Female Judicial Assignment and the Content of Appellate Opinions

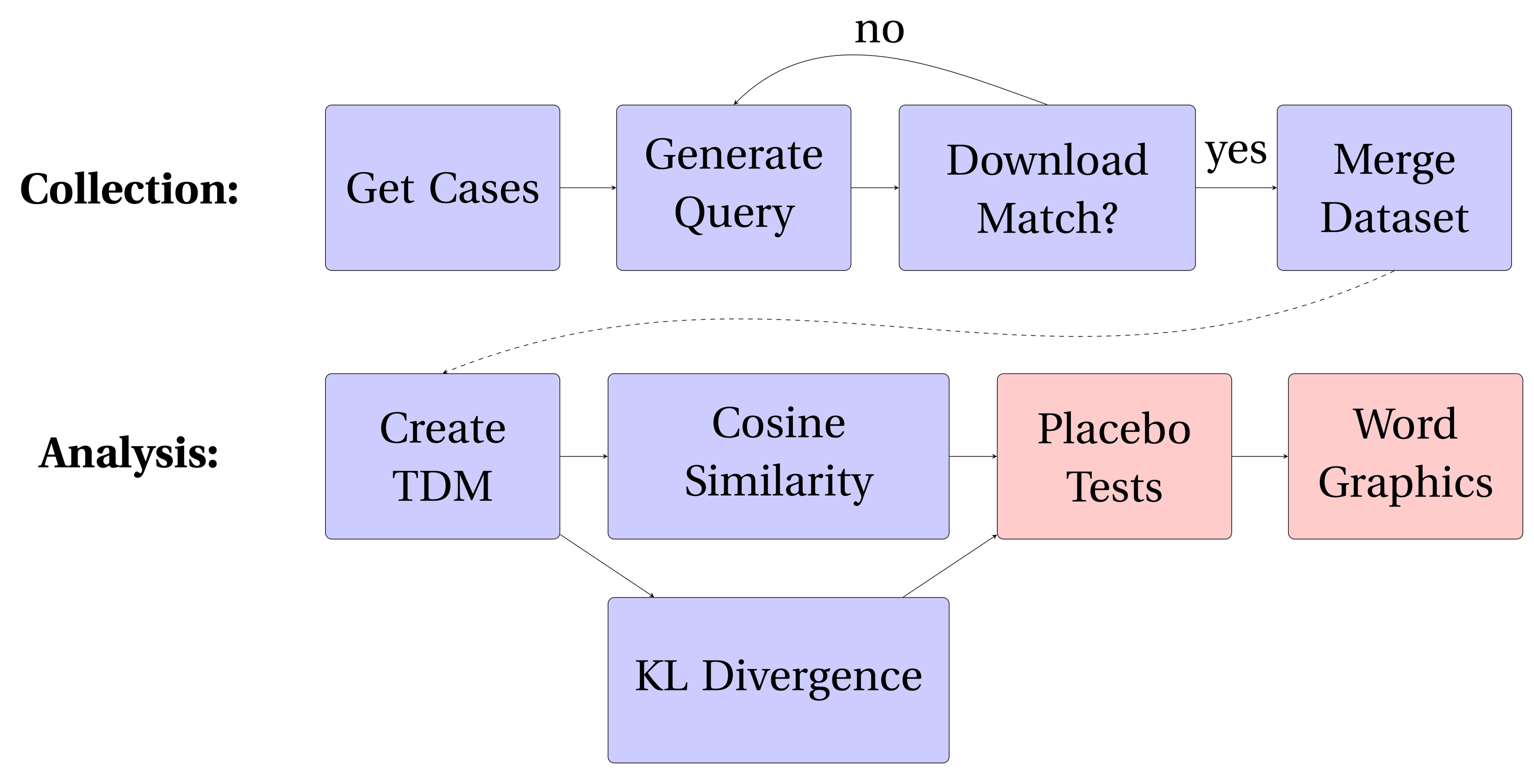
Michael Gill and Andrew Hall

Department of Government, Harvard University
email: mzgill@fas.harvard.edu, hall@fas.harvard.edu

1 Motivation

A significant body of research in social science and empirical law has shown that the random assignment of one or more female judges to appeals court cases affects outcomes. In particular, the presence of female judges makes the appeals court rule more favorably for plaintiffs, especially in cases about sexual discrimination, harassment, abortion, and Title VII. Our project extends this work by examining whether the presence of one or more female judges affects the way the text of the rulings are written. This question is important because the text of published appeals court rulings will inform future cases by precedent. Using automated content analysis, we analyze published appeals court rulings for cases in which all three judges were men, and for cases in which there is at least one female judge. We find that the presence of a female judge changes the manner in which the rulings are written to a significant degree. Furthermore, we show this effect is not the result of female authorship. These results suggest female judges are demonstrably effective representatives for women's issues, not just as judges who rule in a certain way, but as esteemed colleagues who inform the content of a panel's opinion.

2 Workflow



We scraped the texts of over 1200 judicial opinions using `BeautifulSoup` and fuzzy string matching (with `fuzzywuzzy`) in `Python`. The particular cases correspond to data from Boyd, Epstein and Martin (2010) and Sunstein et al. (2006). Each case's text was sorted into one of thirteen legal issue areas, ranging from abortion suits, to capital punishment cases, to those centered around Title VII of the Civil Rights Act. Not all issue areas are represented with equal frequency, and the total number of scraped opinions is somewhat less than the cross-sectional data in Boyd, Epstein and Martin (2010) and Sunstein et al. (2006). A likely reason for this is that the widespread publication of opinions has only been common in recent years, and there is a tradition to publish some issue areas more frequently than others (Hoffman, 1981).[a]

[a] A draft of this paper was presented at the 2011 Annual Meeting of the Society for Political Methodology's poster session. We thank participants of that conference, as well as members of the Harvard Graduate Student Political Economy Workshop, for valuable comments and advice throughout this process.

3 Empirics

We divide the entire corpus into two categories: *treatment*, if at least one female judge is on the panel, and *control*, if the panel is all male. For each group we generate a vector expressing word frequencies for the entire observed vocabulary.

Cosine Similarity

We then calculate the *Cosine Similarity* between the treatment and control groups:

$$similarity = cos(\theta) = \frac{C \cdot T}{||C|| \cdot ||T||}$$

Where T and C are the word-vectors for the treatment and control text corpuses, respectively. To find the significance of our estimates, we employ *placebo tests*. We repeatedly divide the entire corpus into two random groups, equal in size to the treatment and control groups, and calculate $cos(\theta)$.

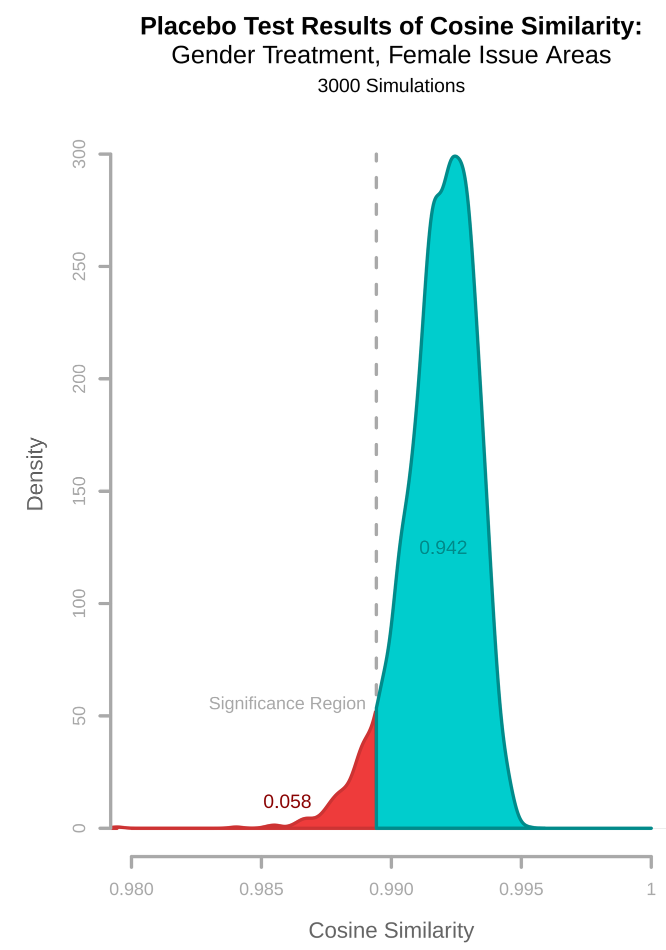
Kullback-Leibler Divergence

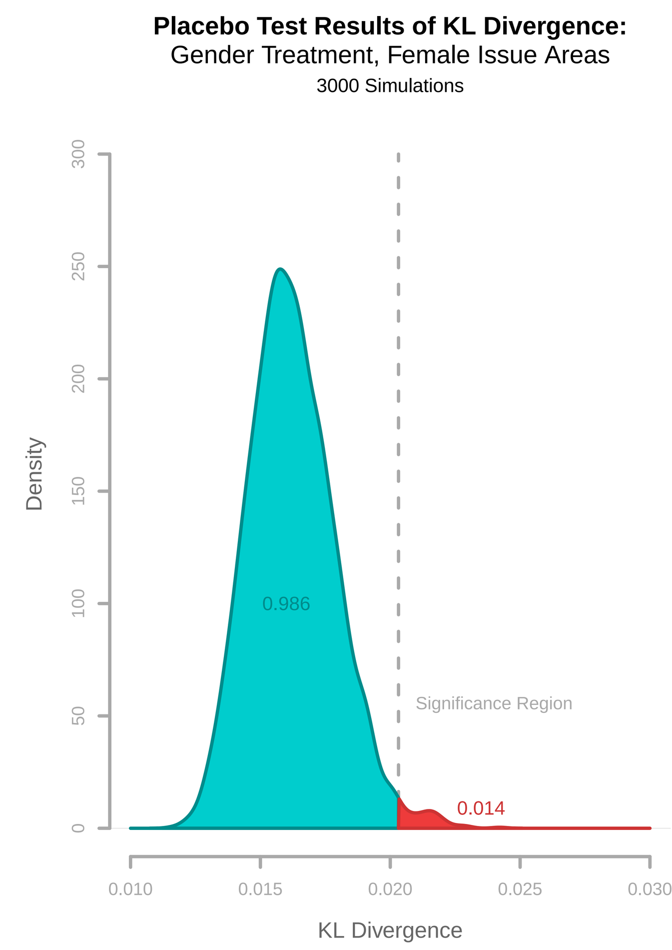
Formally, let V be the set of observed words (i.e. a unique list of every word observed in any document). Continue to let the vectors C and T represent the control and treatment word frequency vectors, thinking of them in terms of probability distributions now. We have:

$$KL(C||T) = \sum_{w \in V} P(w|T) \cdot \log\left(\frac{P(w|T)}{P(w|C)}\right)$$

In words, we take each word w in the vocabulary V and compare how likely it is under each distribution. If the two distributions are identical, then the KLD is 0. If they are completely different, the KLD can range to ∞.

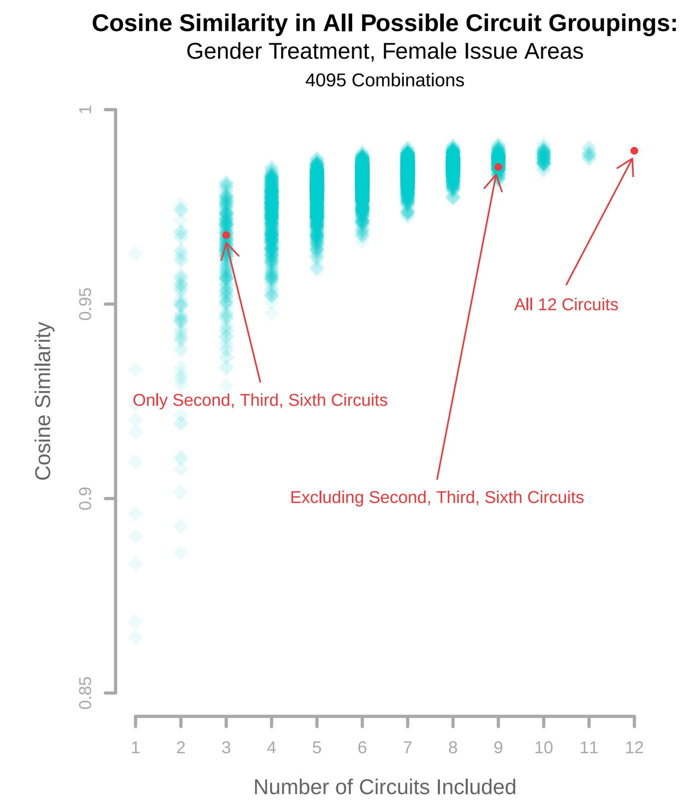
4 Results

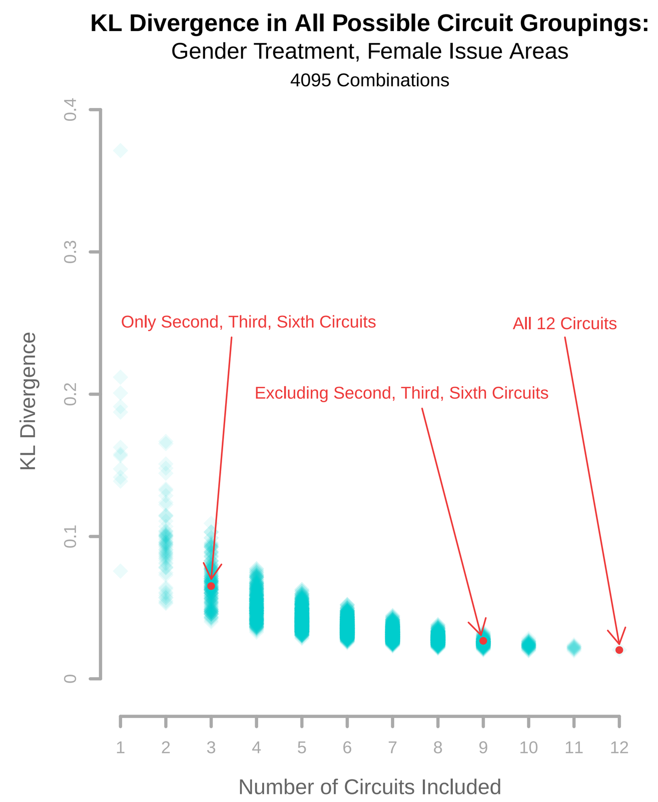




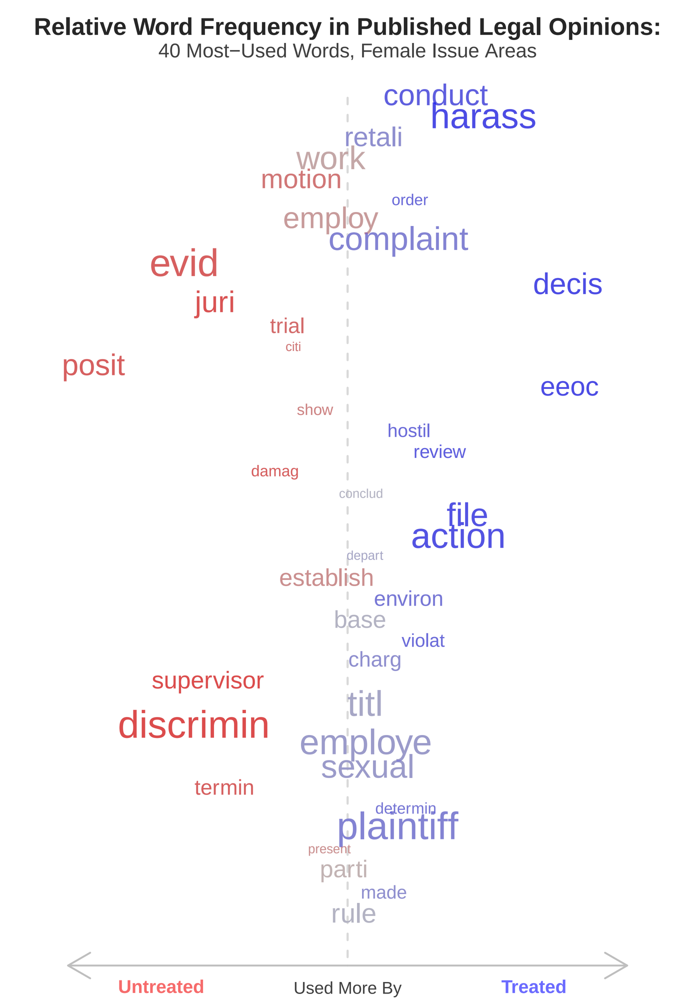
5 Circuit Checks

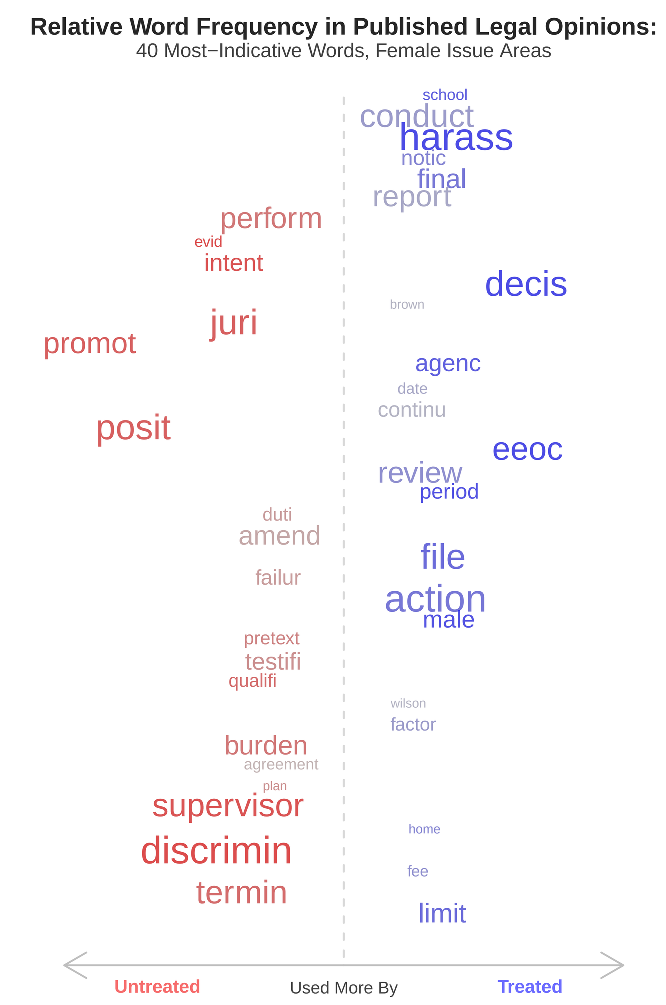
We show how the cosine similarity and KL-divergence change over all possible subsets of circuits. The strong stability gives us reasonable confidence that non-random assignment at the individual circuit level are not driving our results.





6 Word Plots





References

Boyd, CL, L Epstein and AD Martin. 2010. "Untangling the causal effects of sex on judging." *American Journal of Political Science* 54(2):389–411.

Hoffman, DN. 1981. "Nonpublication of Federal Appellate Court Opinions." *Just Sys J* 6:405–434.

Sunstein, Cass R, David Schkade, Lisa M Ellman and Andres Sawicki. 2006. "Are Judges Political? An Empirical Analysis of the Federal Judiciary." pp. 1–190.