In the following document we describe the data and code included in our replication archive, and lay out the steps needed to run all of the code.

**processed\_data:** This folder contains data objects that are used at various stages of the analysis.

1. cleaned\_nodes\_R1. Rds: Node level information on the 4108 legislators in the analysis. Columns include the Twitter ID, Twitter Screen Name, State, Name, Chamber, Party, Race, Gender”, and a Measure of professionalism of the legislators.
2. followers\_adjacencyMatrix.Rds: Adjacency matrix for the Follower network
3. mentions\_adjacencyMatrix.Rds: Adjacency matrix for the Mentions network
4. retweets\_adjacencyMatrix.Rds: Adjacency matrix for the Retweets network
5. state\_similarity\_matrix.Rds: Adjacency matrix indicating if legislators are from the same state
6. party\_similarity\_matrix.Rds: Adjacency matrix indicating if legislators are from the same party
7. chamber\_similarity\_matrix.Rds: Adjacency matrix indicating if legislators are from the same chamber
8. gender\_similarity\_matrix.Rds: Adjacency matrix indicating if legislators are of the same gender
9. same race.Rds: Adjacency matrix indicating if legislators are of the same race
10. contig\_states\_matrix.Rds: Indicator matrix for state contiguity
11. gender\_receiver\_matrix.Rds: Indicator matrix for legislators gender
12. republican\_receiver\_effect.Rds: Indicator matrix for if the legislator receiving a tie is a republican or not
13. democrat\_receiver\_effect.Rds: Indicator matrix for if the legislator receiving a tie is a democrat or not
14. receiver\_profesh.Rds: Matrix with state legislature professionalism scores of the tie receiving legisltor.
15. sender\_profesh.Rds: Matrix with state legislature professionalism scores of the tie sending legislator.
16. receiver\_black.Rds, receiver\_asian.Rds, receiver\_mena.Rds, receiver\_latino.Rds, receiver\_multi, receiver\_native.Rds: Indicator matrices for the tie receiving legislators race.
17. senderer\_black.Rds, sender\_asian.Rds, sender\_mena.Rds, sender\_latino.Rds, sender\_multi, sender\_native.Rds: Indicator matrices for the tie sending legislators’ race.
18. gender\_sender\_matrix: Matrix indicating the gender of the tie sending legislator (female = 1, male=0).
19. gender\_receiver\_matrix: Matrix indicating the gender of the tie receiving legislator.
20. house\_sender\_matrix: Matrix indicating the chamber of the sending legislator (house=1, senate=0)
21. house\_receiver\_matrix: Matrix indicating the chamber of the receiving legislator (house=1, senate=0)
22. profeshScore\_diff.Rds: Adjacency matrix indicating the difference in 2 legislators state legislature professionalism scores.

**Code:** In the following list we describe the scripts that are run, and the order in which they should be run.

1. make\_predicting\_matrix\_for\_QAP: Script to create a single predicting matrix comprising of all the independent variables in the analysis. This script should be run first.
2. network\_descripStats: Create the tables 1 and 2 in the paper (lists the top ten legisltors with the highest in degree and out degree centrality across the 3 networks. Can be run separate from the other scripts.
3. network\_graph\_plots: Creates the network visualizations for all 3 networks. Can be run separate from the other scripts.
4. The QAP folder contains all the scripts which were run on the HPC to create the sampling distributions along with the pbs scripts to run the jobs:
   1. The folders with subscript “DiffState\_analysis\_scripts” use the different state matrix in the interaction (0=legislators are from the same state, 1=legislators are from different states). All scripts in this group can be run in parallel with each other and with those in group (b).
   2. The folders with subscript “SameState\_analysis\_scripts” use the same state matrix in the interaction (0=legislators are from different states, 1=legislators are from the same state)
   3. QAP\_outputs stores the results plots and results tables
   4. QAP\_extract\_dist\_coef\_pval.R script reads in all the data from the QAP runs, creates a sampling distribution and calculates p values for each network. This should be run after the scripts in groups (a) and (b).
   5. QAP\_output\_plots.R: creates the results tables (Figures 4-6 in the paper). This should be run after the scripts in groups (a), (b), and (d).
5. The appendix folder contains scripts and data necessary to reproduce results reported in the appendix.
   1. retweet\_network\_sim.R, follower\_network\_sim.R, mention\_network\_sim.R: Simulate networks for QAP power analysis. Should be run first in this group.
   2. qap\_sim\_mention\*.R, qap\_sim\_retweet\*.R, run\_qap\_on\_sim\_follow\*.R: Run QAP on the simulated networks for power analysis. Should be run after scripts in group (a).
   3. analyze\_power\_results\*.R: summarizes results from power simulations. Should be run after scripts in group (b).
   4. tweet\_trends\_plot.R: plots the overall timeline of tweets posted by state legislators. Calls a data file named nTweetsPerDay.RData. Can be run independent of the other scripts in this folder.
   5. party\_state\_descriptive.R: makes a descriptive table of the cross/within-state and cross/within-party patterns in the data.