The Gender Readings Gap in Political Science Graduate Training[1]

Heidi Hardt
University of California, Irvine

Amy Erica Smith
Iowa State University

Hannah Kim
University of California, Irvine

Philippe Meister
Iowa State University

DRAFT
Please do not cite or distribute without permission.
If you have questions, please contact Amy Erica Smith at aesmith2@iastate.edu.

Do academic institutions influence gender representation in assigned readings during graduate training? Whereas recent studies have identified gender gaps in citations and publications, less is known about the readings used to train future political scientists. Introducing a unique dataset of 88,673 citations from 905 Ph.D. syllabi and reading lists, we find that work by female scholars is underrepresented in all political science subfields. Whereas approximately 28% of instructors in our dataset are female, only 19% of assigned readings have female first authors. We find evidence that subfields and departments affect gender representation. First, math-heavy subfields have lower representation of female-authored readings. Second, department context matters. As departments hire more female instructors, both male and female instructors are more likely to assign scholarship authored by female-only teams. However, context does not affect assignment of mixed-gender work. This research also contributes a comprehensive dataset to study graduate training.

[1] Hardt and Smith were co-PIs on this research and contributed equally to it.

Growing evidence documents gender gaps in political science publications and citations (Teele and Thelen 2017; Zigerell 2015; Maliniak 2013; Mitchell et al. 2013; Mathews and Andersen 2001); these gaps affect hiring and promotions (Hesli et al. 2012). Only recently have studies considered an early influence: graduate training. Students exposed to few female-authored readings may become less likely to cite women in their own scholarship and teaching. Underrepresentation in readings may hurt female graduate student retention (Alper and Gibbons 1993, Mershon and Walsh 2016, 1), given fewer same-gender role models (Gilbert 1985). As in politics, women in leadership roles inspire other women (Hill 1981, Fox and Lawless 2010).

We assess women's underrepresentation in assigned readings across the discipline of political science, and we investigate institutional factors shaping underrepresentation. Recent work finds that female-authored readings are significantly less represented than male-authored readings in syllabi in at least two subfields (International Relations [IR] and American politics) (Colgan 2017; Diament et al. 2017). We ask: *How do academic institutions affect gender representation in political science graduate students' assigned readings?* Gender representation refers to the proportion of readings with female (relative to male-identified) authors. Underrepresentation - thus, a gender readings gap - would be a discrepancy between the proportion of female-authored readings and of female instructors of graduate courses.

Analyzing a new dataset GRADS (the GRaduate Assignments DataSet), with 88,673 readings from 840 syllabi and 65 reading lists, we find a gender readings gap in every subfield. Relative to the 27.6% of Ph.D.-level instructors who are female in the dataset, only 18.6% of the readings' first authors (and 19.1% of all readings' authors) are female. Academic institutions influence the size of the gap. First, women's representation is lower in math-heavy subfields (i.e. methodology and political economy). Second, gender representation is lower in top-ranked

departments and ones with few female faculty. As departments hire more women, male and female faculty assign more diverse readings. Our research advances research on diversity in political science and introduces a new dataset for assessing trends in graduate training.

HYPOTHESES

We assume that implicit gender norms influence student and faculty behavior, and that institutions can become gendered (Lorber 1994, 111, Rivera 2017). Incorrect, implicit biases hold that women are less generally competent (Eagly 1994, Fausto-Sterling 2008, Leeds 2013) and worse at math than men (Spencer et al. 1999, Bell and Burkley 2014, Morrow-Jones and Box-Steffensmeier 2014). By age six, children associate brilliance with being male (Bian et al. 2017). Life-long implicit biases shape implicit norms in academic institutions (Acker 1992).

First, we hypothesize subfield differences. Due to self-stereotyping, men may be more likely than women to select into heavy math subfields (e.g. methodology and political economy). Yet gender norms are also shifting in some subfields. In IR, for example, senior female scholars have established formal networks and initiatives to mentor and retain female junior scholars (e.g. Journeys in World Politics, Pay it Forward), increasing awareness of female-authored work.

H1a: *Syllabi and reading lists in subfields that require more (less) training in mathematics will have higher (lower) representation of female authors in readings.*

H1b: *Controlling for gender composition of subfields, syllabi and reading lists in IR will have higher representation of female authors in readings.*

Second, we expect department-level variation. The presence of women in a department's faculty may affect whether instructors assign female-authored work. Just as women's presence in organizations leads to significant changes in functioning and performance (Post and Byron 2015,

Center et al. 2017), departmental networks and norms change as programs hire more women. Interpersonal relations, even hallway conversations, can raise awareness of women's research and underrepresentation in the discipline. Program prestige may also matter. Instructors in top-ranked programs are disproportionately male, likely affecting citation patterns within networks (Massen et al. 2017). Rank could also exacerbate implicit biases about men's superior academic competence (Moss-Racusin 2012, Knobloch-Westerwick 2013).

H2: *Syllabi and reading lists authored by instructors working in departments with fewer (more) women will have lower (higher) representation of female authors in readings.*

H3: *Syllabi authored by instructors in departments with top 10 political science Ph.D. programs, per US News & World Reports, will have lower representation of female authors in readings.*

EMPIRICAL ANALYSIS

We employed multiple modes of data collection to assemble GRADS: a dataset of optional and required readings from 840 syllabi and 65 reading lists.[2] Our own and others' web searches returned 160 syllabi (132 in comparative politics). Then, in September 2016, a national survey of APSA-member faculty yielded 301 syllabi. The APSA Research and Development Division also shared 38 reading lists from a summer 2016 survey of graduate program directors. Finally, in fall 2016, Ph.D. student project affiliates in 27 of the top 50 Ph.D. programs helped us collect 450 syllabi and 28 reading lists. Duplicates and earlier versions were excluded. Citations were extracted and parsed using manual and machine coding methods (see Online Appendix).

[2] Readings include scholarly articles, books, and other documents. Syllabi are faculty-written documents setting course norms. Reading lists are literature guides written by faculty groups.

As a multilevel dataset, GRADS includes 137,305 (non-unique) authors, 88,673 (non-unique) readings, 606 (unique) instructors, and 95 US-based political science departments. To code authors' gender as male or female,[3] we first created a list of scholars whose genders were likely to be miscoded or uncodable. We coded remaining names using a list of given (i.e. first) names from US and UK censuses and social media data.[4] Subfield(s) and syllabus/reading list year were also coded.[5] Instructor gender was based on survey responses, online bios in the third-person, and names data. Instructors and syllabi are cross-nested, since some instructors had multiple syllabi and some syllabi multiple instructors. Departments were ranked in groups of ten following US News and World Reports (2016). In fall 2017, we determined department gender composition from websites by counting tenure-stream instructors with primary appointments in political science. Our unit of analysis is the syllabus/reading list. The dependent variable – proportion of readings with female first authors – runs from 0 to 1 and is modeled using fractional logistic regression. Results are unchanged with hierarchical models by departments.

We find that readings authored by women are underrepresented. We set a conservative benchmark to assess underrepresentation: the proportion of instructors in our sample who are

[3] We concur with Maliniak et al.'s rationale for not coding transgender identity (2013, 1). Where possible, we relied on survey-based self-reporting of gender.

[4] We excluded 285 corporate authors (e.g. World Bank) and 54 readings where there was no identified author (e.g. Bible). Gender is missing for only 76 authors (a rate of 0.055%).

[5] Missingness: 60 syllabi and 16 reading lists lack year; five instructors lack gender; 21 syllabi/reading lists lack department; four syllabi lack program rank. Some syllabi were coded with more than one subfield. Reading lists were not associated with instructors.

female (27.6%). Thus, instructors proxy for authors of 'assignable' research (Teele and Thelen 2017, 436). Contrasting with our benchmark, 18.6% of first authors (19.1% of all authors) are female – indicating a gender readings gap.[6] Gender representation slowly rose over the time period, from 16.2% in 2004–2012 to 18.5% in 2015–2016 (see Online Appendix).

Figure 1 depicts the average percentage of readings with female first authors by subfield. Among major subfields, Methodology has the lowest representation of female first-authored readings (above 10%); Theory has the highest (just over 30%). In multivariate analysis, in support of H1a, women's representation in readings is significantly lower in math-heavy subfields, including American politics, political economy, and methodology.

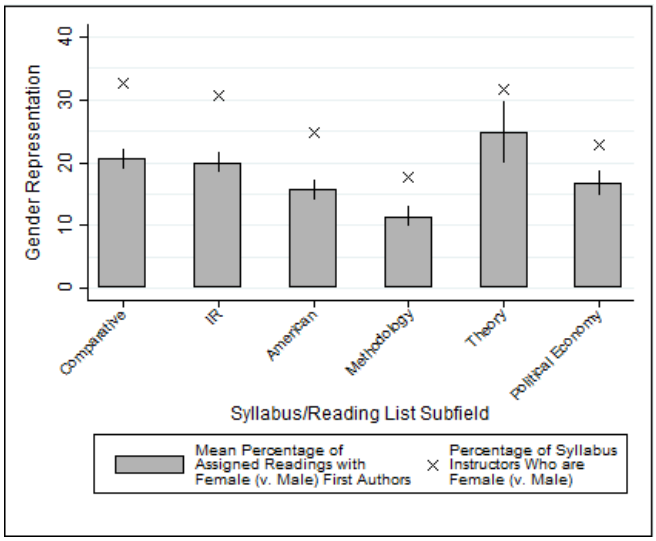
However, subfield effects result in part from varying 'benchmark' gender composition. Instructors assign fewer female-authored works when the pool of women researchers is smaller. Most significant coefficients for subfield disappear after controlling for subfield composition and instructor gender (Online Appendix). Nonetheless, IR and theory instructors assign more work by women than predicted by subfield gender composition, supporting H1b. High gender representation in theory may result from the prominence of feminist scholarship. Such scholarship likely appeared earlier than in other subfields and is predominantly female-authored.

Departments also matter. In support of H2, a department's proportion of women affects syllabi. When only 10% of instructors are female, faculty assign few readings by women – just 10%.[7] In evenly gender balanced departments, about 30% of readings have female first authors. Figure 2 shows that department context shapes *both* men's and women's behavior, yet female

[6] Female authors tend to appear later in author order; women constitute 18.7% of first authors, 19.7% of second, 19.2% of third, 21.2% of fourth, 22.6% of fifth, and 25.7% of sixth authors.

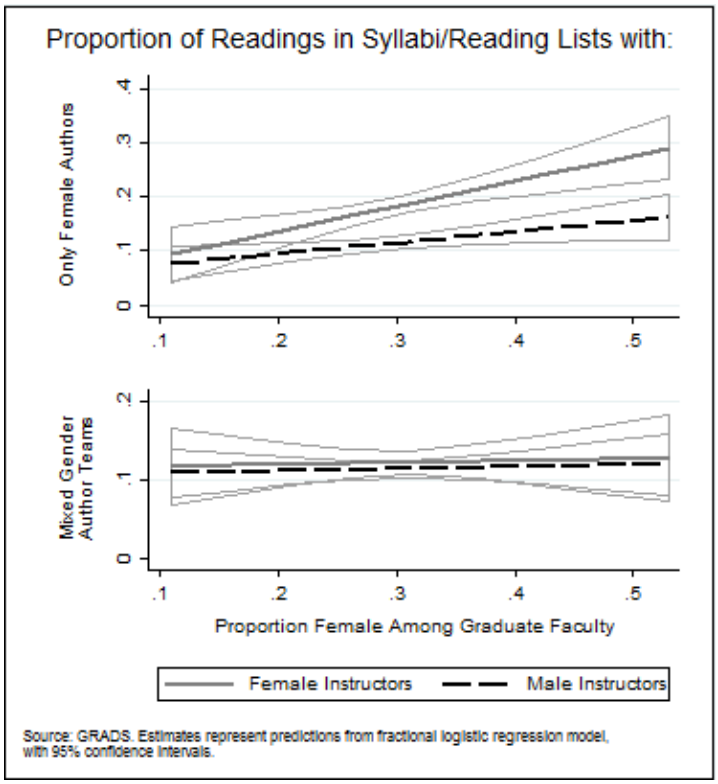
[7] Bivariate predictions from a multilevel linear model (department is a contextual variable).

instructors are more responsive, perhaps due to the role of networks. However, neither instructor gender nor department composition affect the rate of assigning work by mixed-gender teams.



Source: GRADS. Whiskers represent 95% confidence intervals.

FIGURE 1. *Percentage of Assigned Readings with Female First Authors, by Subfield*



Source: GRADS. Predictions and 95% confidence intervals from fractional logistic regression.

FIGURE 2. *Gender Representation in Assigned Readings, by Department Gender Composition*

Support for H3 is mixed. In bivariate analysis, syllabi in departments with top-ranked Ph.D. programs are significantly less likely to include female-authored readings. Again, effects are limited to works written by female-only authorship teams, but not mixed gender teams (Online Appendix). However, program rank and department gender composition are correlated (r=.30); 'top 10' programs have fewer female faculty. Controlling for department gender composition, the effect of program rank disappears. In full multivariate models, the two strongest determinants of gender representation in readings are the benchmark percentage of a subfield that is female and department gender composition. The Online Appendix presents multivariate results, as well as separate models for male and female instructors.[8]

CONCLUSION

This study advances scholarship on gender diversity by revealing evidence of a gender readings gap in graduate training – in and across political science subfields. Female-authored work is underrepresented relative to the proportion of faculty who are female. We introduce what is, to our knowledge, the first comprehensive dataset of readings in graduate training with 88,673 readings from 840 syllabi and 65 comprehensive exam reading lists. Rates of assigning women's work vary significantly across subfields, even after controlling for subfield gender composition. Departments' gender composition also matters significantly for the readings gap. When departments have more female faculty, faculty more frequently assign female-authored readings.

[8] We also find heterogeneity among top journals, ranging from 13.2% female first authors in readings from the *American Political Science Review*, to 21.2% in the *Journal of Politics*, and 35.3% in the *Annual Review of Political Science.*

Our rich, multilevel dataset enables scholars to address related questions. How does the gender readings gap affect the citation gap? How do faculty networks influence citation patterns? How are ethnic minorities represented in readings? If they so choose, instructors can also use our data to diversify their syllabi and publications (see also Beaulieu et al. 2017; Sumner 2018).

REFERENCES

Acker, Joan. 1992. From Sex Roles to Gendered Institutions. *Contemporary Sociology* 21 (5): 565-569.

Alper, Joe and Gibbons, Ann. (1993). The Pipeline is Leaking Women All the Way Along. *Science* 260(5106): 409-412.

Beaulieu, Emily, Amber E. Boydstun, Nadia E. Brown, Kim Yi Dionne, Andra Gillespie, Samara Klar, Yanna Krupnikov, Melissa R. Michelson, Kathleen Searles, and Christina Wolbrecht. 2017. "Women Also Know Stuff: Meta-Level Mentoring to Battle Gender Bias in Political Science." *PS: Political Science & Politics* 50 (3): 779–83.

Bell, Angela, and Melissa Burkley. 2014. Women Like Me Are Bad at Math: The Psychological Functions of Negative Self-Stereotyping. *Social and Personality Psychology Compass.* 8(12): 708-720.

Bian, Lin, Sarah-Jane Leslie, and Andrei Cimpian. (2017) Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, 355 (6323): 389-391.

Colgan, Jeffrey. 2017. Gender Bias in International Relations Graduate Education: Evidence from Syllabi. *PS: Political Science & Politics.* April:1-5.

Diament, Sean, Adam Howat and Matthew Lacombe. 2018. Gender Representation in the American Politics Canon. *PS: Political Science & Politics*: 1-6.

Eagly, Alice and Mladinic, Antonio. (1994) Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European Review of Social Psychology* 5(1):1–35.

Fausto-Sterling, Anne. 2008. *Myths of Gender: Biological Theories about Women and Men*. New York: Basic Books.

Fox, Richard, and Jennifer Lawless. 2010. If only they'd ask: Gender, recruitment, and political ambition. *Journal of Politics*, 72(2): 310-326.

Gilbert, Lucia. 1985. Dimensions of Same-Gender Student-Faculty Role-Model Relationships. *Sex Roles* 12 (1-2): 111–23.

Hesli, Vicki L., Jae Mook Lee, and Sara McLaughlin Mitchell. 2012. Predicting Rank Attainment in Political Science: What Else Besides Publications Affects Promotion? *PS: Political Science & Politics* 45 (03): 475–92.

Hill, David. 1981. Political culture and female political representation. *Journal of Politics*, 43(1): 159-168.

Knobloch-Westerwick, Silvia, Carroll J. Glynn, and Michael Huge. 2013. The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest. *Science Communication* 35 (5): 603–25.

Leeds, Brett Ashley. 2013. Why Is Work by Women Systematically Devalued? *The Monkey Cage* (blog), *The Washington Post*, October 4.

Lorber, Judith. 1994. *Paradoxes of Gender*. New Haven, CT: Yale University Press.

Maliniak, Daniel, Ryan Powers, and Barbara Walter. 2013. The Gender Citation Gap in International Relations. *International Organization* 67 (4): 889–922.

Massen, Jorg, Lisa Bauer, Benjamin Spurny, Thomas Bugnyar, and Mariska E. Kret. 2017. Sharing of science is most likely among male scientists. *Scientific Reports* 7(1): 12927.

Mathews, A., and Kristi Andersen. 2001. A Gender Gap in Publishing? Women's Representation in Edited Political Science Books. *PS: Political Science & Politics* 34 (1): 143–47.

Mershon, Carol and Denise Walsh. 2016. Diversity in Political Science. Why it Matters and How to Get it. *Politics, Groups and Identities, 4 (3)*: 1-5.

Mitchell, Sara, Samantha Lange, and Holly Brus. 2013. Gendered Citation Patterns in International Relations Journals. *International Studies Perspectives* 14 (4): 485–92.

Morrow-Jones, Hazel and Janet M. Box-Steffensmeier. 2014. Implicit Bias and Why It Matters to the Field of Political Methodology. *The Political Methodologist* 21 (2):16-19.

Moss-Racusin, Corinne, John Dovidio, Victoria Brescoll, Mark Graham, and Jo Handelsman. Science faculty's subtle gender biases favor male students. 2012. *Proceedings of the National Academy of Sciences* 109 (41): 16474-16479.

Post, Corinne and Kris Byron. 2015. Women on boards and firm financial performance: A meta-analysis. *Academy of Management Journal* 58(5): 1546-1571.

Spencer, Steven, Claude Steele and Diane Quinn. 1999. Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* 35(1): 4-28.

Sumner, Jane Lawrence. 2018. The GBAT: A Web-Based Tool for Estimating Gender Balance in Syllabi and Bibliographies. *PS: Political Science & Politics*.

Teele, Dawn and Thelen, Kathleen. 2017. Gender in the Journals: Publication Patterns in Political Science. *PS: Political Science & Politics, 50*(2): 433-447.

Zigerell, L. J. 2015. Is the Gender Citation Gap in International Relations Driven by Elite Papers? *Research & Politics* 2 (2): 1-7.

Online Appendix with Supplementary Materials for
'The Gender Readings Gap in Political Science Graduate Training'[1]

I. GRADS (GRaduate Assignments DataSet) Collection of Syllabi and Reading Lists

A. Overview

We employed multiple modes of data collection to assemble a dataset of 840 syllabi and 65 reading lists in Microsoft Word or PDF format. Data collection had five phases:

1. In an exploratory phase prior to beginning this project, in the fall of 2015, 29 graduate syllabi were collected through online searches for the top ten graduate programs in political science, as ranked by US News and World Reports.
2. Professor David Samuels at the University of Minnesota kindly shared with us a collection of 131 comparative politics syllabi that he collected in the fall of 2016, through web searches.
3. The Research and Development Division of the American Political Science Association (APSA) generously shared 38 reading lists collected from a summer 2016 survey of graduate program directors.
4. In September 2016, we collected 301 syllabi through a national survey of faculty disseminated by the APSA Research and Development Division. Only respondents who submitted syllabi are included in GRADS.
5. We invited Ph.D. students from the top 50 political science Ph.D. programs in the U.S. (per US News and World Reports) to serve as project affiliates. With the help of 27 project affiliates, we obtained 450 syllabi and 28 reading lists.

Duplicate syllabi, as well as older versions of the same course taught by the same professor, were removed from the final dataset. In the following two sections, we discuss the representativeness of the survey respondents (phase 4) and the project affiliates (phase 5).

Table A1 presents the breakdown of the sample of syllabi and reading lists by subfield. The subfield of comparative politics is overrepresented due to our collection strategies. Nonetheless, the dataset constitutes the largest collection of graduate syllabi and reading lists to date.

[1] The two lead authors were co-PIs on this research and contributed equally to it in research design and writing. The additional authors executed data collection, supported data analysis and also contributed important ideas and feedback along the way.

TABLE A1. *Characteristics of the GRADS Dataset*

	N Syllabi	N Reading Lists
All documents	840	65
Comparative politics	268	23
International relations	162	10
American politics (including judicial/courts)	163	18
Methods	179	2
Theory	94	7
Political economy	60	1
Public policy/Public administration	15	4
Political psychology	21	0
Gender/Identity	22	4

Note: Counts by subfield sum to more than the total because some syllabi/reading lists are counted in multiple subfields.

B. Representativeness of Respondents to the APSA Faculty Survey

In September 2016, the APSA Research and Development Division disseminated a survey on our behalf to the universe of 2,640 member faculty working in U.S.-based Ph.D. programs. Three reminders were sent. The response rate was low, at 11.8%; 312 individuals started the survey, and 160 submitted one to three syllabi from the past three years.[2] In total, 301 syllabi were submitted. Only respondents who submitted syllabi are included in GRADS.

Out of the 312 individuals who opened the survey, 6 individuals (1.9%) declined to give informed consent.[3] Respondents who gave informed consent received a nine-question questionnaire concerning their gender, age, race/ethnicity, national origin, sexual orientation, academic rank, and country and year of Ph.D.

Of the 306 individuals who gave informed consent, 91 proceeded to the following page to view the survey, but subsequently failed to answer any survey questions. In this section, we report results for the 215 individuals who gave informed consent and for whom we have any demographic data. Out of 215 respondents, 160 uploaded syllabi while 55 did not.

As Figure A1 shows, female respondents were somewhat more likely to upload syllabi than male respondents. While 31.9% of the respondents who uploaded syllabi were female, only 23.6% of those who did not upload syllabi were female. However, in a bivariate logistic regression model, the effect of gender is not statistically significant.

[2] In other recent surveys disseminated by APSA, Hesli, Lee and Mitchell (2012) obtained a 27.0% response rate, while Djupe (2015) obtained 22.3%. We do not have a clear sense of why our response rate was low.

[3] Due to a glitch in the survey, people who failed to give informed consent were allowed to enter the survey. Two of the six people who declined actually responded to the survey and uploaded syllabi, but their responses and syllabi are discarded.

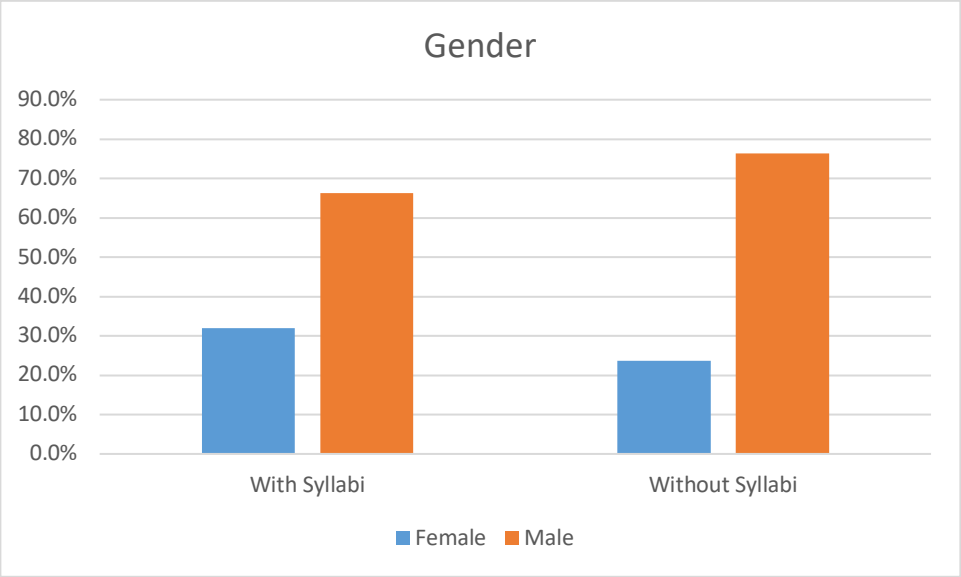


Figure A1. Gender of Survey Respondents, by Whether They Uploaded Syllabi

As Figure A2 shows, those who did not contribute syllabi were substantially younger than those who did contribute syllabi. While nearly half of those aged 20-35 did not upload syllabi, the proportions in other age groups were much smaller. In a bivariate logistic regression model, the effect of age is statistically significant. It seems likely that many respondents aged 20-35 may not have yet taught PhD level courses, leading to their lower rates of uploading syllabi.

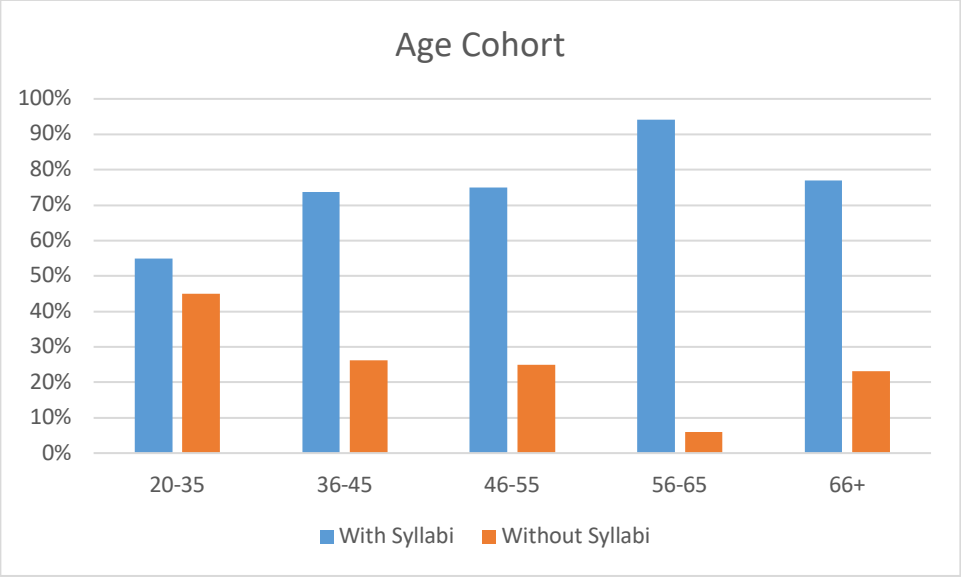


Figure A2. Age of Survey Respondents, by Whether They Uploaded Syllabi

There was also some racial imbalance. White respondents constituted 89.4% of those who uploaded syllabi and 70.9% of those who did not upload syllabi. Hispanic/Latino respondents,

however, constituted more than 10% of those who did not upload syllabi but only 1.3% of those who did. In a bivariate logistic regression model, the effect of race is statistically significant.

Table A1. Race of Survey Respondents, by Whether They Uploaded Syllabi

Race	With Syllabi	Without Syllabi
White	89.4%	70.9%
Hispanic/Latino	1.3%	10.9%
Black	1.9%	1.8%
Asian	2.5%	5.5%
American Indian/Alaskan	0.0%	0.0%
Middle Eastern/North African	0.6%	1.8%
Pacific Islander	0.0%	0.0%
Decline to respond	1.3%	3.6%
Other race	0.6%	0.0%

In addition, those who did not contribute syllabi were much less likely to have been raised in the United States, and less likely to have identified as heterosexual. 89.4% of those who uploaded syllabi said that they grew up in the United States while 76.4% of those who did not upload syllabi said the same. 92.5% of those who uploaded syllabi were heterosexual while 81.8% of those who did not upload syllabi. In bivariate logistic regression models, the effects of country of origin and sexuality are statistically significant.

As Table A3 shows, those who uploaded syllabi were substantially more senior. Full professors constituted more than half (50.6%) of those that uploaded syllabi, while they constituted less than a quarter of those who did not upload syllabi. Meanwhile, most non-tenure stream respondents failed to upload syllabi, perhaps because they do not teach graduate courses.

Table A3. Rank of Survey Respondents, by Whether They Uploaded Syllabi

Academic Rank	With Syllabus	Without Syllabus
Non-tenure Stream	1.9%	27.3%
Assistant Professor	20.6%	16.4%
Associate Professor	26.3%	29.1%
Full Professor	50.6%	23.6%
Decline to respond	0.6%	3.6%

Table A4 presents results of a logistic regression model predicting whether there were systematic differences between those who uploaded syllabi and those who did not. We used dummy variables for female; academic rank; race (white); sexual orientation (heterosexual); national origin (US); and country of PhD. Age cohort and year of PhD were omitted because of multicollinearity with academic rank; academic rank more strongly predicts response rate than the other two variables.

Table A4. *Determinants of Whether a Respondent Uploaded Syllabi (Logistic Regression Model)*

DV = Syllabi	Coefficient	Std. Err.	P
Gender (Female)	0.796	0.449	0.076
Non-Tenure Stream	-3.229	0.776	0.000
Assistant Professor	-0.518	0.497	0.297
Associate Professor	-0.684	0.438	0.118
Race (White)	0.515	0.501	0.304
Country of Origin (United States)	0.429	0.551	0.436
Sexual Orientation (Heterosexual)	0.724	0.553	0.191
Country of PhD (United States)	0.337	1.028	0.743
Constant	-0.223	1.266	0.860
n=212			
Pseudo-R^2 = 0.177			

Note: Omitted category for academic rank corresponds to full professors.

In multivariate analysis, there are two significant determinants of uploading syllabi. By far the most important factor determining whether one contributes syllabi is rank. Non-tenure stream faculty have a predicted probability of .21 of uploading syllabi. By contrast, the predicted probabilities are .77 for assistant professors, .74 for associate professors, and .85 for full professors. In addition, gender is statistically significant at p=.08 once one controls for rank (it is not statistically significant before controlling for rank because women are less likely to be associate or full professors). Controlling for rank, the predicted probability of a female respondent uploading syllabi is .82, while it is .72 for men.

C. Representativeness of Graduate Student Project Affiliates

In the fall of 2016, we attempted to recruit graduate student project affiliates to contribute their departments' syllabi in the 50 universities ranked in the top 50 programs by US News and World Reports. In total, we contacted students at 53 universities, since several were tied in ranking. To identify potential project affiliates, we gathered names of students with an interest in gender and politics—or if none were available, American politics—from department websites. We also contacted department chairs and administrative offices to explain that we would be contacting students, and to ask for help in identifying graduate student leaders. We then emailed the students who had been identified. In the end, 27 Project Affiliates from 27 universities provided syllabi and reading lists. Some potential Project Affiliates responded to initial emails and agreed to participate, but did not end up returning syllabi.

Figure A3 below presents the distribution of responses from contacted departments, by rank. A logistic regression model indicates that there are no statistically significant differences between responding and non-responding departments, by rank.

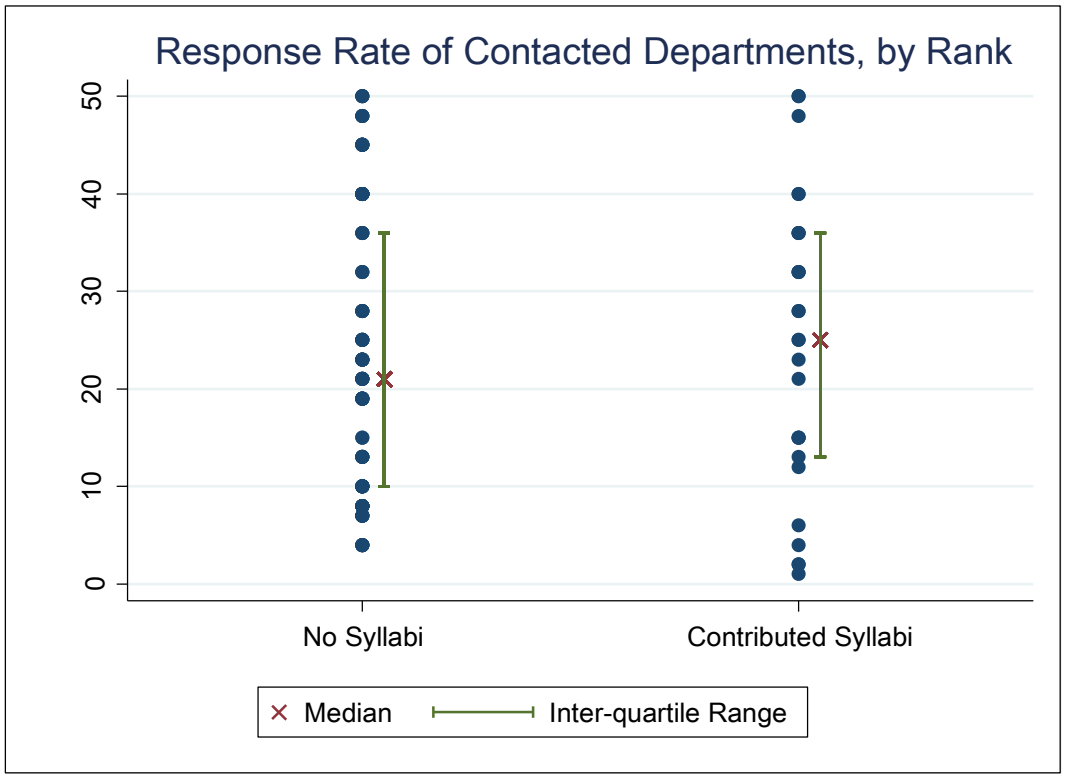


Figure A3. Rank of Responding and Non-Responding Departments

In total, 169 potential project affiliates were contacted, and 27 eventually contributed syllabi: 14 women and 13 men. While female potential project affiliates were slightly less likely to respond to recruitment emails than were male project affiliates, the effect of gender is not statistically significant at $p < .10$. The difference in response rates might be because student leaders (who were more likely to be male) were more likely to respond than were students interested in gender and politics.

II. Assembly of the GRADS Dataset

A. Extraction and Coding of Readings

Syllabi were collected in Microsoft Word or PDF format, and then processed using a combination of manual and machine coding to extract and parse readings. Each syllabus was first converted to a Microsoft Word document, and all syllabus elements except citations were stripped out. Thus, the first phase produced a set of over 900 Word documents comprised of lists of citations.

We intended to utilize Python script to look up each "stripped" citation on the Internet via Google Scholar and retrieve its corresponding BibTeX citation and citation count. However, we discovered that query limits made it impossible to use Google Scholar or Google Books at the

rate/volume that we needed. Various attempts further indicated that it would be impossible to use CrossRef to retrieve accurate results for books.

When automatic searches failed, we decided to write a program to parse citations into tabular (csv) format. The elements of interest we sought to identify in each citation were: *given name* and *surname* for up to six authors; *year*; *title*; and *venue* (e.g. journal, newspaper, magazine, or edited book title). Parsing citations was complicated by the fact that faculty list citations in many citation formats, as well as non-standard/ad hoc formats. We sought to create a training dataset that would enable us to train a natural language processing (NLP)-based algorithm to identify these elements correctly in widely varying syllabi. To do so, we had a team of 15 undergraduate research assistants (URAs) manually color code the desired citation elements in a subset of syllabi (e.g. red for title, yellow for author, etc.). We developed a program to output a single data set from the color-coded text files, with rows corresponding to assigned readings and columns to coded components. When the NLP-based algorithm failed to reach our desired level of accuracy, we ultimately had GRAs manually code every syllabus/reading list. Running the parsing algorithm on the entire set of manually coded syllabi, we ended up with a csv file of over 90,000 citations (some of these would ultimately be eliminated as duplicates – namely, readings that appeared multiple times on one syllabus).

While parsing citations, we simultaneously coded authors' genders. Gender coding was based primarily on names, though we created a list of known scholars whose genders were miscoded or unclassifiable using automated methods. For scholars not on this "known scholars" list, genders of given names were coded using a list assembled from several sources:[4]

- A publicly available list from the 1990 US Census (we coded only names whose genders were identified with 95% certainty).[5]
- This was augmented with a list from the UK Census purchased from GenderChecker. (This listed coded all names with both male and female mentions as "unisex.")
- Finally, names that were still uncodable were queried using *genderize.io* API. To improve our processing speed, names found using the API were subsequently appended to the name/gender list.

We then spent several months cleaning the data set, a task complicated by occasional human error in the manual color coding. Names were cleaned and standardized; we found missing names, corrected typos, and standardized the appearance of each name in the dataset (for instance, "Bob Keohane" and "Robert O. Keohane" were standardized as "Robert Owen Keohane"). In the final data set, 54 names lack an author (these include works such as the Bible and the US Constitution). We manually coded the genders of over 4,000 authors with missing genders using Internet searches for publicly available bios. In the final data set, we have coded the gender of all but 76 individuals, yielding a missingness rate of 0.055%. Another 285 are coded as "corporate authors" (e.g., the World Bank). We likewise searched for missing years; in the final data set we have coded the year for all but 26 references, yielding a missingness rate of

[4] West, Jacquet, King, Correll and Bergstrom (2013) and King, Correll, Jacquet, Bergstrom and West (2015) used the Social Security Administration Baby Names database (https://www.ssa.gov/oact/babynames/#&ht=1). We chose not to use this dataset because it is limited to the top 100 names in each birth year and neglects names uncommon in English.

[5] http://www.census.gov/topics/population/genealogy/data/1990_census/1990_census_namefiles.html.

0.03%. Journal titles were also cleaned and standardized (for instance, "APSR" and "Am Pol Sci Rev" were standardized as "American Political Science Review"). In the final stage of cleaning, titles were debugged.

B. Coding of Data at Other Levels

GRADS is a multilevel dataset that contains not only data on readings and author genders, but also data at the level of documents (i.e. syllabi and reading lists); departments; and instructors. The instructor and syllabi levels are cross-nested, because some instructors submitted multiple syllabi, and some syllabi had multiple instructors. Reading lists are not associated with specific instructors. Some syllabi and reading lists were submitted anonymously or without identifying information in the header, making it impossible to code higher-level attributes.

At the syllabus/reading list level, we coded course year and the syllabus subfield(s). We are missing the year for 60 syllabi and 16 reading lists. Some syllabi corresponded to more than one subfield. At the level of departments, we coded the rank of the Ph.D. program, using 2017 data from the US News and World Reports, and then grouped programs as Ranked 1-10, 11-20, 21-30, 31-40, 41-50, and 51+. We are missing the program rank for 9 syllabi and 16 reading lists. We also coded departmental gender composition—the percentage of tenure-track faculty with primary appointments in political science who are female—based on department websites in fall 2017.

At the instructor level, survey respondents self-reported gender, age, race/ethnicity, sexual orientation, rank, age, country of origin, and country of Ph.D. We manually coded these elements, excluding race/ethnicity and sexual orientation, for the non-survey data. To code age, we assumed that the instructor was 22 at the time of completion of the Bachelor's degree. We used the names data to code instructor gender. If an instructor's name was not listed in the database, we searched for gendered pronouns ("he" or "she") in the faculty member's third-person professional description or bio.

III. Characteristics of the GRADS Dataset

In total, GRADS includes responses from 94 universities, including 82 of the top 100, according to US News & World Reports. We have at least one syllabus from all but two of the top 50 programs. The missing programs in the top 50 are University of California, Davis (rank 17) and University of California, Riverside (rank 48). As Table A5 shows, the data are highly representative of the top 50 programs, but likely to be less so of lower-ranked programs.

Table A5. Distribution of Syllabi and Sample Departments, by Program Ranking

	GRADS Data	
Program Ranking	Number Programs	Number Syllabi/ Reading Lists
1-10	10	230
11-20	11	201
21-30	10	127
31-40	12	105
41-50	5	65
51-60	9	35
61-70	6	38
71-80	7	24
81-90	9	20
91-100	4	11
Over 100	8	17
Unranked programs	3	5
Program unknown		21
TOTAL	94	899

Table A6 presents the subfield distribution. Due to the collection strategies employed, the dataset overrepresents the subfield of comparative politics.

Table A6. Subfield Distribution in GRADS

	N Syllabi	N Reading Lists
All documents	840	65
Comparative politics	268	23
International relations	162	10
American politics (including judicial/courts)	163	18
Methods	179	2
Theory	94	7
Political economy	60	1
Public policy/Public administration	15	4
Political psychology	21	0
Gender/Identity	22	4

Note: Counts by subfield sum to more than the total because some syllabi/reading lists are counted in multiple subfields.

Table A7 presents the distribution of syllabus years.

Table A7. *Distribution of Years in GRADS*

Year	Number of Syllabi
2004	1
2005	1
2006	4
2007	5
2008	7
2009	8
2010	18
2011	29
2012	33
2013	82
2014	153
2015	210
2016	269
2017	2

IV. Supplemental Analyses

A. Alternative Coding of Gender Representation

Figures A4, A5, and A6 present alternative ways to understand gender representation in syllabi. First, Figure A4 presents the same statistic as Figure 1 in the main text, but includes all subfields, including small ones. Second, Figure A5 assesses the mean percentage of *all* authors (not only first authors) assigned in a syllabus who are female. Third, Figure A6 presents results using a trichotomous variable: female only, male only, and mixed gender works.

Is author gender correlated with author order? The results in Figures A4 and A5 are essentially identical, in large part because the majority of works are single-authored. Nonetheless, women are statistically significantly more likely to appear later in the author order, rising from 18.7% of first authors to 25.7% of sixth authors.

At the same time, women are somewhat more likely to appear on publications with high numbers of authors. Women constitute between 18 and 19 percent of authors of publications with four or fewer authors, but 23% of authors of publications with five or more authors. On publications with high numbers of authors, women tend not to be first author, and they are more likely to appear in fifth or sixth place in the author order.

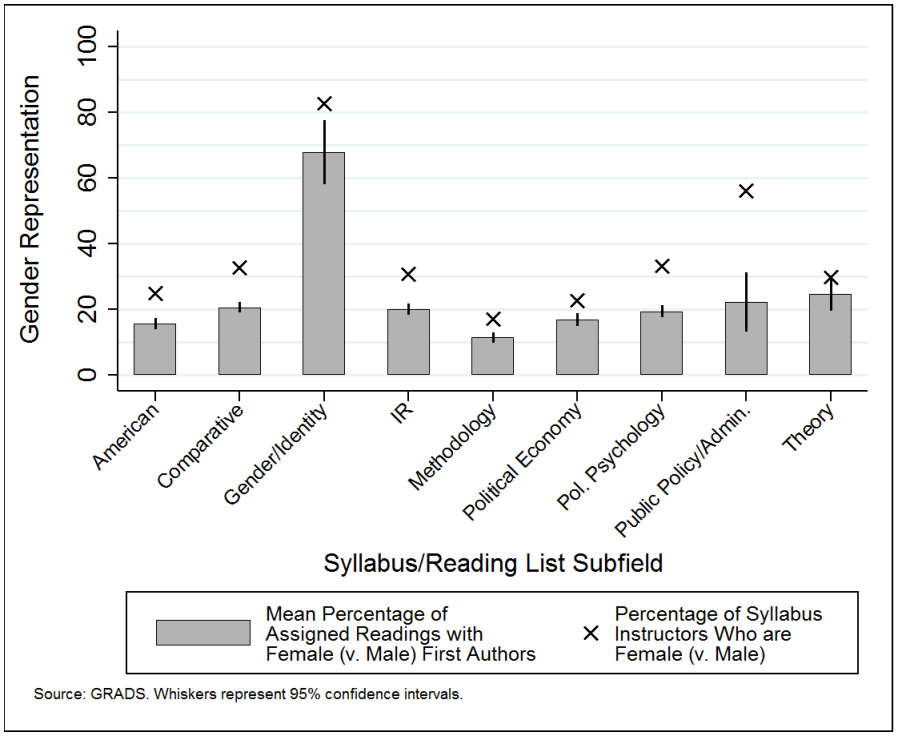


Figure A4. Gender Representation for Entire Set of Subfields

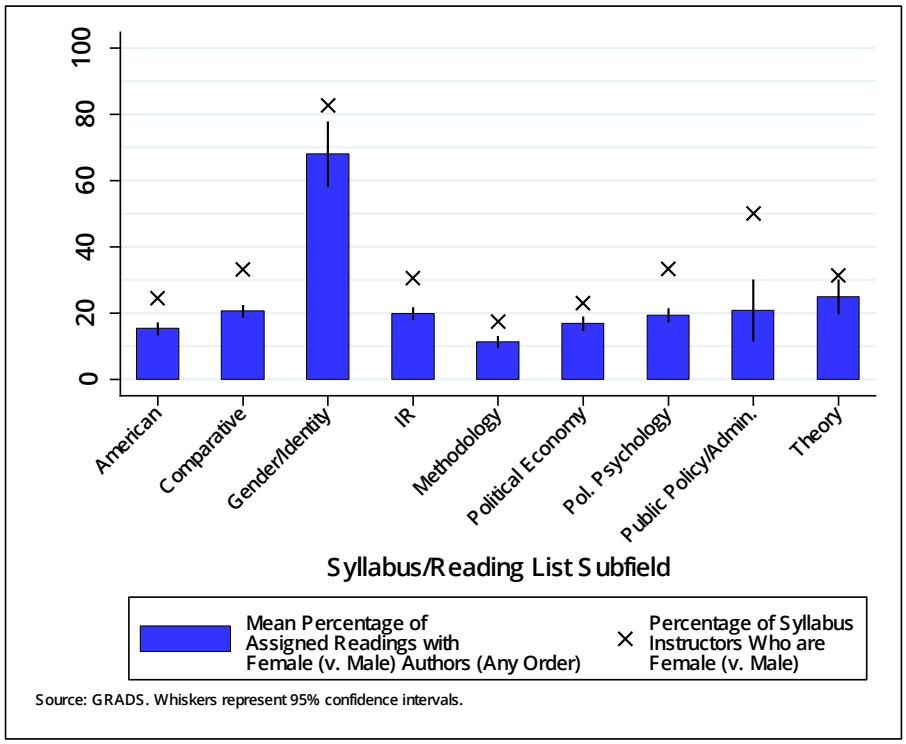


Figure A5. Gender Representation, for All Author Positions (Not Only First Authors)

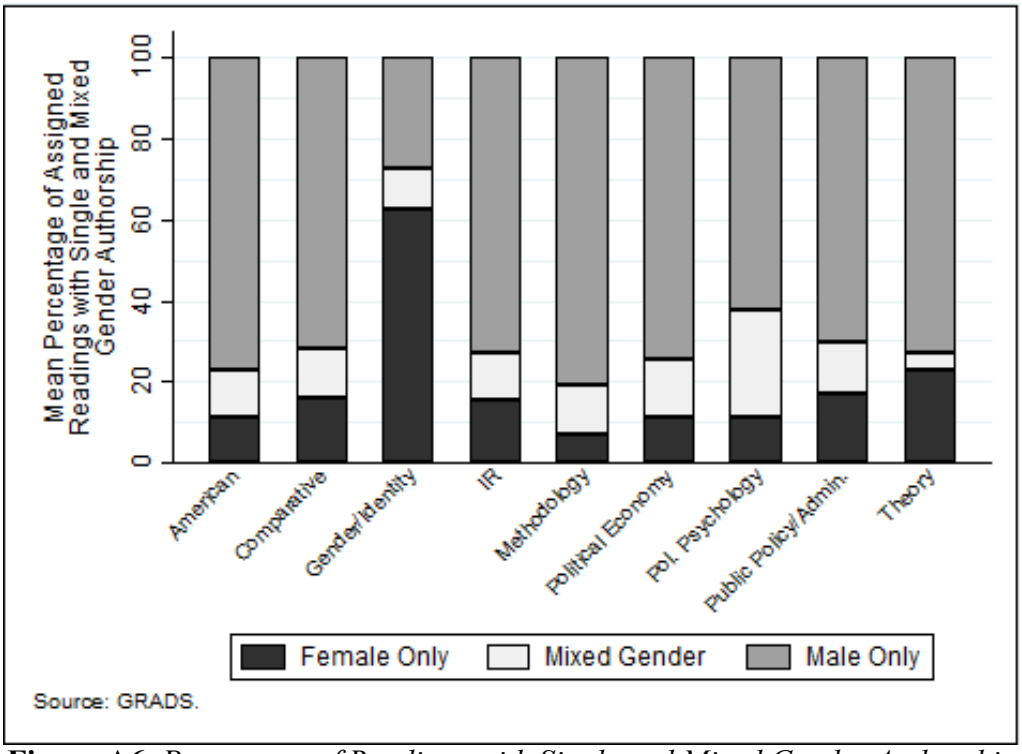
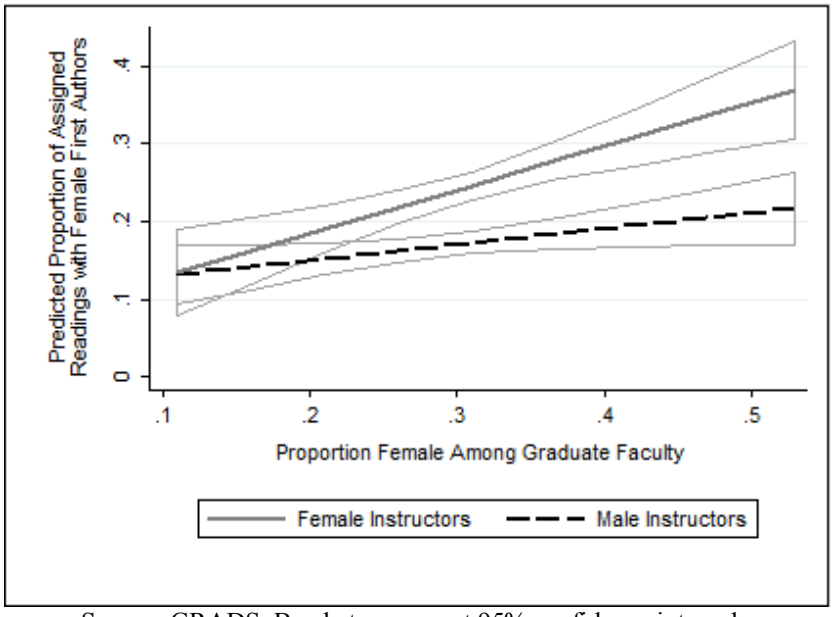


Figure A6. *Percentage of Readings with Single and Mixed Gender Authorship*

B. Supplemental Analyses of Departmental Effects

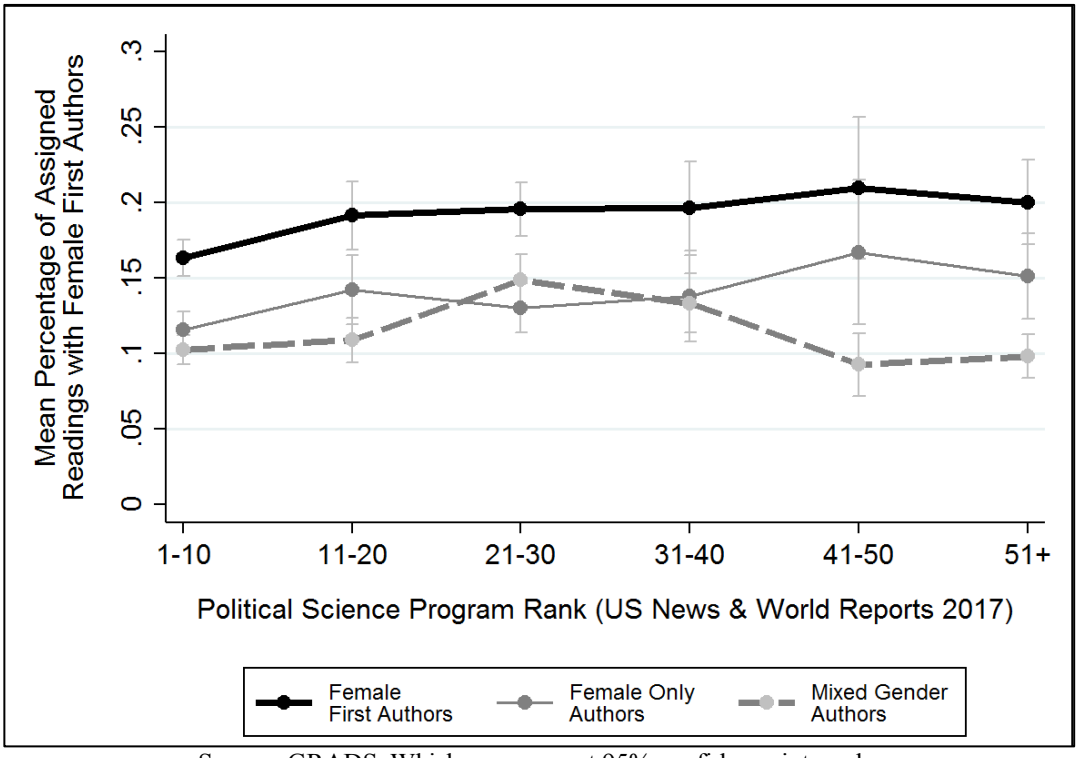
Figure A7 presents an alternative analysis of the effect of department gender composition on the rate of assigning female-authored work. Here, the dependent variable is the proportion of readings on a syllabus/reading list with a female first author.

Figure A8 depicts the relationship between Ph.D. program rank and gender representation. The bold black line depicts the proportion of readings with a female first author; the light gray line depicts the proportion of readings authored exclusively by one or more women; and the dashed line represents the proportion with mixed gender authorship teams. The figure hints at some relationship between program rank and the first two measures of gender representation. However, as seen below in the full multivariate models, the effect of program rank disappears after controlling for gender composition.



Source: GRADS. Brackets represent 95% confidence intervals.

FIGURE A7. *Proportion of Female First-Author Readings, by Department Gender Composition*



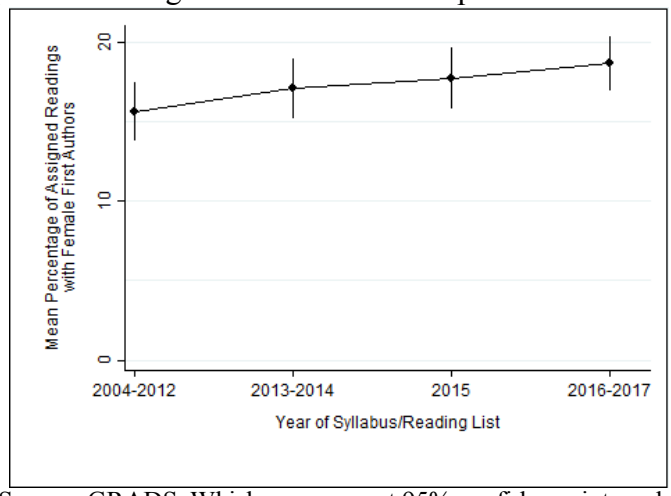
Source: GRADS. Whiskers represent 95% confidence intervals.

FIGURE A8. *Proportion of Female First-Author Readings, by Ph.D. Program Rank*

C. *Temporal Heterogeneity*

Growing awareness of inequalities may change instructors' citation practices. The largest professional associations have gradually increased gender diversity in leadership over the past three decades. Formal women's networks have proliferated. Journeys in World Politics (for female IR and comparative politics scholars) began in 2004; Visions in Methodology conferences (for female methodologists) began in 2006; and ISA's Pay It Forward pre-conference (for female IR scholars) began in 2014. Beginning in early 2016, two social media initiatives promote changing gender norms: Women Also Know Stuff and the online Gender Balance Assessment Tool (Beaulieu et al. 2016, Sumner Forthcoming).

Figure A9 shows that women's representation in readings rose gradually from 2004-2017 (the effect remains statistically significant in multivariate analysis). In addition, the proportion of readings with female authors varies dramatically by the year of the reading, with almost no female-authored readings prior to 1900. We argue that the gradual rise in readings from 2004-2017 is because as gender norms change, women's research becomes more visible, while scholars increasingly perceive citing women's work as important.



Source: GRADS. Whiskers represent 95% confidence intervals.

FIGURE A9. *Proportion of Female First-Author Readings, by Year of Syllabus/Reading List*

D. *Analysis of Journal Effects*

Gender representation may also vary across journals. While there is little evidence of gender differences in rejection rates, women appear to submit work to elite journals less frequently than men do (Breuning 2013, Breuning and Sanders 2007, Ceci et al. 2014, Teele and Thelen 2017). Indeed, women's work is underrepresented in top journals, relative to women's share of department faculty, and the gender citation gap is largest for 'elite articles' from such journals (Teele and Thelen 2017, Zigerell 2015). Table A8 presents the percentage of readings with female first authors in the ten most frequently assigned journals. Overall, readings from top journals are not any less likely to be female-authored. However, we do find great variation among the top ten journals.

TABLE A8. *Mean Percentage of Assigned Readings with Female/Male Authors, by Journal*

	Reading with:		
	Female First Author	Any Female Author	Any Male Author
American Political Science Review	13.2%	21.6%	90.7%
World Politics	14.3%	22.8%	87.7%
International Security	15.1%	20.9%	88.2%
American Journal of Political Science	15.9%	26.2%	91.8%
Journal of Conflict Resolution	17.3%	25.9%	90.0%
Journal of Politics	21.2%	37.9%	88.9%
International Organization	21.6%	29.8%	83.9%
International Studies Quarterly	25.7%	33.5%	81.0%
Comparative Political Studies	28.0%	34.2%	81.0%
Annual Review of Political Science	35.3%	40.3%	71.6%
Other Journals/Venues	20.5%	28.3%	85.1%
Books	17.6%	21.7%	83.9%

Note: Journals represent the ten journals most frequently assigned in syllabi; data are weighted to represent syllabi equally. Source: GRADS.

E. Multivariate Analysis and Analysis by Gender

To facilitate comparison of effects, all institutional independent variables are standardized to run 0–1. In models with both male and female instructors, the benchmark percentage of a subfield that is female and department gender composition are the two strongest determinants of gender representation in readings. In separate models, we confirm that time affects only men's citation practices, while women's citation practices are more sensitive than men's to department gender composition. We also find gender-specific subfield effects; men teaching comparative politics and women teaching theory courses are unexpectedly likely to assign female-authored work. The negative effect of Gender/Identity courses results from the extremely high benchmark. Instructors teaching such courses assign lower proportions of female-authored work than expected based on the proportion of instructors who are female, perhaps aiming for gender diversity.

TABLE A9. *Determinants of Proportion of Readings with Female First Authors*

		Syllabi only		
	Full sample	All Instructors	Men	Women
Comp. Exam Reading List	-0.222**			
	(0.109)			
Female Instructor (Syllabi Only)		0.450**		
		(0.057)		
Syllabus Year	0.177**	0.138*	0.230**	-0.030
	(0.075)	*(0.077)*	*(0.084)*	*(0.157)*
Rank of Department	-0.007	0.010	0.008	-0.031
	(0.081)	*(0.080)*	*(0.086)*	*(0.160)*
Department Gender Composition	0.752**	0.618**	0.415**	0.853**
	(0.180)	*(0.177)*	*(0.179)*	*(0.381)*
Benchmark: % Female in Subfield	9.295**	8.376*	8.641**	7.605
	(4.497)	*(4.618)*	*(3.432)*	*(6.960)*
Comparative Politics	0.139	0.108	0.156*	0.025
	(0.084)	*(0.084)*	*(0.081)*	*(0.137)*
International Relations	0.261**	0.245**	0.268**	0.222*
	(0.074)	*(0.073)*	*(0.083)*	*(0.119)*
American Politics	0.035	0.025	0.004	0.086
	(0.106)	*(0.103)*	*(0.119)*	*(0.158)*
Methods	-0.025	-0.023	-0.015	-0.050
	(0.159)	*(0.159)*	*(0.151)*	*(0.271)*
Political Theory	0.288**	0.287**	0.171	0.434*
	(0.120)	*(0.118)*	*(0.131)*	*(0.235)*
Political Economy	-0.026	-0.021	0.072	-0.279
	(0.093)	*(0.095)*	*(0.086)*	*(0.194)*
Gender/Identity	-6.982	-6.314	-6.792**	-5.557
	(4.342)	*(4.462)*	*(3.343)*	*(6.701)*
Constant	-2.436**	-2.445**	-2.429**	-1.930**
	(0.228)	*(0.223)*	*(0.204)*	*(0.376)*
Number of Observations	810	771	562	209

Note: Results from fractional logistic regression models with robust standard errors. The syllabus is the unit of analysis. Independent variables standardized to run from 0 to 1. Coefficients are statistically significant at * p < .10 or ** p < .05.

TABLE A10. *Determinants of Proportion of Readings with Female-Only Author(s)*

		Syllabi only		
	Full sample	All faculty	Men	Women
Comp. Exam Reading List	-0.056			
	(0.148)			
Female Instructor (Syllabi Only)		0.574**		
		(0.071)		
Syllabus Year	0.152	0.104	0.311**	-0.196
	(0.100)	*(0.104)*	*(0.115)*	*(0.190)*
Rank of Department	0.008	0.032	-0.017	0.053
	(0.107)	*(0.107)*	*(0.115)*	*(0.199)*
Department Gender Composition	0.935**	0.756**	0.635**	0.662
	(0.234)	*(0.226)*	*(0.234)*	*(0.472)*
Benchmark: % Female in Subfield	8.345	6.562	8.680	6.027
	(7.893)	*(8.162)*	*(6.616)*	*(11.794)*
Comparative Politics	0.275**	0.244*	0.261**	0.174
	(0.132)	*(0.136)*	*(0.126)*	*(0.198)*
International Relations	0.386**	0.380**	0.377**	0.390**
	(0.101)	*(0.099)*	*(0.117)*	*(0.146)*
American Politics	-0.009	-0.029	-0.118	0.171
	(0.146)	*(0.140)*	*(0.168)*	*(0.207)*
Methods	-0.230	-0.245	-0.259	-0.111
	(0.262)	*(0.264)*	*(0.264)*	*(0.397)*
Political Theory	0.625**	0.626**	0.486**	0.784**
	(0.135)	*(0.134)*	*(0.151)*	*(0.251)*
Political Economy	-0.137	-0.136	-0.070	-0.282
	(0.137)	*(0.139)*	*(0.130)*	*(0.247)*
Gender/Identity	-5.898	-4.451	-6.765	-3.845
	(7.606)	*(7.869)*	*(6.400)*	*(11.352)*
Constant	-3.009**	-3.010**	-3.104**	-2.245**
	(0.368)	*(0.364)*	*(0.329)*	*(0.587)*
Number of Observations	798	759	549	210

Note: Results from fractional logistic regression models with robust standard errors. The syllabus is the unit of analysis. Coefficients are statistically significant at * $p < .10$ or ** $p < .05$.

TABLE A10. *Determinants of Proportion of Readings with Mixed Gender Authors*

		Syllabi only		
	Full sample	All faculty	Men	Women
Comp. Exam Reading List	-0.631**			
	(0.109)			
Female Instructor (Syllabi Only)		0.055		
		(0.070)		
Syllabus Year	0.158	0.153	0.010	0.490**
	(0.100)	*(0.103)*	*(0.124)*	*(0.188)*
Rank of Department	-0.094	-0.097	0.030	-0.388**
	(0.089)	*(0.093)*	*(0.108)*	*(0.163)*
Department Gender Composition	0.120	0.082	0.078	0.662*
	(0.223)	*(0.229)*	*(0.266)*	*(0.346)*
Benchmark: % Female in Subfield	13.787**	15.483**	16.320**	13.401*
	(4.513)	*(4.691)*	*(5.561)*	*(7.902)*
Comparative Politics	-0.204*	-0.243**	-0.212*	-0.260
	(0.108)	*(0.111)*	*(0.125)*	*(0.214)*
International Relations	-0.086	-0.109	-0.100	-0.099
	(0.103)	*(0.104)*	*(0.124)*	*(0.185)*
American Politics	0.171	0.187	0.287*	-0.023
	(0.132)	*(0.133)*	*(0.159)*	*(0.218)*
Methods	0.393**	0.441**	0.558**	0.228
	(0.186)	*(0.188)*	*(0.210)*	*(0.373)*
Political Theory	-1.137**	-1.145**	-0.954**	-1.497**
	(0.214)	*(0.217)*	*(0.253)*	*(0.324)*
Political Economy	0.277**	0.299**	0.439**	0.012
	(0.094)	*(0.096)*	*(0.114)*	*(0.153)*
Gender/Identity	-12.966**	-14.623**	-14.499**	-13.000*
	(4.355)	*(4.521)*	*(5.406)*	*(7.591)*
Constant	-2.578**	-2.625**	-2.707**	-2.716**
	(0.256)	*(0.258)*	*(0.305)*	*(0.453)*
Number of Observations	798	759	549	210

Note: Results from fractional logistic regression models with robust standard errors. The syllabus is the unit of analysis. Coefficients are statistically significant at * $p < .10$ or ** $p < .05$.

References

Beaulieu, Emily, et al. 2017. "Women Also Know Stuff: Meta-Level Mentoring to Battle Gender Bias in Political Science." *PS: Political Science & Politics* 50 (3): 779–83.

Breuning, Marijke. 2013. Editors and the Gender Gap. *The Monkey Cage* (blog), *The Washington Post*, October 3. https://www.washingtonpost.com/news/monkey-cage/wp/2013/10/03/editors-and-the-gender-gap/.

Breuning, Marijke and Kathryn Sanders. 2007. Gender and Journal Authorship in Eight Prestigious Political Science Journals. *PS: Political Science & Politics* null (02): 347–51.

Ceci, Stephen J., Donna K. Ginther, Shulamit Kahn, and Wendy M. Williams. 2014. Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest* 15 (3): 75–141.

Lu, Hsi-Peng, Kuo-Lun Hsiao. 2009. Gender differences in reasons for frequent blog posting. *Online Information Review* 33 (1): 135-156.

Sumner, Jane Lawrence. 2018. The GBAT: A Web-Based Tool for Estimating Gender Balance in Syllabi and Bibliographies. *PS: Political Science & Politics*.

Teele, Dawn and Thelen, Kathleen. 2017. Gender in the Journals: Publication Patterns in Political Science. *PS: Political Science & Politics*, *50*(2): 433-447.

Zigerell, L. J. 2015. Is the Gender Citation Gap in International Relations Driven by Elite Papers? *Research & Politics* 2 (2): 1-7.