Bagues, Manuel; Campa, Pamela; Etingin-Frati, Giulian

Working Paper

Gender Differences in Cooperation in the U.S. Congress? An Extension of Gagliarducci and Paserman (2022)

I4R Discussion Paper Series, No. 75

Provided in Cooperation with:
The Institute for Replication (I4R)

INSTITUTE for REPLICATION

No. 75
I4R DISCUSSION PAPER SERIES

Gender Differences in Cooperation in the U.S. Congress? An Extension of Gagliarducci and Paserman (2022)

Manuel Bagues

Pamela Campa

Giulian Etingin-Frati

October 2023

INSTITUTE for REPLICATION

I4R Discussion Paper Series

I4R DP No. 75

Gender Differences in Cooperation in the U.S. Congress? An Extension of Gagliarducci and Paserman (2022)

Manuel Bagues[1], Pamela Campa[2], Giulian Etingin-Frati[3]

[1]University of Warwick, Coventry/Great Britain, Centre for Economic Policy Research, London/Great Britain, IZA Bonn/Germany, and J-Pal, Cambridge/USA
[2]SITE – Stockholm School of Economics, Stockholm/Sweden, Mistra Center for Sustainable Markets (Misum), Stockholm/Sweden, Centre for Economic Policy Research, London/Great Britain, J-Pal Europe, Paris/France
[3]Stockholm School of Economics, Stockholm/Sweden

OCTOBER 2023

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

Gender Differences in Cooperation in the U.S. Congress? An Extension of Gagliarducci and Paserman (2022)

Manuel Bagues, Pamela Campa & Giulian Etingin-Frati *

August 18, 2023

Abstract

Gagliarducci and Paserman (2022) study gender differences in cooperative behavior among politicians using information from the U.S. House of Representatives between 1988 and 2010 on (i) the number of co-sponsors on bills and (ii) the share of co-sponsors from the rival party. Through different empirical strategies, they show that women-sponsored bills tend to have more co-sponsors, but the gap is only statistically significant among Republicans. Moreover, Republican women recruit a significantly larger share of co-sponsors from the rival party than Republican men, whereas the opposite is true among Democrats. GP argue that the observed pattern is consistent with a commonality of interest driving cooperation, rather than gender *per se*, since during this period Republican women were ideologically closer to the rival party than their male colleagues, while female Democrats were further away.

We examine the robustness of these findings to (i) the correction of some errors in two control variables of the dataset used by GP and (ii) clustering the standard errors at the individual level, instead of individual-term. These changes have a relatively minor impact on results: most coefficients are still statistically significant and the main conclusions from the analysis are confirmed. Furthermore, we extend the analysis to the 2011-2020 period. The analysis of gender differences in bipartisan cooperation confirms GP's hypothesis that ideological distance plays an important role. However, results are slightly different when we analyze overall cooperation. The gender gap in favor of women is larger in magnitude than in GP and it is statistically significant in several specifications, providing support for the hypothesis that gender also matters for cooperation.

*Bagues: University of Warwick, CEPR, IZA and J-Pal. Campa: SITE, Misum, CEPR and J-Pal. email: pamela.campa@hhs.se. Etingi-Frati: Stockholm School of Economics. We thank Cecilia Smitt Meyer for excellent research assistance.

1 Introduction

Gagliarducci and Paserman (2022), henceforth GP, study gender differences in cooperative behavior in the U.S. House of Representatives between 1988 and 2010. To measure cooperativeness, they compare the number of co-sponsors that women and men recruit on bills that they sponsor as well as the share of these co-sponsorships from the opposite party. GP describe their main results as follows: "We find that among Democrats there is no significant gender gap in the number of co-sponsors recruited, but women-sponsored bills tend to have fewer co-sponsors from the opposite party. On the other hand, we find robust evidence that Republican women recruit more co-sponsors and attract more bipartisan support on the bills that they sponsor." They conclude that this pattern indicates that cooperation is mostly driven by a commonality of interest, rather than gender *per se*, reflecting that during this period Republican female representatives were ideologically closer to Democrats than their male colleagues, whereas Democratic women were ideologically further away from Republicans.[1]

The original study was successfully reproduced by the Institute for Replication's collaborators team and we were also able to successfully reproduce GP's Table 5, which is the target of our reanalysis and is reproduced for reference in Table A1. We checked the corresponding code and have not found any mistakes. In this table GP consider two dependent variables: *number of co-sponsors* and *percent co-sponsors of opposite party*. The analysis is run at the bill level and the coefficient of interest is a dummy for the bill's sponsor being a woman. The authors report results for the overall sample as well as by party. For each analysis they consider five different empirical strategies (e.g. linear model with identification based on observables, regression discontinuity design etc.).

We investigate whether the results reported in this table are robust to (1) correcting some errors in GP's raw data that affect two of the control variables, (2) changing how standard errors are clustered, and (3) extending the analysis one decade more, using data between 2011 and 2020.

[1] GP document this pattern by comparing districts where Republican and Democratic women and men are respectively elected, based on their predicted share of Republican votes. Both among Democrats and Republicans women tend to be elected in less conservative districts than their male colleagues (see Figure 1, panel a).

The data errors concern the control variables *population density* and *median household income*. As we explain in more detail below, the values in GP's dataset for Congresses 108^{th} through 111^{th} are highly discordant from official census statistics. Fortunately, when we re-run GP's analysis with the amended data, point estimates are generally similar to the original ones and the statistical significance is generally unchanged.

In their original analysis, GP cluster their standard errors at the individual-term level. In our re-analysis, we cluster the standard errors at the individual level. Clustering the standard errors by individual rather than individual-term increases slightly the standard errors and, out of the ten coefficients that were statistically significant in GP, three lose the 5% significance level, but the broad conclusions are unchanged.[2]

By extending the analysis to years 2011-2020, we can test the robustness of GP's hypotheses in a context that differs in at least two relevant aspects. The share of women in the House of Representatives is substantially larger, around 21% compared to 13% in GPs dataset, reflecting the increasing presence of women over time in Congress. Moreover, within-party gender differences in ideology seem to have changed compared to previous decades. While Democratic female representatives are still less conservative that Democratic men, among Republicans women became ideologically similar to their male colleagues.[3] Consistent with GP's hypothesis that gender differences in cooperation across parties are driven mainly by ideological distance, we observe that bills sponsored by female Democrats are less likely to have Republican co-sponsors, but we do not observe anymore any gender differences in bipartisan cooperative behavior among Republicans. Our results are slightly different from GP when we analyze the total number of co-sponsors. While GP only found a significant gender gap in favor of women among Republicans (in three out of five specifications), during the last decade we observe that bills from both Republican and Democratic women attract more sponsors than bills from their male colleagues. The effect is larger in magni-

[2] In the remainder of the text, we call "statistically significant" those estimates that are at least 5% statistically significant.

[3] As in GP, we proxy representatives' ideology using information on the ideological leaning of voters in their constituency in the presidential elections. During this decade, (a) Republican women and men are elected in ideologically similar districts, and (b) Democratic women are less likely than Democratic men to be elected in conservative districts.

tude and it is statistically significant at the 1% level in three out of five specifications, compared to only one coefficient significant at the 5% level in GP's overall sample.

While GP conclude that the observed gender differences in cooperation are driven by bipartisan cooperation rather than gender *per se*, our analysis suggests that gender might be also playing a role.

2 Replication

In this section we discuss our replication in greater detail. For transparency, we start explaining how we planned the replication. Initially we decided to assess the robustness replicability of the study by changing how the standard errors were clustered and its direct replicability by adding ten more years to the study sample. We decided to conduct these tests after we read the paper but before we had looked at the codes and data provided by the authors in the replication package. Later on, when we inspected summary statistics from GP data, we noticed some remarkable discrepancies across different years for two control variables and a closer examination of the original data helped to uncover some apparent mistakes. Therefore, in our replication we also verify the robustness of results using the amended dataset.

In our replication exercise we consider all the estimations reported originally by GP in Table 5, which we reproduce in Appendix Table A1. GP consider in this table two outcomes, *number of co-sponsors* (Panel A) and *percent co-sponsors of opposite party* (Panel B), for all the bills together (top row of each panel) and separately by party affiliation of the sponsor (middle and bottom rows). They report estimates from five different specifications. In column (1) they consider a simple OLS regression with controls for an assortment of sponsor, bill and district characteristics. Column (2) presents an RD design, which controls for the margin of victory of female candidates in mixed-gender races using the optimal bandwidth. In column (3) they use a similar specification but, in order to account for potential selection bias when comparing closely elected politicians, they weigh observations by an inverse propensity score based on district characteristics (IPW).

Finally, they run specifications on the full sample (i.e. not only mixed-gender electoral races), and again include IPW. District characteristics and the margin of victory are used in the propensity score in column (4) and sponsor characteristics are added in column (5).

Below we first describe how we collected the data that we use in this analysis and then we show results for each of the tests considered, separately and pooled together.

2.1 Data collection

GP's main data set is based on the Library of Congress' data information system, THOMAS, from which the authors retrieve information on *public* bills submitted from the 101^{st} (elected in 1988) to the 111^{th}(elected in 2008) Congress. They merge this data with additional Congress-member individual characteristics, election statistics, and demographic and economic information on congressional districts[4]. We show summary statistics from GP dataset in Table 1, Panel A.

A preliminary inspection of GP's dataset across years revealed some implausible values in the data for two control variables: *population density* and *household income*. We provide some examples in Table A2, where we show a comparison of data from the 107^{th} and 108^{th} Congresses (from GP's dataset) for some districts, as well as the values that we retrieved from the U.S. census, when available. For instance, in the database used by GP, the area of several States becomes more than one million times larger between the 107^{th} and 108^{th} Congresses and the median household income decreases by around 50%.[5] Further analysis suggested that GP's data for these two variables for Congresses 108^{th} (elected in 2002) to 111^{th} (elected in 2008) are not consistent with official census statistics.

To address this problem, we collected new data for population density and household income

[4]See GP's Online Appendix A for information on data sources. They also use additional bill-level information from Adler and Wilkerson's Congressional Bills Project (http://www.congressionalbills.org) and Fowler, Waugh and Sohn's Cosponsorship Network Data (http://jhfowler.ucsd.edu/cosponsorship.htm).

[5]For the income variable, one hypothesis is that, while generally using median household income, for the years where we find inconsistencies the values are instead for per capita income. However we did not investigate this issue further.

for the years with implausible values (e.g. Congresses 108^{th} to 111^{th}) from the U.S. census website.[6] We downloaded data for area, population and median household income at district level choosing the available series that was closest in time to our period of interest.[7] While there are some missing values and there might be some errors due to redistricting, we expect this dataset to be more accurate than the one used by GP's (see Table 1, Panel B).

We also collected data for many of the other variables used in GP for the years 2011 to 2020, in order to test the robustness of their results to extending the period of analysis. We relied mostly on GP's data sources, with some modifications that we detail below. We downloaded data on bills sponsors and co-sponsors from the Library of Congress' data information system, THOMAS[8]. As in GP, we focus on House bills that are classified as *public* (i.e. they do not cover private issues). The Library of Congress' data does not identify *public* and *private* bills separately, therefore we retrieve this information from Adler and Wilkerson's Congressional Bills Project (http://www.congressionalbills.org), which also contains information on the sponsor's gender and the bills "minor" topics (the latter is used as control variable in GP).[9]

We recover some information on the characteristics of each bill's sponsor from the Biographical Directory of the United States Congress available online at the Library of Congress[10]. GP digitize information from this source on age, gender, tenure in congress, whether the member is a rookie, committee membership, occupation, whether the sponsor has an Ivy League college degree, and whether the sponsor was born in the state of election. We have forgone the digitization and relied only on information available in the downloadable files from the website, which allow identifying each member's age, tenure, committee membership and rookie status. Relying only on this subset of control variables is unlikely to affect our estimates substantially, for a number of reasons.

[6] Accessible at data.census.gov

[7] This resulted in considering data for area as associated to the 108^{th} Congress, population in 2000 using the boundaries of the 110^{th} Congress, and household income in 1999 using the boundaries of the 106^{th} Congress.

[8] Accessible at http://thomas.loc.gov

[9] Since information on whether the bill is on a private or public issue is missing for slightly over one third of the bills, we use the bill *major* topic to infer whether the bill is private. Specifically, in the sample with non-missing information, the *major* topic is coded as "99" for 98% of the *private* bills. Moreover, 95% of the bills with topic "99" are private. We thus classify as *private* those bills with missing information whose topic is 99, and as non-private otherwise. We found no record in GP of a similar data limitation problem.

[10] Accessible at http://bioguide.congress.gov

First, the controls for sponsor characteristics do not appear to affect the main estimates in the original paper: for example, in GP's Table 5, which is our focus, the results are virtually identical including or excluding sponsor-characteristics for the propensity-score matching (columns 4 and 5). Additionally, in Appendix Table A3 we reproduce GP's Table 5 using their dataset but relying only on those control variables for which we collect information in our dataset. With the exception of one coefficient, the estimates are remarkably similar to those in GP and the overall conclusions of the analysis are unchanged.

For election data we rely on David Leip's Atlas of U.S. Presidential Elections, which reports district-level information on votes by party.[11] Finally, we collect district-level data from the U.S. Census Bureau (USCB) on economic, social, and demographic characteristics for the years 2011-2021.[12] The data primarily come from the annualised estimates of the American Community Survey (ACS), a monthly survey conducted by the USCB that complements the decennial census.[13] As in GP, we collected information on the district population density (per square mile), share of Black residents, share of residents over 65, share of foreign-born residents, share of urban residents, and median household income (in nominal terms).

In Table 1 Panel C we show summary statistics for Congresses 112^{th} to 116^{th}. While generally values for bills and sponsor characteristics align with those from GP's data (see Panel A), there are some differences in the district-level information. This is mostly due to the different time-period that our respective datasets span, but, as discussed earlier, we also found and addressed some errors in GP's dataset.

[11] GP report using electoral data from the Office of the Clerk of the House of Representatives (http://clerk.house.gov). Since we have not been able to locate this data on the posted webpage, we decided to purchase David Leip's data instead.

[12] Specifically, the data are from USCB series CP02, CP03, CP05, and P2. These are accessible at data.census.gov.

[13] For the share of the district population that lives in urban areas, we relied on the 2010 decennial census only, since this variable is not available in the ACS. Therefore, the data used reflects the proportion of residents in a district living in urban areas in 2010. For the 111th, 113th, 115th, and 116th Congresses, we use the USCB provided crosswalks to adjust the data for redistricting. For the 112th and 114th Congresses, where the crosswalk is not provided, we use the data for the last available Congress, the 111th and 113th, respectively.

2.2 Replication using the amended dataset

In Table 2, we replicate GP's analysis using the corrected dataset. The resulting estimates are largely similar to those in GP and can be summarized as follows:

1. *Number of co-sponsors*: The finding that Republican women attract more co-sponsors on their bills is generally confirmed, with coefficients largely similar in magnitude to those in GP. Unlike GP, we also find that female Democrats attract significantly more co-sponsors in the OLS specification (column 1), but this gap is not statistically significant in alternative specifications.[14]

2. *Co-sponsors of the opposite party, Democrats*: The estimated gender difference for Democrats is statistically significant in four out of five specifications, compared to three out of five in GP. The magnitude of the coefficients is largely unchanged.

3. *Co-sponsors of the opposite party, Republicans*: The estimated coefficients and standard errors for the outcome *percent of co-sponsors of opposite party* for Republicans are virtually unchanged.

2.3 Clustering of standard errors

We cluster standard errors at individual level (rather than individual-term) to account for non-independence over time for a given Congress member. In other words, we hypothesize that errors could be correlated across legislatures for the same sponsor, especially since many Congress members are re-elected multiple times and might thus forge long-lasting networks and relationships, which might in turn affect their cooperativeness. In the data, we identify the same sponsor across different Congresses by using information on name, gender, party, and state of election.[15] Our findings are reported in Table 3, and can be summarized as follows:

[14] The gender gap becomes negative in column 5.

[15] Our strategy implies that we fail to connect over time those members who change party affiliation, a however highly infrequent event. Specifically, out of 1,117 members-by-party in GP's dataset, we identify twenty potential party changes. Some members appear to change party affiliation within the same Congress.

1. *Number of co-sponsors*: Clustering the standard errors by individual rather than individual-term does not alter the overall conclusions of the analysis on the *number of co-sponsors* (Panel A), although most estimates become slightly less precise. As in GP's paper, there is no statistically significant gender difference in the number of co-sponsors that Democrats attract. The gender difference among Republicans, with women attracting more co-sponsors, is also generally confirmed, but two coefficients that were 1% statistically significant become 5% statistically significant (columns 1 and 5), and one coefficient that was 5% statistically significant becomes insignificant (column 4).

2. *Co-sponsors of the opposite party, Democrats*: Turning to the second outcome, *percent co-sponsors of opposite party*, three out of five specifications return a statistically significant gender difference among Democrats, with women attracting a lower share of co-sponsors from the opposite party. In GP the number of statistically significant coefficients was four out of five.

3. *Co-sponsors of the opposite party, Republicans*: Finally, our estimates confirm that Republican women are significantly more likely to attract co-sponsors from the opposite party than Republican men, although we estimate slightly larger standard errors. Three coefficients that are 1% significant in GP become 5% significant with our choice of clustering.

2.4 Replicating the analysis in period 2011-2020

The dataset in GP spans Congresses 101^{st} (elected in 1988) to 111^{th} (elected in 2008). We extend the analysis to bills presented in more recently-elected Congresses, in office in years 2011 to 2020.[16] This period differs in at least two ways from the one considered in GP. First, the number of female representatives is substantially larger. As documented in Figure 2, the positive trend in the share of women in the U.S. Congress that started in the 1990s has been continuing through the 2000s and 2010s. During the last decade around 21% of House Representatives were women, compared

[16] One of the data sources that we use, the Congressional Bills Project (http://www.congressionalbills.org/) only includes data through May 2020; since we rely on this dataset to control for the bill topic, which is an important control variable, we do not analyse bills that were passed after May 2020.

to 13% in the previous two decades considered by GP. Second, we also find evidence of changes in ideology along party and gender lines. Among Republicans, while in the decades studied by GP women appear to be more progressive than men, there is no evidence of substantial gender differences in the decade that we study. Among Democrats, similarly to GP, we find that women are more likely to be elected in districts with a lower predicted Republican share than men, whereas men are more likely to be elected in more conservative districts, suggesting that Democratic women are more progressive than Democratic men.

Specifically, to proxy representatives' ideology, we follow GP and plot the empirical cumulative distribution function (CDF) of the predicted Republican vote share in districts where men and women respectively won an election, separately by party.[17] We show the four CDF's in Figure 1 (panel b), where we also report the respective CDF's for the previous decades based on GP's data (panel a). We use the Kolmogorov-Smirnov test of equality of distributions to test the hypothesis that there are no gender differences in ideology within the same party. In years 1989-2010, we observe significant gender differences in ideology both for Republicans and Democrats. Instead, when we use the more recent data that we collected (2011-2020), gender differences are only significant among Democrats.

According to GP's hypothesis that gender differences in cooperativeness across parties are driven by commonality of interest, in the decade 2011 to 2020 we should observe that (a) Democratic women attract a lower fraction of co-sponsors from the opposite party and (b) there is no gender difference in bipartisan co-sponsorship among Republicans.

Our findings broadly confirm GP's conclusion, with an important amendment. Our estimates lend support to the claim that commonality of interest drives gender differences in bipartisan cooperation, but we also conclude that women might be overall more cooperative than men. We reach these conclusions based on the following results:

[17] As in GP, we predict the Republican vote share based on OLS regression of the actual Republican vote share on district characteristics, including three region dummies, percentages of black residents, percentage of urban residents, percentage of foreign-born residents, percentage of over-65 residents, log median income and log population density.

1. *Number of co-sponsors*: We find stronger evidence that women from both parties attract more co-sponsors overall on their bills. As in GP, we find that Republican women attract significantly more co-sponsors in three out of five specifications. Moreover, we also estimate three out of five significant coefficients for the overall sample, with women appearing to attract more co-sponsors, whereas in GP none of these coefficients was statistically significant.

2. *Co-sponsors of the opposite party, Democrats*: As in GP, we find that, among Democrats, female-sponsored bills attract a lower percent of co-sponsors from the opposite party as compared to male-sponsored bills. Similarly to GP, the negative point estimate is significant at the 1% level in three out of five specifications. This finding is in line with the evidence that female Democrats are more progressive than their male colleagues, both during GP's and our sample period (see Figure 1).

3. *Co-sponsors of the opposite party, Republicans*: In contrast to GP, we find no gender difference in the likelihood of attracting bipartisan support among Republicans, consistent with ideological differences between female and male Republicans being attenuated in the decade that we study as compared to the decades considered in GP (see Figure 1).

2.5 Clustering by sponsor, extending the time period and correcting the data errors

Finally, we assess the replicability of the coefficients shown in GP's Table 5 when we consider all the changes described in the previous subsections at once. More specifically, we pool together all the Congresses spanned by GP and our datasets (101^{st} to 116^{th}), we correct the data errors and cluster the standard errors by individual. The results are shown in Table 5 and can be summarized as follows:

1. *Number of co-sponsors*: Consistently with GP's findings, there is a significant gender difference in the number of co-sponsors attracted by Republicans, with women being more

successful in cooperating than men. For Democrats, only the OLS specification returns a significant gender difference; however, all the other specifications return positive coefficients that are much closer in magnitude to the OLS estimate than they are in GP. Overall, when we pool the two parties together the weight of the evidence suggests that women tend to attract more co-sponsors than their male colleagues, a 6 to 12% difference (relative to the sample mean) that is statistically significant in two out of five specifications.

2. *Co-sponsors of the opposite party*: When we consider the outcome *co-sponsors of opposite party*, GP's qualitative conclusions are confirmed. The gender difference favors men in the case of Democrats, and women in the case of Republicans. However, given the hypothesis that gender differences in bipartisan cooperation should be driven by ideological differences, and in light of the evidence that ideological differences evolve over time, this outcome is arguably best analysed separately for different decades, as we have done in the analysis above.

Conclusion

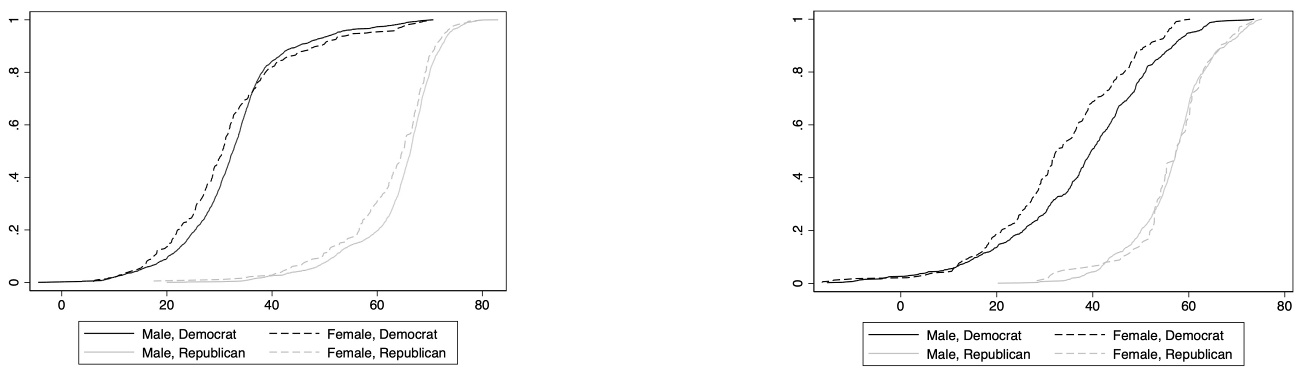
GP study the potential existence of gender differences in cooperation among politicians using data from the House of Representatives between 1989 and 2010. They conclude that the evidence that they produce is consistent mainly with a commonality of interest driving cooperation, rather than gender per se.

Our re-analysis of GP's main results using data for the same period finds that the overall conclusions of their analysis are mostly unaffected by a different choice of clustering and by correcting data errors.

The pattern is slightly different when we replicate GP's analysis using more recent data. We find clear support for GP's hypothesis that commonality of interests explains cooperation with members of the rival party. However, our analysis of gender differences in the overall number of co-sponsors suggests that women from both parties attract more co-sponsors than men, indicating that women might indeed be more cooperative than men.

Figures

Figure 1: Predicted republican share (CDF)

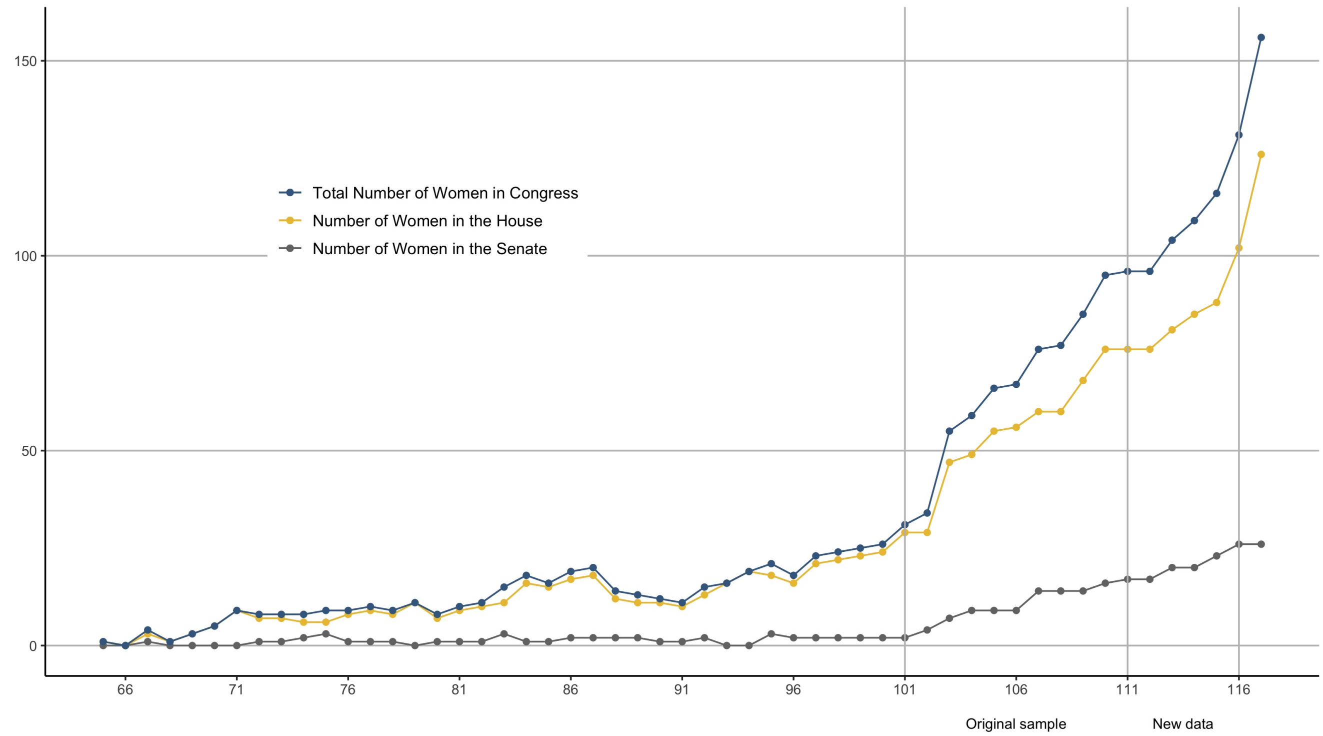


(a) Congresses 101^{st} - 111^{st}

(b) Congresses 112^{th} - 116^{st}

Note: Figure in left panel is reproduced from GP (Congress 101^{st} to 111^{th}). Figure in right panel is based on the same methodology, but more recent data (Congress 112^{th} to 116^{th}). As in GP, we show the cumulative distribution function (CDF) of the predicted Republican vote share in districts represented by male and female representatives, by party. The unit of observation is an individual Congress member. The vote share is predicted using an OLS regression of actual Republican vote shares on district characteristics used elsewhere in the analysis: region dummies, share of Black residents, share of foreign-born residents, share of residents over 65, log median household income, and log population density. A Kolmogorov-Smirnov test for equality of distributions by gender for Congress 101^{st} to 111^{th} (left panel) rejects the null of equal distributions for both Democrats (p-value=0.00) and Republicans (p-value=0.01), indicating within-party gender differences in ideology. Instead, when we consider the more recent years (right panel), the Kolmogorov-Smirnov test rejects the null of equal distributions for Democrats (p-value=0.00) but not for Republicans (p-value=0.65).

Figure 2: Number of women in Congress, 1917-2021



Data is from the Congressional Research Service publication "Women in Congress: Statistics and Brief Overview" (2022), available at: https://crsreports.congress.gov/product/pdf/R/R43244. For all Congresses other than the 117^{th}, data includes turnover during Congresses. For the 117^{th} Congress, data is for the number of women initially elected. GP study ("Original sample") spans Congress 101^{st} to 111^{th}. Our additional sample ("New data"), contains data from the 111^{th} to 116^{th} Congress.

Tables

Table 1: Summary statistics

	Mean	SD	Min	Max	N
Panel A: Congress 101 - 111 (GP's data)					
Number of cosponsors	16.99	35.93	0.00	425.00	61,334
% opposite party cosponsors	14.95	21.29	0.00	100.00	61,331
Sponsor tenure	6.22	3.84	1.00	20.00	61,334
Sponsor age	55.15	10.13	27.00	89.00	61,319
Share Black	0.11	0.14	0.00	0.92	61,334
Share foreign-born	0.10	0.10	0.00	0.59	61,334
Share over 65	0.13	0.04	0.04	0.44	61,334
Share urban	0.74	0.26	0.00	1.00	61,334
Median income	29,265	10,239	8,434	64,199	61,334
Population density	1,746	6,474	0	73,773	61,334
Panel B: Congress 101 - 111 (data corrected)					
Median income	37,124	13,067	8,434	80,000	60,770
Population density	2,827	7,894	0	73,773	61,334
Panel C: Congress 112 - 116 (our data)					
Number of cosponsors	16.56	35.64	0.00	432.00	32,601
% opposite party cosponsors	15.37	22.20	0.00	98.78	32,601
Sponsor tenure	5.95	4.80	1.00	30.00	27,608
Sponsor age	58.48	11.12	30.00	90.00	27,608
Share Black	0.13	0.14	0.00	0.71	32,601
Share foreign-born	0.14	0.11	0.01	0.57	32,601
Share over 65	0.15	0.04	0.06	0.38	32,601
Share urban	0.83	0.19	0.24	1.00	32,601
Median income	60,942	18,049	23,504	149,375	32,601
Population density	2,473	6,919	1	62,032	32,601

Notes: Unit of analysis is bill. Sponsor tenure and sponsor age refer to the bill's sponsor. Share black, Share foreign born, Share over 65, Share urban, Median income and Population density are measured in the district where the sponsor was elected.

Table 2: Regression results: Data corrected

	(1) OLS-full sample	(2) RD-optimal bandwidth	(3) RD-optimal bandwidth with inverse PS-weighting	(4) Inverse PS weighting-full sample	(5) Inverse PS weighting-full sample
Panel A: number of co-sponsors					
All	1.370**	2.054	0.597	0.392	-0.205
SE	(0.638)	(2.655)	(2.323)	(0.611)	(0.691)
No. bills	60106	4732	4732	55108	54444
No. sponsors*term	4698	393	393	4355	4310
Optimal bandwidth		25	25		
Democrats	1.149	2.551	-1.148	0.330	0.210
SE	(0.765)	(3.704)	(3.525)	(0.892)	(0.951)
No. bills	32520	2245	2245	29233	29037
No. sponsors*term	2467	186	186	2253	2240
Optimal bandwidth		30	30		
Republicans	3.161***	5.630	5.111	1.785**	2.899***
SE	(1.019)	(4.876)	(4.031)	(0.859)	(1.017)
No. bills	27434	1215	1188	25852	23581
No. sponsors*term	2221	99	97	2098	1930
Optimal bandwidth		13	13		
Panel B: % co-sponsors of opposite party					
All	-0.021	1.206	2.911	0.883	0.382
SE	(0.508)	(3.905)	(3.073)	(0.744)	(0.875)
No. bills	60103	2709	2709	55106	54442
No. sponsors*term	4698	227	227	4355	4310
Optimal bandwidth		16	16		
Democrats	-1.663***	-3.913**	-5.727***	-1.007*	-1.361**
SE	(0.422)	(1.921)	(2.030)	(0.592)	(0.532)
No. bills	32519	1880	1880	29232	29036
No. sponsors*term	2467	160	160	2253	2240
Optimal bandwidth		24	24		
Republicans	3.727***	13.275**	5.653	2.939***	2.853***
SE	(0.921)	(6.679)	(3.740)	(0.809)	(0.875)
No. bills	27432	1031	1004	25851	23580
No. sponsors*term	2221	87	85	2098	1930
Optimal bandwidth		11	11		
Sponsor characteristics	Yes	No	No	No	No
Bill characteristics	Yes	No	Yes	Yes	Yes
District characteristics	Yes	No	No	No	No
Propensity score			Distr.	Distr.+MV	Distr.+MV+Spon

Notes: Entries in the table represent the coefficient on the female sponsor dummy. Robust SEs, clustered at the sponsor-Congress level, in parentheses. The unit of observation is a bill. All estimates include Congress fixed effects. Bill characteristics include dummies for the committee of referral and dummies for the topic. Sponsor characteristics include: age, tenure in Congress, a dummy for whether the sponsor is a rookie, committee memberships (i.e. chair or ranking member), and the total number of bills sponsored within Congress. District characteristics include: three macro area dummies, the share of Black, over-65, foreign and urban residents, the logarithm of the median household income and the logarithm of the population density.

Table 3: Regression results: Clustering by sponsor

	(1) OLS-full sample	(2) RD-optimal bandwidth	(3) RD-optimal bandwidth with inverse PS-weighting	(4) Inverse PS weighting-full sample	(5) Inverse PS weighting-full sample
Panel A: number of co-sponsors					
All	1.395	2.308	1.138	0.491	-0.083
SE	(1.025)	(2.586)	(2.332)	(0.952)	(0.970)
No. bills	60670	4871	4871	55651	55008
No. sponsors	1106	280	280	1101	1090
Optimal bandwidth		25	25		
Democrats	1.172	2.182	-1.600	0.429	0.301
SE	(1.223)	(3.929)	(3.544)	(1.333)	(1.333)
No. bills	32847	2343	2343	29560	29364
No. sponsors	588	128	128	578	575
Optimal bandwidth		30	30		
Republicans	3.149**	6.101	6.734*	1.912	3.124**
SE	(1.359)	(4.865)	(4.083)	(1.235)	(1.295)
No. bills	27671	1227	1200	26089	23818
No. sponsors	516	84	82	522	477
Optimal bandwidth		13	13		
Panel B: % co-sponsors of opposite party					
All	0.181	0.822	1.884	0.976	0.651
SE	(0.959)	(3.979)	(3.102)	(1.447)	(1.517)
No. bills	60667	2781	2781	55649	55006
No. sponsors	1106	182	182	1101	1090
Optimal bandwidth		16	16		
Democrats	-1.506**	-3.351*	-5.297***	-0.929	-1.208
SE	(0.669)	(1.864)	(1.983)	(0.871)	(0.828)
No. bills	32846	1978	1978	29559	29363
No. sponsors	588	117	117	578	575
Optimal bandwidth		24	24		
Republicans	3.666**	12.514*	4.428	2.905**	2.827**
SE	(1.494)	(6.493)	(3.764)	(1.453)	(1.424)
No. bills	27669	1043	1016	26088	23817
No. sponsors	516	75	73	522	477
Optimal bandwidth		11	11		
Sponsor characteristics	Yes	No	No	No	No
Bill characteristics	Yes	No	Yes	Yes	Yes
District characteristics	Yes	No	No	No	No
Propensity score			Distr.	Distr.+MV	Distr.+MV+Spon

Notes: Entries in the table represent the coefficient on the female sponsor dummy. Robust SEs, clustered at the sponsor level, in parentheses. The unit of observation is a bill. All estimates include Congress fixed effects. Bill characteristics include dummies for the committee of referral and dummies for the topic. Sponsor characteristics include: age, tenure in Congress, a dummy for whether the sponsor is a rookie, committee memberships (i.e. chair or ranking member), and the total number of bills sponsored within Congress. District characteristics include: three macro area dummies, the share of Black, over-65, foreign and urban residents, the logarithm of the median household income and the logarithm of the population density.

Table 4: Regression results: Congresses 112-116

	(1) OLS-full sample	(2) RD-optimal bandwidth	(3) RD-optimal bandwidth with inverse PS-weighting	(4) Inverse PS weighting-full sample	(5) Inverse PS weighting-full sample
Panel A: number of co-sponsors					
All	2.509***	2.467	3.050	1.932***	3.358***
SE	(0.675)	(3.017)	(2.832)	(0.683)	(0.810)
No. bills	27596	2576	2576	32598	27596
No. sponsors	1886	183	183	2205	1886
Optimal bandwidth		17	17		
Democrats	1.647*	5.872	8.182*	1.210	2.049**
SE	(0.842)	(3.725)	(4.538)	(0.845)	(0.966)
No. bills	14363	749	749	16865	14363
No. sponsors	883	51	51	1030	883
Optimal bandwidth		9	9		
Republicans	2.573**	5.925	10.384	2.458**	2.561**
SE	(1.186)	(4.866)	(7.509)	(1.177)	(1.230)
No. bills	13233	872	872	15733	13233
No. sponsors	1003	71	71	1175	1003
Optimal bandwidth		14	14		
Panel B: % co-sponsors of opposite party					
All	-2.057***	-2.253	-2.591	-2.802***	-1.307*
SE	(0.545)	(4.086)	(4.110)	(0.585)	(0.716)
No. bills	27596	2325	2325	32598	27596
No. sponsors	1886	165	165	2205	1886
Optimal bandwidth		16	16		
Democrats	-2.016***	-7.205*	-5.517*	-2.324***	-2.312***
SE	(0.580)	(3.857)	(3.129)	(0.636)	(0.730)
No. bills	14363	1044	1044	16865	14363
No. sponsors	883	66	66	1030	883
Optimal bandwidth		13	13		
Republicans	1.133	-5.773	3.723	-0.371	0.024
SE	(0.868)	(6.970)	(7.601)	(0.911)	(0.941)
No. bills	13233	625	625	15733	13233
No. sponsors	1003	51	51	1175	1003
Optimal bandwidth		11	11		
Sponsor characteristics	Yes	No	No	No	No
Bill characteristics	Yes	No	Yes	Yes	Yes
District characteristics	Yes	No	No	No	No
Propensity score			Distr.	Distr.+MV	Distr.+MV+Spon

Notes: Entries in the table represent the coefficient on the female sponsor dummy. Robust SEs, clustered at the sponsor-Congress level, in parentheses. The unit of observation is a bill. All estimates include Congress fixed effects. Bill characteristics include dummies for the committee of referral and dummies for the topic. Sponsor characteristics include: age, tenure in Congress, a dummy for whether the sponsor is a rookie, committee memberships (i.e. chair or ranking member), and the total number of bills sponsored within Congress. District characteristics include: three macro area dummies, the share of Black, over-65, foreign and urban residents, the logarithm of the median household income and the logarithm of the population density.

Table 5: Regression results: Clustering by sponsor, Congresses 101 -116, data corrected

	(1) OLS-full sample	(2) RD-optimal bandwidth	(3) RD-optimal bandwidth with inverse PS-weighting	(4) Inverse PS weighting-full sample	(5) Inverse PS weighting-full sample
Panel A: number of co-sponsors					
All	1.846***	1.989	1.600	1.062	1.839**
SE	(0.670)	(2.115)	(1.833)	(0.702)	(0.738)
No. bills	88336	6768	6768	87685	82674
No. sponsors	2086	407	407	2243	2070
Optimal bandwidth		19	19		
Democrats	1.716**	0.817	2.118	0.804	1.473
SE	(0.807)	(2.824)	(2.352)	(0.951)	(0.979)
No. bills	47079	3084	3084	46098	43596
No. sponsors	1046	175	175	1109	1033
Optimal bandwidth		20	20		
Republicans	2.488**	5.837*	12.618***	2.176**	2.604***
SE	(1.013)	(3.316)	(4.449)	(0.871)	(0.998)
No. bills	41105	2278	2243	41585	39076
No. sponsors	1038	163	160	1133	1036
Optimal bandwidth		14	14		
Panel B: % co-sponsors of opposite party					
All	-0.934	-1.000	-0.548	-0.324	1.995*
SE	(0.661)	(3.030)	(2.833)	(0.977)	(1.109)
No. bills	88333	5027	5027	87683	82672
No. sponsors	2086	323	323	2243	2070
Optimal bandwidth		15	15		
Democrats	-1.852***	-2.709	-1.489	-1.441**	-0.764
SE	(0.482)	(2.569)	(2.538)	(0.654)	(0.728)
No. bills	47078	2262	2262	46097	43595
No. sponsors	1046	137	137	1109	1033
Optimal bandwidth		15	15		
Republicans	2.310**	4.836	5.793	1.667	2.513**
SE	(1.114)	(5.274)	(5.286)	(1.113)	(1.191)
No. bills	41103	1363	1306	41584	39075
No. sponsors	1038	104	97	1133	1036
Optimal bandwidth		9	9		
Sponsor characteristics	Yes	No	No	No	No
Bill characteristics	Yes	No	Yes	Yes	Yes
District characteristics	Yes	No	No	No	No
Propensity score			Distr.	Distr.+MV	Distr.+MV+Spon

Notes: Entries in the table represent the coefficient on the female sponsor dummy. Robust SEs, clustered at the sponsor level, in parentheses. The unit of observation is a bill. All estimates include Congress fixed effects. Bill characteristics include dummies for the committee of referral and dummies for the topic. Sponsor characteristics include: age, tenure in Congress, a dummy for whether the sponsor is a rookie, committee memberships (i.e. chair or ranking member), and the total number of bills sponsored within Congress. District characteristics include: three macro area dummies, the share of Black, over-65, foreign and urban residents, the logarithm of the median household income and the logarithm of the population density.

Appendix

A1 Additional Tables

Table A1: GP Table 5

	(1) OLS-full sample	(2) RD-optimal bandwidth	(3) RD-optimal bandwidth with inverse PS-weighting	(4) Inverse PS weighting-full sample	(5) Inverse PS weighting-full sample
Panel A: number of co-sponsors					
All	1.395**	2.308	1.138	0.496	-0.083
SE	(0.628)	(2.597)	(2.333)	(0.608)	(0.673)
No. bills	60670	4871	4871	55672	55008
No. sponsors*term	4746	403	403	4403	4358
Optimal bandwidth		25	25		
Democrats	1.172	2.182	-1.600	0.429	0.301
SE	(0.746)	(3.609)	(3.459)	(0.892)	(0.950)
No. bills	32847	2343	2343	29560	29364
No. sponsors*term	2492	193	193	2278	2265
Optimal bandwidth		30	30		
Republicans	3.149***	6.101	6.734*	1.912**	3.124***
SE	(1.017)	(4.796)	(3.946)	(0.852)	(1.040)
No. bills	27671	1227	1200	26089	23818
No. sponsors*term	2244	100	98	2121	1953
Optimal bandwidth		13	13		
Panel B: % co-sponsors of opposite party					
All	0.181	0.822	1.884	0.982	0.651
SE	(0.507)	(3.876)	(3.071)	(0.732)	(0.860)
No. bills	60667	2781	2781	55670	55006
No. sponsors*term	4746	232	232	4403	4358
Optimal bandwidth		16	16		
Democrats	-1.505***	-3.351*	-5.297***	-0.929	-1.208**
SE	(0.419)	(1.910)	(2.000)	(0.570)	(0.525)
No. bills	32846	1978	1978	29559	29363
No. sponsors*term	2492	167	167	2278	2265
Optimal bandwidth		24	24		
Republicans	3.666***	12.514*	4.428	2.905***	2.827***
SE	(0.916)	(6.673)	(3.737)	(0.820)	(0.868)
No. bills	27669	1043	1016	26088	23817
No. sponsors*term	2244	88	86	2121	1953
Optimal bandwidth		11	11		
Sponsor characteristics	Yes	No	No	No	No
Bill characteristics	Yes	No	Yes	Yes	Yes
District characteristics	Yes	No	No	No	No
Propensity score			Distr.	Distr.+MV	Distr.+MV+Spon

Notes: Entries in the table represent the coefficient on the female sponsor dummy. Robust standard errors, clustered at the individual-Congress level, in parentheses. The unit of observation is a bill. All estimates include Congress fixed effects. Bill characteristics include 33 dummies for the committee of referral, and 226 dummies for the topic. Sponsor characteristics include: age, tenure in Congress, a dummy for whether the sponsor is a rookie, a committee leader (chair or ranking member) or black, a party dummie, 5 occupational dummies, a dummy for whether the sponsor has an Ivy League college degree, a dummy for whether the sponsor was born in the state of election, and the total number of bills sponsored within the congress. District characteristics include: 3 macro area dummies, the percentage of black, over-65, foreign and urban residents, the logarithm of the median income, and the logarithm of the population density.

Table A2: Comparison of GP's dataset for the 107th and 108th Congresses and our dataset for the 112th Congress, along with online census data.

State	District	Dataset	Congress	Area (sq. miles)	Population	Income
Alaska	At Large	GP	107	570,373	629,099	46,581
		GP	108	1,481,000,000,000	626,932	25,776
		Ours	112	665,381	722,718	67,825
		Census	—	665,384	732,673	77,845
Alabama	District 1	GP	107	6,785	577,630	27,360
		GP	108	16,360,000,000	635,495	20,844
		Ours	112	7,183	693,871	43,144
		Census	—	7,182	727,212	52,278
New York	District 5	GP	107	151	581,073	57,915
		GP	108	171,600,000	654,253	27,182
		Ours	112	85	671,449	61,638
		Census	—	52	822,717	73,628
Wyoming	At Large	GP	107	97,104	475,503	32,216
		GP	108	251,488,665,361	493,782	19,763
		Ours	112	97,812	568,158	56,322
		Census	—	97,914	578,803	65,204
Delaware	At Large	GP	107	1,954	666,168	40,252
		GP	108	5,059,704,780	783,600	25,910
		Ours	112	2,489	907,135	58,814
		Census	—	2,489	1,003,384	71,091
Vermont	At Large	GP	107	9,247	562,758	34,780
		GP	108	23,956,228,057	608,827	21,497
		Ours	112	9,616	626,431	52,776
		Census	—	9,616	645,570	72,431

Congress refers to the session of Congress. The 107th Congress ran from January 2001 to January 2003, the 108th from January 2003 to 2005, and the 112th from January 2011 to January 2013. Income refers to the median household income in a district. This is not deflated and is presented in the USD of the first year of the respective Congress. The census income figures are for 2021, presented in 2021 USD.

Table A3: Regression results: Fewer controls, Congresses 101 - 111

	(1) OLS-full sample	(2) RD-optimal bandwidth	(3) RD-optimal bandwidth with inverse PS-weighting	(4) Inverse PS weighting-full sample	(5) Inverse PS weighting-full sample
Panel A: number of co-sponsors					
All	1.395**	2.240	1.117	0.496	0.915
SE	(0.613)	(2.563)	(2.282)	(0.608)	(0.670)
No. bills	61304	5032	5032	55672	55642
No. sponsors*term	4787	413	413	4403	4399
Optimal bandwidth		25	25		
Democrats	1.452**	2.093	-1.583	0.429	0.736
SE	(0.732)	(3.572)	(3.417)	(0.892)	(1.032)
No. bills	33043	2366	2366	29560	29560
No. sponsors*term	2505	194	194	2278	2278
Optimal bandwidth		30	30		
Republicans	2.136**	5.936	6.877*	1.912**	2.337**
SE	(0.985)	(4.619)	(3.922)	(0.852)	(0.946)
No. bills	28109	1318	1291	26089	26080
No. sponsors*term	2272	106	104	2121	2120
Optimal bandwidth		13	13		
Panel B: % co-sponsors of opposite party					
All	0.032	-0.484	0.901	0.982	3.064***
SE	(0.504)	(4.061)	(3.185)	(0.732)	(0.853)
No. bills	61301	2895	2895	55670	55640
No. sponsors*term	4787	239	239	4403	4399
Optimal bandwidth		16	16		
Democrats	-1.584***	-3.057	-5.138**	-0.929	-0.243
SE	(0.414)	(1.932)	(1.993)	(0.570)	(0.596)
No. bills	33042	2001	2001	29559	29559
No. sponsors*term	2505	168	168	2278	2278
Optimal bandwidth		24	24		
Republicans	3.070***	10.922	3.499	2.905***	3.917***
SE	(0.884)	(6.744)	(3.931)	(0.820)	(0.849)
No. bills	28107	1134	1107	26088	26079
No. sponsors*term	2272	94	92	2121	2120
Optimal bandwidth		11	11		
Sponsor characteristics	Yes	No	No	No	No
Bill characteristics	Yes	No	Yes	Yes	Yes
District characteristics	Yes	No	No	No	No
Propensity score			Distr.	Distr.+MV	Distr.+MV+Spon

Notes: Entries in the table represent the coefficient on the female sponsor dummy. Robust SEs, clustered at the sponsor-Congress level, in parentheses. The unit of observation is a bill. All estimates include Congress fixed effects. Bill characteristics include dummies for the committee of referral and dummies for the topic. Sponsor characteristics include: age, tenure in Congress, a dummy for whether the sponsor is a rookie, committee memberships (i.e. chair or ranking member), and the total number of bills sponsored within Congress. District characteristics include: three macro area dummies, the share of Black, over-65, foreign and urban residents, the logarithm of the median household income and the logarithm of the population density.

References

Gagliarducci, S. and M. D. Paserman (2022). Gender differences in cooperative environments? Evidence from the us congress. *The Economic Journal 132*(641), 218–257.