No Party for Women? Unequal Policy Responsiveness by Gender Persists Under Democratic and Republican Control

Nicholas Conroy, Janae Jackson, Christopher Vazquez,
Replication by Jing Ling Tan

3 November, 2024

Abstract

Original Abstract: Existing research demonstrates that policy outcomes in the United States tend to favor the preferences of men. Does this gender gap depend on which party controls the branches of government? Using a comprehensive dataset of policy preferences and enacted outcomes from 1964 to 2008, we investigate the effect of party control on policy responsiveness by gender. We find that public policy outcomes remain almost equally biased toward the preferences of men under both political parties. These results run counter to conventional wisdom, suggesting that greater descriptive representation of women within the Democratic Party does not lead to greater substantive representation.

1 Strengths of the Paper

- The research puzzle is convincingly motivated and explained to make the null gender*party effect an interesting and curious observation.
- Replication code is well documented and runs smoothly. I managed to replicate all the figures and got the same regression results.
- Great visualization to convey the paper's key arguments, especially Figures 2 and 4.

2 Comment 1 - Clarify Independent Variables

- The concept of a gender gap is intuitive but it wasn't clear how you constructed the independent variable in the paper and thus how to interpret the coefficient for that variable. From what I understood, it is the absolute difference in policy support between men and women for each policy, based on survey data. Why is there a logit transformation in the code, was it used, and how does it change the structure/distribution of the data? I believe this is documented in the paper you originally replicated but it will be helpful if this is conveyed early in the paper and in the interpretation of the results.
- Building from the previous point, I am confused about how to interpret the interaction term because the magnitude of the gender term and party control terms are constructed differently. The party control variable is a scale from 0 to 1 in increments of 0.25 but the gender variable is a log-odds term (or a positive integer, depending on which one you used).. It'll be good to clarify: what does a 1 unit change in gender difference * party control mean substantively?
- One suggestion that may help with interpreting the interaction term would be to standardize the original gender difference variable rather than log-transform (unless you really need to transform it, log-odds interaction terms just seem unnecessarily difficult to interpret). I ran this partial standardization as an extension code and found that your results hold.
- After standardization, I believe you can interpret the interaction term as: the effect of a one-standard-deviation increase in gender difference on policy responsiveness varies across levels of party control. I find this easier to understand than the original version.

Effect of Standardized Gender Difference and Non-Standardized Party Control on Policy Responsiveness

	Dependent variable:
	Policy Outcome
Standardized Gender Difference	0.154**
	(0.077)
Party Control Scale	0.218
	(0.139)
Interaction: Gender Difference * Party Control	0.148
	(0.143)
Constant	-0.897***
	(0.076)
Observations	2,207
Log Likelihood	-1,355.094
Akaike Inf. Crit.	2,718.188
Note:	*p<0.1; **p<0.05; ***p<0.01

3 Comment 2 - Are your units independent?

- It's clear what your quantity of interest is conceptually but I have some reservations about how it is measured, especially the part on row duplication. From what I understood, you duplicated rows for policies enacted over multiple years, with each row representing a different year so that each year can be considered independently.
- But I'm concerned that this ends up violating the independence of each policy in your dataset. After duplication, each row is not actually independent from one another and you need to control for those correlations which can bias standard errors. I believe you added weights which may have addressed this issue but some added explanation (maybe in an appendix) would help as this can change the results drastically.
- Some suggestions: would not duplicating rows be a concern? Other approaches to avoid this duplication issue, or to run as robustness, could be a survival analysis or time hazard model.
- I tried replicating your analysis without row duplication but keeping all the other steps and found that: (1) your interaction term is still null, (2) your gender term ends up being non-significant, (3) while the party control term ends up being significant. It does show that your main conclusion is robust but you probably should discuss this issue. (See table at end of document)

4 Extension - Explore Temporal rather than issue heterogeneity

- I think Table 2's analysis of policy responsiveness by issue area doesn't add to your main argument. I'm concerned about two things in that table: (a) the interaction term is slightly significant for economic policy which goes against your overall null finding, (b) you have flipped coefficient signs for model 4 compared to the others...which shows that there's a lot of noise when you splice the data up this way.
- Instead of running that, I went back to the original paper and found that they coded DV with up to a 4 year time lag while your paper uses a 2 year time lag. Why did they choose 4 and your

paper use 2? Instead of trying to justify that, I think explaining the variation across time could be interesting as well.

- I ran this analysis (without the row duplication) for 0 to 4 years of time lag and show the results in the table below.
- I think the significant results for 0 lag years and the drastic change when you introduce lags would be interesting to discuss.

Effect of Party Control on Gender Responsiveness Across Different Time Lags

	Dependent variable:				
	Policy Adoption	outcome_dv_1	outcome_dv_2	outcome_dv_3	outcome_dv_4
	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4
	(1)	(2)	(3)	(4)	(5)
Gender Difference	1.277***	0.138	0.045	0.057	0.000
	(0.408)	(0.333)	(0.225)	(0.111)	(0.000)
Party Control	-0.160***	-0.247***	-0.112***	-0.023	0.000
	(0.053)	(0.043)	(0.029)	(0.014)	(0.000)
Gender Difference * Party Control	-2.147***	-0.235	0.152	-0.266	-0.000
	(0.708)	(0.579)	(0.390)	(0.192)	(0.000)
Constant	0.592***	0.895***	0.967***	0.993***	1.000***
	(0.030)	(0.025)	(0.017)	(0.008)	(0.000)
Observations	709	709	709	709	709
R2	0.030	0.045	0.021	0.008	0.500
Adjusted R2	0.026	0.041	0.017	0.004	0.498
Residual Std. Error (df = 705)	0.493	0.403	0.272	0.134	0.000
F Statistic (df = 3; 705)	7.232***	11.112***	5.128***	1.971	235.055***
Note:					*p<0.1; **p<0.05; ***p<0.01

5 Suggestion - Run a power analysis?

- Could null findings simply mean that your data is not powered to capture any potential significant effects for the interaction term? I.e. there could be a gender * party interaction effect but you don't have the right sample size in your data to capture that.
- Perhaps you could run a power analysis to be able to specify the conditions under which you can rule this out.

6 Overall

This is a really fascinating set of results. It has a promising and punchy puzzle and key takeaway. My comments should be fairly quick fixes and will probably address themselves when you write out your methods section. Lastly, thank you all for the clarity and ease of replication!