates that a term appears a lot in document d, but rarely in the corpus.

In the preprocessing stage, we removed HTML tags, control characters, non-alphanumeric characters, capitalization, punctuation and stopwords for all corpora, and ran a stemming algorithm to reduce the dimensionality of the matrix.

A.5 Cross-validation

We assessed the out-of-sample predictive performance of the SVM classifier with (a) k-fold cross-validation, and (b) randomly-repeated cross-validation. Each method examines the extent to which SVM is able to replicate the hand coding decisions, by splitting the full training set into subsamples, some of which it uses to fit the SVM model, and others it temporarily withholds for out-of-sample prediction.

k-fold cross-validation splits the data into k equally-sized sub-samples, sets one subset aside for out-of-sample tests, and fits the models to the remainder of data. The procedure then repeats for each of the other $k-1$ subsets, and averages predictive performance across the k test sets. The second approach randomly splits the dataset into an in-sample (for model-fitting) and out-sample (for prediction), and repeats m times, for different random splits of the same size. It then averages the SVM's predictive performance across out-samples in the m random splits. The main difference between the procedures is that k-fold cross-validation uses a one-time random split into k parts, while the randomly repeated method randomly splits the data again for each iteration. In the examples below, we used $k = 10$ folds and $m = 10$ random splits, with a 75/25 in-sample/out-sample mix for the latter routine.

We considered two measures of predictive performance. The first is *accuracy*, or proportion of hand-coded values correctly predicted,

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP and FP are the number of true and false positives, respectively, and TN, FN are the number of true and false negatives.

The second measure is the area under the receiver-operator characteristic (ROC) curve, or AUC,

$$\text{AUC} = \frac{E[rank(\hat{\pi}_d^c)|y_d^c = 1] - \frac{(\sum_d^D y_d^c)+1}{2}}{D - \sum_d^D y_d^c}$$

where $\hat{\pi}_d^c$ is the SVM-predicted probability that document d belongs to category c, and y_d^c is the hand-coded value for that document. An intuitive interpretation of the AUC is that it represents the probability that $\hat{\pi}$ is greater for a document with $y = 1$ than $y = 0$.

Table A.3 reports the results of the cross-validation exercise. Figure A.6 visualizes these results as a series of ROC curves, which plots the true positive rate ($\frac{TP}{TP+FP}$, vertical axis) against the false positive rate $\left(1 - \frac{TN}{TN+FN}\right)$. The solid line represents in-sample predictive performance, averaged over the $m = 10$ random splits, and the dashed line represents out-of-sample performance. The diagonal line represents $AUC = .5$, or "as-good-as-even" prediction.

Out-of-sample predictive performance is generally strong, with between 80 and 90 percent of labels correctly predicted in the test sets, on average. The AUC statistics suggest near-perfect classification performance in-sample, and .74 to .78 out-of-sample. Across

all tests and measures, performance was strongest for "victim-blaming language" and "implication of consent." It is lower – but still respectable – for "empathy for the accused."

Figure A.7 compares cross-validation results for the pooled training set against individual training sets. In most cases, pooling the training data (as described in Section A.3) improves predictive performance, relative to the average individual training set. Averaged over all variables, out-of-sample classification accuracy from k-fold cross-validation is 86.3 for the pooled training set, and 84.9 for individual training sets. Randomly-repeated cross-validation yields similar results, at 86.2 versus 84.7. AUC statistics reveal an even starker advantage for pooling. Averaged over all variables, in-sample AUC is .994 (k-fold) and .995 (randomly-repeated) for the pooled training set, and .853 (k-fold) and .857 (randomly-repeated) for individual training sets. For out-of-sample AUC, pooled training data scored .765 (k-fold) and .753 (randomly-repeated) on average, while individual training sets scored .653 (k-fold) and .657 (randomly-repeated). For some individual training sets, predictive accuracy was *worse* than random chance, with AUC's below .5. Pooling the training data helped overcome these problems.

Table A.3: Classification accuracies for SVM. Cross-validation results. Average statistics for $k = 10$ folds (left) or $m = 10$ random slits (right). 95% confidence intervals in parentheses.

method:	k-fold cross validation			Randomly-repeated cross validation		
measure:	Accuracy	Area under ROC curve		Accuracy	Area under ROC curve	
in/out sample:	out-sample	in-sample	out-sample	out-sample	in-sample	out-sample
Victim-blaming	88.83	0.993	0.78	88.64	0.996	0.78
	(88.58,89.04)	(0.992,0.994)	(0.73,0.82)	(88.38,88.87)	(0.993,0.997)	(0.75,0.79)
Empathy for accused	80.39	0.988	0.75	80.39	0.99	0.74
	(80.16,80.69)	(0.987,0.99)	(0.72,0.77)	(79.91,80.67)	(0.989,0.992)	(0.73,0.75)
Implication of consent	90.08	0.997	0.78	90.01	0.998	0.78
	(89.97,90.22)	(0.996,0.998)	(0.76,0.82)	(89.78,90.32)	(0.997,0.999)	(0.77,0.8)
Incredulity toward victim	86.03	0.996	0.75	85.83	0.997	0.75
	(85.82,86.23)	(0.996,0.997)	(0.73,0.78)	(85.54,86.08)	(0.995,0.998)	(0.74,0.76)

Figure A.6: Receiver-operator characteristic curves. Randomly-repeated cross-validation.

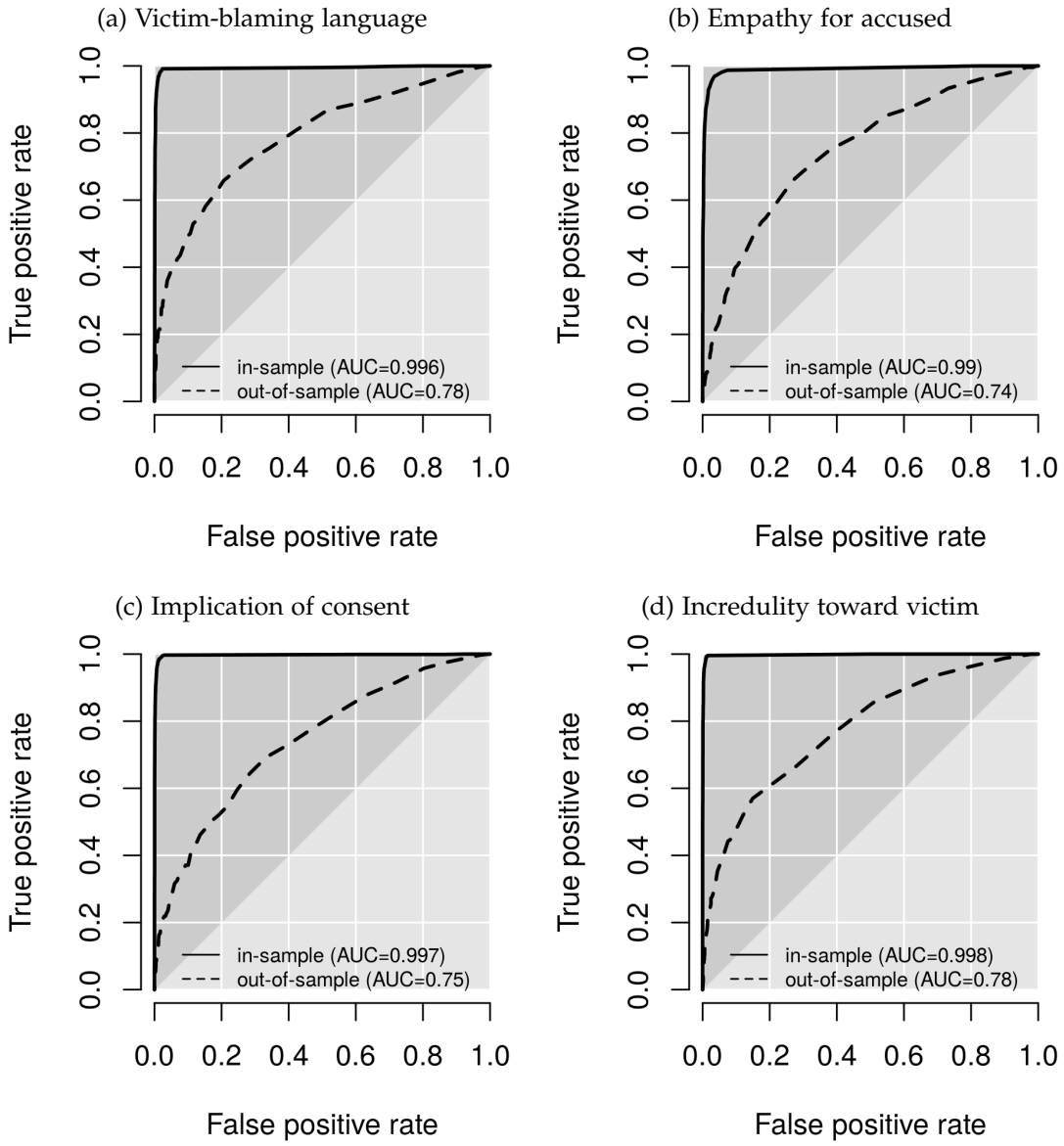
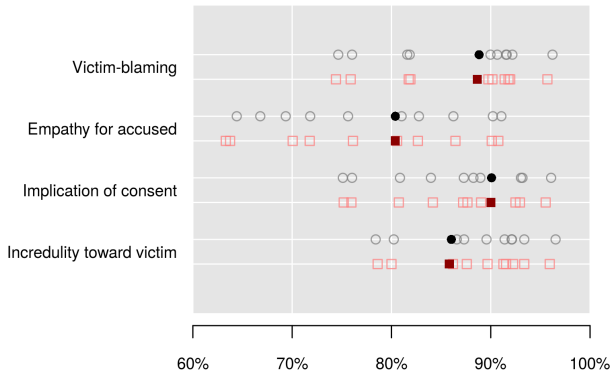
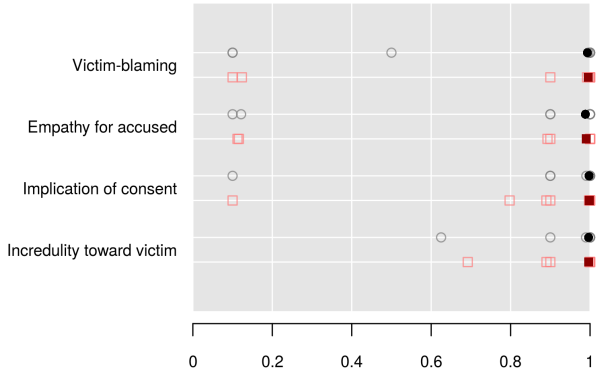


Figure A.7: Cross-validation results for pooled vs. individual training data. Black circles represent results for k-fold cross-validation. Red squares are results for randomly-repeated cross-validation. Open circles/squares represent predictive performance from individual training sets. Closed circles/squares are the same from the pooled training set.

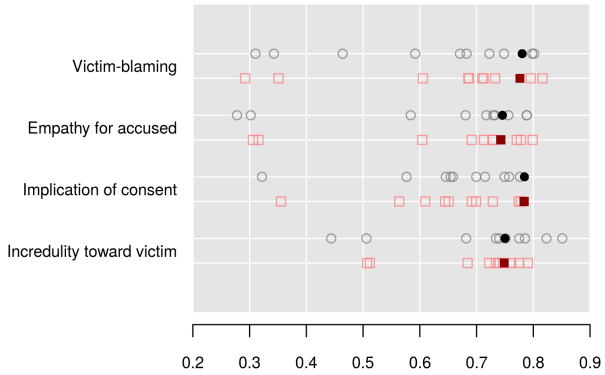
(a) Accuracy



(b) In-sample AUC



(c) Out-of-sample AUC



A.6 Classification results

The current section summarizes the article-level results of SVM classification.

1. Table A.4 presents the proportion of newspaper articles that SVM classified into each category. Formally, $P(y_d^c = 1)$, where c is the category and d is the document (article). According to the table, just 3.2 percent of newspaper articles on rape contained any of the four categories of rape culture. The proportion of articles that included all four main categories was a tiny fraction of 1 percent.

2. Figure A.8 displays wordclouds for the main categories, with terms vertically positioned according to tf-idf weights.

3. Figure A.9 presents a similar term-category matrix as reported in the main text, but with points sized according to average in-category td-idf weights.

4. Figure A.10 reports the relative probability of terms appearing articles of each category of rape culture (x-axis), relative to documents not in the category (y-axis). Terms that are in the lower triangle are thus more likely than not to appear in articles with rape culture.

Table A.4: Summary statistics, SVM-classified articles.

Category (c)	$P(y_d^c = 1)$	$SD(y_d^c = 1)$
Victim-blaming language	0.013	0.115
Empathy for accused	0.008	0.091
Implication of consent	0.005	0.070
Incredulity toward victim	0.010	0.097
No rape culture	0.968	0.175
Any rape culture category	0.032	0.175
All 4 rape culture categories	0.00004	0.007

Figure A.8: Relative frequencies of words in newspaper articles, by category. Font size and vertical position of words organized by average term frequency - inverse document frequency (tf-idf) weight of word in articles classified by the SVM algorithm as belonging to each category (a) - (f).

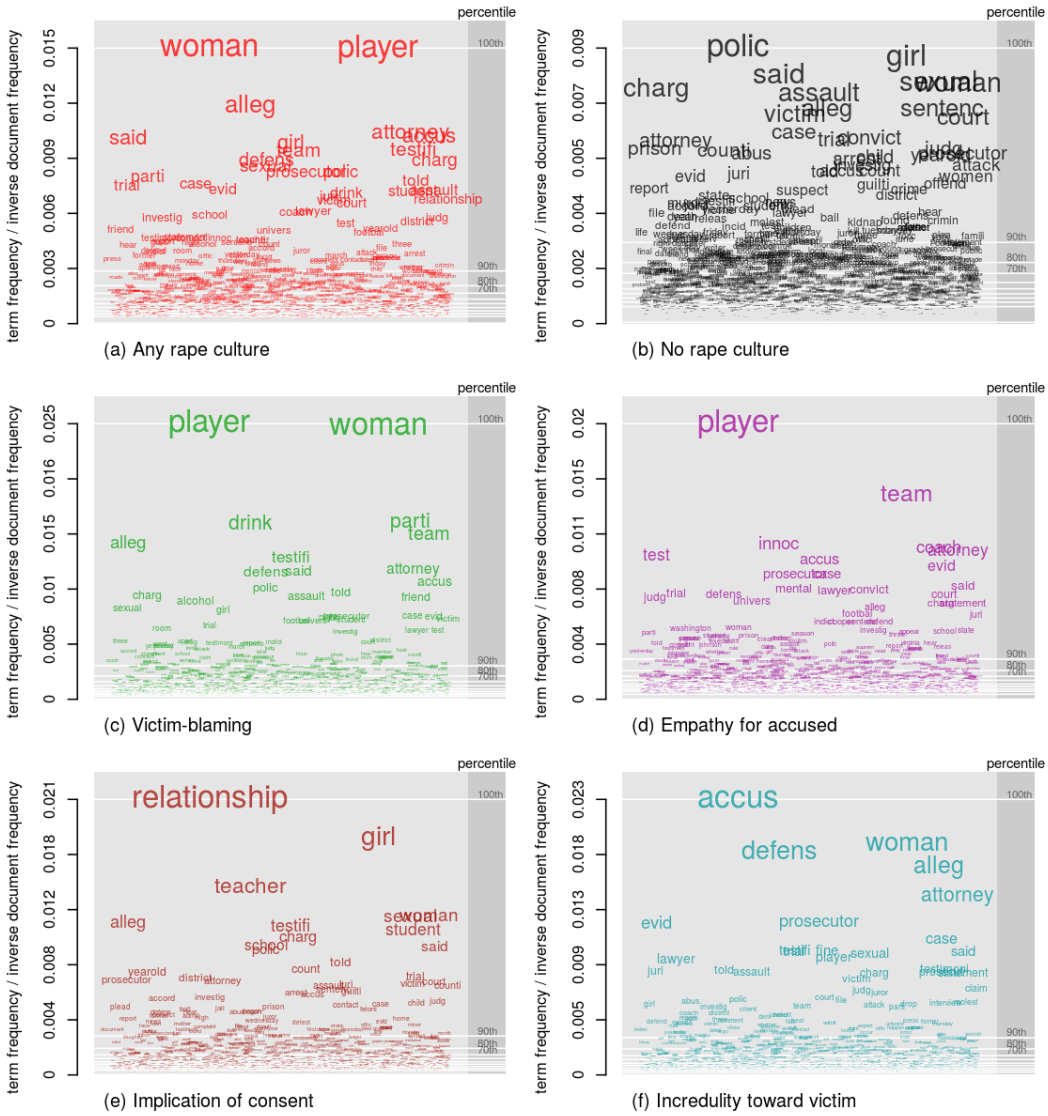


Figure A.9: Relative frequencies of words in newspaper articles, by category. Symbol size proportional to average term frequency - inverse document frequency (tf-idf) weight of word in articles classified by the SVM algorithm as belonging to each category (a) - (f). List restricted to subset of words in top 95th percentile by tf-idf weights.

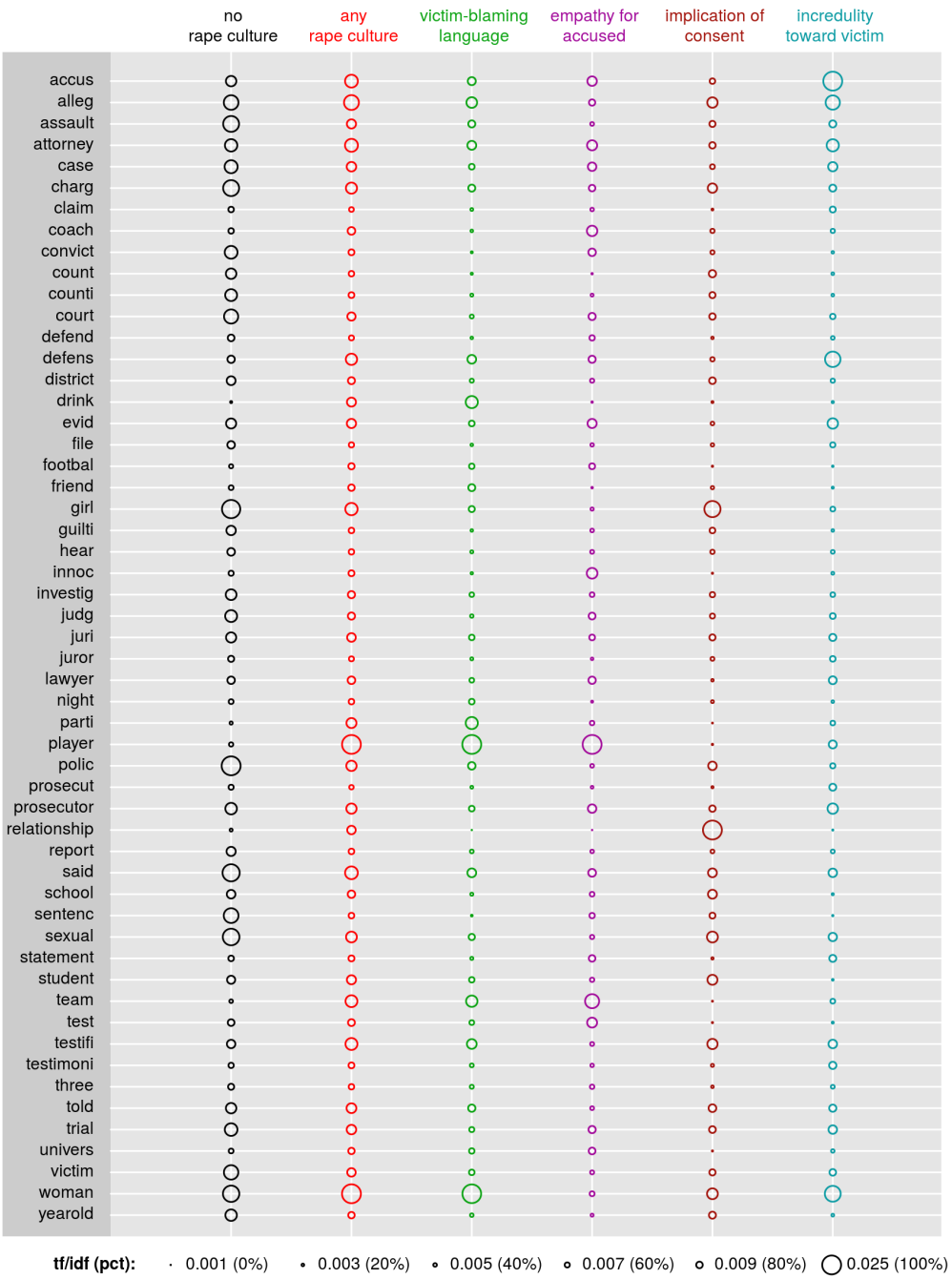
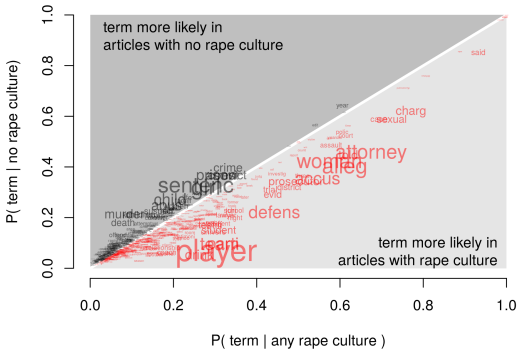
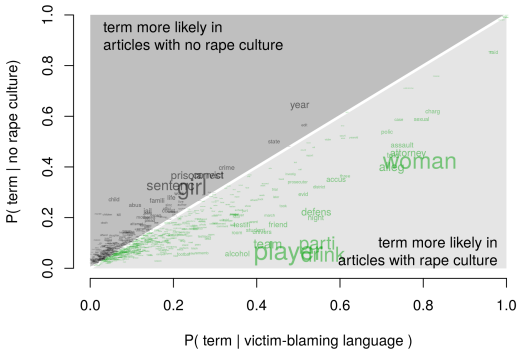


Figure A.10: Relative probability of words in newspaper articles, by category. x-axis represents probability that an article in each SVM-classified category of rape culture contains term. y-axis represents probability that an article not in that category contains term. Font size proportional to average term frequency - inverse document frequency (tf-idf) weight.

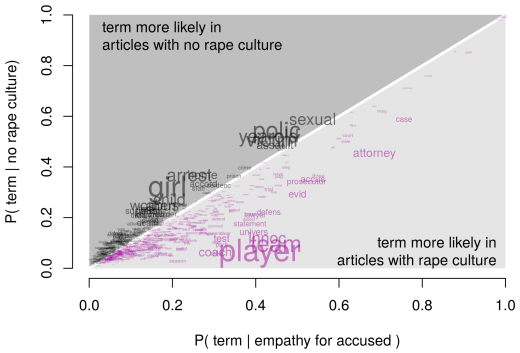
(a) Any rape culture



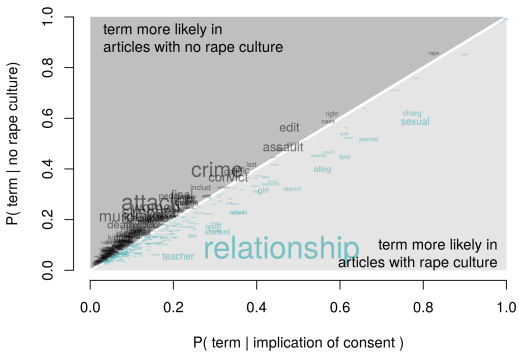
(b) Victim-blaming



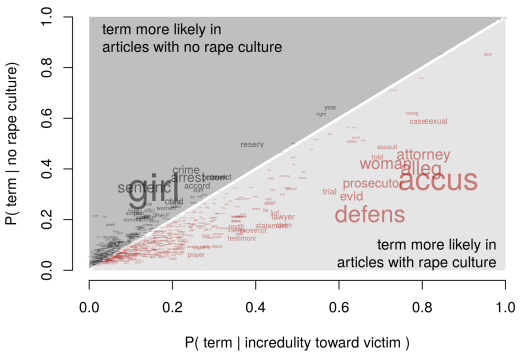
(c) Empathy for accused



(d) Implication of consent



(e) Incredulity toward victim



A.7 Aggregation to county-years

We aggregated SVM-classified article-level indicators to the county-year level, as local proportions of newspaper stories in the local media market containing each category of rape culture. To establish a definition of "local news," we matched newspaper articles to counties in two ways:

1. *Content producers* weights newspapers based on the geographic proximity of each outlet's main bureau to the county center. Formally,

$$x_{i,t}^{ALL} = \sum_{k}^{K} w_{ik} x_{k,t}^{ALL} \tag{A.1}$$

$$x_{i,t}^{RC} = \sum_{k}^{K} w_{ik} x_{k,t}^{RC} \tag{A.2}$$

$$x_{i,t} = \frac{x_{i,t}^{RC}}{x_{i,t}^{ALL}} \tag{A.3}$$

   where $x_{i,t}$ is the proportion of local newspaper articles about rape with content suggestive of rape culture (equation A.3).

   $x_{k,t}^{ALL}$ (in equation A.1) is the total number of articles about rape, published by newspaper k in year t. $x_{k,t}^{RC}$ (in equation A.2) is the number of those articles that the SVM algorithm classified as containing any of the four main categories of rape culture (victim-blaming language, empathy for the accused, implication of consent, incredulity toward victim). $x_{i,t}^{ALL}$ and $x_{i,t}^{RC}$ are weighted sums of these two measures, at the level of county-year.

   The newspaper weights (w_{ik}) are based on a geographic nearest-neighbor search:

$$w_{ik} = \begin{cases} 1 & \text{if } d(i,k) \leq d_{(r)}(i,k) \\ 0 & \text{if } d(i,k) > d_{(r)}(i,k) \end{cases} \tag{A.4}$$

   where $d(i,k)$ is the distance between a country center and main bureau of k, and $d_{(r)}(i,k)$ is the r-th nearest bureau to i. All newspapers with $d(i,k) \leq d_{(r)}(i,k)$ receive equal weight. These weights ensure that all counties, including ones without a nearby bureau, have the same number of newspapers.

   In the main text, we calculated these weights based on the $r = 5$ nearest newspapers. In Section C.1, we report a sensitivity analysis for all $r \in \{1, \ldots, 20\}$.

2. *Content consumers* weights newspapers based on their relative market shares in the county. Calculation of the local proportion is the same as in equations A.1-A.3. The newspaper weights (w_{ik}) take a different form:

$$w_{ik} = \frac{\text{circulation}_{ik}}{\sum_{k}^{K} \text{circulation}_{ik}} \tag{A.5}$$

where circulation_{ik} is newspaper k's total circulation in county i, according to annual data from the Alliance for Audited Media.

A.8 Summary statistics at county-year level

Table A.5 reports summary statistics at the county-year level, following each aggregation process as described above.

Table A.5: Summary statistics, county-year level data.

	N	*Mean*	*Std.Dev.*	*Min*	*Max*
Rape culture variables ('producers')					
Any rape culture	43454	0.028	0.020	0	0.269
Victim-blaming	43454	0.011	0.014	0	0.172
Empathy for accused	43454	0.007	0.009	0	0.212
Implication of consent	43454	0.009	0.009	0	0.114
Victim's credibility	43454	0.004	0.006	0	0.104
Rape culture variables ('consumers')					
Any rape culture	43463	0.007	0.016	0	0.161
Victim-blaming	43463	0.002	0.008	0	0.143
Empathy for accused	43463	0.002	0.007	0	0.117
Implication of consent	43463	0.002	0.007	0	0.121
Victim's credibility	43463	0.001	0.004	0	0.059
Crime					
Reported rapes per 1,000 people	43463	0.689	3.752	0	201.257
Rape arrests per 1,000 people	43463	0.224	1.310	0	69.161
Police vigilance (arrests minus reports per 1,000 people)	43463	-23.323	68.415	-2193	703
Control variables					
Median personal income	43463	31574.452	8694.114	8579	132728
Percent urban population	43463	41.385	31.414	0	100
Percent female population	43463	50.315	2.042	0	58
Percentage of workers unemployed	43463	6.227	2.854	0.900	29
Percent population in religious congregation	43463	52.833	20.807	0.099	100
Percent presidential vote for Republican	43463	58.476	13.338	5.980	93.290
Instrumental variable					
Circulation change	40358	-1431	59464	-1052650	1053244

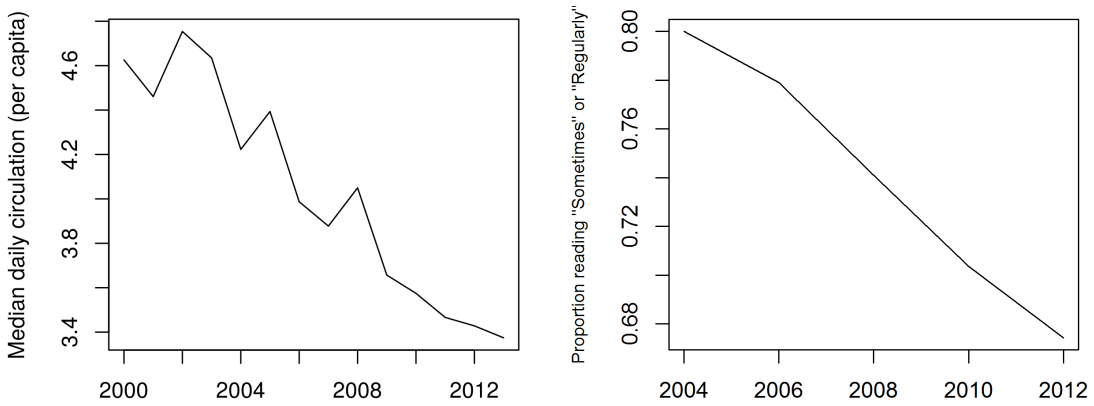
B Analysis

B.1 Instrumental variable design and newspaper readership

The following section substantiates some of the claims made in the main paper in support of the instrumental variable design. In particular, we show that (a) the print media market has seen an overall decline in readership in the past decade, and (b) the demographic group among which the print media market has seen the greatest relative growth – single women – is less likely to read newspapers regularly where rape culture is high.

Figure B.1 reports the overall recent decline in newspaper circulation and readership. Over the period 2000-2012, the annual decline in daily circulation of local newspapers in each U.S. county was -1364 (median -439.5).[2] In the Biennial Media Consumption Survey conducted by the Pew Research Center, which collects data on newspaper readership every two years, the proportion of respondents who reported reading a newspaper "Sometimes" or "Regularly" has declined by an average of 4.18 percent between each two-year period. The proportion that read a newspaper "Yesterday" has declined by an average of 5.69 percent.

Figure B.1: Decline in newspaper circulation and readership.



(a) Median per capita daily circulation of newspapers in local media market

(b) Proportion of Pew respondents reading newspaper "Sometimes" or "Regularly"

How have these trends affected newspaper coverage of rape? As the cross-tabulation in Table B.1 reports, rape culture is lower than we would expect by chance where the average frequency of newspaper readership declined from year-to-year, according to Pew (upper-right cell of 2 by 2 table). We find a similar result when using an alternative measure of newspaper readership, where respondents report reading a newspaper "Yesterday" (Table B.2) as opposed to "Sometimes" or "Regularly." However, these latter results are only marginally significant, at $p<.10$.[3]

[2] Local media market is here defined as the 5 newspapers with head bureaus closest to each county center.

[3] Note that similar relationships also emerge when we employ newspaper readership estimates from the Pew surveys in place of our circulation data.

Table B.1: Rape culture and newspaper readership. Decline in readership defined as an average decline in proportion of county residents who read a newspaper frequently (i.e. "Regularly" to "Sometimes", "Sometimes" to "Hardly ever", "Hardly ever" to "Never"). N=988 counties with media coverage and Pew data on all covariates, matched by zip code. Numbers in parentheses are expected cell frequencies.

		Rape culture		
		Below average	Above average	*Total*
Newspaper readership	Decline	214 (201.6)	141 (153.4)	355
	Increase/no change	347 (359.4)	286 (273.6)	633
	Total	561	427	988

Pearson's Chi-squared statistic = 2.77, d.f. = 1, p = 0.096

Table B.2: Rape culture and newspaper readership (alternate). Decline in readership is as an average decline in proportion of county residents who read a newspaper "Yesterday." N=1088 counties. Numbers in parentheses are expected cell frequencies.

		Rape culture		
		Below average	Above average	*Total*
Newspaper readership	Decline	168 (154.6)	103 (116.4)	271
	Increase/no change	407 (420.4)	330 (316.6)	737
	Total	575	433	1088

Pearson's Chi-squared statistic = 3.7, d.f. = 1, p = 0.054

To show this result slightly more robustly, we report the full first-stage IV regression results in Tables B.4 (producers measure) and B.5 (consumers measure). The instrumental variable here is annual change in circulation by county. The coefficient is positive and highly significant as a predictor of both "Any rape culture" and its four components. The tables also report the statistics of tests for underidentification, overidentification and weak instruments. Each of the test statistics for these tests in Table B.4 falls within conventional bounds of significance. The F statistic for the Kleibergen-Paap rk LM underidentification test is high for all models, indicating that the circulation instrument is correlated with all measures of rape culture. The Kleibergen-Paap rk Wald F statistic further allows us to reject the null hypothesis that the circulation instrument is only weakly correlated with rape culture. In Table B.5, only the "any rape culture" and "empathy" variables produced significant estimates.

What is the mechanism behind this result? Why would newspapers respond to a decline in readership by reducing gendered biases in their coverage of rape? Part of the answer may lie in Table B.6, which breaks down the Pew survey results by demographic group. Single women (i.e. never married or co-habiting with partner) are, by far, the fastest-growing demographic group among respondents who reported reading a newspaper "Sometimes" or "Regularly." Given the pronounced trend toward greater newspaper reading by this demographic group – which stands in stark contrast to most other groups – we would anticipate their increased centrality to newspapers' bottom lines to result in increased concern among publishers to avoid alienating them (Hamilton, 2004). We may

Table B.3: Average bi-annual change in newspaper readership across demographic groups. Change defined as bi-annual difference in proportion of Pew survey respondents who reported reading a newspaper "Sometimes" or "Regularly."

Group	Average change (%)
Single women	8.63
Over 100K income	5.95
Liberal	4.95
Single (all)	3.95
Single men	2.65
Men (all)	1.06
Married men	0.85
Women (all)	0.25
Married (all)	-0.88
College-educated	-1.06
Married women	-1.14
Newspaper readers	-4.18

further expect single women, to a greater extent than other groups, to be particularly attentive to gendered biases in news coverage.

Some indirect evidence substantiating this last expectation is available in a 2012 survey by CBS News.[4] One question in the survey asked respondents about the following controversial comments by Missouri Republican Representative and Senate nominee Todd Akin: "It seems to me, from what I understand from doctors, that's really rare. If it's a legitimate rape, the female body has ways to try to shut that whole thing down. But let's assume that maybe that didn't work or something: I think there should be some punishment, but the punishment ought to be of the rapist, and not attacking the child."[5] The survey question was as follows: "How much have you heard or read about the controversy over comments made by Missouri Congressman Todd Akin (AY-kin) regarding abortion and women who have been raped – a lot, some, or not much?"

We can employ the extent of awareness of Akin's comments regarding "forcible rape" as a rough proxy for the salience of rape and compare the extent of awareness of Akin's comments among single women relative to married individuals of either gender (note that single vs. married men are statistically indistinguishable). A t-test indicates that single Republican women were statistically significantly more aware of Akin's comments than married respondents of either gender (by .52 standard deviations, $p<.03$).[6] Though by no means definitive, this pattern appears consistent with our expectation that single women are likely to be particularly attentive to gender biases in the news, at least with respect to rape.

The Pew data also support our expectations. As the cross-tabulation in Table B.6

[4]CBS News/60 Minutes/Vanity Fair Poll: 2012 Presidential Election/Economy. August 22-26, 2012 (USCBS2012-08B).

[5]Source: http://www.nytimes.com/2012/08/20/us/politics/todd-akin-provokes-ire-with-legitimate-rape-comment.html (accessed 9/23/15).

[6]Because this was a highly partisan issue, leading many Democrats to tune in for potentially orthogonal reasons, we focus on Republicans. We also exclude several ambiguous, and sparsely populated, categories (widowed, separated, divorced).

Table B.4: First-stage instrumental variable regression results (producers set)

	Dependent variable:				
	(1)	(2)	(3)	(4)	(5)
	Any rape culture	Victim-blaming	Empathy for accused	Implication of consent	Questioning victim's credibility
Circulation change, lagged	0.048***	0.018***	0.055***	0.021***	0.043***
	(0.0044)	(0.0041)	(0.0053)	(0.0054)	(0.0055)
Median personal income	0.0080	0.016	-0.032′	-0.014	-0.023
	(0.017)	(0.016)	(0.017)	(0.019)	(0.018)
Percent female population	0.027**	0.032***	-0.0069	0.040***	-0.016′
	(0.0094)	(0.0082)	(0.0088)	(0.012)	(0.0089)
Percentage of workers unemployed	-0.22***	-0.15***	-0.092***	-0.18***	-0.081***
	(0.014)	(0.014)	(0.012)	(0.015)	(0.013)
Percent population in religious congregation	0.25***	0.13**	0.38***	-0.13*	0.31***
	(0.048)	(0.040)	(0.055)	(0.060)	(0.042)
Percent presidential vote for Republican	-0.31***	-0.13***	-0.23***	-0.35***	0.039′
	(0.024)	(0.017)	(0.026)	(0.026)	(0.020)
Constant	-0.042**	-0.38***	0.027′	0.61***	-0.086***
	(0.016)	(0.014)	(0.016)	(0.022)	(0.019)
Observations	37,244	37,244	37,244	37,244	37,244
R-squared	0.187	0.254	0.263	0.079	0.224
Number of CL_GEOID	3,105	3,105	3,105	3,105	3,105
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
LL	-47734	-46078	-46162	-49154	-47070

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

shows, rape culture was significantly lower that we would expect by chance in counties where single women represented a higher-than-average proportion of newspaper readers ("Sometimes" or "Regularly"). As Table B.7 further shows, this relationship among women overall (married or single) – among whom readership has remained steady – is in the same direction, but not statistically significant. These patterns suggest that single women are becoming increasingly important to newspapers' bottom lines, and the content of newspaper coverage is potentially reflecting this influence where this group represents a sizeable share of the local newspaper consumer base.

In the 5% of U.S. counties that have highly competitive local newspaper markets (i.e. where there are at least two newspapers with head bureaus located within the county's borders) the inverse relationship between single female readers and the prevalence of rape culture in the news is stronger still. The reason, presumably, is that as competition for readers rises, the economic incentive to cater to a demographic group, like single women, who comprise an increasingly consequential proportion of the readership base, also heightens. Table B.8 reports the results of logit regression models of higher-than-average rape culture (producers measure) on several covariates, including indicators of whether the proportion of single female Pew respondents who reported reading a newspaper "Sometimes" or "Regularly" was higher than the national average, and whether the county saw an average annual increase in the proportion of Pew respondents who reported having read a newspaper yesterday. We fit the same model to data on all counties for which both Pew and media coverage data were available (N=990, Model 1) and the subset of counties where at least two newspapers were based and Pew data were available (N=46, Model 2). Figures B.2 and B.3 report simulations from these models.

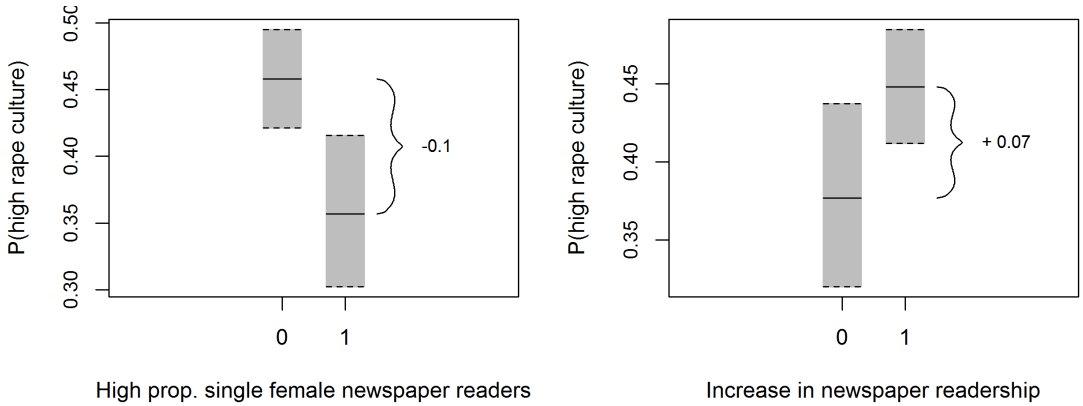
Table B.5: First-stage instrumental variable regression results (consumers set)

	Dependent variable:				
	(1)	(2)	(3)	(4)	(5)
	Any rape culture	Victim-blaming	Empathy for accused	Implication of consent	Questioning victim's credibility
Circulation change, lagged	0.012**	-0.0069	0.031***	-0.00085	-0.0029
	(0.0042)	(0.0059)	(0.0049)	(0.0030)	(0.0068)
Median personal income	0.031′	0.096***	-0.041**	0.055*	-0.036*
	(0.016)	(0.019)	(0.015)	(0.022)	(0.014)
Percent female population	0.010	0.029*	-0.0076	-0.0027	0.011
	(0.0085)	(0.012)	(0.0086)	(0.0086)	(0.0088)
Percentage of workers unemployed	-0.024′	-0.0083	-0.11***	0.025	0.067***
	(0.013)	(0.014)	(0.016)	(0.016)	(0.015)
Percent population in religious congregation	0.12**	-0.051	0.35***	0.021	-0.064
	(0.040)	(0.040)	(0.060)	(0.040)	(0.043)
Percent presidential vote for Republican	-0.048**	-0.038*	-0.0035	-0.12***	0.18***
	(0.016)	(0.018)	(0.022)	(0.018)	(0.024)
Constant	-0.15***	-0.28***	-0.0054	0.0053	-0.15***
	(0.015)	(0.016)	(0.016)	(0.021)	(0.015)
Observations	37,253	37,253	37,253	37,253	37,253
R-squared	0.048	0.065	0.053	0.014	0.058
Number of CL_GEOID	3,105	3,105	3,105	3,105	3,105
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
LL	-40786	-46648	-43230	-47279	-47416

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

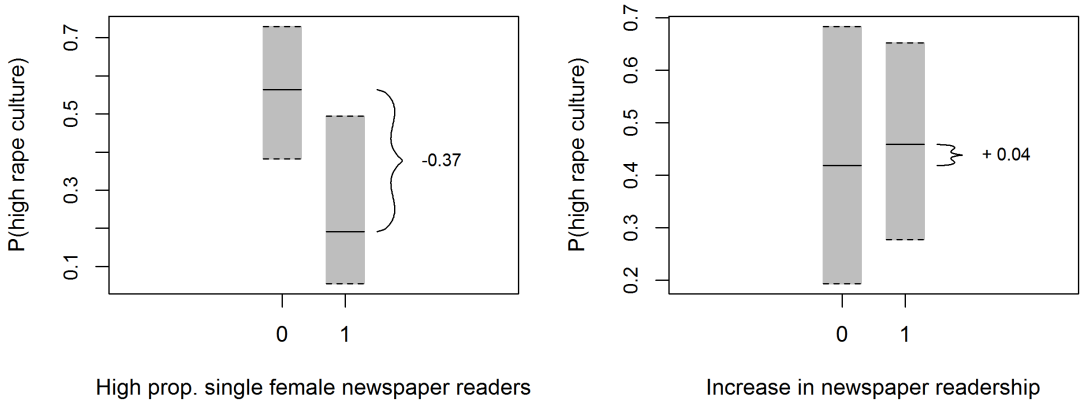
As the simulations show, counties with an above-average proportion of single female newspaper readers were significantly less likely to feature high levels of rape culture in the press, after controlling for other potential confounding factors, like the average education, age and religiosity of the county's residents. The size of this negative impact was almost four times greater in the 46 counties with highly-competitive local newspaper markets, than in U.S. counties overall.

Figure B.2: Single female newspaper readership and rape culture in the press. "High proportion of single female newspaper readers" defined as county where the proportion of single female Pew respondents who read a newspaper "Sometimes" or "Regularly" is higher than average. Simulations based on Model 6 in Table B.8.



(a) Single female news consumers (b) Newspaper readership change

Figure B.3: Single female readership and rape culture (highly-competitive local newspaper markets). "High proportion of single female newspaper readers" is a county where the proportion of single female Pew respondents who read a newspaper "Sometimes" or "Regularly" was higher than average. Simulations based on Model 1 in Table B.8.



(a) Single female news consumers (b) Newspaper readership change

Table B.6: Rape culture and single female newspaper readership. Rows indicate whether the proportion of newspaper readers ("Sometimes" or "Regularly") who were single women was higher than the national average. N=984 counties with media coverage and Pew data on all relevant covariates, matched by zip code. Numbers in parentheses indicate expected frequencies per cell.

		Rape culture		
		Below average	Above average	*Total*
Proportion single female	Below average	382 (400.8)	321 (302.2)	703
newspaper readers in county	Above average	179 (160.2)	102 (120.8)	281
	Total	561	423	984

Pearson's Chi-squared statistic = 7.18, d.f. = 1, p = 0.007

Table B.7: Rape culture and ALL female newspaper readers. Rows indicate whether the proportion of newspaper readers ("Sometimes" or "Regularly") who were women (single or married) was higher than the national average. N=984 counties with media coverage and Pew data on all relevant covariates, matched by zip code. Numbers in parentheses indicate expected frequencies per cell.

		Rape culture		
		Below average	Above average	*Total*
Proportion female	Below average	300 (303.9)	233 (229.1)	533
newspaper readers in county	Above average	261 (257.1)	190 (193.9)	451
	Total	561	423	984

Pearson's Chi-squared statistic = 0.25, d.f. = 1, p = 0.617

Table B.8: Logit regression results, high (above-average) rape culture (producers measure). Level of analysis is county. Highly competitive local market is defined as a county home to at least two newspapers' main bureaus.

	Dependent variable:	
	(1)	(2)
	High rape culture	High rape culture (highly-competitive markets)
High proportion single female newspaper readers	-0.420***	-1.696**
	(0.148)	(0.845)
Increase in newspaper readership	0.295**	0.165
	(0.149)	(0.702)
College-educated	-0.489	5.865
	(0.737)	(4.698)
Proportion age 20-34	-7.732**	-9.357
	(3.730)	(16.836)
Proportion population in religious congregation	0.065*	5.408
	(0.037)	(5.768)
Constant	1.149	0.072
	(0.730)	(3.757)
Observations	990	46
Log Likelihood	-670.894	-34.250

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

B.2 Monte Carlo study: circulation shocks and newspaper weights

Because one of our two measures of rape culture uses a subscription-based weighted average (*consumers*), one may worry that shocks to overall local newspaper circulation, as defined by our instrumental variable, may affect the measure of local news content in a linear, mechanistic way. In this section, we perform a simple simulation to show that this is not the case, and that we can reasonably attribute variation in the independent variable to changes in news content, rather than changes in weights due to local market shocks.

If we recall equation A.5, we can think of our instrumental variable as the first difference in the denominators of the weights from year to year (total local market size). This first difference, however, does not directly affect, in a consistent manner, each newspaper's relative local subscriber base (local market share). It is possible, for instance, that a newspaper's market share could rise in a declining market, or fall in a growing market.

To illustrate this, we performed a Monte Carlo study, in which we hold (unweighted) newspaper coverage constant, expose our newspaper weights to hypothetical shocks to the local media market, and re-aggregate the article-level classification results to the level of the county-year, per equation A.5.

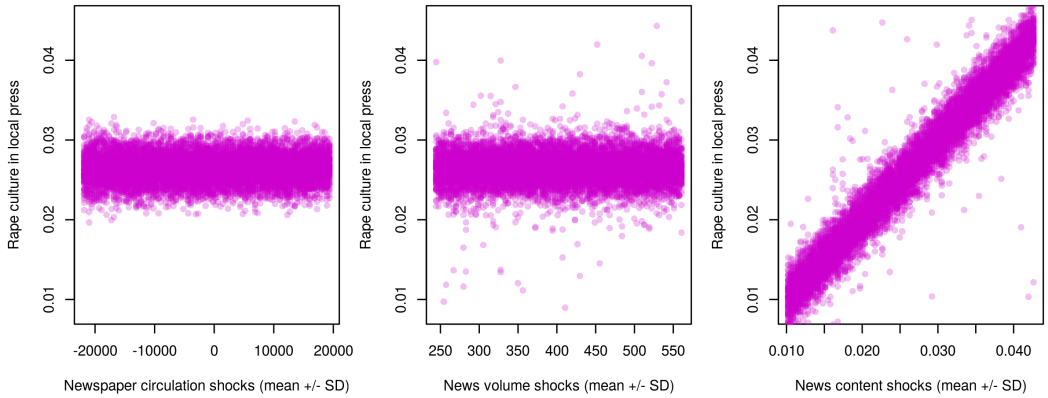
In each of 10,000 rounds, our simulation did the following:

- each newspaper k prints $x_k^{ALL} \sim N\left(\mu_k\left(x^{ALL}\right), \sigma_k^2\left(x^{ALL}\right)\right)$ articles on rape
- some latent proportion $\rho_k \sim N\left(\mu_k(\rho), \sigma_k^2(\rho)\right)$ of these will contain rape culture, where $\rho_k = \frac{x_k^{RC}}{x_k^{ALL}}$
- there is a shock of size $S_i \sim N\left(\mu_i(S), \sigma_i^2(S)\right)$ to local circulation in county i, where $S_i = \text{circulation}_{it} - \text{circulation}_{it-1}$
- based on the randomly-drawn values of x_k^{ALL}, x_k^{RC} (from ρ_k) and S_i, calculate circulation weights $\hat{w}_i$ and circulation-weighted average number of articles about rape per county-year ($\hat{x}_i^{ALL}$), and the circulation-weighted average number of articles with rape culture ($\hat{x}_i^{RC}$)
- based on the simulated values of $\hat{x}_i^{ALL}, \hat{x}_i^{RC}$, calculate the local prevalence of rape culture in news content: $\hat{x}_i = \frac{\hat{x}_i^{RC}}{\hat{x}_i^{ALL}}$

To seed the simulation, we took initial values for $\mu_k\left(x^{ALL}\right), \sigma_k^2\left(x^{ALL}\right), \mu_k(\rho), \sigma_k^2(\rho), \mu_i(S), \sigma_i^2(S)$ directly from the data. Using this basic algorithm, we examined how the simulated values of x_i vary across three scenarios:

1. How does $\hat{x}_i$ depend on circulation shocks ($E[S_i] \pm SD(S_i)$)?
2. How does $\hat{x}_i$ depend on news events (number of articles per paper, $E[x_k^{ALL}] \pm SD(x_k^{ALL})$)?
3. How does $\hat{x}_i$ depend on actual rape culture ($E[\rho_k] \pm SD(\rho_k)$)?

Figure B.4: Simulation results: circulation shocks and local news content ($\hat{x}_i$). Each dot represents the outcome of one of 10,000 simulations.



The results of the simulation are in Figure B.4. The figure shows that shocks to local circulation do not mechanistically increase or decrease the local proportion of news stories with rape culture. As we vary the size of hypothetical shocks from one standard deviation below to one standard deviation above the mean in S_i, the simulated values of $\hat{x}_i$ remain essentially constant (left pane). Shocks to news volume increase the uncertainty of $\hat{x}_i$, but not the mean (middle pane). The only discernible change to the mean follows shocks to the underlying newspaper-level content variable (right pane).

In sum, the simulation shows that if we do observe changes to news content (*consumer* measure) following circulation shocks – as the first-stage regressions in Table B.4 suggest – then these changes in x_i are due to the impact of circulation shocks on underlying news content, and are not an artifact of the weighting scheme.

B.3 Main results

The current section reports full results for the main analysis in the text. Table B.9 reports the main regression results, both in reduced form (Models 1-2, 5-6) and instrumented (Models 3-4, 7-8), for the 'producers' (Models 1,3,5,7) and 'consumers' (Models 2,4,6,8) county-year measures of rape culture. The parameters of these models were used for graphs in Figure 3 of the main text, and most other statistics cited in the empirical section.

Table B.9: Main models. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable:							
	Reported rapes per 1000 residents				Police vigilance (rape arrests minus rape reports)			
	Rape culture measure:							
	Producers (1)	Consumers (2)	Producers (3)	Consumers (4)	Producers (5)	Consumers (6)	Producers (7)	Consumers (8)
Any rape culture, lagged	0.019***	0.011***	0.51***	1.43**	-0.0071***	-0.0037	-0.28*	-0.31**
	(0.0043)	(0.0028)	(0.11)	(0.47)	(0.0020)	(0.0027)	(0.13)	(0.10)
Median personal income	0.55***	0.55***	0.60***	0.56***	-0.015′	-0.015′	-0.029**	-0.021*
	(0.063)	(0.063)	(0.021)	(0.032)	(0.0089)	(0.0090)	(0.011)	(0.0087)
Percent female population	-0.047**	-0.047**	-0.062***	-0.062***	0.00039	0.00023	-0.0059	0.0057
	(0.017)	(0.017)	(0.012)	(0.017)	(0.0037)	(0.0037)	(0.0058)	(0.0053)
Percentage of workers unemployed	0.034*	0.030*	0.16***	0.086***	-0.0076	-0.0061	-0.032*	-0.030**
	(0.015)	(0.015)	(0.029)	(0.025)	(0.0096)	(0.0096)	(0.014)	(0.010)
Percent population in religious congregation	0.17*	0.17*	0.0039	-0.041	-0.0019	-0.0029	0.091	0.035
	(0.078)	(0.078)	(0.063)	(0.095)	(0.024)	(0.024)	(0.057)	(0.030)
Percent presidential vote for Republican	0.087***	0.083***	0.27***	0.18***	-0.061***	-0.059***	-0.16***	-0.083***
	(0.022)	(0.022)	(0.044)	(0.044)	(0.015)	(0.015)	(0.045)	(0.012)
Constant	1.19***	1.19***			-0.58***	-0.58***		
	(0.045)	(0.045)			(0.018)	(0.018)		
IV	NO	NO	YES	YES	NO	NO	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	40,349	40,358	37,244	37,253	40,349	40,358	37,244	37,253
Number of counties	3,105	3,105	3,105	3,105	3,105	3,105	3,105	3,105
LL	-51506	-51519	-52673	-64384	-14035	-14043	-21089	-19992
Craig-Donald F			105	16.2			13.6	30.1
Kleiberger-Popp F			105	16.2			13.6	30.1
Anderson-Rubin F			26.2***	21.6***			7.01**	12.8***
Stock-Wright LM S			26.2***	21.6***			7.02**	12.8***
Kleiberger-Popp LM			105	16.2			13.6	30.1

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

B.4 Rape culture and incidence of rape (Models in Figure 3.a)

Table B.10: Rape culture (producers measure) and reported rapes per 1,000 residents. Models in Figure 3.a. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Reported rapes per 1000 residents				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Producers	Producers	Producers	Producers	Producers
	(1)	(2)	(3)	(4)	(5)
Any rape culture, lagged	0.513***				
	(0.111)				
Victim-blaming, lagged (any)		1.357**			
		(0.424)			
Empathy for accused, lagged (any)			0.446***		
			(0.0945)		
Implication of consent, lagged (any)				1.188***	
				(0.357)	
Victim's credibility, lagged (any)					0.571***
					(0.125)
Median personal income	0.595***	0.578***	0.613***	0.615***	0.612***
	(0.0206)	(0.0305)	(0.0202)	(0.0288)	(0.0208)
Percent female population	-0.0617***	-0.0908***	-0.0445***	-0.0949***	-0.0382**
	(0.0123)	(0.0218)	(0.0117)	(0.0217)	(0.0122)
Percentage of workers unemployed	0.163***	0.251***	0.0930***	0.267***	0.0979***
	(0.0291)	(0.0666)	(0.0183)	(0.0685)	(0.0194)
Percent population in religious congregation	0.00394	-0.0431	-0.0358	0.286**	-0.0426
	(0.0628)	(0.0977)	(0.0654)	(0.0920)	(0.0683)
Percent presidential vote for Republican	0.268***	0.282***	0.215***	0.522***	0.0886**
	(0.0441)	(0.0670)	(0.0349)	(0.130)	(0.0278)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,244	37,244	37,244	37,244	37,244
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-52673	-66299	-51827	-64853	-52993
Craig-Donald F	105.4	16.48	151.8	18.22	88.30
Kleiberger-Popp F	105.4	16.48	151.8	18.22	88.30
Anderson-Rubin F	26.21	26.21	26.21	26.21	26.21
Anderson-Rubin p	3.08e-07	3.08e-07	3.08e-07	3.08e-07	3.08e-07
Stock-Wright LM S	26.20	26.20	26.20	26.20	26.20
Stock-Wright p	3.07e-07	3.07e-07	3.07e-07	3.07e-07	3.07e-07
Kleiberger-Popp LM	105.2	16.48	151.2	18.22	88.11

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

Table B.11: Rape culture (consumers measure) and reported rapes per 1,000 residents. Models in Figure 3.a. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Reported rapes per 1000 residents				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Consumers	Consumers	Consumers	Consumers	Consumers
	(1)	(2)	(3)	(4)	(5)
Any rape culture, lagged	1.433**				
	(0.468)				
Victim-blaming, lagged (any)		-3.657			
		(2.834)			
Empathy for accused, lagged (any)			0.604***		
			(0.146)		
Implication of consent, lagged (any)				4.224	
				(3.803)	
Victim's credibility, lagged (any)					-3.587
					(2.785)
Median personal income	0.555***	0.949***	0.624***	0.367	0.471***
	(0.0318)	(0.278)	(0.0215)	(0.224)	(0.121)
Percent female population	-0.0620***	0.0599	-0.0431***	-0.0365	-0.00663
	(0.0169)	(0.0921)	(0.0121)	(0.0461)	(0.0503)
Percentage of workers unemployed	0.0861***	0.0215	0.117***	-0.0530	0.293
	(0.0252)	(0.0583)	(0.0230)	(0.113)	(0.195)
Percent population in religious congregation	-0.0406	-0.0546	-0.0817	0.0382	-0.101
	(0.0953)	(0.233)	(0.0765)	(0.229)	(0.256)
Percent presidential vote for Republican	0.179***	-0.0270	0.112***	0.622	0.757
	(0.0438)	(0.137)	(0.0275)	(0.473)	(0.511)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,253	37,253	37,253	37,253	37,253
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-64384	-96574	-53070	-102158	-96643
Craig-Donald F	16.21	1.816	80.13	1.316	1.812
Kleiberger-Popp F	16.21	1.816	80.13	1.316	1.812
Anderson-Rubin F	21.61	21.61	21.61	21.61	21.61
Anderson-Rubin p	3.36e-06	3.36e-06	3.36e-06	3.36e-06	3.36e-06
Stock-Wright LM S	21.60	21.60	21.60	21.60	21.60
Stock-Wright p	3.35e-06	3.35e-06	3.35e-06	3.35e-06	3.35e-06
Kleiberger-Popp LM	16.21	1.817	79.98	1.316	1.812

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

B.5 Rape culture and police vigilance (Models in Figure 3.b)

Table B.12: Rape culture (producers measure) and police vigilance in rape cases. Models in Figure 3.b. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Police vigilance (rape arrests minus reports)				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Producers	Producers	Producers	Producers	Producers
	(1)	(2)	(3)	(4)	(5)
Any rape culture	-0.276*				
	(0.128)				
Victim-blaming (any)		0.967			
		(0.951)			
Empathy for accused (any)			-0.180*		
			(0.0742)		
Implication of consent (any)				-0.151*	
				(0.0612)	
Victim's credibility (any)					-0.185*
					(0.0770)
Median personal income	-0.0289**	0.00351	-0.0112	-0.0303**	-0.0182*
	(0.0105)	(0.0270)	(0.00808)	(0.00951)	(0.00791)
Percent female population	-0.00588	0.00554	0.000209	-0.00217	-0.00765
	(0.00584)	(0.0120)	(0.00452)	(0.00456)	(0.00563)
Percentage of workers unemployed	-0.0318*	0.0319	-0.0187*	-0.0162*	-0.000889
	(0.0135)	(0.0410)	(0.00792)	(0.00722)	(0.00679)
Percent population in religious congregation	0.0908	-0.258	0.0661	-0.0259	0.0545
	(0.0573)	(0.239)	(0.0419)	(0.0213)	(0.0382)
Percent presidential vote for Republican	-0.162***	0.0349	-0.120***	-0.136***	-0.0543***
	(0.0446)	(0.105)	(0.0233)	(0.0288)	(0.0121)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,244	37,244	37,244	37,244	37,244
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-21089	-49041	-16628	-16048	-17015
Craig-Donald F	13.64	1.205	34.12	43.80	31.50
Kleiberger-Popp F	13.64	1.205	34.12	43.80	31.50
Anderson-Rubin F	7.014	7.014	7.014	7.014	7.014
Anderson-Rubin p	0.00809	0.00809	0.00809	0.00809	0.00809
Stock-Wright LM S	7.016	7.016	7.016	7.016	7.016
Stock-Wright p	0.00808	0.00808	0.00808	0.00808	0.00808
Kleiberger-Popp LM	13.64	1.205	34.10	43.76	31.49

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

Table B.13: Rape culture (consumers measure) and police vigilance in rape cases. Models in Figure 3.b. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Police vigilance (rape arrests minus reports)				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Consumers (1)	Consumers (2)	Consumers (3)	Consumers (4)	Consumers (5)
Any rape culture	-0.180*				
	(0.0860)				
Victim-blaming (any)		-0.218*			
		(0.111)			
Empathy for accused (any)			-0.165*		
			(0.0780)		
Implication of consent (any)				1.637	
				(3.331)	
Victim's credibility (any)					-0.0813*
					(0.0366)
Median personal income	-0.0101	0.00334	-0.0197*	-0.00598	-0.00518
	(0.00732)	(0.00917)	(0.00940)	(0.0305)	(0.00677)
Percent female population	0.00390	0.00893	0.00212	0.00727	0.00145
	(0.00441)	(0.00609)	(0.00417)	(0.0221)	(0.00393)
Percentage of workers unemployed	-0.0218*	-0.0277*	-0.0327*	-0.0613	-0.00379
	(0.00873)	(0.0118)	(0.0130)	(0.111)	(0.00572)
Percent population in religious congregation	-0.0133	-0.0579**	0.0359	-0.0696	-0.0559**
	(0.0254)	(0.0219)	(0.0438)	(0.0962)	(0.0191)
Percent presidential vote for Republican	-0.0648***	-0.0619***	-0.0501***	0.225	-0.0389**
	(0.0102)	(0.0104)	(0.00991)	(0.574)	(0.0120)
Arrest to report difference per 1,000 people, lagged	0.618***	0.623***	0.614***	0.623***	0.621***
	(0.00902)	(0.00931)	(0.00981)	(0.0372)	(0.00831)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,253	37,253	37,253	37,253	37,253
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-13394	-15853	-13107	-67423	-11235
Craig-Donald F	29.77	15.57	31.68	0.254	105.5
Kleiberger-Popp F	29.77	15.57	31.68	0.254	105.5
Anderson-Rubin F	5.165	5.165	5.165	5.165	5.165
Anderson-Rubin p	0.0231	0.0231	0.0231	0.0231	0.0231
Stock-Wright LM S	5.167	5.167	5.167	5.167	5.167
Stock-Wright p	0.0230	0.0230	0.0230	0.0230	0.0230
Kleiberger-Popp LM	29.76	15.57	31.66	0.254	105.2

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

B.6 Other types of crime (Models in Figure 3.c)

Table B.14: Rape culture (producers measure) and police vigilance in murder cases. Models in Figure 3.c. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Police vigilance (murder arrests minus reports)				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Producers	Producers	Producers	Producers	Producers
	(1)	(2)	(3)	(4)	(5)
Any rape culture	0.285				
	(0.275)				
Victim-blaming (any)		-0.998			
		(1.281)			
Empathy for accused (any)			0.186		
			(0.175)		
Implication of consent (any)				0.156	
				(0.144)	
Victim's credibility (any)					0.191
					(0.180)
Median personal income	0.000738	-0.0327	-0.0175	0.00216	-0.0103
	(0.0175)	(0.0342)	(0.0137)	(0.0176)	(0.0125)
Percent female population	0.0121	0.000328	0.00583	0.00827	0.0139
	(0.00821)	(0.0123)	(0.00548)	(0.00593)	(0.00928)
Percentage of workers unemployed	0.0123	-0.0534	-0.00119	-0.00378	-0.0196′
	(0.0249)	(0.0551)	(0.0129)	(0.0111)	(0.0113)
Percent population in religious congregation	-0.0869	0.273	-0.0614	0.0335	-0.0495
	(0.119)	(0.319)	(0.0938)	(0.0369)	(0.0831)
Percent presidential vote for Republican	0.0802	-0.123	0.0367	0.0530	-0.0310′
	(0.0972)	(0.136)	(0.0550)	(0.0689)	(0.0187)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,244	37,244	37,244	37,244	37,244
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-33495	-53580	-31342	-31075	-31467
Craig-Donald F	13.64	1.205	34.12	43.80	31.50
Kleiberger-Popp F	9.149	1.125	12.38	47.87	21.82
Anderson-Rubin F	1.190	1.190	1.190	1.190	1.190
Anderson-Rubin p	0.275	0.275	0.275	0.275	0.275
Stock-Wright LM S	1.356	1.356	1.356	1.356	1.356
Stock-Wright p	0.244	0.244	0.244	0.244	0.244
Kleiberger-Popp LM	9.130	1.128	12.74	44.69	22.65

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

Table B.15: Rape culture (consumers measure) and police vigilance in murder cases. Models in Figure 3.c. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Police vigilance (murder arrests minus reports)				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Consumers (1)	Consumers (2)	Consumers (3)	Consumers (4)	Consumers (5)
Any rape culture	0.0696				
	(0.228)				
Victim-blaming (any)		0.0845			
		(0.277)			
Empathy for accused (any)			0.0633		
			(0.208)		
Implication of consent (any)				-0.635	
				(2.321)	
Victim's credibility (any)					0.0315
					(0.103)
Median personal income	-0.00550	-0.0107	-0.00184	-0.00706	-0.00738
	(0.0126)	(0.0176)	(0.0200)	(0.0162)	(0.0121)
Percent female population	0.00433	0.00237	0.00502	0.00301	0.00527
	(0.00676)	(0.0119)	(0.00562)	(0.0127)	(0.00537)
Percentage of workers unemployed	-0.00643	-0.00411	-0.00227	0.00893	-0.0134
	(0.0180)	(0.0250)	(0.0308)	(0.0753)	(0.0105)
Percent population in religious congregation	0.00939	0.0265	-0.00937	0.0311	0.0258
	(0.0563)	(0.0385)	(0.111)	(0.0562)	(0.0375)
Percent presidential vote for Republican	-0.0134	-0.0145	-0.0191	-0.126	-0.0234
	(0.0196)	(0.0174)	(0.0155)	(0.397)	(0.0243)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,253	37,253	37,253	37,253	37,253
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-30094	-30206	-30088	-43295	-29985
Craig-Donald F	30.09	15.58	32.36	0.254	106
Kleiberger-Popp F	16.10	8.610	15.58	0.332	70.88
Anderson-Rubin F	0.0937	0.0937	0.0937	0.0937	0.0937
Anderson-Rubin p	0.759	0.759	0.759	0.759	0.759
Stock-Wright LM S	0.110	0.110	0.110	0.110	0.110
Stock-Wright p	0.740	0.740	0.740	0.740	0.740
Kleiberger-Popp LM	16.20	8.590	15.54	0.331	70.58

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

Table B.16: Rape culture (producers measure) and police vigilance in robbery cases. Models in Figure 3.c. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Police vigilance (robbery arrests minus reports)				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Producers	Producers	Producers	Producers	Producers
	(1)	(2)	(3)	(4)	(5)
Any rape culture	-0.00509				
	(0.157)				
Victim-blaming (any)		0.0178			
		(0.551)			
Empathy for accused (any)			-0.00332		
			(0.103)		
Implication of consent (any)				-0.00279	
				(0.0860)	
Victim's credibility (any)					-0.00342
					(0.105)
Median personal income	-0.0313***	-0.0307*	-0.0310***	-0.0313***	-0.0311***
	(0.00856)	(0.0133)	(0.00672)	(0.00916)	(0.00566)
Percent female population	0.00139	0.00160	0.00150	0.00145	0.00135
	(0.00466)	(0.00439)	(0.00314)	(0.00341)	(0.00545)
Percentage of workers unemployed	0.00954	0.0107	0.00978	0.00983	0.0101
	(0.0139)	(0.0236)	(0.00735)	(0.00630)	(0.00646)
Percent population in religious congregation	0.0325	0.0261	0.0320	0.0303*	0.0318
	(0.0654)	(0.136)	(0.0518)	(0.0143)	(0.0454)
Percent presidential vote for Republican	-0.0616	-0.0579	-0.0608′	-0.0611	-0.0596***
	(0.0557)	(0.0576)	(0.0321)	(0.0409)	(0.00998)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,244	37,244	37,244	37,244	37,244
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-5520	-5572	-5519	-5517	-5518
Craig-Donald F	13.64	1.205	34.12	43.80	31.50
Kleiberger-Popp F	9.149	1.125	12.38	47.87	21.82
Anderson-Rubin F	0.00105	0.00105	0.00105	0.00105	0.00105
Anderson-Rubin p	0.974	0.974	0.974	0.974	0.974
Stock-Wright LM S	0.00123	0.00123	0.00123	0.00123	0.00123
Stock-Wright p	0.972	0.972	0.972	0.972	0.972
Kleiberger-Popp LM	9.130	1.128	12.74	44.69	22.65

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

Table B.17: Rape culture (consumers measure) and police vigilance in robbery cases. Models in Figure 3.c. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Police vigilance (robbery arrests minus reports)				
	Independent variable:				
	Any rape culture	Victim-blaming language	Empathy for accused	Implication of consent	Incredulity toward victim
	Rape culture measure:				
	Consumers (1)	Consumers (2)	Consumers (3)	Consumers (4)	Consumers (5)
Any rape culture	-0.0865				
	(0.129)				
Victim-blaming (any)		-0.105			
		(0.159)			
Empathy for accused (any)			-0.0786		
			(0.117)		
Implication of consent (any)				0.789	
				(1.813)	
Victim's credibility (any)					-0.0391
					(0.0574)
Median personal income	-0.0331***	-0.0266**	-0.0376***	-0.0311*	-0.0307***
	(0.00622)	(0.00970)	(0.0106)	(0.0151)	(0.00595)
Percent female population	0.00297	0.00540	0.00211	0.00460	0.00180
	(0.00395)	(0.00688)	(0.00335)	(0.0116)	(0.00320)
Percentage of workers unemployed	0.00332	0.000440	-0.00185	-0.0158	0.0119'
	(0.00997)	(0.0142)	(0.0171)	(0.0596)	(0.00617)
Percent population in religious congregation	0.0453	0.0240	0.0686	0.0184	0.0250
	(0.0284)	(0.0167)	(0.0607)	(0.0472)	(0.0157)
Percent presidential vote for Republican	-0.0638***	-0.0624***	-0.0567***	0.0756	-0.0514***
	(0.0112)	(0.0100)	(0.00838)	(0.310)	(0.0130)
IV	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	37,253	37,253	37,253	37,253	37,253
Number of counties	3,105	3,105	3,105	3,105	3,105
LL	-6559	-7476	-6436	-42198	-5879
Craig-Donald F	30.09	15.58	32.36	0.254	106
Kleiberger-Popp F	16.10	8.610	15.58	0.332	70.88
Anderson-Rubin F	0.467	0.467	0.467	0.467	0.467
Anderson-Rubin p	0.494	0.494	0.494	0.494	0.494
Stock-Wright LM S	0.542	0.542	0.542	0.542	0.542
Stock-Wright p	0.461	0.461	0.461	0.461	0.461
Kleiberger-Popp LM	16.20	8.590	15.54	0.331	70.58

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

B.7 Impact of news volume

In addition to the content of news stories, which we analyzed in Table B.9, we examined the impact of the volume of news stories on rape in a county-year. Table B.18 replicates Models 3-4 and 7-8 from our main results in Table B.9, substituting the average number of articles on rape (per newspaper in each county-year) for our previous, proportional measure of rape culture.

Table B.18: Impact of rape news volume on crime. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable:			
	Reported rapes per 1000 residents		Police vigilance (rape arrests minus rape reports)	
	Rape culture measure:			
	Producers	Consumers	Producers	Consumers
	(1)	(2)	(3)	(4)
Rape articles per paper, lagged	0.49***	2.67**	0.097**	0.53*
	(0.10)	(0.84)	(0.037)	(0.24)
Median personal income	0.60***	0.43***	-0.016*	-0.050**
	(0.020)	(0.062)	(0.0072)	(0.018)
Percent female population	-0.063***	-0.071***	-0.0028	-0.0044
	(0.012)	(0.019)	(0.0044)	(0.0054)
Percentage of workers unemployed	0.044**	0.079**	-0.0081	-0.0013
	(0.016)	(0.025)	(0.0058)	(0.0073)
Percent population in religious congregation	0.22***	-0.35*	-0.0038	-0.12*
	(0.057)	(0.17)	(0.021)	(0.050)
Percent presidential vote for Republican	0.19***	0.073′	-0.053***	-0.076***
	(0.031)	(0.041)	(0.011)	(0.012)
IV	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	37,253	37,253	37,253	37,253
Number of counties	3,105	3,105	3,105	3,105
R-squared	0.128	-0.994	0.188	-0.149
LL	-51033	-66439	-13897	-20362
Craig-Donald F	191	15.5	191	15.5
Kleiberger-Popp F	191	15.5	191	15.5
Anderson-Rubin F	26.2***	26.2***	6.84**	6.84**
Stock-Wright LM S	26.2***	26.2***	6.84**	6.84**
Kleiberger-Popp LM	190	15.5	190	15.5

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

B.8 Article-level analysis

We investigate article-level variation in rape culture with the following model:

$$\begin{aligned}\text{RapeCulture}_{djk} =& \text{logit}^{-1}(X_d\beta + X_j\gamma + X_k\delta \\ &+ \text{Newspaper}_j + \text{County}_k + \text{Year}_d + \text{Month}_d)\end{aligned} \tag{B.1}$$

where d indexes the document (i.e. newspaper article), j indexes the newspaper, k indexes the county and t indexes the day. For notational simplicity, we suppress the temporal index t (the publication date of article d).

The dependent variable, RapeCulture_{djk}, is the SVM-generated document class, equal to 1 if article d by newspaper j in county k contains any of the four main categories of rape culture, and 0 otherwise. In addition to the combined measure, we repeat this analysis separately for each of the four categories (victim-blaming, empathy for accused, implication of consent, incredulity toward victim).

On the right-hand side are matrices of article-level (X_d), newspaper-level (X_j) and county-level covariates (X_k). Article-level covariates include:

- *Weekend*: equal to 1 if article d was published on a Saturday or Sunday, 0 otherwise.
- *Major case*: equal to 1 if article d mentions a high-profile sexual assauly case that dominated multiple news cycles in 2000-2013, and 0 otherwise. These cases include the Central Park jogger case (whose convictions were overturned in 2002), the Steubenville, Ohio rape case, and cases involving Dominique Strauss Kahn, Jameis Winston and Ariel Castro.
- *Criminal justice stage: arrest*: equal to 1 if article d mentions police investigations, arrests and other law enforcement activities related to a rape case, and 0 otherwise.
- *Criminal justice stage: prosecution*: equal to 1 if article d mentions courtroom activity, jury selection, witness testimony, cross-examination, oral arguments and other activities related to the prosecution and legal defense of rape suspects, and 0 otherwise.
- *Criminal justice stage: corrections*: equal to 1 if article d mentions sentencing, incarceration and other activities following the conviction of rape suspects, and 0 otherwise.
- *Number of rape articles in newspaper, past week*: number of articles about rape published in newspaper j in the week prior to publication of article d.

Newspaper-level covariates (X_j) include:

- *National paper*: equal to 1 if newspaper j has either national distribution (e.g. USA Today) or expanded distribution outside its home metropolitan area (e.g. New York Times), and 0 otherwise.
- *Trade publication*: equal to 1 if newspaper j publishes primarily industry-specific content, and relatively little general-audience information (e.g. Lawyers Weekly, Variety), and 0 otherwise.

County-level covariates (X_k) include:

- *Median personal income* in county k during d's year of publication
- *Percent urban population* in county k during d's year of publication
- *Percent female population* in county k during d's year of publication
- *Percentage of workers unemployed* in county k during d's year of publication
- *Percent population in religious congregation* in county k during d's year of publication
- *Percent presidential vote for Republican* in county k during the presidential election most recent to d's year of publication.

We also account for several additional sources of variation in news coverage. These include common shocks across time (dummies for the Year_d of d's publication), and months of the year (Month_d dummies). To account for time-invariant differences between localities where newspaper j's home bureau is located, we include dummies for each County_k. Finally, to account for editorial differences across individual newspapers, we include dummies for each newspaper (Newspaper_j). We introduce these terms iteratively, beginning with year and month (Model 1), then county (Model 2) and newspaper (Model 3).

Table B.19 reports summary statistics for all article-level variables used in the analysis. Table B.20 reports coefficient estimates for the model in equation (B.20). Table B.21 repeats this analysis for the four sub-components of rape culture.

The results suggest that content consistent with rape culture is most likely to appear in newspaper articles during coverage of law enforcement activity and court proceedings. The coefficients on *Criminal justice stage: prosecution* are consistently positive and highly significant, across all specifications and components of rape culture. The estimates for *Criminal justice stage: arrest* are of similar sign and significance, with two exceptions: coverage of arrests and police investigations does not correlate with empathy toward the accused or incredulity toward victims. By contrast, rape culture is significantly less likely to appear in articles about the corrections stage of the criminal justice process, after a suspect had been tried and convicted of rape. The coefficients on *Criminal justice stage: corrections* are negative and highly significant, with one exception: such coverage does not affect language that implies the victim's consent.

Our models also suggest that rape culture is more likely following a relatively high frequency of news reports about rape in the same newspaper. The coefficients on *Number of articles in paper, past week* are consistently positive and significant. There also appears to be slightly more rape culture in articles covering *Major cases*, like the Steubenville, Ohio gang rape. However, this coefficient estimate is not stable across specifications. In one case (implication of consent), the coefficient is actually negative, indicating that news stories about high-profile rape cases are less likely to suggest that a consensual, romantic relationship existed between victim and perpetrator. Rape culture is also less likely to appear in articles published on the weekend, although this estimate exceeds conventional standards of statistical significance only for some sub-categories (i.e. victim-blaming, implication of consent).

The timing and substantive focus of newspaper articles are far stronger predictors of rape culture than aggregate, county-level differences. Consistent with the logic behind our instrumental variable design, rape culture is less likely to appear in counties with a high proportion of female residents. However, this negative coefficient is only marginally significant, and becomes more uncertain once we account for county-level fixed effects – suggesting that increases in a county's female population over time are less predictive of rape culture than differences in the gender balance between counties. By the same token, rape culture is somewhat more likely to appear in newspapers within poorer, more religious and more Republican counties, but most of these estimates fail to reach acceptable levels of statistical significance. Finally, rape culture does not strongly vary according to a newspaper's type (national vs. local, general audience vs. trade).

Taken together, these results reinforce the centrality of the criminal justice process (and perceptions thereof) to victims' and perpetrators' decision-making. Rape culture is most likely to appear in the press when a criminal case reaches the point in its "life cycle" that is arguably most consequential for victims' pursuit of justice: when the level of public attention is high (when news coverage is highly frequent, or focused on a high profile case), when law enforcement is in the process of investigating and arresting perpetrators, and especially when those perpetrators are on trial.

The theoretical implication, of course, is not that potential perpetrators and victims simply read newspaper articles about court cases, and change their behavior accordingly. Rape culture in the press is ultimately a reflection of local community norms. Yet these article-level analyses reveal important insights about when and where these norms are most likely to come to the surface and be publicly reinforced – due to perceived consumer demand or editorial and journalistic discretion. If victims know that heightened media attention to rape cases invites increased scrutiny of victims' accounts – especially when (and if) these cases come to trial – they may be less likely to come forward. If perpetrators expect public sympathies to shift in favor of the accused during court proceedings, they may see the likelihood of arrest and prosecution as relatively low, especially if these same forces also deter a victim from reporting the crime. Victims and perpetrators do not necessarily need to read a newspaper article to reach these conclusions – chances are, if a newspaper prints a "victim-blaming" story, such sentiments already exist on the ground. News coverage does not create these sentiments, but it does amplify them and makes them more visible in the public debate.

Table B.19: Summary statistics, article level

Variable	Mean	Std. Dev.	Min.	Max.	N
Article-level					
Any rape culture	0.032	0.175	0	1	143820
Victim-blaming (any)	0.013	0.115	0	1	143820
Empathy for accused (any)	0.008	0.091	0	1	143820
Implication of consent (any)	0.01	0.097	0	1	143820
Victim's credibility (any)	0.005	0.069	0	1	143820
Major case	0.019	0.137	0	1	143820
Criminal justice stage: arrest	0.764	0.425	0	1	143820
Criminal justice stage: prosecution	0.846	0.361	0	1	143820
Criminal justice stage: corrections	0.562	0.496	0	1	143820
Number of articles in paper, past week	6.053	6.972	0	87	143820
Year	2006.893	4.004	2000	2013	143820
Month	6.344	3.409	1	12	143820
Weekend	0.236	0.424	0	1	143820
Newspaper-level					
National newspaper	0.253	0.435	0	1	143556
Trade publication	0.004	0.063	0	1	143556
County-level					
Percent female population	51.508	1.217	44.9	55.328	143819
Median personal income	51486.726	25569.729	20933	120790	143819
Percent urban population	94.464	9.448	1.21	100	143819
Percentage of workers unemployed	6.059	2.338	2.1	18.2	143819
Percent population in religious congregation	47.184	16.557	0.448	91.868	143819
Percent presidential vote for Republican	34.98	16.906	6.53	77.3	143819

Table B.20: Determinants of rape culture in newspaper articles.

VARIABLES	(1) Any rape culture	(2) Any rape culture	(3) Any rape culture
Article-level			
Weekend	-0.078	-0.067	-0.073
	(0.048)	(0.049)	(0.049)
Major case	0.35′	0.40′	0.42*
	(0.19)	(0.20)	(0.20)
Criminal justice stage: arrest	0.34***	0.33***	0.31***
	(0.071)	(0.071)	(0.069)
Criminal justice stage: prosecution	1.02***	1.00***	0.97***
	(0.094)	(0.099)	(0.10)
Criminal justice stage: corrections	-0.39***	-0.39***	-0.40***
	(0.053)	(0.052)	(0.052)
Number of articles in paper, past week	0.023***	0.022***	0.021***
	(0.0030)	(0.0030)	(0.0030)
Newspaper-level			
National newspaper	0.13	0.35	
	(0.12)	(0.30)	
Trade publication	-0.68	-0.81*	
	(0.42)	(0.41)	
County-level			
Percent female population	-0.075′	0.0018	-0.0075
	(0.044)	(0.065)	(0.067)
Median personal income	-0.095*	-0.45	-0.20
	(0.042)	(0.30)	(0.26)
Percent urban population	0.030	-0.024	
	(0.037)	(0.22)	
Percentage of workers unemployed	-0.019	-0.0089	-0.026
	(0.068)	(0.089)	(0.094)
Percent population in religious congregation	0.048	-0.081	0.041
	(0.037)	(0.29)	(0.36)
Percent presidential vote for Republican	0.020	-0.11	0.10
	(0.054)	(0.22)	(0.29)
Constant	-4.96***	-4.76***	-4.10***
	(0.15)	(0.41)	(0.40)
Newspaper FE	NO	NO	YES
County FE	NO	YES	YES
Year FE	YES	YES	YES
Month FE	YES	YES	YES
Clustered SE	YES	YES	YES
Observations	143,556	142,951	142,294
Log-likelihood	-19418	-19180	-19081
Num. of clusters	209	183	132
LRT χ^2: Mod 1 vs. 2,3		474.5***	673.4***
LRT χ^2: Mod 2 vs. 3			198.9***

Logit coefficients reported. Clustered robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

Table B.21: Determinants of rape culture in newspaper articles, additional outcome measures.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	Any rape culture		Victim-blaming language		Empathy toward accused		Implication of consent		Incredulity toward victim	
Article-level										
Weekend	-0.078	-0.073	-0.19**	-0.18*	0.12	0.083	-0.19**	-0.17*	-0.052	-0.072
	(0.048)	(0.049)	(0.073)	(0.074)	(0.099)	(0.11)	(0.072)	(0.073)	(0.11)	(0.11)
Major case	0.35′	0.42*	0.56′	0.60′	0.69*	0.75**	-0.93***	-0.87***	0.17	0.17
	(0.19)	(0.20)	(0.33)	(0.35)	(0.29)	(0.29)	(0.26)	(0.25)	(0.27)	(0.26)
Criminal justice stage: arrest	0.34***	0.31***	0.76***	0.72***	-0.15′	-0.14′	0.37***	0.35**	-0.014	-0.042
	(0.071)	(0.069)	(0.13)	(0.13)	(0.078)	(0.080)	(0.097)	(0.11)	(0.13)	(0.13)
Criminal justice stage: prosecution	1.02***	0.97***	1.19***	1.16***	1.30***	1.13***	0.50***	0.48***	1.98***	1.94***
	(0.094)	(0.10)	(0.16)	(0.17)	(0.21)	(0.20)	(0.10)	(0.11)	(0.23)	(0.23)
Criminal justice stage: corrections	-0.39***	-0.40***	-0.78***	-0.76***	-0.14′	-0.16*	0.022	0.0063	-0.66***	-0.69***
	(0.053)	(0.052)	(0.091)	(0.090)	(0.080)	(0.078)	(0.077)	(0.078)	(0.11)	(0.11)
Number of articles in paper, past week	0.023***	0.021***	0.019***	0.020***	0.023***	0.020***	0.026***	0.027**	0.022***	0.012′
	(0.0030)	(0.0030)	(0.0043)	(0.0060)	(0.0061)	(0.0057)	(0.0077)	(0.0087)	(0.0045)	(0.0066)
Newspaper-level										
National newspaper	0.13		0.10		0.44′		-0.031		0.29	
	(0.12)		(0.17)		(0.23)		(0.17)		(0.20)	
Trade publication	-0.68		-1.59		-0.55		-1.58*		0.38	
	(0.42)		(1.15)		(0.51)		(0.69)		(0.56)	
County-level										
Percent female population	-0.075′	-0.0075	-0.098	0.15	0.011	-0.17	-0.13*	-0.056	-0.087	-0.022
	(0.044)	(0.067)	(0.073)	(0.093)	(0.071)	(0.12)	(0.063)	(0.12)	(0.078)	(0.16)
Median personal income	-0.095*	-0.20	-0.095′	-0.56′	-0.16′	0.25	-0.031	-0.57	-0.14*	-0.61
	(0.042)	(0.26)	(0.054)	(0.33)	(0.094)	(0.30)	(0.084)	(0.52)	(0.072)	(0.39)
Percent urban population	0.030		-0.0011		0.0056		0.092*		0.0073	
	(0.037)		(0.055)		(0.062)		(0.046)		(0.064)	
Percentage of workers unemployed	-0.019	-0.026	0.018	0.071	-0.061	-0.12	-0.0055	-0.0023	-0.21′	-0.28
	(0.068)	(0.094)	(0.11)	(0.14)	(0.15)	(0.17)	(0.086)	(0.12)	(0.13)	(0.25)
Percent population in religious congregation	0.048	0.041	0.065	0.84	0.095	-0.039	-0.012	-0.31	0.036	-0.32
	(0.037)	(0.36)	(0.051)	(0.93)	(0.071)	(0.65)	(0.060)	(0.48)	(0.073)	(1.08)
Percent presidential vote for Republican	0.020	0.10	0.0052	0.26	-0.060	0.84	0.19**	-0.61	-0.24*	-0.093
	(0.054)	(0.29)	(0.081)	(0.45)	(0.095)	(0.58)	(0.063)	(0.46)	(0.098)	(0.73)
Constant	-4.96***	-4.10***	-6.51***	-5.11***	-6.33***	-3.68***	-6.03***	-6.10***	-7.48***	-6.74***
	(0.15)	(0.40)	(0.33)	(0.85)	(0.25)	(0.68)	(0.26)	(0.78)	(0.40)	(1.04)
Newspaper FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
County FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Clustered SE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	143,556	142,294	143,556	137,652	143,556	135,774	143,556	141,088	143,556	129,920
Log-likelihood	-19418	-19081	-9389	-9071	-6499	-6291	-7578	-7413	-4125	-3989
Number of clusters	209	132	209	111	209	111	209	118	209	91

Logit coefficients reported. Clustered robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

C Robustness checks

The current section reports several additional regression results, which we omitted from the main text due to space constraints.

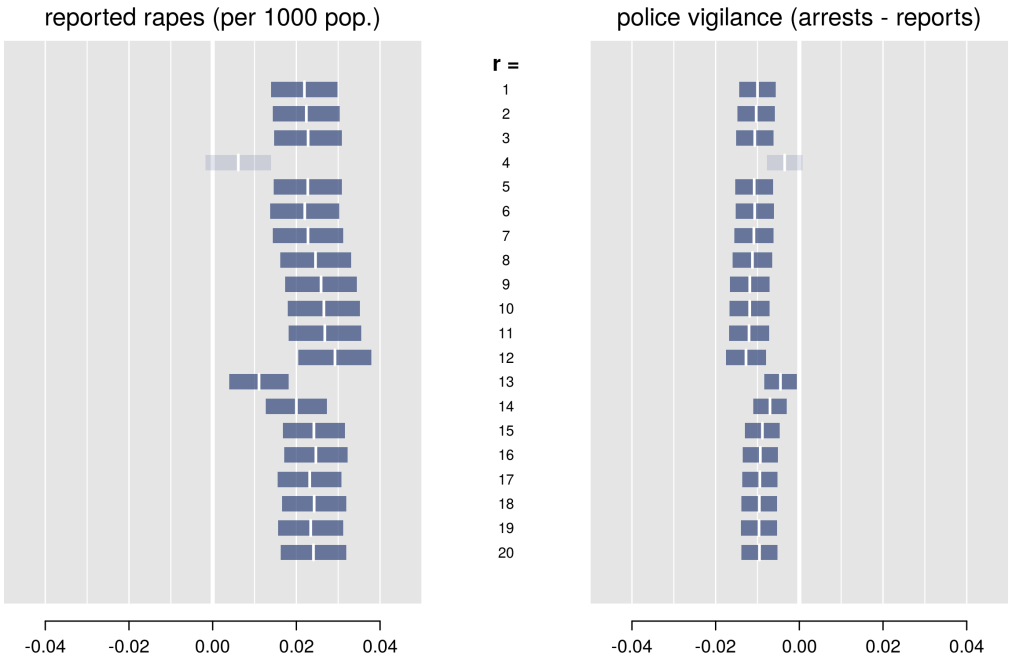
C.1 Alternate "content producers" weights

We employ two measures of local rape culture in the press, differentiated by the aggregation method used to link newspapers to counties: *content producers* (weighted by geographic proximity of newspapers' main bureaus), and *content consumers* (weighted by newspapers' relative market shares in county). We summarized the methodology behind these measures in online appendix A.7.

As described in equation A.4, the first of these weights uses a search parameter (r) to define the number of nearest geographic neighbors (i.e. newspapers). In the main analyses, we used $r = 5$ (i.e. 5 closest newspapers to county center).

To show that our results are not specific to this parameter value, we replicate our two main fixed effects models, with alternate values of r, from 1 to 20. As Figure C.1 shows, the resulting coefficients on the rape culture variable are consistent in magnitude, direction and (with one exception) significance to those achieved with the original $r = 5$.

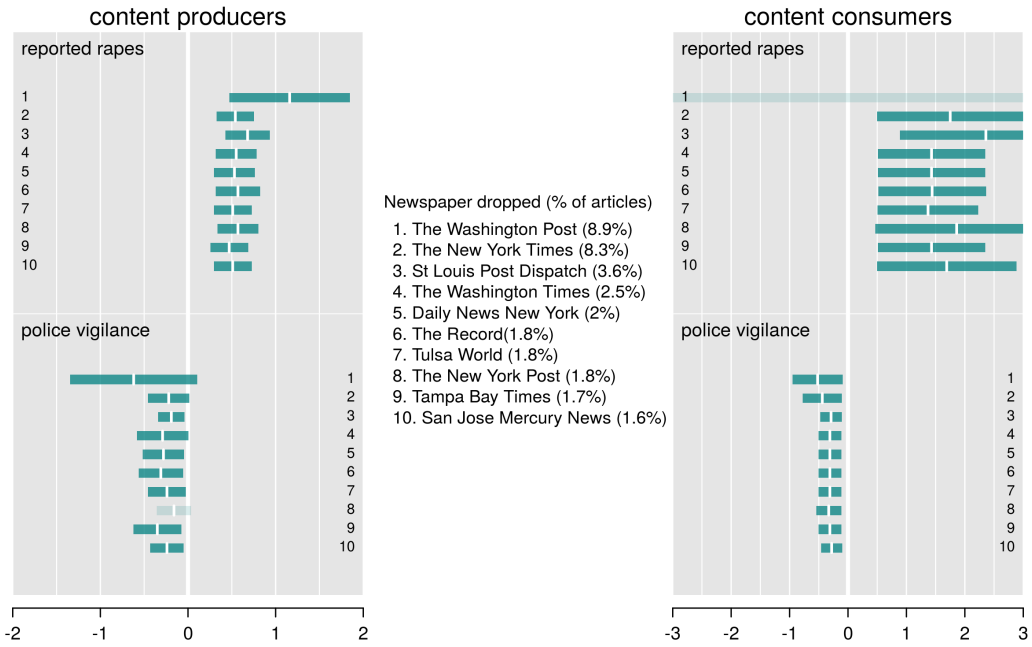
Figure C.1: Robustness check: estimated impact of rape culture on crime (producers measure), with alternate values of r in weights. Values reported are standardized coefficients and 95% confidence intervals from Models 1 (left) and 3 (right) from Table B.9.



C.2 High-volume newspapers

To account for the potentially disproportionate influence of major newspapers like the *Washington Post* and *New York Times* – which together account for 17 percent of the articles in our dataset – we replicated our analyses on restricted versions of the dataset, iteratively excluding each of the top ten newspapers by volume. The results, which we report in Figure C.2, are consistent with those from the full dataset. Removing the Washington Post from our sample, predictably, has the most disruptive impact on our results – increasing the magnitude of the estimate, but also the uncertainty around it. Overall, however, the estimates remain consistently positive for rape reports and negative for police vigilance.

Figure C.2: Robustness check: estimated impact of rape culture on crime, with top ten highest-volume newspapers iteratively excluded from data sample.



C.3 Dynamic panel data analysis

Our main analysis employed the following core model specification:

$$y_{it} = \rho' \mathbf{C}_{it-1} + \beta' \mathbf{X}_{it} + \alpha_i + u_t + \epsilon_{it} \tag{C.1}$$

where y_{it} is the number of reported rapes per 1,000 residents in county i in year t (or the difference between rape arrests and reports in i, t), $\mathbf{C}_{it-1}$ is the proportion of local newspaper articles on rape published in the previous year containing one or more of the rape culture categories, $\mathbf{X}_{it}$ is a matrix of exogenous covariates, and α_i and u_t are fixed effects. We also considered an expanded model, where we instrumented $\mathbf{C}$ with local newspaper circulation shocks.

The model in equation C.1 omits a lagged dependent variable because of the structure of our panel data: relatively few time periods ($T = 14$) and many individual units ($N > 3000$). In the context of a fixed effect estimator, including a lagged y for "small T, large N" data can – through the demeaning process – induce a correlation between regressors and the error term. The resulting correlation biases estimates of the coefficient on the lagged dependent variable, and potentially other regressors if they are correlated with the lagged dependent variable (Nickell, 1981). Omitting the temporal lag, however, risks overlooking the dynamic nature of our dependent variable, which may depend on its own past realizations in an autoregressive manner.

To address this concern, we estimated a series of additional models for dynamic panel data analysis. The first of these was the Anderson and Hsiao (1982) estimator, which removes unit fixed effects through first-differencing:

$$\Delta y_{it} = \gamma' \Delta y_{it-1} + \rho' \Delta \mathbf{C}_{it-1} + \beta' \Delta \mathbf{X}_{it} + \Delta u_t + \Delta \epsilon_{it} \tag{C.2}$$

this approach then instruments for the lagged dependent variable, using the second and third lags of y, which – assuming ϵ_{it} are i.i.d. – will be highly correlated with the difference of y_{it-1}, but uncorrelated with the composite error process.

We also considered the Arellano and Bond (1991) estimator, which similarly takes the first difference of the regression equation, and uses deeper lags of y as instruments for Δy_{it-1}. Unlike Anderson-Hsiao, however, Anderson-Bond uses a more asymptotically efficient generalized method of moments (GMM) estimator.

Finally, we considered the the Arellano and Bover (1995) forward orthogonal deviations (FOD) transformation, which avoids serial correlation of the transformed error terms by subtracting the average of all available future observations from the current value, rather than simple demeaning or subtracting previous observations from current ones.

Tables C.1 and C.2 report the results of these models, omitting all coefficient estimates other than those on our main variable of interest (rape culture at $t - 1$). These estimates are generally consistent with those in the static panel data models in the main text, but vary in magnitude across the estimators. For models of reported rape cases in Table C.1, Anderson-Hsiang estimates a relatively large and positive coefficient on the rape culture variable, while the other two estimators report a coefficient that is similarly positive and significant, although far smaller. The dynamic police vigilance models in Table C.2 pro-

duce estimates that are more uncertain than those we reported in the main text, but in the same general direction.

Table C.1: Dynamic panel data models: reported rape cases.

	(1) Anderson-Hsiang	(2) Arellano-Bond	(3) Arellano-Bover	(4) Anderson-Hsiang	(5) Arellano-Bond	(6) Arellano-Bover
	Producers			Consumers		
Any rape culture	0.376***	0.0183*	0.0311**	0.796***	0.0287***	0.0261*
$t-1$	(0.0803)	(0.00766)	(0.0103)	(0.234)	(0.00788)	(0.0105)
Covariates	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	34,137	37,242	37,244	34,147	37,252	37,253
Number of counties	3,105	3,105	3,105	3,105	3,105	3,105
LL	-43064			-53416		
Craig-Donald F	44.36			12.51		
Kleiberger-Popp F	33.96			8.132		
Anderson-Rubin F	20.27***			20.16***		
Stock-Wright LM S	36.54***			36.36***		
Kleiberger-Popp LM	63.09			16.41		
Z Rank		163	162		163	162
Chi2		93586	64572		93703	64542
RSS		19795			19864	
Sargan			10159			8761

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

Table C.2: Dynamic panel data models: police vigilance (rape arrests minus rape reports).

	(1) Anderson-Hsiang	(2) Arellano-Bond	(3) Arellano-Bover	(4) Anderson-Hsiang	(5) Arellano-Bond	(6) Arellano-Bover
	Producers			Consumers		
Any rape culture	-0.00241	-0.0101***	-0.0261***	-0.00249	-0.00696'	-0.00794'
$t-1$	(0.00164)	(0.00304)	(0.00406)	(0.00183)	(0.00355)	(0.00416)
Covariates	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	34,137	37,242	37,244	34,147	37,252	37,253
Number of counties	3,105	3,105	3,105	3,105	3,105	3,105
LL	-12863			-12905		
Craig-Donald F	6472			6458		
Kleiberger-Popp F	678.4			676.2		
Anderson-Rubin F	9.576**			10.12**		
Stock-Wright LM S	9.116**			9.651**		
Kleiberger-Popp LM	1054			1053		
Z Rank		163	162		163	162
Chi2		21147	9436		21023	9417
RSS		6849			6878	
Sargan			19797			19772

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

C.4 Media market effects

In addition to the country-level fixed effects we considered in the main analysis, we sought to account for unobserved heterogeneity at the level of the media market. Nielsen Media Research's Designated Market Areas (DMA) generally include several counties, some of which may come from adjacent states. Within a DMA, local residents tend to receive similar media offerings, primarily in television and radio, but also in terms of newspaper circulation. Although our main empirical analysis sought to account for common shocks to media markets by clustering robust standard errors on DMAs, we now consider additional models that more directly account for time-invariant market characteristics:

$$y_{it} = \rho' \mathbf{C}_{it-1} + \beta' \mathbf{X}_{it} + \alpha_m + u_t + \epsilon_{it} \tag{C.3}$$

where α_m are DMA-level fixed effects.

Table C.3 replicates our main regression results, substituting DMA-level fixed effects for the original county-level fixed effects. Because counties are typically smaller than media markets, DMA-level intercepts drop of out the original models. As these results show, our original results are robust to this model extension.

Table C.3: Main models, with media-market (DMA) fixed effects.

	Dependent variable:							
	Reported rapes per 1000 residents				Police vigilance (rape arrests minus rape reports)			
	Rape culture measure:							
	Producers (1)	Consumers (2)	Producers (3)	Consumers (4)	Producers (5)	Consumers (6)	Producers (7)	Consumers (8)
Any rape culture, lagged	0.019***	0.0042	0.54***	1.88*	0.0062	-0.015**	-0.32	-0.36*
	(0.0042)	(0.0028)	(0.11)	(0.80)	(0.0058)	(0.0059)	(0.24)	(0.17)
Median personal income	0.098***	0.098***	0.11***	0.058*	-0.13***	-0.13***	-0.13***	-0.13***
	(0.015)	(0.015)	(0.0084)	(0.025)	(0.0097)	(0.0097)	(0.0071)	(0.0072)
Percent urban population	-0.095***	-0.095***	-0.093***	-0.16***	-0.24***	-0.24***	-0.25***	-0.23***
	(0.0082)	(0.0082)	(0.0072)	(0.029)	(0.0045)	(0.0045)	(0.0061)	(0.0078)
Percent female population	-0.021***	-0.020***	-0.025***	-0.028*	-0.030***	-0.030***	-0.032***	-0.027***
	(0.0051)	(0.0051)	(0.0062)	(0.011)	(0.0034)	(0.0034)	(0.0055)	(0.0051)
Percentage of workers unemployed	-0.0063	-0.0080	0.046**	0.021	0.046***	0.044***	0.033**	0.031**
	(0.0087)	(0.0087)	(0.014)	(0.021)	(0.013)	(0.013)	(0.012)	(0.011)
Percent population in religious congregation	0.026***	0.026***	0.020**	0.033**	0.0053	0.0049	0.012	0.0034
	(0.0076)	(0.0076)	(0.0069)	(0.012)	(0.0049)	(0.0049)	(0.0076)	(0.0056)
Percent presidential vote for Republican	0.0033	0.0029	0.022*	0.033'	0.18***	0.18***	0.17***	0.17***
	(0.0076)	(0.0076)	(0.0086)	(0.018)	(0.0085)	(0.0084)	(0.0087)	(0.0073)
Constant	-0.41***	-0.20	1.92***	-1.35	0.014	-0.22***	0.37	0.97**
	(0.043)	(0.16)	(0.31)	(1.38)	(0.048)	(0.042)	(0.26)	(0.35)
IV	NO	NO	YES	YES	NO	NO	YES	YES
DMA FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	40,349	40,358	37,244	37,253	40,349	40,358	37,244	37,253
LL	-52928	-52941	-54274	-75197	-48388	-48404	-46955	-46917
Craig-Donald F			111	7.35			15.3	33
Kleiberger-Popp F			111	7.35			15.3	33
Anderson-Rubin F			29.1	20.9			2.03	4.85
Anderson-Rubin p			6.8e-08	4.9e-06			0.15	0.028
Stock-Wright LM S			29.3	21			2.04	4.88
Stock-Wright p			6.1e-08	4.5e-06			0.15	0.027
Kleiberger-Popp LM			112	7.41			15.4	33.2

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

C.5 Placebo tests with temporal leads

To account for the possibility that our coefficient estimates on the rape culture variable may be capturing anticipatory effects or pre-existing trends, we conducted a series of placebo tests with leads of the rape culture variable on the right-hand side:

$$y_{it} = \rho' \mathbf{C}_{it+1} + \beta' \mathbf{X}_{it} + \alpha_i + u_t + \epsilon_{it} \tag{C.4}$$

where $\mathbf{C}_{it+1}$ is a first-order temporal lead of the rape culture variable. If increases in crime indeed follow changes in local rape culture, these leads should be insignificant. As the results in Table C.4 report, this is indeed what we find.

Table C.4: Placebo tests, with leads $(t+1)$ of rape culture. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable:							
	Reported rapes per 1000 residents				Police vigilance (rape arrests minus rape reports)			
	Rape culture measure:							
	Producers	Consumers	Producers	Consumers	Producers	Consumers	Producers	Consumers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any rape culture, lead	0.0015	0.00067	-0.0059	0.031	0.011	0.0056	0.023	-0.031
	(0.020)	(0.024)	(0.025)	(0.030)	(0.047)	(0.030)	(0.055)	(0.035)
IV	NO	NO	YES	YES	NO	NO	YES	YES
County FE	NO	NO	NO	NO	YES	YES	YES	YES
First differences	YES	YES	YES	YES	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	31,033	31,043	31,033	31,043	34,139	34,148	34,139	34,148
Number of counties	3,104	3,105	3,104	3,105	3,105	3,105	3,105	3,105
LL	11971	11983	5718	5322	9865	9935	4394	4341

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

C.6 Heterogeneous time trends

To control for variation in sexual crime due to potentially heterogeneous regional trends, we ran another set of models that interacted yearly fixed effects with media market-level (equation C.5) and state-level dummies (equation C.6):

$$y_{it} = \rho' \mathbf{C}_{it-1} + \beta' \mathbf{X}_{it} + \sum_m d_m t + \alpha_i + u_t + \epsilon_{it} \tag{C.5}$$

$$y_{it} = \rho' \mathbf{C}_{it-1} + \beta' \mathbf{X}_{it} + \sum_s d_s t + \alpha_i + u_t + \epsilon_{it} \tag{C.6}$$

where d_m are DMA-level dummies and d_s are state-level dummies.

The results, which we report in Tables C.5 and C.6, are consistent in magnitude and significance with those reported in the main text.

Table C.5: Robustness check: media market-level time trends. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable:							
	Reported rapes per 1000 residents				Police vigilance (rape arrests minus rape reports)			
	Rape culture measure:							
	Producers	Consumers	Producers	Consumers	Producers	Consumers	Producers	Consumers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any rape culture, lagged	0.012*	0.010***	0.54***	1.88*	-0.0052**	-0.32***	-0.0049′	-0.36*
	(0.0051)	(0.0031)	(0.11)	(0.80)	(0.0019)	(0.0026)	(0.061)	(0.17)
Median personal income	0.47***	0.46***	0.11***	0.058*	-0.021*	-0.021*	-0.13***	-0.13***
	(0.064)	(0.064)	(0.0084)	(0.025)	(0.0083)	(0.0083)	(0.0069)	(0.0072)
Percent female population	-0.058***	-0.058***	-0.025***	-0.028*	-0.0029	-0.0029	-0.032***	-0.027***
	(0.016)	(0.016)	(0.0062)	(0.011)	(0.0036)	(0.0036)	(0.0051)	(0.0051)
Percentage of workers unemployed	0.22***	0.21***	0.046**	0.021	-0.038***	-0.038***	0.033***	0.031**
	(0.025)	(0.025)	(0.014)	(0.021)	(0.0096)	(0.0096)	(0.0090)	(0.011)
Percent population in religious congregation	0.032	0.034	0.020**	0.033**	0.087**	0.086**	0.012*	0.0034
	(0.080)	(0.080)	(0.0069)	(0.012)	(0.028)	(0.028)	(0.0057)	(0.0056)
Percent presidential vote for Republican	0.22***	0.22***	0.022*	0.033′	-0.057**	-0.056**	0.17***	0.17***
	(0.035)	(0.034)	(0.0086)	(0.018)	(0.017)	(0.017)	(0.0067)	(0.0073)
Constant	-3.41	-2.73	-3,522***	-3,582***	-25.9***	-26.3***	1,033	1,238***
	(11.4)	(11.5)	(71.7)	(156)	(5.34)	(5.46)	(0)	(44.4)
IV	NO	NO	YES	YES	NO	NO	YES	YES
State trends	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	40,349	40,358	37,244	37,253	40,349	40,358	37,244	37,253
Number of counties	3,105	3,105			3,105	3,105		
LL	-49835	-49843	-54269	-75168	-12013	-12018	-46971	-46929
Craig-Donald F			111	7.36			15.3	32.9
Kleiberger-Popp F			111	7.36			15.3	32.9
Anderson-Rubin F			29.1	20.9			2.04	4.86
Anderson-Rubin p			6.8e-08	4.9e-06			0.15	0.028
Stock-Wright LM S			29.3	21			2.05	4.89
Stock-Wright p			6.1e-08	4.5e-06			0.15	0.027
Kleiberger-Popp LM			112	7.42			15.4	33.2

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

Table C.6: Robustness check: state-level time trends. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable:							
	Reported rapes per 1000 residents				Police vigilance (rape arrests minus rape reports)			
	Rape culture measure:							
	Producers	Consumers	Producers	Consumers	Producers	Consumers	Producers	Consumers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any rape culture, lagged	0.011*	0.0094**	0.53***	0.90**	-0.0054**	-0.0057′	-0.24***	-0.74*
	(0.0050)	(0.0032)	(0.11)	(0.30)	(0.0020)	(0.0029)	(0.062)	(0.32)
Median personal income	0.54***	0.54***	0.12***	0.11***	-0.030**	-0.030***	-0.11***	-0.12***
	(0.070)	(0.070)	(0.0081)	(0.0097)	(0.0093)	(0.0092)	(0.0068)	(0.0085)
Percent female population	-0.072***	-0.072***	-0.027***	-0.034***	0.00082	0.00080	-0.043***	-0.029***
	(0.016)	(0.016)	(0.0061)	(0.0080)	(0.0036)	(0.0036)	(0.0051)	(0.0081)
Percentage of workers unemployed	0.20***	0.20***	0.050***	0.049**	-0.042***	-0.041***	0.047***	0.010
	(0.027)	(0.026)	(0.013)	(0.016)	(0.0072)	(0.0072)	(0.0086)	(0.020)
Percent population in religious congregation	0.12	0.12	0.023***	0.042***	0.082***	0.081***	0.017**	0.0033
	(0.092)	(0.092)	(0.0064)	(0.0092)	(0.023)	(0.023)	(0.0054)	(0.0080)
Percent presidential vote for Republican	0.19***	0.19***	0.027***	0.040**	-0.058**	-0.057**	0.15***	0.12***
	(0.030)	(0.029)	(0.0079)	(0.012)	(0.018)	(0.018)	(0.0062)	(0.013)
Constant	93.3**	94.0**	-3,521***	-3,448***	-6.93	-7.21	1,078	1,225***
	(31.7)	(31.6)	(72.5)	(80.1)	(5.13)	(5.05)	(0)	(56.0)
IV	NO	NO	YES	YES	NO	NO	YES	YES
State trends	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	40,349	40,358	37,244	37,253	40,349	40,358	37,244	37,253
Number of counties	3,105	3,105			3,105	3,105		
LL	-50802	-50811	-55030	-60635	-13185	-13201	-48565	-55400
Craig-Donald F			108	23.2			16.9	14.7
Kleiberger-Popp F			108	23.2			16.9	14.7
Anderson-Rubin F			26.4	15.3			1.19	8.45
Anderson-Rubin p			2.8e-07	0.000090			0.28	0.0036
Stock-Wright LM S			26.5	15.4			1.19	8.47
Stock-Wright p			2.7e-07	0.000089			0.27	0.0036
Kleiberger-Popp LM			108	23.2			16.9	14.7

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ′ p<0.1

C.7 Logarithmic transformations

Our measure of the local prevalence of rape – "reported rapes per 1,000 residents" (Models 1-4 in Table B.9, or Table 2 of main text) – is highly skewed, with a long right tail (Figure C.3). It is also universally non-negative, unlike our second dependent variable – "police vigilance," or the difference between rape arrests and reports – which can take any value on the real line. To account for these distributional features, we re-ran Models 1-4, with the dependent variable logged rather than in levels:

$$\ln(y_{it}) = \rho' \ln(\mathbf{C}_{it-1}) + \beta' \ln(\mathbf{X}_{it}) + \alpha_i + u_t + \epsilon_{it} \tag{C.7}$$

Figure C.3 shows the distribution of the variable, before and after the transformation. While taking the log succeeded in "pulling in" extreme values, it did not transform the variable to full symmetry. The results, which we report in Table C.7, are substantively the same as before: coefficients on the rape culture variable remain positive and statistically significant.

Figure C.3: Histograms of first dependent variable ("reported rapes per 1,000 residents") before and after logarithmic transformation.

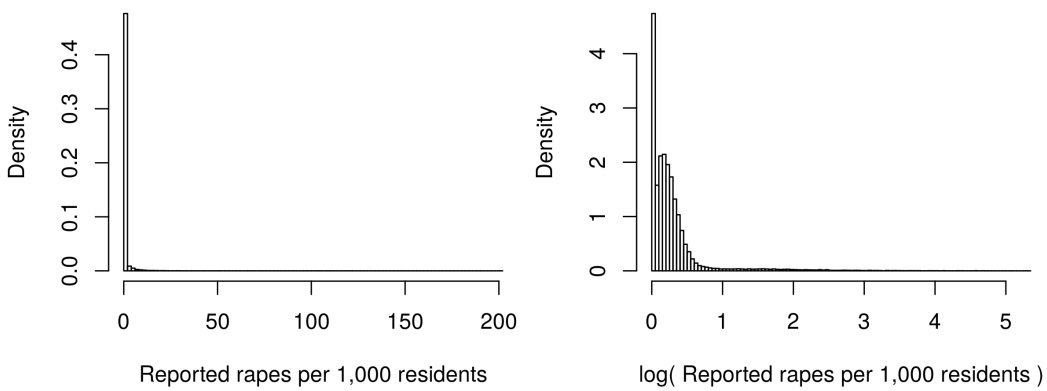


Table C.7: Robustness check: logged dependent variable. Standardized coefficients, robust standard errors in parentheses.

	Dependent variable: Reported rapes per 1000 residents (log)			
	Rape culture measure:			
	Producers (1)	Consumers (2)	Producers (3)	Consumers (4)
Any rape culture, lagged, log	0.019***	0.013***	0.52***	1.39***
	(0.0033)	(0.0032)	(0.088)	(0.40)
Median personal income	0.39***	0.39***	0.42***	0.38***
	(0.025)	(0.025)	(0.016)	(0.027)
Percent female population	0.0033	0.0037	-0.011	-0.0100
	(0.010)	(0.010)	(0.0093)	(0.014)
Percentage of workers unemployed	0.041***	0.037***	0.16***	0.083***
	(0.011)	(0.011)	(0.023)	(0.021)
Percent population in religious congregation	0.40***	0.41***	0.29***	0.25**
	(0.041)	(0.041)	(0.048)	(0.082)
Percent presidential vote for Republican	0.21***	0.21***	0.41***	0.31***
	(0.017)	(0.017)	(0.034)	(0.037)
IV	NO	NO	YES	YES
County FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	40,358	40,358	37,253	37,253
Number of counties	3,105	3,105	3,105	3,105
LL	-36540	-36550	-42366	-58492
Craig-Donald F			97.6	16.3
Kleiberger-Popp F			97.6	16.3
Anderson-Rubin F			52.9	43.1
Anderson-Rubin p			0	5.4e-11
Stock-Wright LM S			52.8	43
Stock-Wright p			0	5.4e-11
Kleiberger-Popp LM			97.4	16.3

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ' p<0.1

References

Anderson, Theodore Wilbur and Cheng Hsiao. 1982. "Formulation and estimation of dynamic models using panel data." *Journal of econometrics* 18(1):47–82.

Arellano, Manuel and Olympia Bover. 1995. "Another look at the instrumental variable estimation of error-components models." *Journal of econometrics* 68(1):29–51.

Arellano, Manuel and Stephen Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The review of economic studies* 58(2):277–297.

Hamilton, James. 2004. *All the news that's fit to sell: How the market transforms information into news*. Princeton University Press.

Nickell, Stephen. 1981. "Biases in dynamic models with fixed effects." *Econometrica: Journal of the Econometric Society* pp. 1417–1426.