INSTITUTE for REPLICATION

No. 69
I4R DISCUSSION PAPER SERIES

Talking: *Talking Shops*

Blake Lee-Whiting

Lewis Krashinsky

William Roelofs

September 2023

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 69

Talking: *Talking Shops*

Blake Lee-Whiting[1], Lewis Krashinsky[2], William Roelofs[1]

[1]University of Toronto/Canada
[2]Princeton University, Princeton/USA

SEPTEMBER 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

Talking: *Talking Shops*

Blake Lee-Whiting*, Lewis Krashinsky†, & William Roelofs‡

October 26, 2023

Abstract

In *Talking Shops: The Effects of Caucus Discussion on Policy Coalitions*, Zelizer analyzes the causal effect of caucus deliberations on legislative policy coalitions. In practice, political scientists have little empirical evidence on how policy discussions actually work among sitting legislators and whether these discussions have an effect on policy making and policy opinion. Taking on this challenge, Zelizer conducted two field experiments in an American state legislature. In short, the experiments randomized whether a bill was selected for discussion among a bi-partisan legislative caucus. The paper then measures and reports the corresponding effects of that discussion around the bill. Zelizer finds that deliberation increased the amount of co-sponsorship for a given bill, among both co-partisans and counter-partisans, but deliberation did not effect whether a bill was passed by the legislature or whether the bill received more amendments.

We conduct a robustness replication of the main results of *Talking Shops*. Specifically, we reproduce Tables 3 and 4 of the paper under alternative specifications. We find that the main results of the paper are reproducible and robust to multiple alternative specifications.

*University of Toronto. No conflict of interest or relationship with the original author. Corresponding Author's Email: Blake.lee.whiting@mail.utoronto.ca

†Princeton University. No conflict of interest or relationship with the original author. Email: krashinsky@princeton.edu

‡University of Toronto. No conflict of interest or relationship with the original author. Email: william.roelofs@utoronto.ca

1 Introduction

In *Talking Shops: The Effects of Caucus Discussion on Policy Coalitions*, Zelizer (2022) writes about the causal effect of policy deliberation among legislators on the final formation of policy coalitions. Group discussion, as Zelizer points out, is a pillar of modern representative democracy. In theory, legislators may shift their opinions on certain policies, or be more willing to compromise during a given debate, in response to the presentation of new information about a given policy area from a colleague in their legislature. Yet, we know little about how discussion and deliberation actually work in practice among sitting legislators. Furthermore, in the context of partisan polarization within American legislatures, there is a fundamental question of whether deliberation even exists. This is a challenging topic to study and one where there is no obvious method of data collection that stands above others. With a creative and original empirical design, however, Zelizer tackles this challenge and provides important insights on the effects of deliberation on policy coalitions and legislative behavior.

Zelizer's evidence comes from two field experiments that he launched in successive years within an American state legislature. Specifically, within a bipartisan caucus (the freshmen caucus for new legislators in the state), Zelizer effectively randomized whether or not a specific bill was brought up for discussion. Legislators themselves put forward a number of bills that they said they would be willing to present to the caucus and argue in support of, but the actual bills selected for deliberation were random. Accordingly, the paper then tests what effects this deliberation had relative to bills that were not discussed within the caucus.

In short, Zelizer has several major findings. First, deliberation in caucus increased the amount of co-sponsorship for a given bill. This result holds both for co-partisans and counter-partisans alike, showing that deliberation effects are not limited to within-party groups. Zelizer also shows however that bill-level outcomes are not effected by deliberation. Bills that were selected for deliberation were no more

likely to pass the legislature, nor receive more amendments.

2 Reproducibility

The work is computationally reproducible using the code and data provided by the original investigator . We should note that the work was published in the American Journal of Political Science (AJPS), which also conducts this reproducibility check.

3 Replication

We conducted a robustness reproduction of Zelizer's Talking Shops. Using the data provided by the author, we proceeded to test whether the main results of the paper stood up to various specifications and alternative constructions than the original empirical design. We first replicated Table 3 of the paper both as it was presented originally with weighted means and alternatively with unweighted means. Interestingly, we found that with unweighted means the evidence is even more supportive of Zelizer's finding that co-sponsorship increased when a bill was deliberated. Second, we then replicated Table 4 of the paper, but with alternative model specifications and alternative calculations of standard errors. Notably, none of these alternative specifications produced substantively different results than the original findings of the paper. Our replication validates the author's modeling decisions.

Note that in the Table 3 replication (Table 2) and in some of the Table 4 replications (Table 3, models 3 and 6), we drop the weights that Zelizer used. The weights account for the fact that different legislators introduced different numbers of bills and, thus, the probability that each bill was assigned to treatment varied by legislator. While the author rightfully included the weights to avoid bias (Gerber and Green 2012), we nonetheless drop them in a few model specifications to see whether the estimates and significance are impacted. We find no evidence that they are.

3.1 Replicating Table 3

Zelizer's main analysis of his field experiment begins with his third table. This table shows the weighted average co-sponsorship rates for the control and treatment groups of the legislators separately. Zelizer breaks down each of these tables by legislators who were counter-partisans or co-partisans, and those who attended the caucus meeting and those who did not.

We reproduced table 3, but instead of showing the results separately for control and treatment groups, we calculated the difference between the means for each different entry from the treatment and control group, respectively. Additionally, we conducted a two-sided t-test of the difference-in-means analysis, in order to assess the statistical significance of each difference, which was not done originally by Zelizer.

Below, we display Table 1 which reports the difference-in-means in weighted co-sponsorship rates between the control and treatment groups, by the same group breakdowns that Zelizer reports in the paper. Of note, we successfully replicated the exact results that Zelizer reports in the paper. The differences between the treatment and control groups in weighted co-sponsorship rates were clear and were much larger among those who actually attended the meetings. The results of the t-tests we conducted provide further support for the paper's findings. Out of the nine t-tests we conducted, in eight of the tests, we found that the difference-in-means was statistically significant. The only difference that was not significant was among those who did not attend the caucus meeting and who were counter-partisans, which makes clear intuitive sense and is consistent with the theory advanced by the author.

Table 1: Difference in means (weighted) between treatment and control groups

	Did not attend	Attended	Total
Counter-partisan	0.32	**6.55**	**1.33**
Co-partisan	**2.70**	**4.65**	**3.00**
Total	**1.66**	**5.53**	**2.26**
Bold indicates that $p < 0.05$			

We next reproduced Table 3 without weighting the observations. In other words,

we first calculated the simple average rate of cosponsorship for each of the subgroups and then calculated the difference in means between observations in treatment and observations in control. Table 2 shows the results of this analysis. The results are robust to Zelizer's original findings in Table 3. In fact, the difference between the treatment and control groups actually *increases* for each of the subgroups and the level of statistical significance is unchanged.

Table 2: Difference in means (unweighted) between treatment and control groups

	Did not attend	Attended	Total
Counter-partisan	1.15	**7.04**	**2.18**
Co-partisan	**4.25**	**6.13**	**4.53**
Total	**3.20**	**6.55**	**3.73**
Bold indicates that $p < 0.05$			

3.2 Replicating Table 4

We replicated Table 4 according to six models, each of which has similar and consistent results to Zelizer's analysis.

The first three models utilize intention to treat (treatment assignment 1 = assigned to treatment; 0 = assigned to control). Model 1 in Table 3 is a verbatim replication of the model in *Table 4*. Model 2 removes observations with pre-treatment co-sponsorship (1 = cosponsor; 0 = not cosponsor); the complete model was N=6,633, dropping pre-treatment co-sponsors removes 149 observations (N=6,484). Model 3 is a simple regression model which drops all of the controls: pre-treatment co-sponsorship, legislator effects, sponsorship effects, and weighting.

Models 4, 5, and 6 repeat the approaches above, substituting treatment assignment (z) for actual treatment administration (d). Model 4 reproduces the original model from *Table 4* using actual treatment administration. Model 5 removes observations with pre-treatment co-sponsorship. Model 6 is the same simple regression model as Model 3 which again removes controls: pre-treatment co-sponsorship, legislator effects, sponsorship effects, and weights.

Table 3

	Estimated Deliberation Effects on Cosponsorship (in pp)					
	1	**2**	**3**	**4**	**5**	**6**
Sponsor's Party: No Attended Meetings: No	0.4	0.4	1.1	0.8	0.8	1
(SE)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
Sponsor's Party: No Attended Meetings: Yes	5.9*	6.0*	7.0*	6.3*	6.3*	7.1*
(SE)	(2.9)	(2.9)	(2.9)	(2.9)	(2.9)	(2.9)
Sponsor's Party: No Total	1.3	1.3	2.2	1.7	1.7	2.0
(SE)	(1.4)	(1.4)	(1.3)	(1.4)	(1.4)	(1.3)
Sponsor's Party: Yes Attended Meetings: No	2.2	2.2	4.2	3.1*	3.2*	3.3
(SE)	(1.9)	(2.0)	(2.3)	(1.9)	(2.0)	(2.3)
Sponsor's Party: Yes Attended Meetings: Yes	4.2	4.4	6.1	5.2	5.5	7.4*
(SE)	(3.9)	(3.9)	(4.0)	(3.9)	(3.9)	(4.0)
Sponsor's Party: Yes Total	2.5	2.5	4.5	3.4*	3.5*	3.9
(SE)	(2.1)	(2.1)	(2.4)	(2.1)	(2.1)	(2.4)
Attended Meetings: No Total	1.5	1.4	3.2	2.1	2.2	2.6
(SE)	(1.5)	(1.5)	(1.6)	(1.5)	(1.5)	(1.6)
Attended Meetings: Yes Total	4.9*	5.1*	6.6*	5.6*	5.7*	7.3*
(SE)	3.0	(3.1)	(3.1)	(3.0)	(3.1)	(3.1)
Overall Total	2.0	2.0	3.7	2.7*	2.7*	3.3
	(1.6)	(1.6)	(1.7)	(1.6)	(1.6)	(1.7)

Note: Significance indicated at p <0.05 (*) one-sided.
Standard errors and p-values obtained from randomization inference with 1,000 simulated treatment assignments.
'Total' includes all observations in each row, column, or the entire table, respectively.

Finally, we reproduced Zelizer's Table 4 using alternative specifications and a new software package and functions in R. The coefficients Zelizer reports in his Table 4 are the same ones we report in the row *Treatment bill*. We include a separate model for each subset of the data, which correspond to the cells Zelizer uses in Table 4.

In Table 4, we fit the same models as Zelizer except we use the function feols from the fixest package and use heteroskedasticity-robust standard-errors. In Table

5, we use again use feols and cluster standard errors around each bill. Finally, in Table 6 we use the function glmer from the lme4 package to fit a model where legislator and bill sponsor are included as random effects rather than fixed effects.

In Tables 4 and 6 we find the difference between the treatment and control groups is statistically significant for all subsets of the data except one – legislators who are not in the same party as the sponsor and did not attend the caucus meeting. In Table 5, the standard errors are somewhat inflated and the difference is only statistically significant in three of the nine models. However, Zelizer only found significant results in two of the nine models, so our replication validates his modeling decisions.

Table 4: Zelizer's Table 4 models fit with heteroskedasticity-robust standard-errors

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
Treatment bill	0.0421*	0.0586*	0.0225*	0.0044	0.0255*	0.0131*	0.0146*	0.0494*	0.0200*
	(0.0174)	(0.0205)	(0.0063)	(0.0048)	(0.0060)	(0.0055)	(0.0041)	(0.0139)	(0.0042)
Original cosponsor	0.9750*	0.8264*	0.9272*	0.9658*	0.9274*	0.9508*	0.9431*	0.9326*	0.9394*
	(0.0362)	(0.0679)	(0.0133)	(0.0156)	(0.0124)	(0.0155)	(0.0093)	(0.0262)	(0.0087)
Fixed-effects									
sponsor	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
leg	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample									
Same party?	Yes	No	Yes	No	Yes	No	All	All	All
Attended?	Yes	Yes	No	No	All	All	No	Yes	All
Fit statistics									
Observations	556	475	3,160	2,442	3,716	2,917	5,602	1,031	6,633
R^2	0.36392	0.33659	0.50632	0.37729	0.47992	0.33689	0.46280	0.29678	0.42540
Within R^2	0.25530	0.13979	0.42146	0.30672	0.39640	0.25318	0.40051	0.20437	0.36103

Heteroskedasticity-robust standard-errors in parentheses
$^{*}p < 0.05$

Table 5: Zelizer's Table 4 models fit with standard-errors clustered on bills

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
Treatment bill	0.0421	0.0586*	0.0225	0.0044	0.0255	0.0131	0.0146	0.0494*	0.0200*
	(0.0260)	(0.0218)	(0.0129)	(0.0085)	(0.0138)	(0.0096)	(0.0101)	(0.0207)	(0.0042)
Original cosponsor	0.9750*	0.8264*	0.9272*	0.9658*	0.9274*	0.9508*	0.9431*	0.9326*	0.9394*
	(0.0372)	(0.0790)	(0.0248)	(0.0162)	(0.0253)	(0.0197)	(0.0182)	(0.0285)	(0.0087)
Sample									
Same party?	Yes	No	Yes	No	Yes	No	All	All	All
Attended?	Yes	Yes	No	No	All	All	No	Yes	All
Fixed-effects									
sponsor	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
leg	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics									
Observations	556	475	3,160	2,442	3,716	2,917	5,602	1,031	6,633
R^2	0.36392	0.33659	0.50632	0.37729	0.47992	0.33689	0.46280	0.29678	0.42540
Within R^2	0.25530	0.13979	0.42146	0.30672	0.39640	0.25318	0.40051	0.20437	0.36103

Bill clustered standard-errors in parentheses
*$p < 0.05$

Table 6: Zelizer's Table 4 models fit with random effects

	1	2	3	4	5	6	7	8	9
(Intercept)	0.04	0.04	0.03*	0.02*	0.03*	0.02*	0.02*	0.03*	0.03*
	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Treatment bill	0.04*	0.06*	0.02*	0.00	0.03*	0.01*	0.01*	0.05*	0.02*
	(0.02)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Original cosponsor	0.96*	0.86*	0.94*	0.97*	0.94*	0.96*	0.95*	0.93*	0.94*
	(0.07)	(0.10)	(0.02)	(0.03)	(0.02)	(0.03)	(0.01)	(0.06)	(0.01)
Sample									
Same party?	Yes	No	Yes	No	Yes	No	All	All	All
Attended?	Yes	Yes	No	No	All	All	No	Yes	All
Random effects									
sponsor	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
leg	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AIC	−78.02	1.93	−1574.91	−3014.93	−1666.42	−2586.48	−4198.01	−96.67	−4084.53
BIC	−52.09	26.91	−1538.56	−2980.12	−1629.10	−2550.61	−4158.22	−67.04	−4043.74
Log Likelihood	45.01	5.03	793.45	1513.46	839.21	1299.24	2105.00	54.33	2048.27
Num. obs.	556	475	3160	2442	3716	2917	5602	1031	6633
Num. groups: leg	29	29	117	117	137	137	117	29	137
Num. groups: sponsor	26	26	26	26	26	26	26	26	26
Var: leg (Intercept)	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Var: sponsor (Intercept)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Var: Residual	0.09	0.09	0.07	0.03	0.07	0.04	0.05	0.09	0.06

*$p < 0.05$

4 Conclusion

In summary, we found that the main findings of Zelizer's Talking Shops are robust to multiple alternative specifications. Generally, we found that the author was conservative in their estimation strategy. Some of our alternative specifications, including reproducing Table 3 with unweighted means or substituting intention to treat for actual treatment administration produced even stronger effects than those originally reported in the paper. We did not replicate estimated deliberation effects on roll call voting (*Table 6*), but this could be a starting point for other replicators.

References

Gerber, A. S. and Green, D. P.: 2012, Field experiments: Design, analysis, and interpretation, *W.W. Norton* .

Zelizer, A.: 2022, Talking shops: The effects of caucus discussion on policy coalitions, *American Journal of Political Science* **66**(4), 902–917.