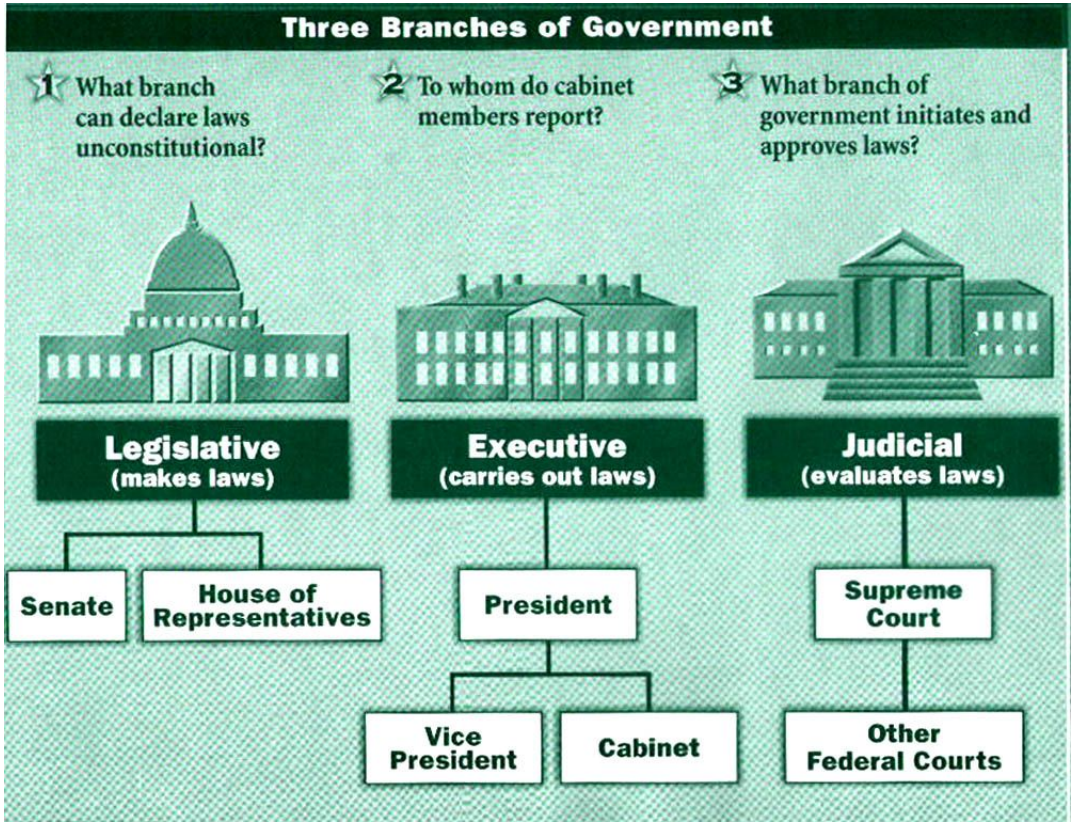
# Does Popularity Matter?

## Analysis of cosponsorship and election results in the US Senate

Unni Kurumbail,  Siyu Nan, Matt Yetto

# Quick Intro to Our Highly Efficient Government




- We'll focus on the legislative branch
- Responsible for writing legislation and presenting it to the President for approval
- Senate- Comprised of 100 members; 2 per state, voted into office by citizens
- Typically senators hail from the 2 major political parties

Image source: http://historyfor353.weebly.com/

# Why senate?

- Senators can co-sponsor each other's legislation as a show of support ➔ Creates an interesting network
- Co-sponsor a lot; 52-83 unique co-sponsors per legislator from 93rd-108th senate (1)
- Filibusters, long-term relationships, and the amendment process force collaboration in the senate (2)
    - How do these processes affect the network?
- Many bills (10000+) per 2-year term ➔ Lots of data points


(1) Fowler, James H. "Legislative Cosponsorship Networks in the US House and Senate." *Social Networks* 28.4 (2006): 454-65. *Elsevier*. Web.

(2) Rippere, Paulina S. "Polarization Reconsidered: Bipartisan Cooperation through Bill Cosponsorship." *Polity* 48.2 (2016): 243-78. Web. 9 Oct. 2016.

# Fundamental Questions

- What's the connection between the typical number of cosponsorships a senator gains for his/her bill and the senator's election success?

- Do bipartisan cosponsorships (cosponsorships between a member of one party and one of  the other) matter in particular?

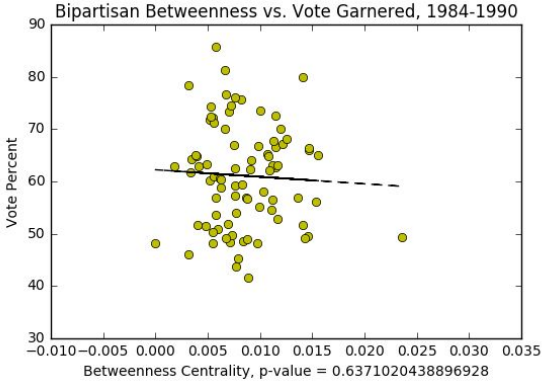
- Prediction: More central senators are more well connected and therefore are more likely to get elected because of more endorsements and senate activity

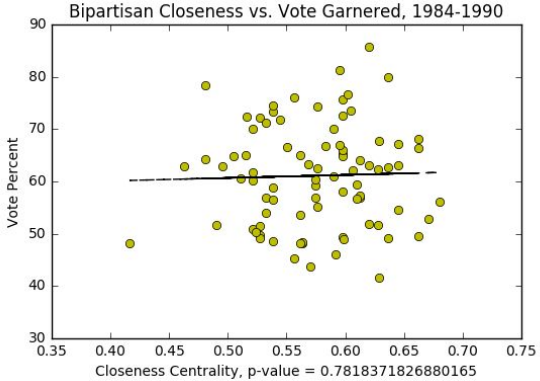
# Network Centrality Analysis

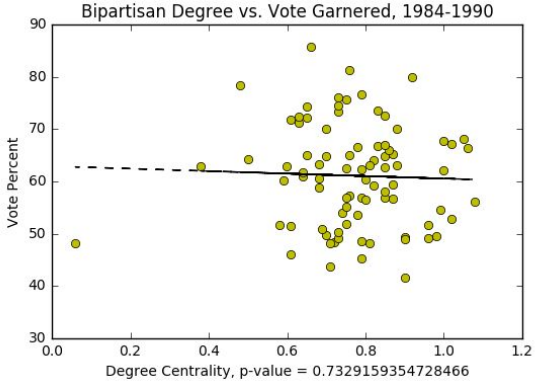
Methods:

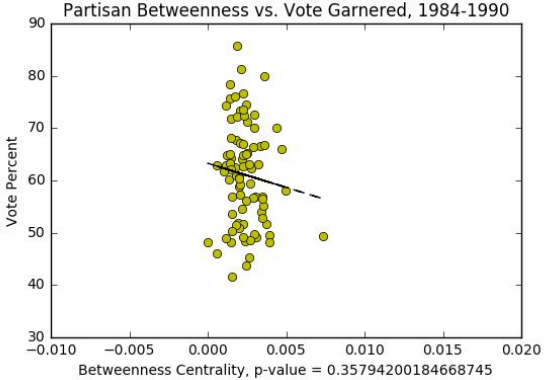
- Collect cosponsorship data for 98th through 110th congresses
- For each congress, construct directed cosponsorship network, whole and bipartisan variants
- Calculate Betweenness Centrality, Closeness Centrality, and Degree Centrality
- Collect vote percentage data from Federal Election Commission Archive
- Plot each centrality against percentage of vote garnered in the senator's election year.  Three 2-year terms are combined on each plot
- Run linear regression on each plot

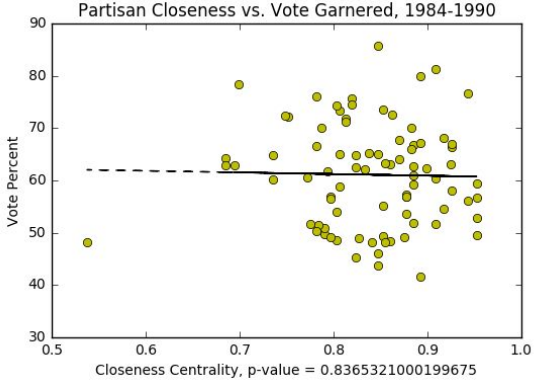
# 98th, 99th, 100th Congresses

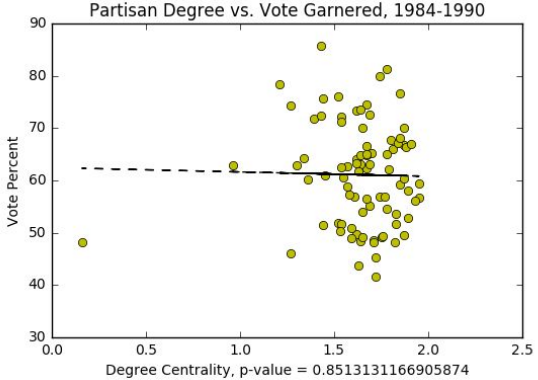


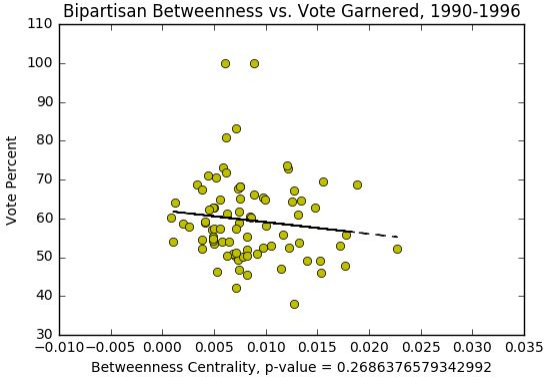



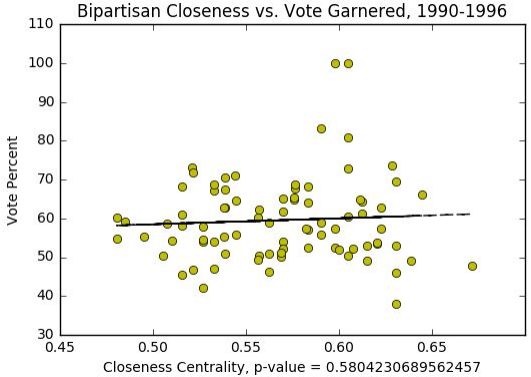



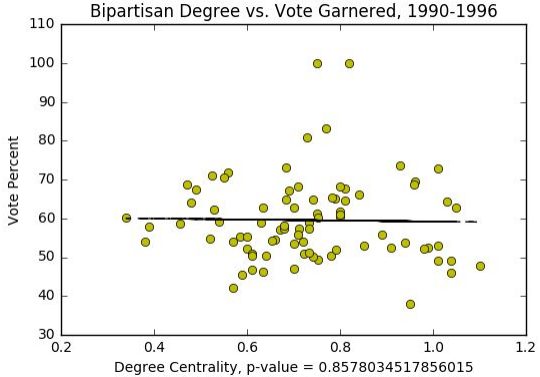



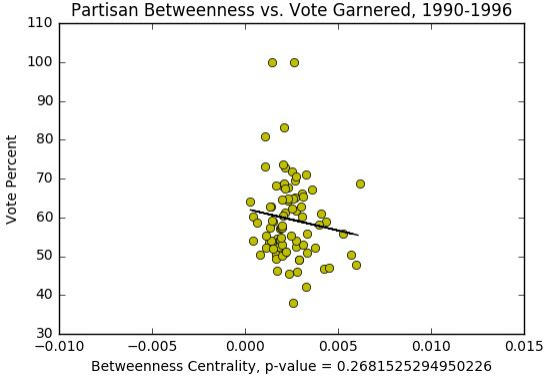



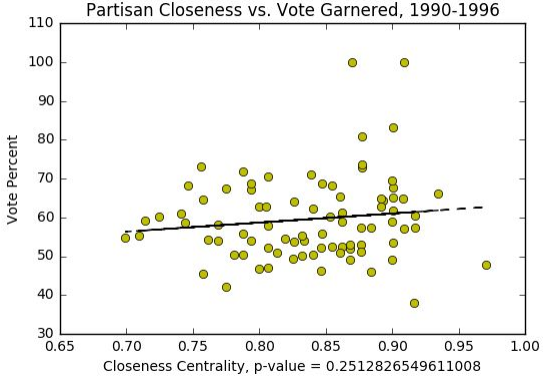
# 101st, 102nd, 103rd Congress

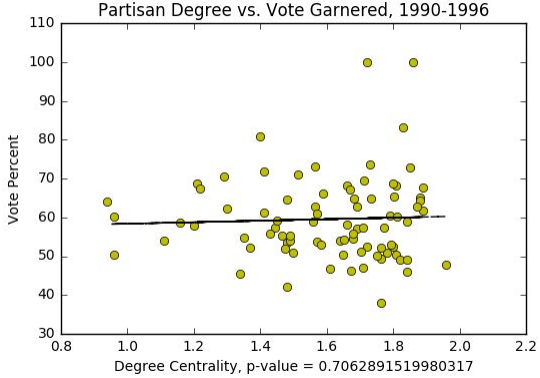


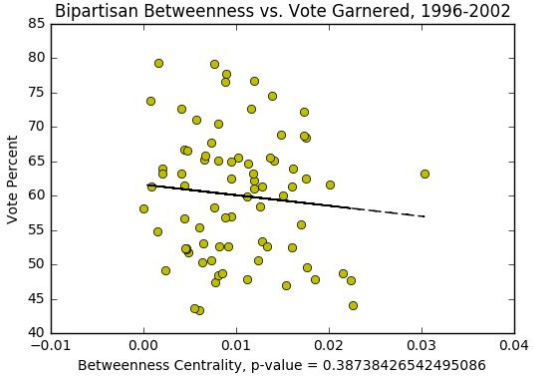



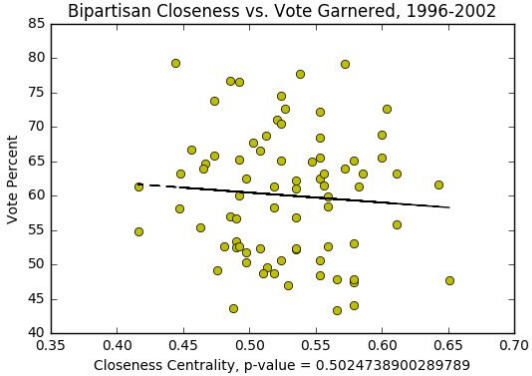



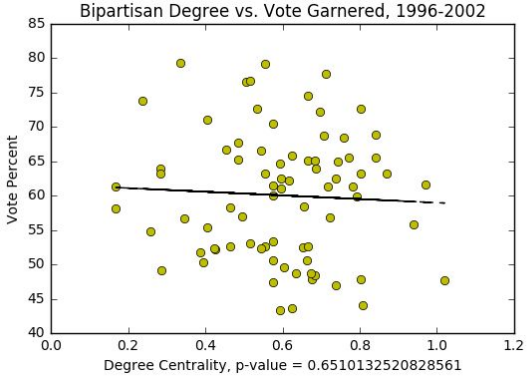



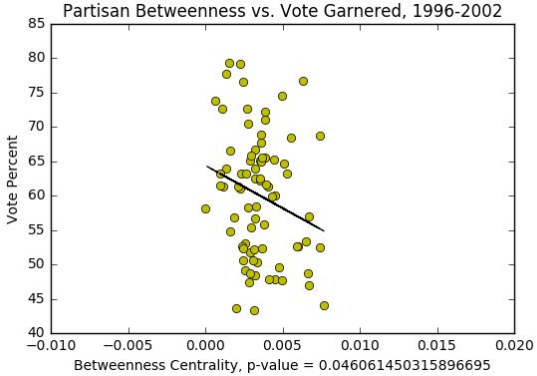



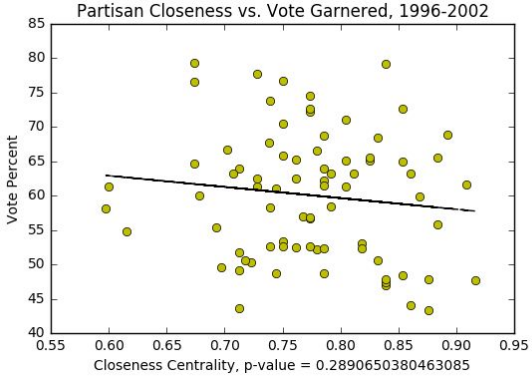
# 104th, 105th, 106th Congress

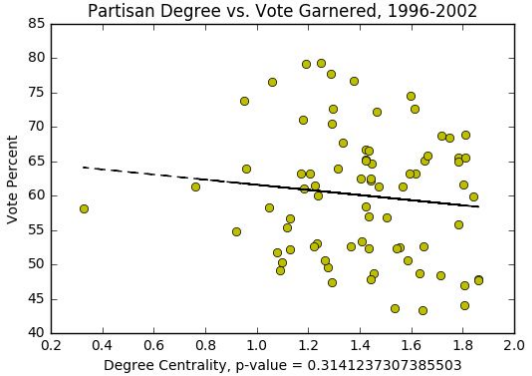


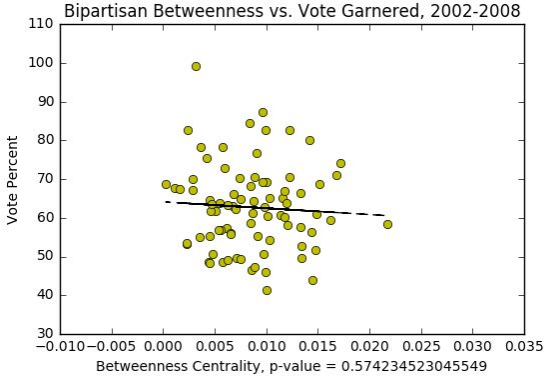



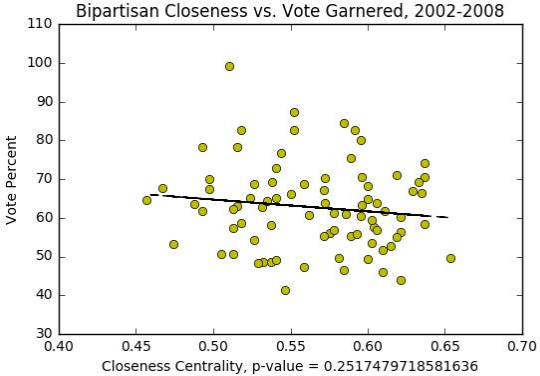



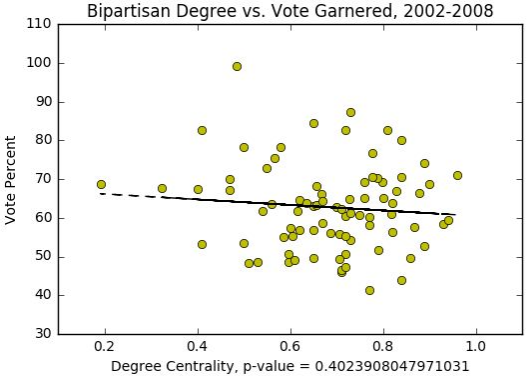



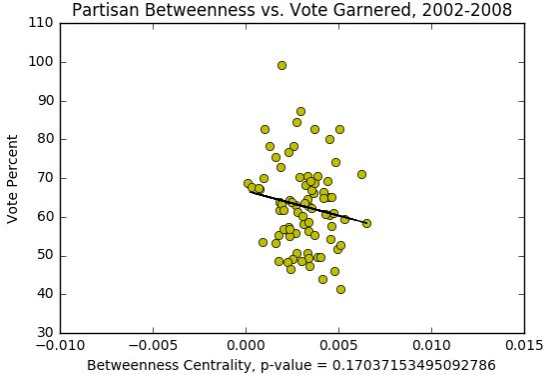



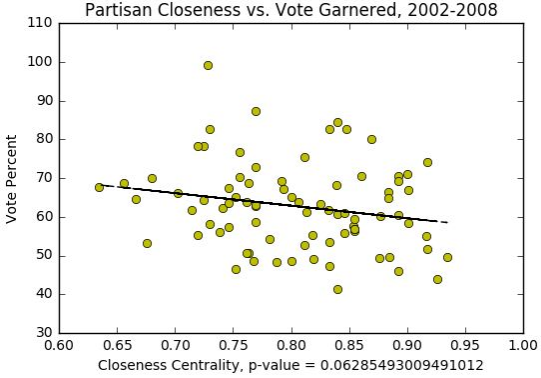
# 107th, 108th, 109th Congress

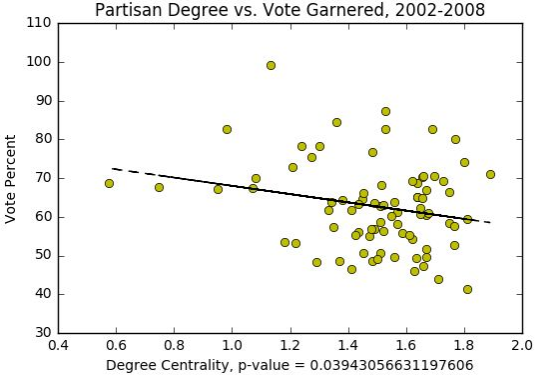


# Network Centrality Analysis

In (almost) all plots, there is no significant correlation between the centrality measure and vote percentage

Possible explanations without debunking centrality as a vote success marker:

- Noise in data due to structure of dataset
- Strongly red and strongly blue states’ voting tendency outway effect from centrality on voting turnout

# Methods to Test Problems

To reduce noise:

- Create a threshold for cosponsorship between two senators
- Reconstruct directed network, recalculate centrality measures

To reduce effect of strongly red or strongly blue states:

- Compare centrality measure to binary win/loss rather than continuous vote percentage

# Idea of threshold

# Data Distribution--Bipartisan(98th congress)

- 65% of the elements are 0.

| Cosponsorship number | Percentage |
|---|---|
| 1-2 | 52% |
| 3-4 | 25% |
| >4 | 23% |
| >10 | 3.4% |

⇒ Implement a threshold to de-noise

# Original

- Closeness centrality: y= -38x+86 (p=0.36, r=-0.18)
- Betweenness centrality: y= -749x+71 (p=0.23, r=-0.23)
- Degree centrality: y= -21x+81 (p=0.18, r=-0.26)

# Threshold

- One or two co-sponsorship from one senator to another doesn’t give us much information, those may be noise for our centrality measurement.
- 3 as Threshold: Change all elements 1 or 2 in our matrix into 0
- 5 as Threshold: Change all elements 1 to 4 in our matrix into 0
- Calculate Closeness centrality, betweenness centrality and degree centrality with new matrix
- Run regression x: centrality  y: election result

# 3 as Threshold

- Closeness centrality: y= -56.2x+89.42 (p=0.24, r=-0.23)
- Betweenness centrality: y= -130.58x+66.4 (p=0.44, r=-0.15)
- Degree centrality: y= -17.56x+71.1 (p=0.18, r=-0.26)
- All p-value are too high ---- no relation

| | | | | |
|---|---|---|---|---|
| 10 | 5 | 2 | 4 | 1 |
| 11 | 0 | 1 | 1 | 2 |
| 12 | 7 | 16 | 1 | 2 |
| 13 | 2 | 0 | 0 | 68 |

| | | | | |
|---|---|---|---|---|
| 10 | 5 | 0 | 4 | 0 |
| 11 | 0 | 0 | 0 | 0 |
| 12 | 7 | 16 | 0 | 0 |
| 13 | 0 | 0 | 0 | 68 |

# 5 as Threshold

- Closeness centrality: y= -18.8x+70.27 (p=0.53, r=-0.12)
- Betweenness centrality: y= -97.46x+66.22 (p=0.32, r=-0.19)
- Degree centrality: y= -26.64x+69.6(p=0.12, r=-0.29)
- All p-value are too high ---- no relation

| | | | | |
|---|---|---|---|---|
| 10 | 5 | 2 | 4 | 1 |
| 11 | 0 | 1 | 1 | 2 |
| 12 | 7 | 16 | 1 | 2 |
| 13 | 2 | 0 | 0 | 68 |

| | | | | |
|---|---|---|---|---|
| 10 | 5 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 |
| 12 | 7 | 16 | 0 | 0 |
| 13 | 0 | 0 | 0 | 68 |

# Data Distribution--Partisan(98th congress)

- 24% of the elements are 0.

| Cosponsorship number | Percentage |
|---|---|
| 1-2 | 46% |
| 3-4 | 26% |
| >4 | 28% |
| >10 | 4.7% |

⇒ Implement a threshold to de-noise

# Original

- Closeness centrality: y= -56x+111 (p=0.07, r=-0.34)
- Betweenness centrality: y= -3876x+74 (p=0.07, r=-0.34)
- Degree centrality: y= -23x+102 (p=0.05, r=-0.37)

# 3/5 as Threshold

3:

- Closeness centrality: y= -76x+113 (p=0.03, r=-0.39)
- Betweenness centrality: y= -423x+68 (p=0.12, r=-0.29)
- Degree centrality: y= -12x+76 (p=0.06, r=-0.35)

5:

- Closeness centrality: y= -81x+104 (p=0.05, r=-0.37)
- Betweenness centrality: y= -147x+66 (p=0.19, r=-0.25)
- Degree centrality: y= -16x+72 (p=0.03, r=-0.40)

⇒ Is it the case for all 13 congresses we have?

# Regression on 6 years (3 terms)

- **Unfortunately adding the threshold doesn't work on the 6-year time frame.**
- Implemented both 3 and 5 co-sponsors as thresholds; **only 3 regressions (out of 48) have a significant p-value,** the same plots as those that had a significant p value before threshold.

- In fact, **only 98th congress has multiple significant p values in regression**. Other than that, 110th congress has a significant p value in degree centrality. All the rest have p-value > 0.1 for all three measurements' regression.

**⇒ Threshold idea doesn't work. Data distribution is similar through all time but there's no specific threshold that works for all datasets as a denoiser.**

# Win/Loss Analysis

# Win/Loss Analysis

- Do centrality measures correlate with winning/losing the election?
  - In other words, do more connected senators end up winning more elections?

- First step: Do bipartisan cosponsorships correlate with election results?

# Win/Loss Analysis: Wilcoxon Sum of Ranks

Bipartisan cosponsorships:

Cosponsorships to the other party only

number of losers:  44
number of winners:  312
closeness centrality p_value: 0.1714
betweenness centrality p_value: 0.8078
degree centrality p_value: 0.1946
in degree centrality p_value: 0.2683
out degree centrality p_value: 0.3809

Two groups: election winners and election losers

# Win/Loss Analysis

- Do centrality measures correlate with winning/losing the election?
  - In other words, do more connected senators end up winning more elections?
- First step: Do bipartisan cosponsorships correlate with election results?
  - **No correlation**
- What about overall cosponsorships?

# Win/Loss Analysis: Wilcoxon Sum of Ranks

## Bipartisan cosponsorships:

(Cosponsorships to the other party only)

number of losers: 44
number of winners: 312
closeness centrality p_value: 0.1714
betweenness centrality p_value: 0.8078
degree centrality p_value: 0.1946
in degree centrality p_value: 0.2683
out degree centrality p_value: 0.3809

## Overall Cosponsorships

(ignoring party)

number of losers: 44
number of winners: 312
closeness centrality p_value: 0.0885
betweenness centrality p_value: 0.8812
degree centrality p_value: 0.0214
in degree centrality p_value: 0.0248
out degree centrality p_value: 0.0885

Two groups: election winners and election losers

# Win/Loss Analysis

- Do centrality measures correlate with winning/losing the election?
  - In other words, do more connected senators end up winning more elections?
- First step: Do bipartisan cosponsorships correlate with election results?
  - No correlation
- What about overall cosponsorships?
  - **Significant relationship between degree centrality and whether a senator wins or loses**

- Is the effect of degree centrality on election results because of a senator's party?

# Win/Loss Analysis: Wilcoxon Sum of Ranks

## Bipartisan Cosponsorships

number of democrats: 652
number of republicans: 660
closeness centrality p_value: 0.00000014
betweenness centrality p_value: 0.1842
degree centrality p_value: 0.7221
in degree centrality p_value: 0.00000018
out degree centrality p_value: 0.00000001

## Overall Cosponsorships

number of democrats: 652
number of republicans: 660
closeness centrality p_value: 0.0048
betweenness centrality p_value: 0.4078
degree centrality p_value: 0.5474
in degree centrality p_value: 0.0000
out degree centrality p_value: 0.0048

Two groups: Democrats and Republicans

# Win/Loss Analysis

- Do centrality measures correlate with winning/losing the election?
  - In other words, do more connected senators end up winning more elections?
- First step: Do bipartisan cosponsorships correlate with election results?
  - No correlation
- What about overall cosponsorships?
  - Significant relationship between degree/closeness centrality and whether a senator wins or loses
- Is the effect of degree centrality on election results mediated by a senator's party?
  - **Maybe. Use multiple linear regression to explore this**

# Win/Loss Analysis: OLS Multiple Regression

- Win/loss in election significantly dependent on party *and* degree centrality independently

| Variable | Value | Std Error | P-Value |
|---|---|---|---|
| **Party** | 0.1171 | 0.034 | 0.001 |
| **Degree Centrality** | -0.1346 | 0.062 | 0.030 |
| **const** | 1.0385 | 0.102 | 0.000 |

Source Code for regression: http://stackoverflow.com/questions/21234539/statsmodels-ols-function-for-multiple-regression-parameters

# Win/Loss Analysis: OLS Multiple Regression

- Win/loss in election significantly dependent on party *and* in-degree centrality independently

| Variable | Value | Std Error | P-Value |
|---|---|---|---|
| **Party** | 0.1236 | 0.035 | 0.000 |
| **In-Degree Centrality** | -0.3224 | 0.138 | 0.020 |
| **const** | 1.0769 | 0.111 | 0.000 |

Source Code for regression: http://stackoverflow.com/questions/21234539/statsmodels-ols-function-for-multiple-regression-parameters

# Conclusions/Future Study

- Overall connection between centrality and election success is weak: tough to find direct correlation between data and election success (lots of noise)
- In-degree centrality *negatively* correlated with success in elections
  - Surprising: more cosponsors on one's bills associated with poorer election results
  - Perhaps people who actively seek cosponsors are trying harder because they are not doing as well in the polls?
- In-degree centrality correlation with election success is not due to party affiliation (some other mechanism at play)
- Future studies: Analyze late-term vs. early-term senators to see differences in centrality measures