ENDOGENOUS NETWORK FORMATION IN CONGRESS ONLINE APPENDIX

NATHAN CANEN AND FRANCESCO TREBBI

Appendix I. Online Appendix - Reduced Form Evidence

In this Appendix, we provide further reduced form evidence on the existence of positive impacts of networks (and influential cosponsors) on the approval of bills.

In our structural model, cosponsorships work as proxies for socialization, which is deemed to positively impact the probability of approving a bill. The model does not focus on bill quality or targeted cosponsors/socialization, which are often discussed in the political science literature.

Here, we connect these two by showing that socialization does impact bill approval, even when controlling for (ex-ante) bill quality, number of cosponsors and for multiple bills (potentially strategically introduced) by the same sponsor, validating the more parsimonious approach in the model.

We do so by using variations of identical bills across the Senate and the House of Representatives, where the bills are the same, but the only differences are the networks of support each bill has. We use two different specifications: one that looks at pairs of bills that are sponsored in the House and in the Senate by the same pair of sponsors, but where one bill is approved and the other is not in the House; and an identical bill which is approved in one chamber but not in the other.

We begin by looking at specifications which show the correlation between cosponsors of a bill and whether the bill is approved or not. In our model, cosponsorship can only help bill approval as it extends the network (endogenously formed).

A first approach to this problem is to test whether networks do impact approval of bills in Congress. To do so, we can check whether the number of cosponsors of a bill and the extended network of those cosponsors are relevant in explaining passing rates in Congress (as in our structural model). To do so, consider the following regression:

$$billpass_{i,k} = \beta_1 cosponsors_{i,k} + \beta_2 avgcospbill_{i,k} + X_i\gamma + \varepsilon_{i,k} \tag{I.1}$$

where *cosponsors* represents the number of cosponsors of bill k (proposed by sponsor i); and *avgcospbill* represents the average number of cosponsors that cosponsors of this bill have (in their own bills). The latter captures the influence, or additional order network effects of

Date: August 11, 2016.

those agents. X represents a series of politician level controls, such as the sponsor's ideology, tenure, party.

(I.1) implies that having additional cosponsors (captured by β_1) and those cosponsors being more influential/with larger networks (β_2) directly affects approval of legislation.

One may expect the OLS estimates of (I.1) to be inconsistent. First, it is possible that certain sponsors/politicians are more politically able and/or have better bills, and so would attract more cosponsors and better networks. In our model, higher types/returns α_i socialize more and have larger and more influntial networks, and hence would be observed to cosponsor more on average.

To control for that, consider the fixed effects regression:

$$billpass_{i,k} = \alpha_i + \beta_1 cosponsors_{i,k} + \beta_2 avgcospbill_{i,k} + \varepsilon_{i,k} \tag{I.2}$$

where α_i is a fixed effect for the politician who sponsors the bill. This effect captures the above problem, and would use the following variation: different bills by the same sponsor can have different number of cosponsors/extended network. The differences in their outcomes in Congress would then be attributed to the different (observed proxies for) networks.

A threat to identification in (I.2) is that we are not controlling for bill quality. The same sponsor can have some bills which are better than others, which by themselves attract more cosponsors. To deal with this issue, one can increase the set of controls, for instance focus on the specific characteristics of the Senate sponsor of the House bill.

Under this interpretation, consider the next specification:

$$billpass_{i,j,k} = \alpha_i + \gamma_j + \beta_1 cosponsors_{i,j,k} + \beta_2 avgcospbill_{i,j,k} + \varepsilon_{i,j,k} \tag{I.3}$$

where α_i, γ_j represents a fixed effect for the House sponsor (i) and Senate sponsor (j) pair. The bills studied here are those present in both chambers.

Our preferred specification further controls for bill type. Although the above intuitively should do so, there is still a threat that part of the bill quality is not being captured by having the same sponsors in both chambers.

For that reason, consider the within bill variation model:

$$billpass_{i,j,k,h} = \delta_k + \beta_1 cosponsors_{i,j,k,h} + \beta_2 avgcospbill_{i,j,k,h} + \varepsilon_{i,j,k,h} \tag{I.4}$$

$$billpass_{i,j,k,s} = \delta_k + \beta_1 cosponsors_{i,j,k,s} + \beta_2 avgcospbill_{i,j,k,s} + \varepsilon_{i,j,k,s} \tag{I.5}$$

In this version, we are using variation in outcomes for the identical bills across chambers (h for House, s for Senate). We posit that the same bill, if it faces different results in separate chambers, must have that due to differential (networks) supporting it. It cannot be coming from the bill quality, as it is the same bill in both scenarios. It cannot be coming from different politician abilities, as these are spanned by δ_k. The difference in outcomes is due to network effects.

Identification in (I.4)-(I.5) is due to the availability of bills that switch status across chambers.

I.1. **Data.** The data available is from the 93rd (1973-1975) to the 110th Congress (2007-2009), and is originally from the Library of Congress, and used in Fowler (2006). The data includes all bills (both House and Senate) in these periods, with data for the politicians in each Congress (such as tenure, party, ideology measure), the cosponsoring decisions for each bill in each Congress and Senate (i.e. who sponsored and cosponsored each one) and the outcomes for each (passed house, passed Senate, was vetoed or not, and so forth).

With this data, it is possible to construct network variables such as: the number of cosponsors for each bill, average number of cosponsors for a politician's own bills, a network graph using cosponsorship decisions as links. The focus is on House bills. Table 1 presents the summary statistics.

We will also use the definitions of identical bills in the Senate, as defined by the Library of Congress. This is done by checking for identical bills in the Senate (under related bills) for all house bills in Congresses 93-110. Table 2 shows that there are bills that switch status across chambers, which is key to our identification. These constitute around 20% of the sample.

I.2. **Results.** Table 3 presents the results across our various specifications (I.1), (I.2), (I.3) and (I.4)-(I.5).

As can be seen, the estimates of β_1 and β_2 are positive and very significant across specifications (OLS, OLS with controls, House Sponsor Fixed Effects, House and Senate Sponsor Fixed Effects and within bill variation). The number of cosponsors positively impacts the approval of bills. So does their influence through the congressional network.

The estimate of β_1 is between 0.0003 and 0.0005. This represents that an additional cosponsor increases the probability of approval (directly) by 0.05%. This is a small, but non negligible amount, as bills usually have many cosponsors. The coefficient for β_2 amplifies this effect, and is estimated to be around 3 times as large as β_1 (in Columns (1)-(4)). This implies that adding a cosponsor who has on average 10 cosponsors on his own bill, leads to an average increase of $0.000541 + 10 \times 0.00144 = 0.0149$, or a 1.49 point increase in the percentage probability of approval.

Table 4 allows for heterogeneity in the effects for the House and the Senate, for the specification of (I.4)-(I.5). The results confirm the positive and significant effects in the House, and shows that the influence term β_2 is really important in the House, although not so much from the Senate, which presents noisy estimates.

I.2.1. *Discussion of alternative models: the Co-authorship Model.* Our results indicate that it is advantageous to have additional cosponsors. In the context of the structural model, this means there are gains in having larger networks and more connections. We should hence, observe denser networks in Congress. This seems to be the case in our structural model. It also seems to hold in evidence in Fowler (2006) and Cho and Fowler (2010). This suggests that models with sparse equilibrium interconnections would not provide a good fit for Congressional activity.

For example, Jackson and Wolinsky (1996), Jackson (2008) present a co-authorship model. There, agents have a number of projects. They are said to be linked to other agents if they co-author those.

Agents have a set availability of time, and the less projects they co-author, the more time/effort they can devote to their own works. Adapting to our scenario, less cosponsorships by members he is linked to imply more time spent by them in his own bill. We assume that more effort into a bill enhances its approval rate.

Therefore, one could interpret this as a competing mechanism to the complementarity in production of socialization and effort of those in your network. There is synergy when working together. However, on a network, agents prefer to focus on their own bills. The more bills your links cosponsor, the less time they have on your own bill. The more bills the sponsor chooses to be in, the less time (or effort) he puts into his own. It follows that this model proposes negative relationships between number of cosponsors and passing of rates of the bills in Congress, and between the extended influence/participation of one's cosponsors in other projects and the approval of a bill.

As could be seen in the reduced form evidence, this is not the case. Clearly there is a positive and significant relation between number of cosponsors and the passing of those bills.

These results reject the co-authorship model. Finally, the density of the network would be severely inefficient in this world: as shown in Jackson and Wolinsky (1996), it would be efficient to have isolated pairs of nodes. This is not often observed in Congress.

All of these point towards sustaining the model of complementarity of efforts of your network as support for your own bill against the co-authorship model.

References

Cho, W. K. T. and Fowler, J. H. (2010). Legislative success in a small world: Social network analysis and the dynamics of congressional legislation. *The Journal of Politics*, 72(01):124–135. 4

Fowler, J. H. (2006). Connecting the congress: A study of cosponsorship networks. *Political Analysis*, 14(4):456–487. 3, 4

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press, Princeton, NJ, USA. 4

Jackson, M. O. and Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74. 4

Table 1. (Appendix) Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Pass	139021	0.077	0.267	0	1
Party	138986	60.32% Democrat			
		39.37 % Republican			
Ideology	137426	-0.069	0.388	-.757	1.685
Tenure	138986	5.974	4.096	1	27
Number of cosponsors	139021	10.311	27.084	0	406
Avg. cosponsors of cosp.	139021	6.239	8.55	0	175

Table 2. (Appendix) Bills with "Switching" Outcomes

	(1) Outcome $\Delta_{h-s} pass$	(2) Frequency	(3) Percent
Panel A: All identical bills			
	Pass Senate, Not Pass in House	1,073	8.30
	Same Outcome in Both	10,478	81.02
	Pass House, Not Pass in Senate	1,381	10.68
N: 12852			
Panel B:$ncosponsors > 0$ in both			
	Pass Senate, Not Pass in House	356	7.01
	Same Outcome in Both	4,077	80.33
	Pass House, Not Pass in Senate	642	12.65
N: 5045			

Panel A: All bills with paired observations. Panel B: Only those with number of cosponsors bigger than zero in both the House and Senate.

Table 3. (Appendix) Main Results

	(1) OLS	(2) OLS with Controls	(3) House Sp. FE	(4) House+Sen.Sp. FE	(5) Within bill
N. of Co-sp.	0.000589***	0.000554***	0.000594***	0.000541***	0.000339***
	(0.0000540)	(0.0000547)	(0.0000527)	(0.000107)	(0.0000713)
Avg.cosp. of cosp.	0.00275***	0.00140***	0.00227***	0.00144***	-0.0000591
	(0.000214)	(0.000209)	(0.000208)	(0.000547)	(0.000350)
Constant	0.0536***	-0.0742***	0.0566***	0.0646***	0.114***
	(0.00250)	(0.00772)	(0.00135)	(0.0105)	(0.00326)
N	137703	137426	137703	12852	12932
R^2	0.015	0.035	0.010	0.042	0.002

Standard errors in parentheses, clustered at the House Sponsor level (first 4 columns) and Senate Sponsor (Column (5), due to lack of data to cluster at the House sponsor). Individual controls in Column (2) include Tenure, Party, Ideology and Congress.* $p < .1$, ** $p < .05$, *** $p < .01$. The first column is the OLS regression, the second puts controls (described above), the third is fixed effects at the House Sponsor level, the fourth has fixed effects of both House and Senate sponsor. Column (5) is the specification with within bill variation. N for Column (5) is the number of bills we have pairs of observations. It is larger than (4) because it does not use information on the id code of the sponsor in the House.

Table 4. (Appendix) Heterogeneity

	(1) pass	(2) pass	(3) pass	(4) pass
Congress Outcome (Indicator)	0.0244***	-0.0335***	0.0682***	-0.0324*
	(0.00505)	(0.00705)	(0.00716)	(0.0186)
N. Cosponsors	0.000304***	-0.000173	0.000286***	-0.000157
	(0.0000706)	(0.000108)	(0.0000845)	(0.000130)
Avg Co-sp. of cosponsors	0.000352	-0.00125***	0.00222***	-0.000183
	(0.000352)	(0.000381)	(0.000718)	(0.000819)
Interaction: House × N.Cosponsors		0.000722***		0.000713***
		(0.000165)		(0.000207)
Interaction: House × Avg.Cosponsors		0.00482***		0.00469***
		(0.000550)		(0.000969)
Constant	0.0981***	0.124***	0.0471***	0.107***
	(0.00428)	(0.00473)	(0.0143)	(0.0165)
N	12932	12932	5075	5075
R^2	0.005	0.017	0.020	0.029

Standard errors in parentheses, clustered at the senate sponsor level. Tests reject the hypothesis that the coefficients of the interactions are the same as those without. Columns (1) and (2) focus on all bills with paired observations. Columns (3) and (4) only on bills with positive number of cosponsors in both the House and the Senate. N is the number of bills (each bill has 2 observations). * $p < .1$, ** $p < .05$, *** $p < .01$