Received: 15 January 2024 | Accepted: 8 September 2025

DOI: 10.1111/ecin.70031

ORIGINAL ARTICLE

Economic Inquiry

Gender differences in cooperation in Congress: Replicating Gagliarducci and Paserman (2022)

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

[1]University of Warwick, CEPR, IZA, and J-Pal, Coventry, UK

[2]SITE - Stockholm School of Economics, Misum, CEPR, and J-Pal, Stockholm, Sweden

[3]KOF Swiss Economic Institute, ETH Zurich, Zurich, Switzerland

Correspondence
Pamela Campa, SITE, Misum, CEPR, and J-Pal. Stockholm School of Economics, Stockholm, 113 83, Sweden.
Email: pamela.campa@hhs.se

Abstract

Gagliarducci and Paserman study gender differences in cooperation among politicians using data from the U.S. House of Representatives in 1988–2010. The evidence is consistent with commonality of interest driving cooperation, rather than gender per se. We show that GP's results are robust to the correction of some errors in the control variables and to clustering standard errors at the individual level, instead of individual-term. Additional data from 2011 to 2020 confirms the relevance of the ideological distance between male and female representatives, but also indicates that in recent years female politicians tend to recruit more co-sponsors for the bills that they sponsor.

KEYWORDS

gender differences in politics, legislative cooperation, US congress

JEL CLASSIFICATION

D72, J16, Z18

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, since during this period Republican female representatives were ideologically closer to Democrats than their male colleagues, whereas Democratic women were ideologically further away from Republicans.

The paper is novel in testing gender differences in cooperative behavior in a real-world and high-stakes setting, building on a large experimental literature on the subject (see Balliet et al. (2011)). It also has important practical

Managing Editor: Abel Brodeur

implications, by challenging, as the authors observe, the commonly held view that an increase in female representation in the US Congress would help solve partisan gridlock.

We conduct verification, reanalysis, and extension tests on GP's original study (Clemens, 2017).[1] More precisely, we investigate whether the main results reported in GP 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 using data between 2011 and 2020.

The data errors concern the control variables *population density* and *median household income*, where the values in GP's dataset for Congresses 108 through 111 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 largely unchanged. We also replicate GP's analysis clustering the standard errors at the individual level, instead of at the individual-term level, to account for non-independence over time for a given Congress member. Standard errors are slightly larger and, out of the 10 coefficients that were statistically significant in GP, three lose the 5% significance level, but the broad conclusions are unchanged.[2]

Finally, by extending the analysis to 2011–2020, we can test the robustness of GP's hypotheses in a context that differs in at least two relevant aspects. First, during this period the share of women in the House of Representatives became substantially larger. The selection and the behavior of women as political leaders might change as their representation grows and they become more experienced (see, e.g., Chaudhuri et al. (2024) on changing gender gaps in dishonesty). Moreover, within-party gender differences in ideology changed compared to previous decades. While Democratic female representatives are still less conservative than Democratic men, among Republicans women became ideologically more similar to their male colleagues.[3,4] 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, and we do not observe any gender differences in bipartisan cooperative behavior among Republicans. Instead, the pattern is slightly different from the results in 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), we observe that during the last decade bills from both Republican and Democratic women attracted more sponsors than bills from their male colleagues. In the 2011–2020 sample, we find that women tend to recruit 15% more co-sponsors and this effect is statistically significant at the 1% level in three out of five specifications, compared to a gender gap in favor of women of around 8% in GP, with only one coefficient significant at the 5% level.[5]

In sum, the novel evidence from the 2011–2020 period confirms that cooperation with members of the other party is driven mainly by ideological proximity rather than gender per se but, in terms of the overall number of co-sponsors recruited, we observe that in recent years women might have been more cooperative.

2 | REPLICATION

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 10 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.

In our replication exercise we consider all the estimates reported originally by GP in Table 5 (reproduced in Table 1). GP consider two outcomes in this table, *number of co-sponsors* (Panel A) and *percent co-sponsors of opposite party* (Panel B), for all the bills together and separately by party affiliation of the sponsor. They report estimates from five different specifications: (1) a simple OLS regression with controls for sponsor, bill and district characteristics, (2) an RD design, controlling for the margin of victory of female candidates in mixed-gender races using the optimal bandwidth, (3) a similar RD specification that also uses inverse propensity score based on district characteristics (IPW) to account for potential selection bias when comparing closely elected politicians (4) a full sample IPW specification (i.e., not restricted to mixed-gender electoral races), based on district characteristics and margin of victory, and (5) the same specification but with sponsor characteristics also included in the computation of the IPW score. Below we first describe how we collected our data and then we show results for each of the robustness checks considered, separately and pooled together. We summarize all the main results in Table 2, where we also report GP's corresponding estimates for comparison. More estimation details are shown in the supplementary material.[6]

TABLE 1 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	60,670	4871	4871	55,672	55,008
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	32,847	2343	2343	29,560	29,364
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	27,671	1227	1200	26,089	23,818
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	60,667	2781	2781	55,670	55,006
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	32,846	1978	1978	29,559	29,363
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	27,669	1043	1016	26,088	23,817
No. sponsors × term	2244	88	86	2121	1953
Optimal bandwidth		11	11		

(Continues)

TABLE 1 (Continued)

	(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
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

Note: 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 dummy, 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.

2.1 | Data collection

GP's main data set is based on the Library of Congress' data information system, THOMAS, from which they retrieve information on *public* bills submitted from the 101st (elected in 1988) to the 111th (elected in 2008) Congresses. They merge this data with additional Congress-member individual characteristics, election statistics, and demographic and economic information on congressional districts.[7] We show summary statistics from GP's dataset in Table 3, 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 4, where we show a comparison of data from the 107th and 108th Congresses (from GP's dataset) for some districts, as well as the values that we retrieved from the U.S. census. For instance, in the database used by GP, the area of several States becomes more than one million times larger between the 107th and 108th Congresses and the median household income decreases by around 50%. Further analysis suggested that values in GP's data for these two variables for Congresses 108th (elected in 2002) to 111th (elected in 2008) are not consistent with official census statistics.

To address this problem, we collected new data for population density and household income for the years with implausible values (e.g., Congresses 108–111) from the U.S. census website,[8] choosing the available series that was closest in time to our period of interest.[9] 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 (see Table 3, Panel B).

We also collected data for many of the other variables used in GP for the years 2011–2020, 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 explain in more detail in Appendix A.

2.2 | Replication using the amended dataset

We replicate GP's analysis using the corrected dataset. Correcting for these errors might potentially help to obtain a more consistent estimate of the size of the gender difference. As GP notice, controlling for district characteristics attenuates some of their estimates, although the broad conclusions of their analysis are mostly unaffected.[10]

We summarize the estimates in Table 2 in the rows labeled "GP + Corrected data", and we also report more estimation details in Supporting Information S2: Table S1. The resulting estimates closely align with those in GP. The finding that Republican women attract more co-sponsors on their bills is confirmed, with coefficients of comparable magnitude to the original study. When we consider co-sponsorship with MPs from the opposite party, 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. For Republicans, the estimated coefficients and standard errors are also virtually unchanged.

TABLE 2 Combined table of estimates.

	(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—Outcome variable: number of co-sponsors					
Sample: Both parties					
GP original results (years 1988–2010)	1.395**	2.308	1.138	0.496	−0.083
	(0.628)	(2.597)	(2.333)	(0.608)	(0.673)
GP + corrected data	1.370**	2.054	0.597	0.392	−0.205
	(0.638)	(2.655)	(2.323)	(0.611)	(0.691)
GP + Spons.-clustered	1.395	2.308	1.138	0.491	−0.083
	(1.025)	(2.586)	(2.332)	(0.952)	(0.970)
Extension: years 2011–2020	2.509***	2.467	3.050	1.932***	3.358***
	(0.675)	(3.017)	(2.832)	(0.683)	(0.810)
All years (1988–2020)	1.846***	1.989	1.600	1.062	1.839**
	(0.670)	(2.115)	(1.833)	(0.702)	(0.738)
Sample: Democrats					
GP original results (years 1988–2010)	1.172	2.182	−1.600	0.429	0.301
	(0.746)	(3.609)	(3.459)	(0.892)	(0.950)
GP + corrected data	1.149	2.551	−1.148	0.330	0.210
	(0.765)	(3.704)	(3.525)	(0.892)	(0.951)
GP + Spons.-clustered	1.172	2.182	−1.600	0.429	0.301
	(1.223)	(3.929)	(3.544)	(1.333)	(1.333)
Extension: years 2011–2020	1.647*	5.872	8.182*	1.210	2.049**
	(0.842)	(3.725)	(4.538)	(0.845)	(0.966)
All years (1988–2020)	1.716**	0.817	2.118	0.804	1.473
	(0.807)	(2.824)	(2.352)	(0.951)	(0.979)
Sample: Republicans					
GP original results (years 1988–2010)	3.149***	6.101	6.734*	1.912**	3.124***
	(1.017)	(4.796)	(3.946)	(0.852)	(1.040)
GP + corrected data	3.161***	5.630	5.111	1.785**	2.899***
	(1.019)	(4.876)	(4.031)	(0.859)	(1.017)
GP + Spons.-clustered	3.149**	6.101	6.734*	1.912	3.124**
	(1.359)	(4.865)	(4.083)	(1.235)	(1.295)
Extension: years 2011–2020	2.573**	5.925	10.384	2.458**	2.561**
	(1.186)	(4.866)	(7.509)	(1.177)	(1.230)
All years (1988–2020)	2.488**	5.837*	12.618***	2.176**	2.604***
	(1.013)	(3.316)	(4.449)	(0.871)	(0.998)

(Continues)

TABLE 2 (Continued)

	(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 B—Outcome variable: % co-sponsors of opposite party					
Sample: Both parties					
GP original results (years 1988–2010)	0.181	0.822	1.884	0.982	0.651
	(0.507)	(3.876)	(3.071)	(0.732)	(0.860)
GP + corr. Data	−0.021	1.206	2.911	0.883	0.382
	(0.508)	(3.905)	(3.073)	(0.744)	(0.875)
GP + Spons.-clustered	0.181	0.822	1.884	0.976	0.651
	(0.959)	(3.979)	(3.102)	(1.447)	(1.517)
Extension: years 2011–2020	−2.057***	−2.253	−2.591	−2.802***	−1.307*
	(0.545)	(4.086)	(4.110)	(0.585)	(0.716)
All years (1988–2020)	−0.934	−1.000	−0.548	−0.324	1.995*
	(0.661)	(3.030)	(2.833)	(0.977)	(1.109)
Sample: Democrats					
GP original results (years 1988–2010)	−1.505***	−3.351*	−5.297***	−0.929	−1.208**
	(0.419)	(1.910)	(2.000)	(0.570)	(0.525)
GP + corrected data	−1.663***	−3.913**	−5.727***	−1.007*	−1.361**
	(0.422)	(1.921)	(2.030)	(0.592)	(0.532)
GP + Spons.-clustered	−1.506**	−3.351*	−5.297***	−0.929	−1.208
	(0.669)	(1.864)	(1.983)	(0.871)	(0.828)
Extension: years 2011–2020	−2.016***	−7.205*	−5.517*	−2.324***	−2.312***
	(0.580)	(3.857)	(3.129)	(0.636)	(0.730)
All years (1988–2020)	−1.852***	−2.709	−1.489	−1.441**	−0.764
	(0.482)	(2.569)	(2.538)	(0.654)	(0.728)
Sample: Republicans					
GP original results (years 1988–2010)	3.666***	12.514*	4.428	2.905***	2.827***
	(0.916)	(6.673)	(3.737)	(0.820)	(0.868)
GP + corrected data	3.727***	13.275**	5.653	2.939***	2.853***
	(0.921)	(6.679)	(3.740)	(0.809)	(0.875)
GP + Spons.-clustered	3.666**	12.514*	4.428	2.905**	2.827**
	(1.494)	(6.493)	(3.764)	(1.453)	(1.424)
Extension: years 2011–2020	1.133	−5.773	3.723	−0.371	0.024
	(0.868)	(6.970)	(7.601)	(0.911)	(0.941)
All years (1988–2020)	2.310**	4.836	5.793	1.667	2.513**
	(1.114)	(5.274)	(5.286)	(1.113)	(1.191)

TABLE 2 (Continued)

	(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
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

Note: Entries in the table represent the coefficient on the female sponsor dummy. 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 3 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	8434	64,199	61,334
Population density	1746	6474	0	73,773	61,334
Panel B: Congress 101–111 (data corrected)					
Median income	37,124	13,067	8434	80,000	60,770
Population density	2827	7894	0	73,773	61,334
Panel C: Congress 112–116					
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

(Continues)

TABLE 3 (Continued)

	Mean	SD	Min	Max	*N*
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	2473	6919	1	62,032	32,601

Note: Unit of analysis is a 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 4 Comparison of GP's dataset for the 107th and 108th Congresses.

State	District	Congress	Area (sq. miles)	Population	Income
Alaska	At large	107	570,373	629,099	46,581
		108	1,481,000,000,000	626,932	25,776
Alabama	District 1	107	6785	577,630	27,360
		108	16,360,000,000	635,495	20,844
New York	District 5	107	151	581,073	57,915
		108	171,600,000	654,253	27,182
Wyoming	At large	107	97,104	475,503	32,216
		108	251,488,665,361	493,782	19,763
Delaware	At large	107	1954	666,168	40,252
		108	5,059,704,780	783,600	25,910
Vermont	At large	107	9247	562,758	34,780
		108	23,956,228,057	608,827	21,497

Note: Congress refers to the session of Congress. The 107th Congress ran from January 2001 to January 2003, the 108th from January 2003 to January 2005. Income refers to the median household income in a district.

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. 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 build long-lasting networks, 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.[11] Our findings are summarized in Table 2 in the rows labeled "GP + Spons.-clustered", and we also report more estimation details in Supporting Information S2: Table S2.

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). Turning to the second outcome, *percent of co-sponsors of opposite party*, two 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 three out of five. 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 | Extending the analysis to period 2011–2020

The dataset in GP spans Congresses 101 (elected in 1988) to 111 (elected in 2008). We extend the analysis to bills presented in more recently-elected Congresses, in office from 2011 to 2020.[12] This period differs in at least two important ways from the one considered in GP. First, the number of female representatives is substantially larger, a trend which started in the 1990s and continued through the 2000s and 2010s (see Figure 1). During the last decade, around 21% of Representatives were women, compared to 13% in the previous 2 decades considered by GP. Second, we also find evidence of changes in ideology along party and gender lines. While in the decades studied by GP Republican women appear to be more progressive than Republican men, there are no 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, 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.[13] We show the four CDFs in Figure 2 (Panel B), where we also report the respective CDFs 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 data for 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 from 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 this hypothesis, with an important amendment. Our estimates lend support to the claim that commonality of interest drives gender differences in bipartisan cooperation, but we also find that women in recent years might have been overall more cooperative than men. We reach these conclusions based on a number of results, summarized in Table 2 in the rows labeled "Extension: years 2011–2020" (see Supporting Information S2: Table S3 for estimation details).[14]

First, we find stronger evidence that women from both parties attract more co-sponsors 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

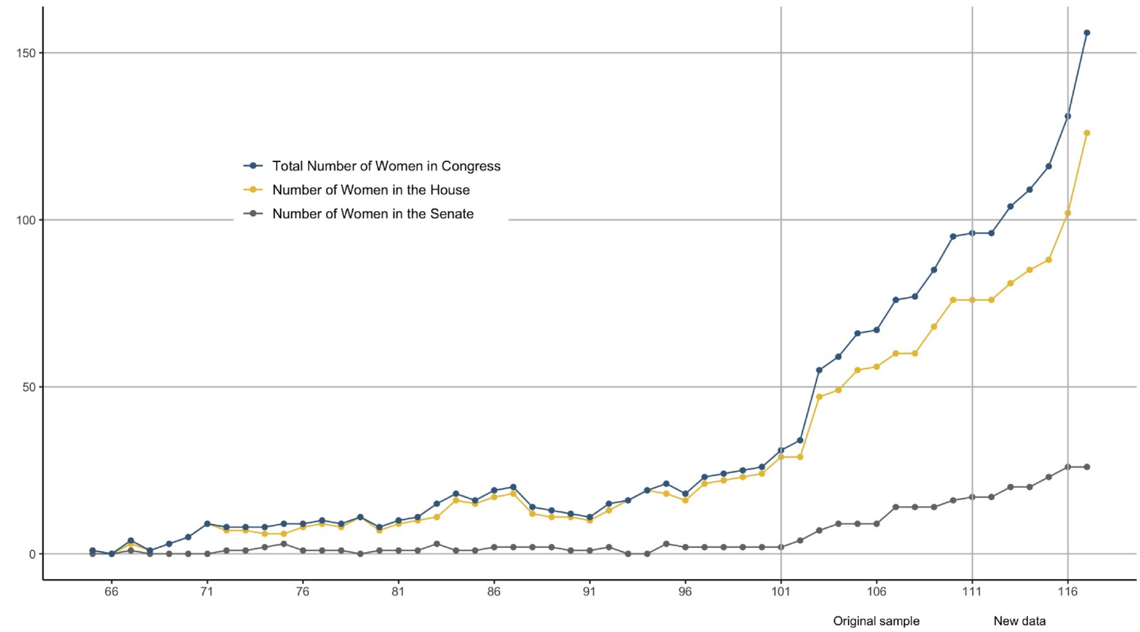
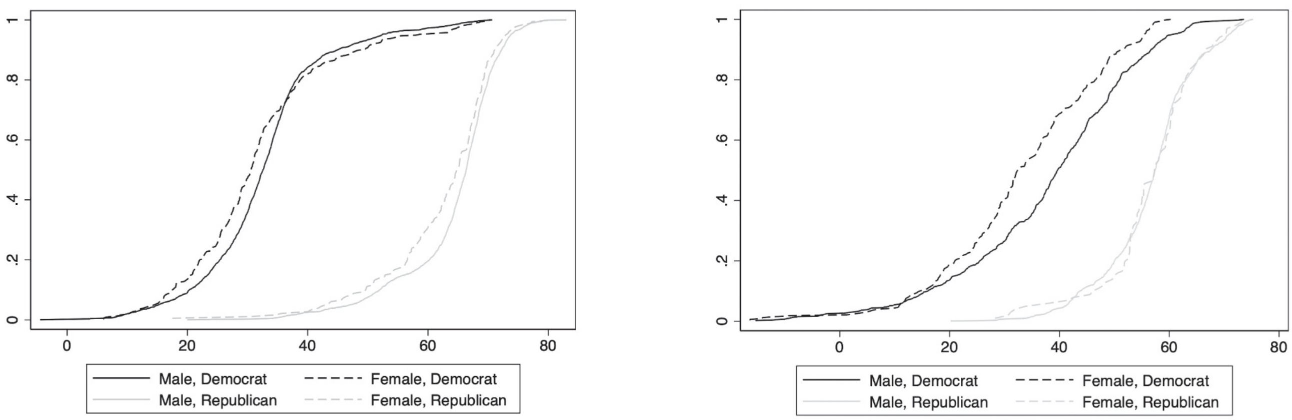


FIGURE 1 Number of women in Congress, 1917–2021. Data is from the Congressional Research Service publication "Women in Congress: Statistics and Brief Overview" (2022), available at: https://crsreports.congress.gov/product/pdf/R/R43244. For all Congresses other than the 117th, data includes turnover during Congresses. For the 117th Congress, data is for the number of women initially elected. GP study ("Original sample") spans Congress 101st to 111th. Our additional sample ("New data"), contains data from the 111th to 116th Congress.



(a) Congresses 101^{st} - 111^{th} **(b)** Congresses 112^{th} - 116^{th}

FIGURE 2 Predicted Republican share (CDF). Figure in left panel is reproduced from GP (Congresses 101–111). Figure in right panel is based on the same methodology, but more recent data (Congresses 112–116). As in GP, we show the 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 Congresses 101–111 (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). CDF, cumulative distribution function.

find significant results in three out of the five specifications considered for the overall sample, whereas in GP only one of these coefficients was statistically significant. Second, as in GP, we find that among Democrats female-sponsored bills attract a lower percentage 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. Finally, in contrast to the previous period, 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 2).

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

Finally, we assess the replicability of the analysis reported by GP in Table 5 taking into account 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 summarized in Table 2 in the rows labeled "All years (1988–2020)" and we also report more estimation details in Supporting Information S2: Table S4.

Consistent 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%–12% difference (relative to the sample mean) that is statistically significant in two out of five specifications. 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 analyzed separately for different decades, as we have done in the analysis above.

3 | 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 during this period 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 extend GP's analysis to 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 in more recent years women from both parties attract more co-sponsors than men.

ACKNOWLEDGMENTS

We thank Cecilia Smitt Meyer for excellent research assistance and two anonymous referees for very useful comments and suggestions. Financial support from the Mistra Center for Sustainable Markets (MISUM) is gratefully aknowledged.

DATA AVAILABILITY STATEMENT

Most of the data that support the findings of this study are openly available in OpenICPSR at https://doi.org/10.3886/E222321V4. Restrictions apply to the availability of election data, which were used under license for this study. Data are available at selectionatlas.org. See Bagues et al. (2025) for details on how to obtain the election data from the third party.

ENDNOTES

[1] Clemens (2017) categorizes replication exercises into four areas, namely verification, reproduction, reanalysis, and extension. A verification attempts to reproduce the results of the original paper using the same empirical strategy and the same data. A reproduction test resamples from the same population and then applies the same empirical strategy of the original study. A reanalysis uses the same data of the original work but utilizes an alternative empirical strategy while an extension replicates the analysis on a different population.

[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.

[4] Another if not the major development in recent decades in US politics is of course the increased polarization (Boxell et al., 2024). This development could have major implications for gender differences in cooperation if it were reflected in less co-sponsorship with the opposite party and overall. However, data show that neither the number of cosponsors nor the share of opposite party cosponsors differ substantially between GP's and our period of analysis (see Table 3).

[5] To quantify the magnitude of the gender gap, we divide the point estimates reported in the first column of Table 2 by the average number of co-sponsors, around 17 in both periods.

[6] In addition to the main robustness checks outlined above, we also show the sensitivity of GP's RD estimates by running a version of the same specifications with twice the optimal bandwidth as well as a version with a quadratic term included in the regression. The results are shown in the supplementary material. The finding that Democrats women are less likely to co-sponsor with the opposite party does not hold when we use a bandwidth that is twice the optimal one, while it remains unchanged when we control for a quadratic polynomial. Finally, in the supplementary material we show the standard RDD plots for both GP's data and the full dataset including both GP's corrected data and our additional data. We note that the evidence in the plots is not always aligned with GP's estimates, likely due to their sensitivity to the bandwidth (in the graphs we basically consider a bandwidth equal to 100.) Nevertheless, our assessment is that the general conclusions of the analysis do not change, especially considering that most often the RDD estimates are not statistically significant.

[7] 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.

[8] Accessible at data.census.gov.

[9] 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.

[10] GP's analysis also shows that selection on district characteristics, including income and population density, is especially large in the sample used for the Regression Discontinuity Design conditional on party, which is one of the strategies that the authors employ.

[11] Our strategy implies that we fail to connect over time those members who change party affiliation, a however highly infrequent event. Specifically, out of 1117 members-by-party in GP's dataset, we identify 20 potential party changes. Some members appear to change party affiliation within the same Congress.

[12] 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 analyze bills that were passed after May 2020.

[13] 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.

[14] We separately verify that our extended dataset conforms to GP's RDD diagnostics. We do this by replicating the McCrary test as in GP's Figure 5 as well as the balance tables in GP's Table 2A,B. In each case, we produce a version with GP's data only, our data only, and the two datasets combined. GP's data fail the McCrary test for Republicans only. In contrast, our data does not fail the McCrary test for either party or for the parties taken together. The full dataset follows GP's original dataset in failing the McCrary test for Republicans only. For the balance tables, we replicate the coefficients for the variables that are common between GP's dataset and our own. In the original GP table (considering only the variables for which we also have data), there are 184 coefficients in total. In the GP data, 47 of these (25.5%) breach the 0.25 threshold. In our data, 36 coefficients do (19.6%). In the combined data, 30 coefficients do (16.3%). All of the relevant figures and tables can be found in the supplementary material.

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

[16] http://www.congressionalbills.org.

[17] 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's *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.

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

[19] 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.

[20] Specifically, the data are from USCB series CP02, CP03, CP05, and H2. These are accessible at data.census.gov.

[21] 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.

REFERENCES

Bagues, M., Campa, P. & Etingin-Frati, G. (2025) *ECIN replication package for "Gender differences in cooperation in the U.S. Congress? Replicating Gagliarducci and Paserman (2022)"*. Inter-university Consortium for Political and Social Research [distributor]. Available from: https://doi.org/10.3886/E222321V4

Balliet, D., Li, N.P., Macfarlan, S.J. & Van Vugt, M. (2011) Sex differences in cooperation: a meta-analytic review of social dilemmas. *Psychological Bulletin*, 137(6), 881–909. Available from: https://doi.org/10.1037/a0025354

Boxell, L., Gentzkow, M. & Shapiro, J.M. (2024) Cross-country trends in affective polarization. *The Review of Economics and Statistics*, 106(2), 557–565. Available from: https://doi.org/10.1162/rest_a_01160

Chaudhuri, A., Iversen, V., Jensenius, F.R. & Maitra, P. (2024) Time in office and the changing gender gap in dishonesty: evidence from local politics in India. *American Journal of Political Science*, 68(1), 106–122. Available from: https://doi.org/10.1111/ajps.12733

Clemens, M.A. (2017) The meaning of failed replications: a review and proposal. *Journal of Economic Surveys*, 31(1), 326–342. Available from: https://doi.org/10.1111/joes.12139

Gagliarducci, S. & Paserman, M.D. (2022) Gender differences in cooperative environments? Evidence from the US Congress. *The Economic Journal*, 132(641), 218–257. Available from: https://doi.org/10.1093/ej/ueab069

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bagues, M., Campa, P. & Etingin-Frati, G. (2025) Gender differences in cooperation in Congress: replicating Gagliarducci and Paserman (2022). *Economic Inquiry*, 1–13. Available from: https://doi.org/10.1111/ecin.70031

APPENDIX A

Data collection

We downloaded data on bills sponsors and co-sponsors from the Library of Congress' data information system, THOMAS.[15] As in GP, we focus on House bills that are classified as *public*. 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,[16] which also contains information on the sponsor's gender and the bill's “minor” topics (the latter is used as control variable in GP).[17]

We recover some information on the characteristics of each bill's sponsor from the Biographical Directory of the United States Congress.[18] GP digitized information from this source on age, gender, and various other personal and tenure-based characteristics. We have forgone the digitization and relied only on information available from the website, which allows 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. 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 Supporting Information S2: Table S5 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.[19] Finally, we collect district-level data from the U.S. Census Bureau (USCB) on economic, social, and demographic characteristics for the years 2011–2021.[20] The data primarily come from the annualized estimates of the American Community Survey (ACS), a monthly survey conducted by the USCB that complements the decennial census.[21]