

Lobbying in the 111th Congress (2009-2011)

Matthew Meyers

December 17th, 2023

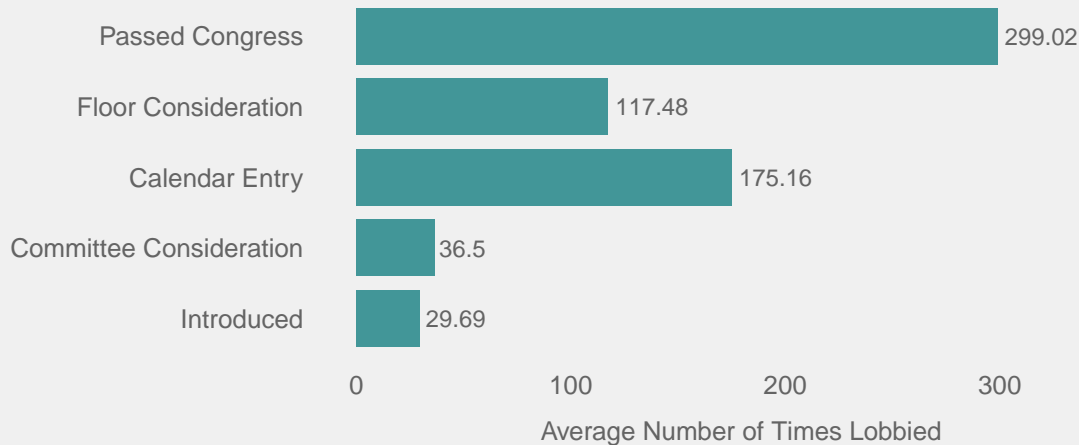
Abstract

This paper is an attempt to learn what types of bills draw the most interest from lobbyists and how lobbying data can be used to predict how far in the legislative process a bill will go. It uses the 111th Congress, which spans 2009 to 2011, as a case study for these questions. Data came from two sources. The non-profit organization OpenSecrets provided data about how frequently a bill was lobbied in the 111th Congress which they drew from Lobbying Disclosure Act Reports, and congress.gov provided information about the bills in the OpenSecrets data including the legislative actions, the topics, Congressional Research Services Summaries, sponsors, cosponsors, committees, and related bills. To understand what types of bills were most lobbied, models where the amount of times a bill was lobbied was the response variable were fit. One such model modeled the log counts of lobbying with a linear model while another modeled the raw counts of lobbying with a negative binomial model. To understand which types of bills advance the furthest in Congress, an ordinal logistic models with the most advanced state of the bill was fit. In all three models, the LASSO was used in order to perform variable selection. Bills that advanced further in Congress were more likely to be lobbied and bills with higher amounts of lobbying advanced further in Congress, but the direction of the causal relationship or relationships was unclear. The data also seems to suggest that issues of high salience tend to be heavily lobbied but less likely to advance in Congress and vice versa. Other issues, such as appropriations, were both heavily lobbied and frequently advanced in Congress.

Executive Summary

Lobbying in the 111th Congress varied widely with the most lobbied bills being lobbied up to 10,747 times with an average number of times lobbied of 53.64352. The most lobbied bills included several of the most high profile legislative battles of the 111th Congress including the American Recovery and Reinvestment Act, the American Clean Energy and Security Act, the Affordable Care Act, and the Dodd-Frank Wall Street Reform Act. The bills that were lobbied the most on average were economics bills, environmental bills, bills from the Appropriations Committees, Bills from the House Budget Committee, and bills from the Banking, Housing, and Urban Affairs Act. Bills that advanced further in Congress were lobbied more frequently, as the chart below demonstrates; however, this paper was not able to determine whether lobbying causes bills to advance, whether bills advancing causes them to be lobbied more, both, or whether there is an as-yet unknown underlying variable.

Average Number of Times Lobbied by Bill State



This paper fit three different models that predicted two different response variables. Two types of models were fit in order to predict the number of times a bill was lobbied and another was fit to predict the latest step of the legislative process that a bill would reach.

The two models to predict the number of times that a bill was lobbied had similar coefficients for the predictive variables and made similar predictions about the number of times a bill would be lobbied. In the selected model, the most significant legislative state was passing committee, suggesting that the step from receiving consideration in a committee to passing committee is either one that catches the attention of lobbyists or one that additional influence from lobbyists can prove crucial to a bill achieving. The issue areas with the highest predictors in the selected model, in order, were commerce; finance and financial sector; Environmental protection; science, technology, and communications; and energy. The committees with the highest positive predictors of lobbying were Appropriations; Budget; Agriculture, Nutrition and Forestry; Finance; and Health, Education, Labor, and Pensions. Other variables that corresponded with more lobbying were more associated legislative subjects, more cosponsors, an earlier introduction date, more related bills, and being introduced by the Democratic Party. The cross validated root mean square error of the selected model was 261.7.

In the model for the final legislative state of the bill, more lobbying was associated with bills reaching a later state in the process. Log lobbying had a coefficient of 0.154, which means that increasing the log lobbying by 1 (equivalent to multiplying the total amount of lobbying by e) increases the odds of a bill reaching a later state by a factor of 1.17. Somewhat paradoxically, despite this relationship, many of the issue areas and committees with the highest predictors of reaching a later state had a negative relationship with lobbying. These included issues like Arts, Culture, and Religion and Sports and Recreation as well as committees such as the Small Business Committee. Similarly, issues like Finance and Financial Sector and committees like Finance and Agriculture, Nutrition and Forestry were strong negative predictors of a law advancing even though they were strong positive predictors of a law receiving more lobbying. This suggests an interaction between lobbying and issue area and committee that future work should explore. Other factors that made a bill more likely to advance in this model include more legislative subjects, more cosponsors, being introduced earlier in the session, having more related bills, and being introduced by a Democrat. The model had an out-of sample classification error rate of 30.3%.

While more work is necessary to determine the direction or directions of the relationship between lobbying and bills advancing as well as to understand how lobbying affects the advancement of bills on multiple issues, there are several key takeaways from this analysis. In general, it appears that bills on highly salient issues were more likely to be lobbied, but less likely to advance in Congress. Conversely, bills on low salience issues

were less likely to be lobbied but more likely to advance in Congress. Despite this inverse relationship, in general bills that were lobbied more advanced further in Congress. Bills on issues of high salience that did not directly affect economic industries, such as civil rights bills, did not attract as much lobbying as bills of issues of high salience that had direct effect on industry, such as regulatory bills.

Section 1: Introduction

Lobbying is the practice of companies, individuals, and interest groups seeking to influence government officials to make a decision (National Conference of State Legislators, 2021). Members of Congress are frequent targets of lobbying because of their power to influence which laws are passed. Lobbying is often criticized for its potential to distort the legislative process by giving certain unrepresentative interest groups unfair influence and creating the conditions for regulatory capture. The American political system is noted for an unusually high prevalence of lobbying (Wallach, 2015). The Lobbying Disclosure Act requires that most organizations who lobby Congress and most lobbyists file a quarterly report, but this data is rarely explored and therefore represents a major missed opportunity for increased government transparency (Public Citizen). Among the most relevant issues to understand is how lobbying influences which bills are passed. It is also important to understand which issues tend to attract the most lobbyist interest. It is possible that legislation in Congress is being distorted by the high amounts of lobbying. Beyond just understanding lobbying, it is important to understand what bills are passed in Congress. Very few bills that Congress proposes each year are passed, so any patterns in which bills are passed is highly relevant to the way that Congress works.

The 111th Session of Congress, which lasted from January 3, 2009 until January 3, 2011, was selected as a case study. This is a particularly interesting session of Congress. It was the main session of the first half of Obama's first term, which involved the passage of landmark legislation such as the Affordable Care Act, the Dodd-Frank Wall Street Reform Act, and the repeal of Don't Ask Don't Tell (Library of Congress, Public Laws of the 111th Congress). It also featured the high-profile failed attempt to pass a cap-and-trade system for carbon emissions (Broder, 2010) and to pass a public option for healthcare (Fisher, 2010). This session of Congress is also when Congress passed the bulk of its response to the Financial Crisis of 2008, including the American Recovery and Reinvestment Act of 2009 and subsequent financial regulation (Zeballos-Roig, 2020). The 111th Congress may not be representative of other sessions of Congress. Democrats held both branches of Congress by large margins and the President was also a Democrat (Ballotpedia, 111th United States Congress). Because of this, it may have seen unusually high productivity on liberal policy goals and an unusual focus on Democrats by lobbyists.

The data for this project comes from two sources. First, information about which bills were the most heavily lobbied comes from OpenSecrets, a 501(c)3 non-profit organization that tracks and publicizes data about money in Politics including political donations, the activity of dark money groups, and lobbying (Open Secrets, Our vision and mission: Inform, empower & advocate). They make their data available to the public through their bulk data program. This paper used the "lobby_bills.txt" file from their collection of Lobbying data. The file contains a row for each appearance of a bill in an LDA report. The number of times that a bill was lobbied can be calculated by counting the number of unique report identifiers that a bill appears in.

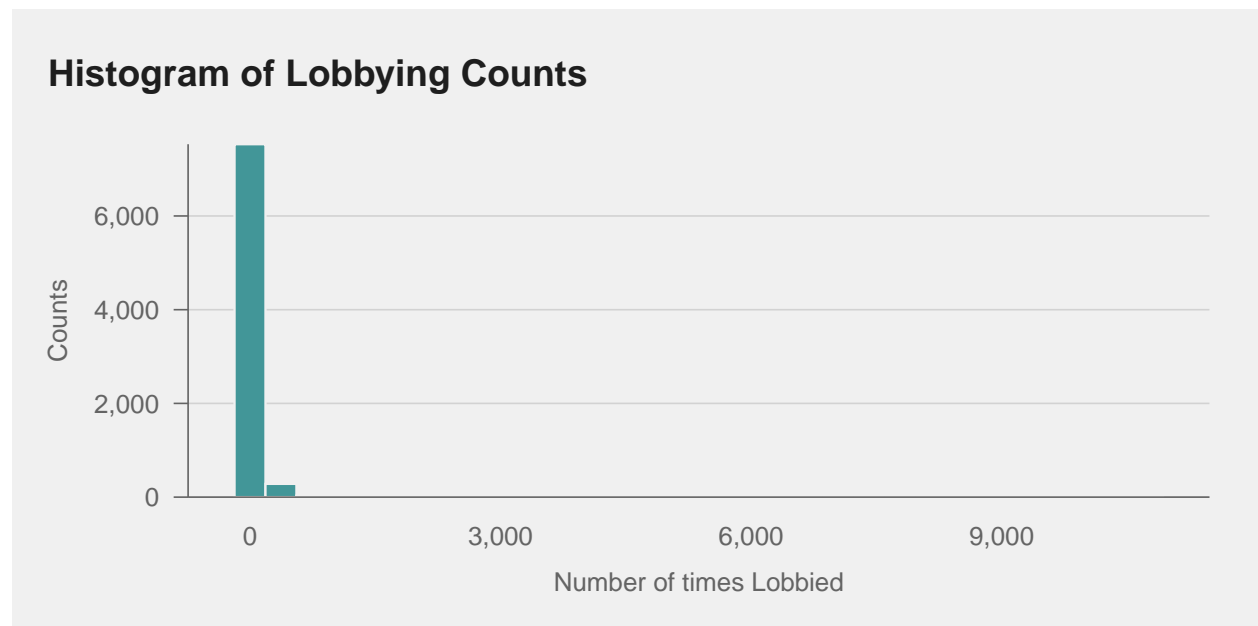
The second data source is the Congress.gov API, which contains detailed information about every bill that has been introduced in Congress since 1973. The information for each bill that the API contains can be found on the Github page for the API. (Library of Congress, API Documentation Bill Endpoints). For each bill, there is a url for a JSON file containing data about the bill and additional urls that a user can download more specialized data from. For each bill, we collected data from the main JSON file as well as 5 of the nested JSON files. We collected the information on each bill that I had complete lobbying data for in the 111th Congress. We excluded bills that were not of the "H.R." or "S" type because those bills typically do not contain substantive policy. We did not collect data for bills that did not appear in the OpenSecrets data because We did not want to infer that they were never lobbied and because it is challenging to query large amounts of data from the API due to a restrictive rate limit of 1000 queries per hour (each bill requires six

queries) and each query is fairly slow. If we had included bills that were not in the OpenSecrets data, there would have been 15055 bills.

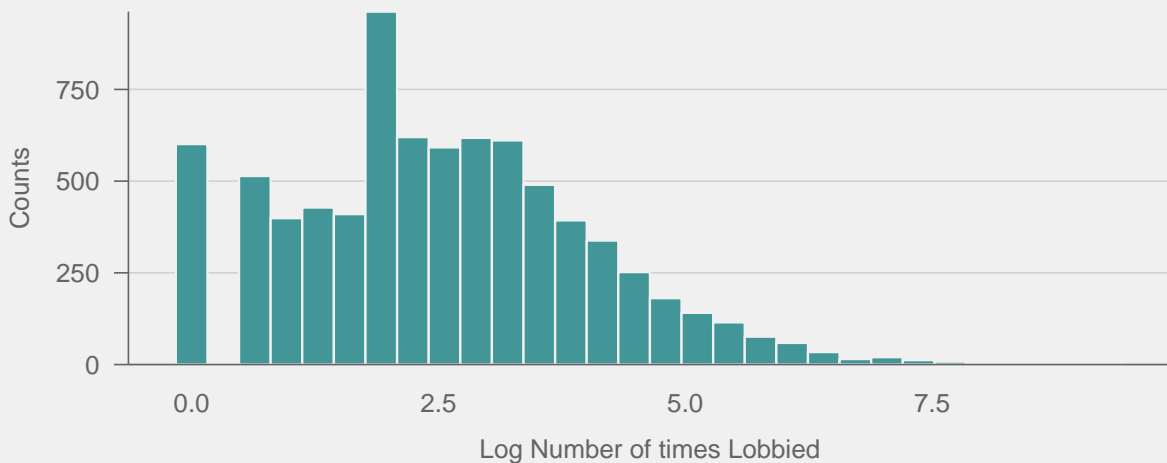
Section 2 contains data visualizations that explore the number of times bills were lobbied and demonstrates relationships between the amount of times that a bill was lobbied and several potential explanatory variables including: the date the bill was introduced, the chamber the bill was introduced into, the party of the sponsor, the policy area of the bill, the number of legislative subjects assigned to the bill by the Library of Congress, the number of cosponsors the bill has, the committee the bill was first assigned to, the number of committees the bill was assigned to, the stage of the legislative process the bill reached, and the number of official actions taken on the bill. Section 3 will create two models to predict the number of times bills are lobbied. Section 4 will create a model of the final legislative stage of the bill. Section 5 will discuss these models and highlight some key implications and create visualizations about the Congressional Research Services Summaries and Legislative Subjects for key subsets of the data.

Section 2: Data exploration and visualization

This section will explore the lobbying data, primarily by comparing it to several potential explanatory variables. It will also explore how several potential explanatory variables compare to how far a bill gets in Congress.

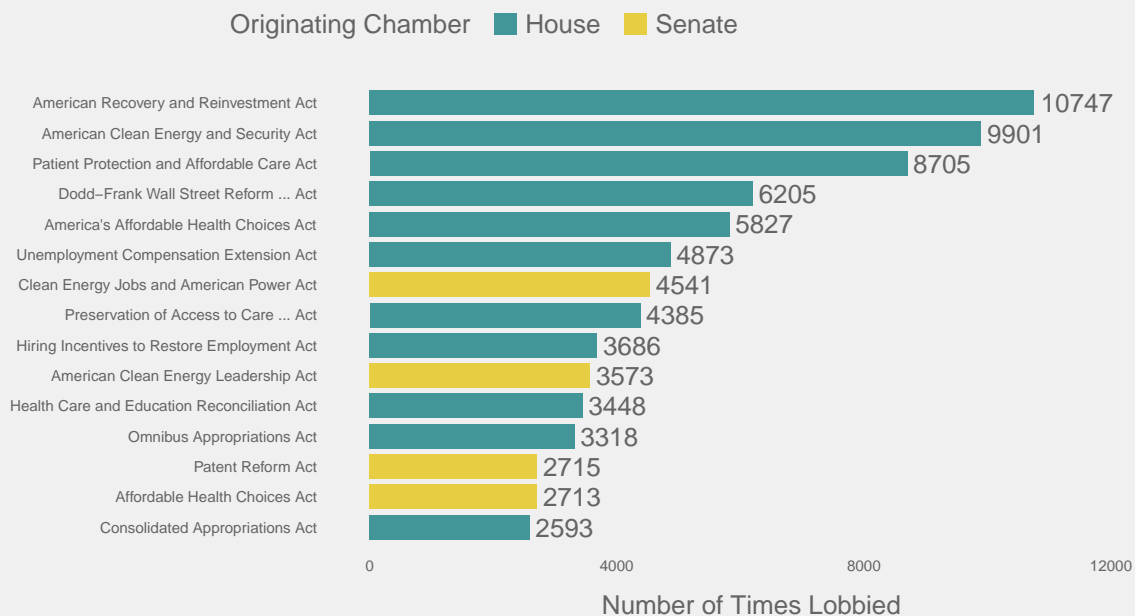


Histogram of Log Lobbying Counts



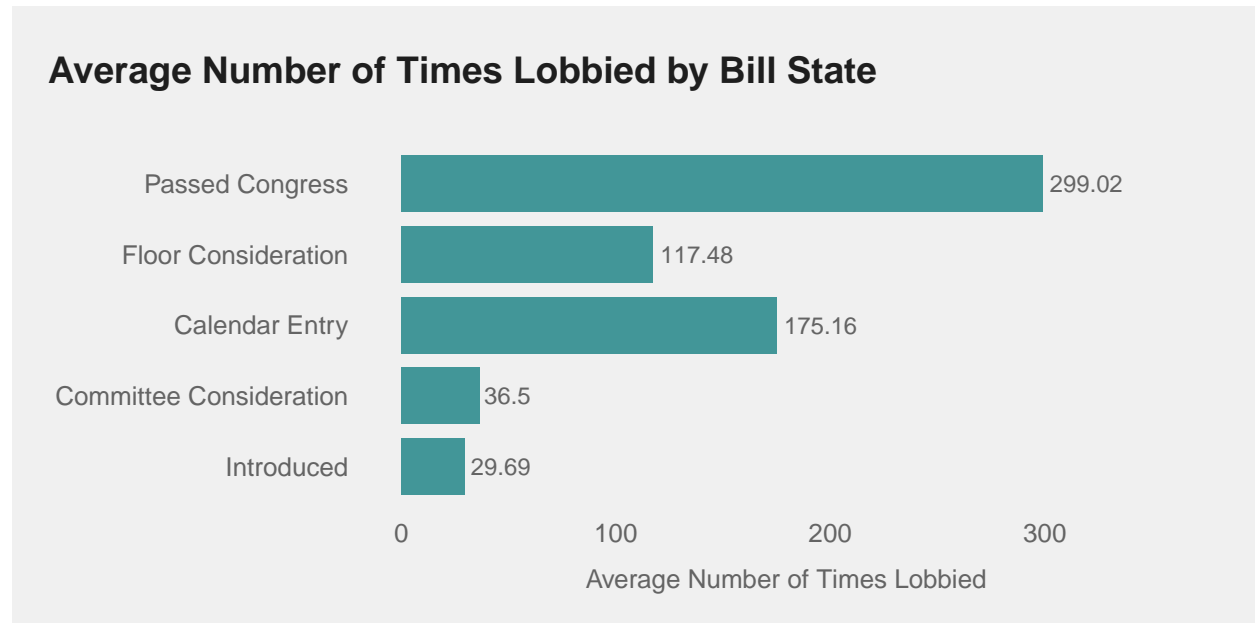
First, it is important to visualize the shape of the lobbying variable. The first chart shows a histogram of the untransformed lobbying counts. It is clearly highly right skewed. To attempt to correct the right skew, the second histogram shows a log transformed version of the variable. While it appears more normally distributed, a slight right skew remains. As such, it may be appropriate to apply a box-cox transformation to linear models to see if a stronger transformation would be necessary.

Most Lobbied Bills in 111th Congress



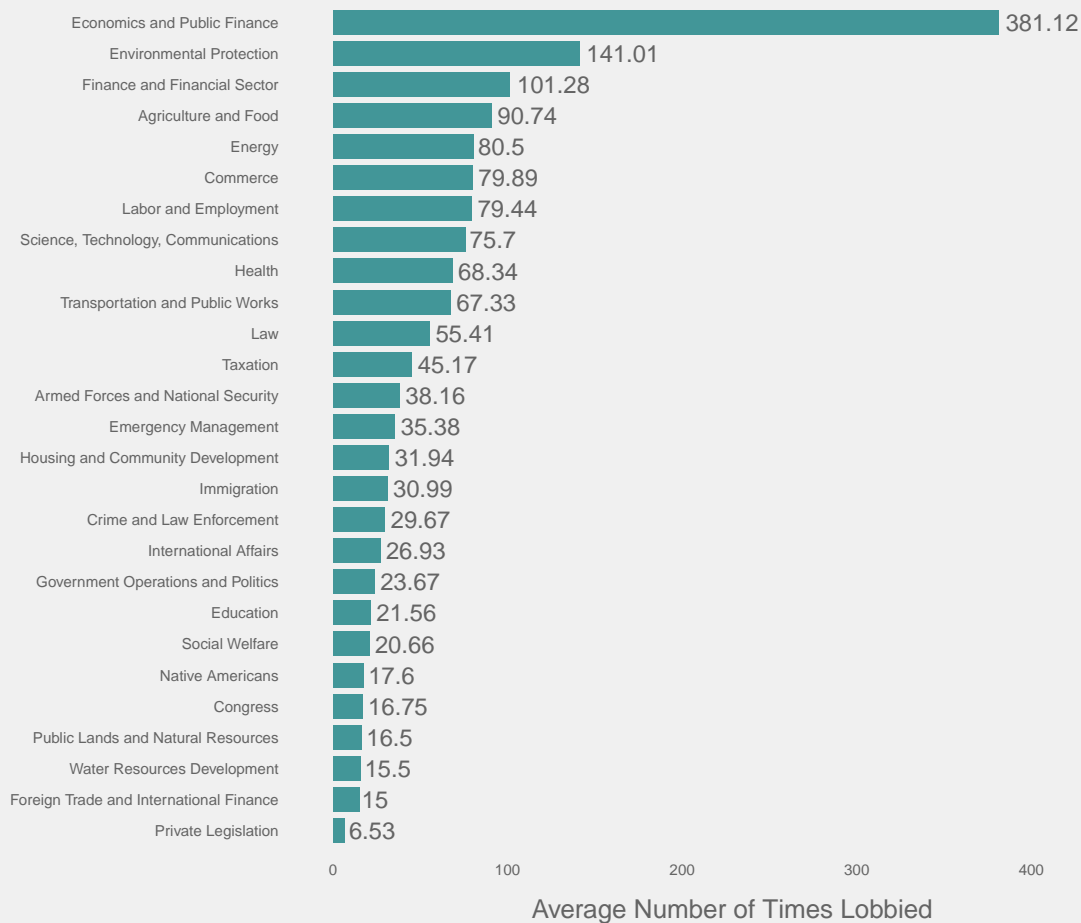
This chart shows the 15 bills that were lobbied the most in the 111th Congress. It primarily includes the most significant bills and the bills that drew the most media attention. The first four bills are the post-

recession stimulus package, the proposed cap-and-trade system, the bill that established Obamacare, and the Dodd-Frank post-recession Wall Street regulation package (Library of Congress, Public Laws of the 111th Congress), (Broder, 2010), (Zeballos-Roig, 2020). The top 15 also includes Appropriations bills, which often contain funding for specific projects that serve lobbies in the states and districts of critical votes (Saturno & Lynch, 2023). Other bills include healthcare and energy bills, both of which have industries that are highly engaged in lobbying (OpenSecrets, *Lobbying Industries*).



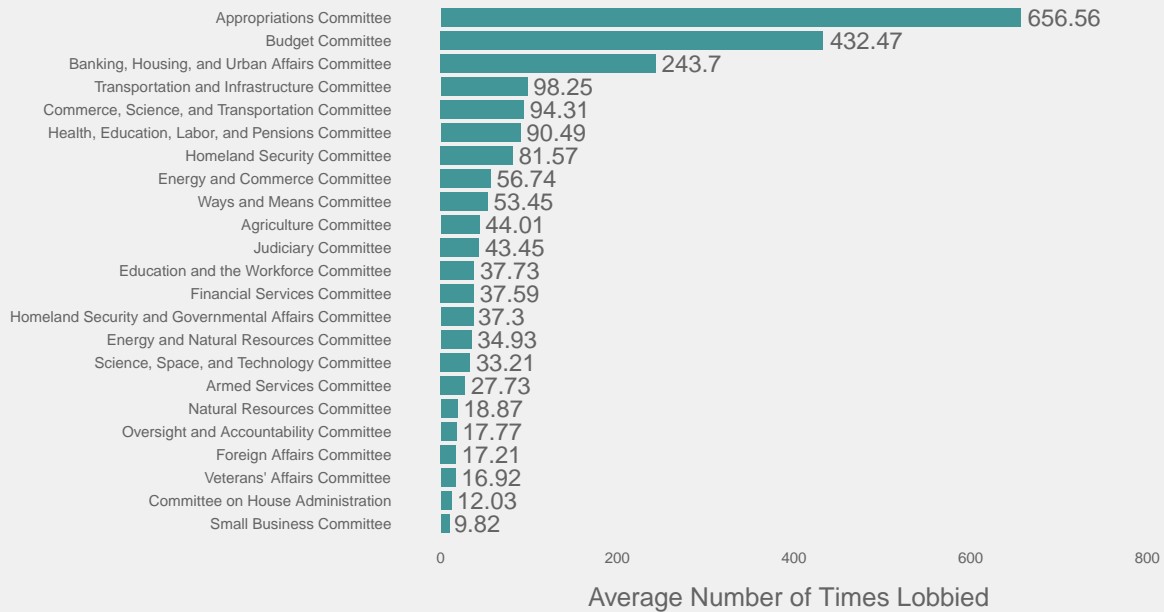
This chart shows that bills that go further in the legislative process draw more lobbying attention. This makes sense because bills that go further would appear more significant to lobbyists and the influence of lobbyists may make legislators more likely to take action on one bill over another. It is unclear which effect is more significant from this visualization.

Lobbying by Policy Area

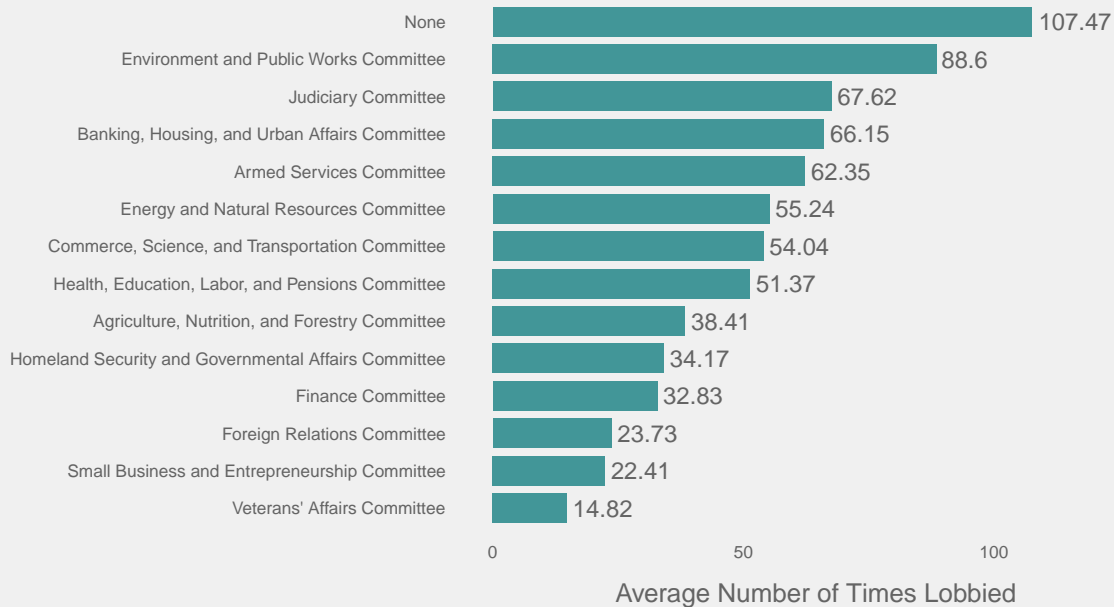


This chart shows which issues generate the most interest from lobbyists. By far the issue are with the most interest from lobbyists was Economics and Public Fiance. Financial entities are large and powerful and thus have the capacity to heavily lobby, especially in a session like the 111th when there was significant legislative attention (OpenSecrets, *Lobbying Industries*). Other issues in the top of the chart were highly salient in the 111th Congress and have active industries. The issues at the bottom of the list typically do not have large industries with interests in the legislation about this subject (OpenSecrets, *Lobbying Industries*).

Lobbying by House Committee



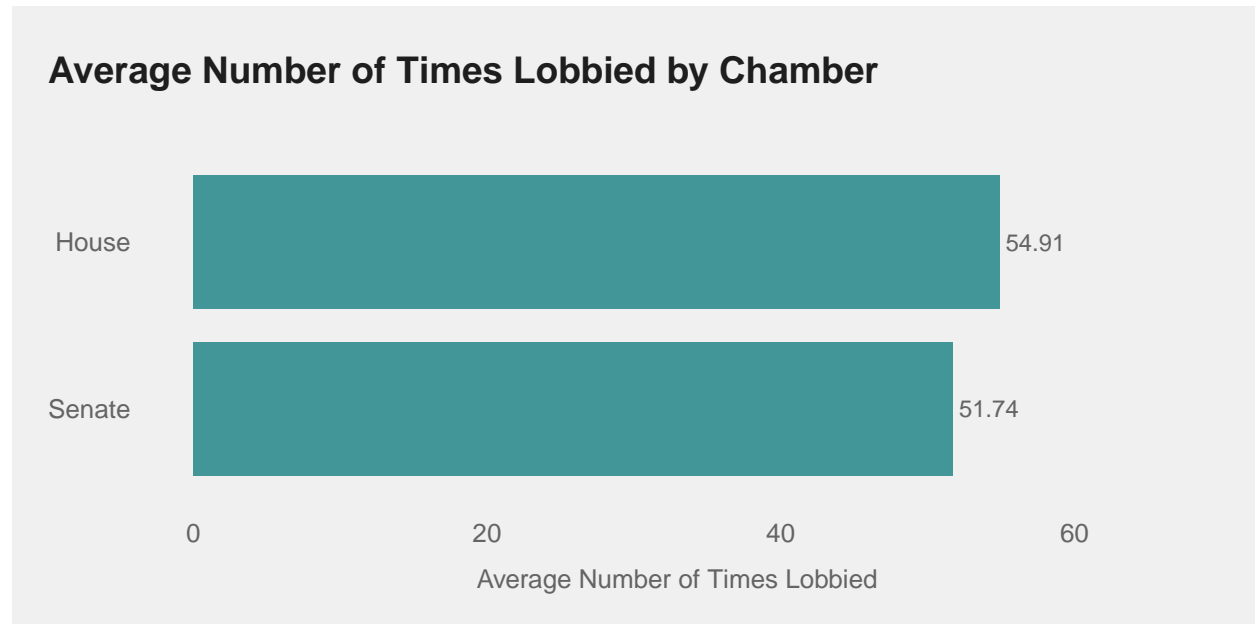
Lobbying by Senate Committee



These charts show the lobbying activity by bills from various House and Senate Committees that had at least 30 bills in the data originate in them. By far, the most lobbying attention went to the Appropriations committee, which may be due to the potential for funding for personal projects and other similar “pork

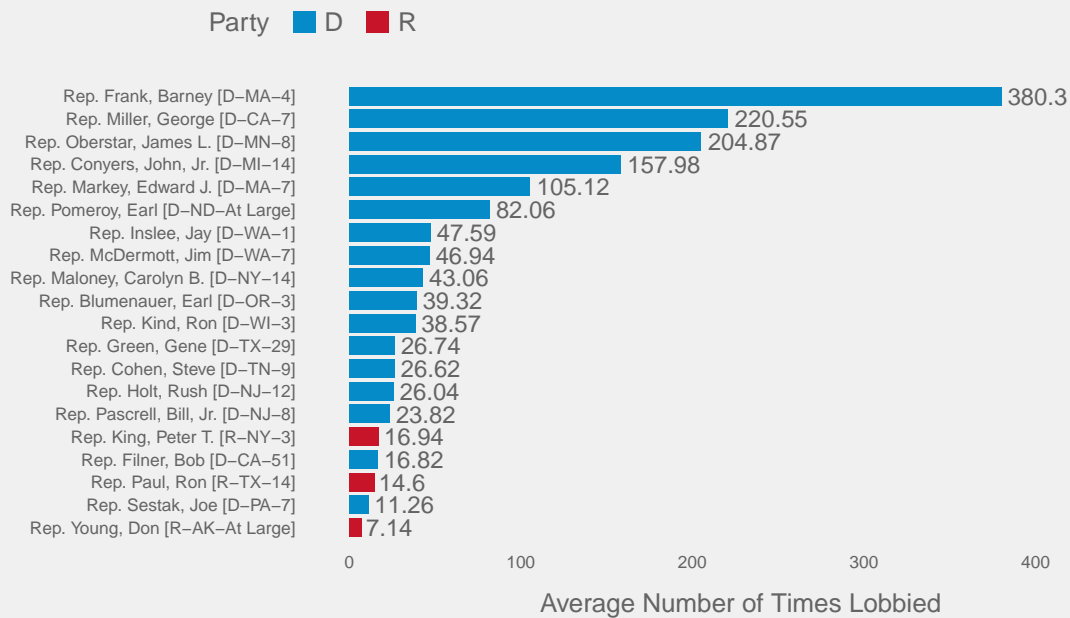
barrel” legislation (Citizens against Government Waste, 2023). The Budget and Finance Committees in the House also drew significant attention. The reason for the attention on the Budget may have been similar to the attention on Appropriations. The attention on the Finance committee may have been due to the focus on financial issues following the recession and the high lobbying activity of the financial industry.

On the Senate side, it is interesting how the Finance Committee did not draw as much attention from lobbyists. Instead, bills that were not sent to committee had the highest lobbying activity. This is potentially because these bills are typically those that leadership has decided to fast-track.

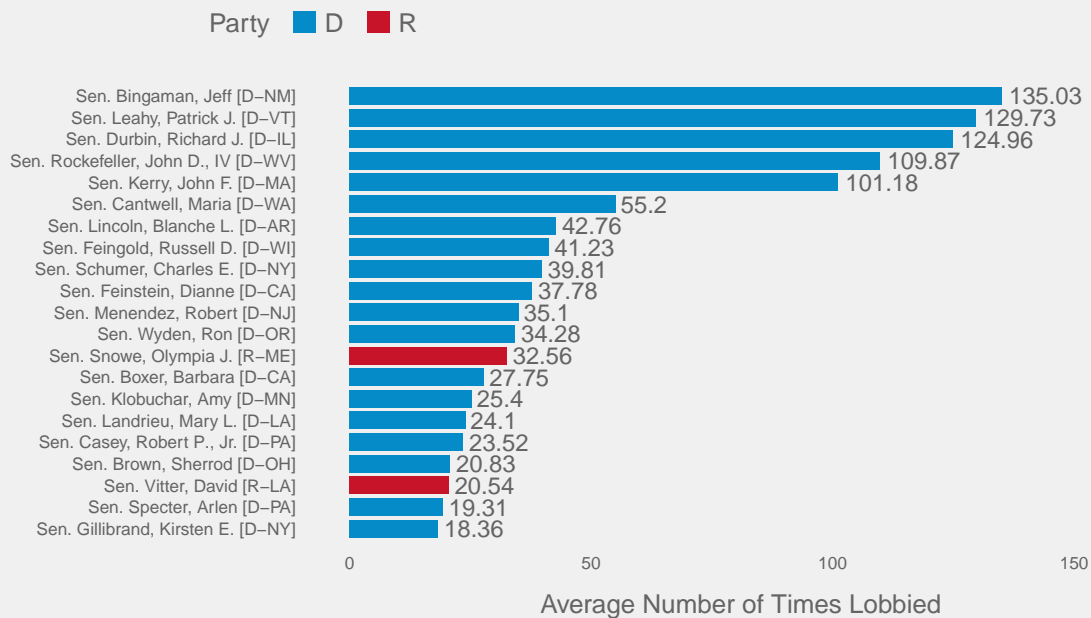


Somewhat surprisingly, House and Senate bills are lobbied similar amounts, even though more of the highly lobbied bills come from the House and several House committees draw the most significant lobbying.

Lobbying by House Sponsor



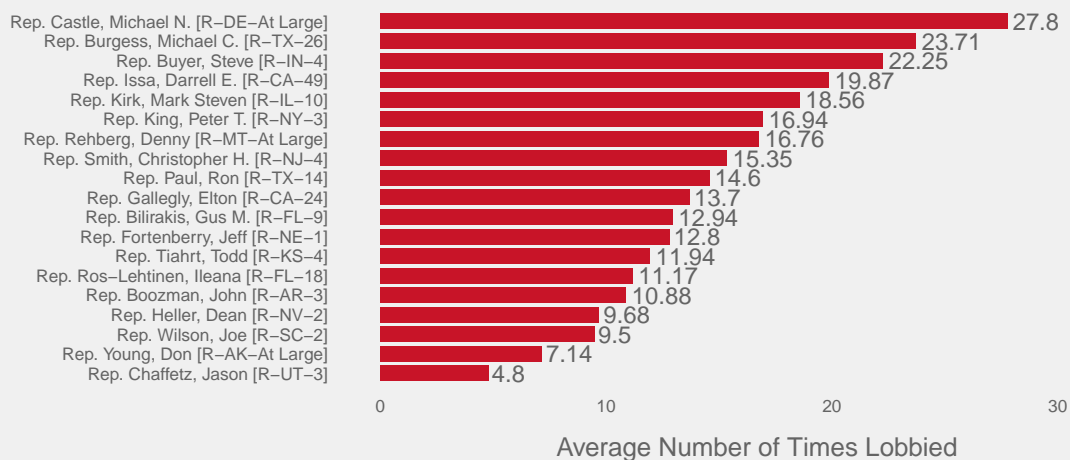
Lobbying by Senate Sponsor



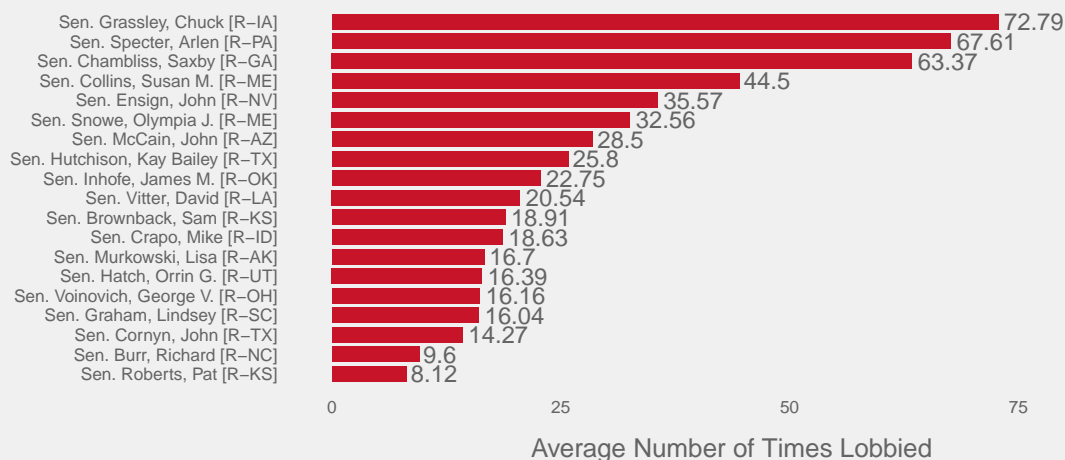
These charts show the average lobbying for the Senators and Representatives that sponsored the most bills in the 111th Congress. Most notably, the most prolific legislators in both chambers are primarily Democrats. This could be because Democrats tend to be more interventionist than Republicans, but it might also be

because they were in power and thus had greater ability to pass legislation which made them more likely to propose it. Barney Frank, who was very involved with Wall Street regulation drew the most interest from lobbyists (Ray, 2023). The Senate lobbyists were most drawn to five Senators: Bingaman, Leahy, Durbin, Rockefeller, and Kerry. These Senators were all long-serving senators (all five were in Congress by 1985) and may have had long-standing relationships with lobbyists and interest groups (Library of Congress, *Senator Jeff Bingaman*), (Library of Congress, *Senator Patrick J. Leahy*), (Library of Congress, *Senator Richard J. Durbin*), (Library of Congress, *Senator John D. Rockefeller IV*), (Library of Congress, *Senator John F. Kerry*). They were all also in either party or committee leadership (Bingaman chaired Energy, Leahy chaired judiciary, Durbin was majority Whip, Rockefeller chaired Commerce, and Kerry chaired Foreign Relations), which made them powerful and valuable to lobbyists.

Lobbying by House Sponsor (GOP)

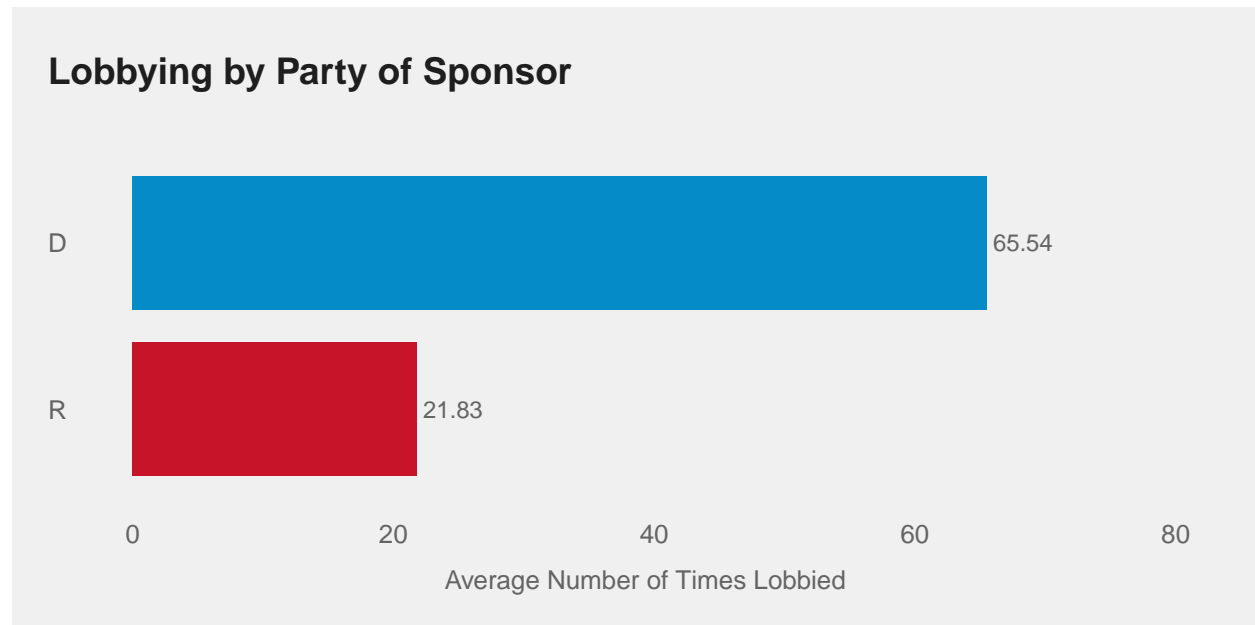


Lobbying by Senate Sponsor (GOP)

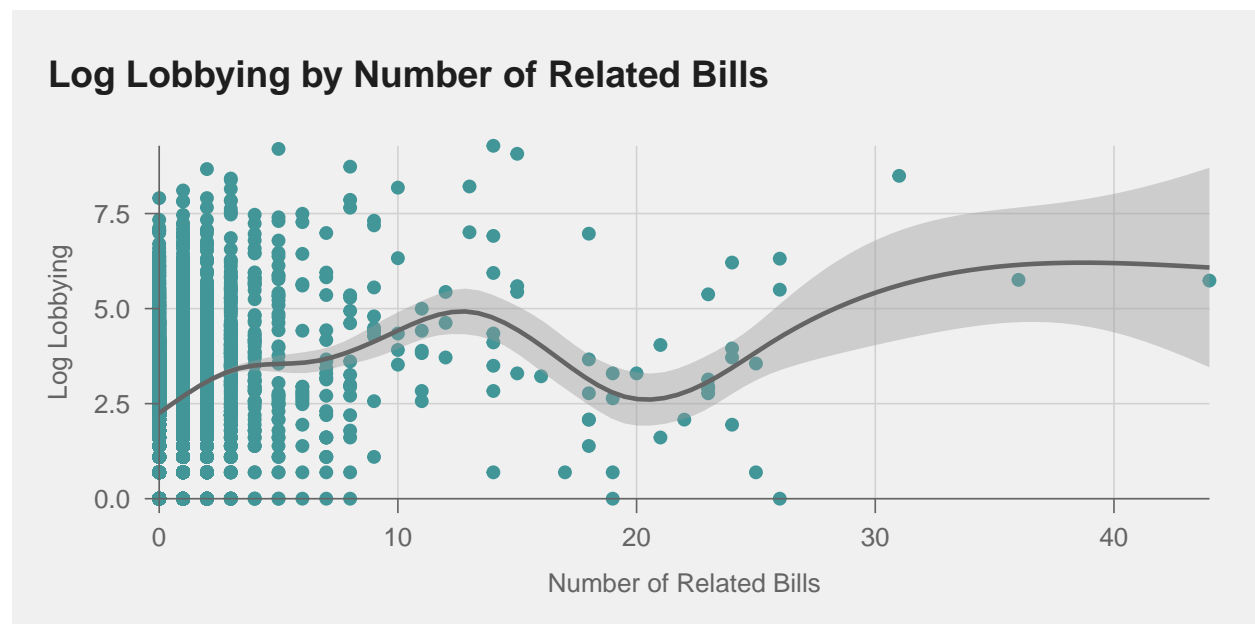


Because the other graph contained mostly Democrats, these graphs show only the Republicans who sponsored the most bills in the 111th Congress. There are no standouts in the Republican House, but Michael Castle,

the representative who drew the most lobbying on this list was known as a moderate and may have been a key swing vote on Democratic-sponsored legislation (Ferraro, 2009). On the Senate side, Grassley and Specter have similar profiles to the five most lobbied Democrats as long-serving members of Congress in leadership positions (Library of Congress, *Senator Chuck Grassley*), (Library of Congress, *Senator Arlen Specter*). They also had moderate reputations which may have brought them to work with Democrats to pass some bipartisan legislation. Chambliss, however, has a different profile as he had only joined Congress in 1995 and was not known for a moderate voting record (Library of Congress, *Senator Saxby Chambliss*).

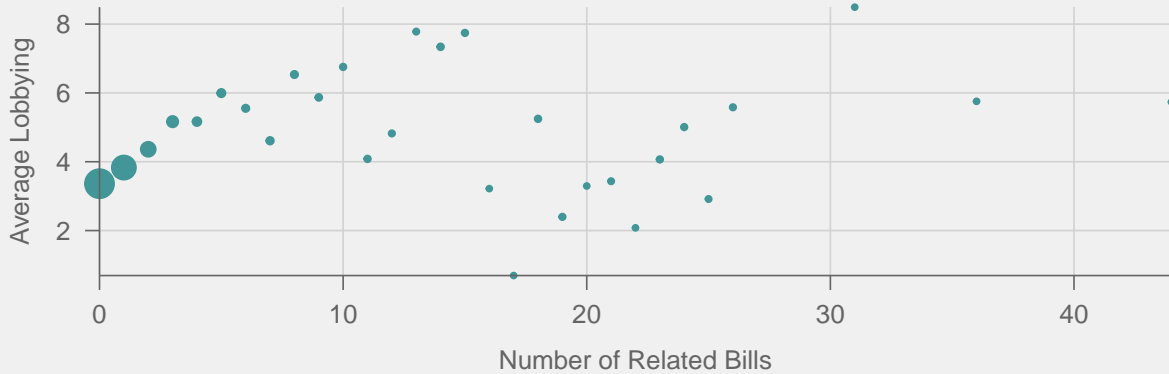


This graph demonstrates how Democrat-sponsored bills were lobbied significantly more than Republican lobbied bills. This is likely a reflection of Democrats controlling both chambers and the presidency meaning their laws tended to have a higher probability of passing.



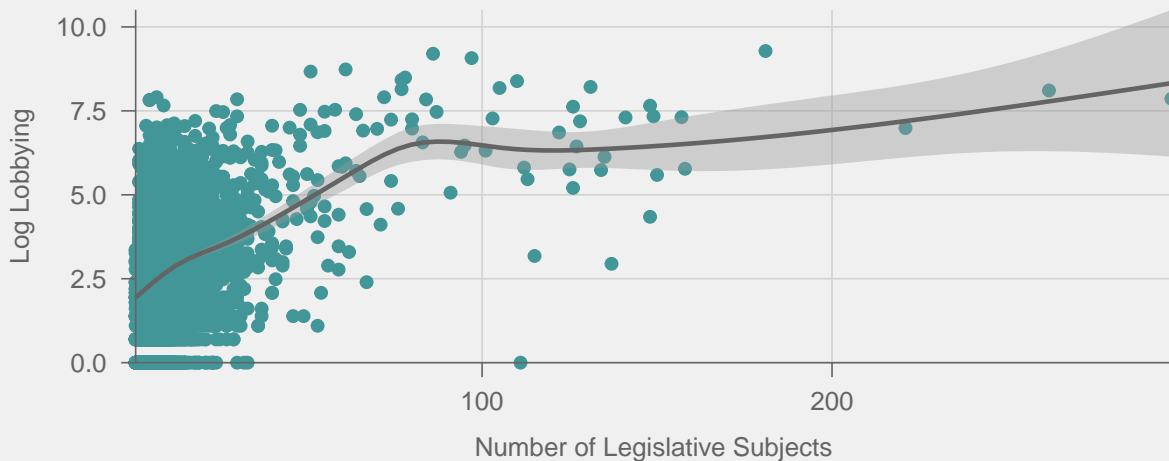
Number of Times Lobbied by Binned Number of Related Bills

size ● 1000 ● 2000 ● 3000 ● 4000



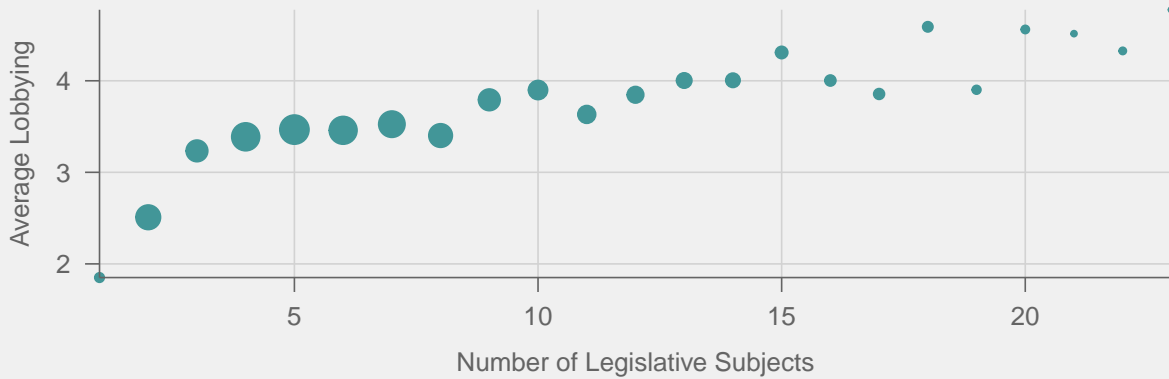
Some bills are marked as related to each other in Congress. This is often when the same or a similar bill is introduced in the other chamber or when multiple related bills are packaged together. Most bills have very few related bills but some have very many. It may be appropriate to attempt to log-transform it. That said, the scatterplot and loess regression line indicates a potential positive relationship, that is easier to see when results are binned in the second graph. Bills with more related bills may be those that are more involved in active negotiation.

Log Lobbying Lobbied by Number of Legislative Subjects



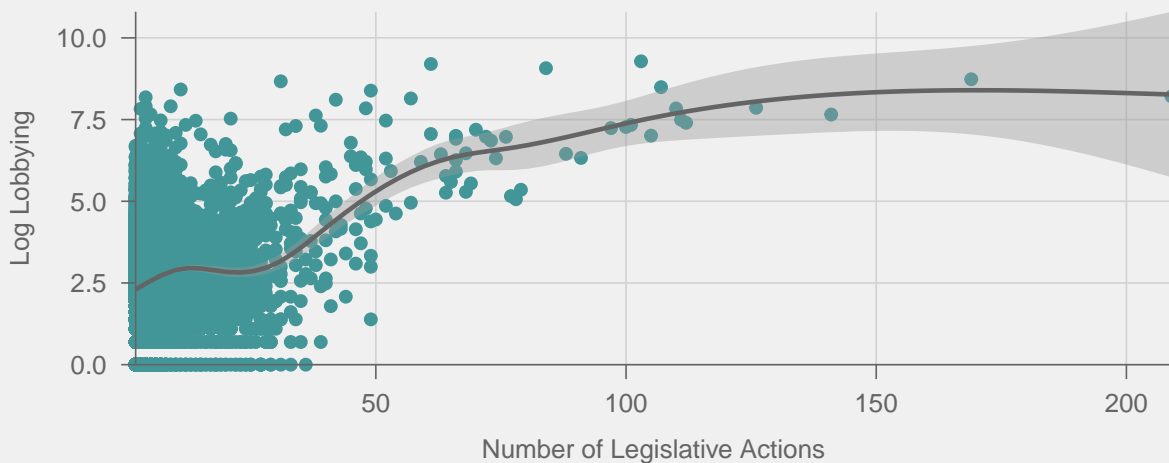
Lobbying by Binned Number of Legislative Subjects

size ● 250 ● 500 ● 750 ● 1000

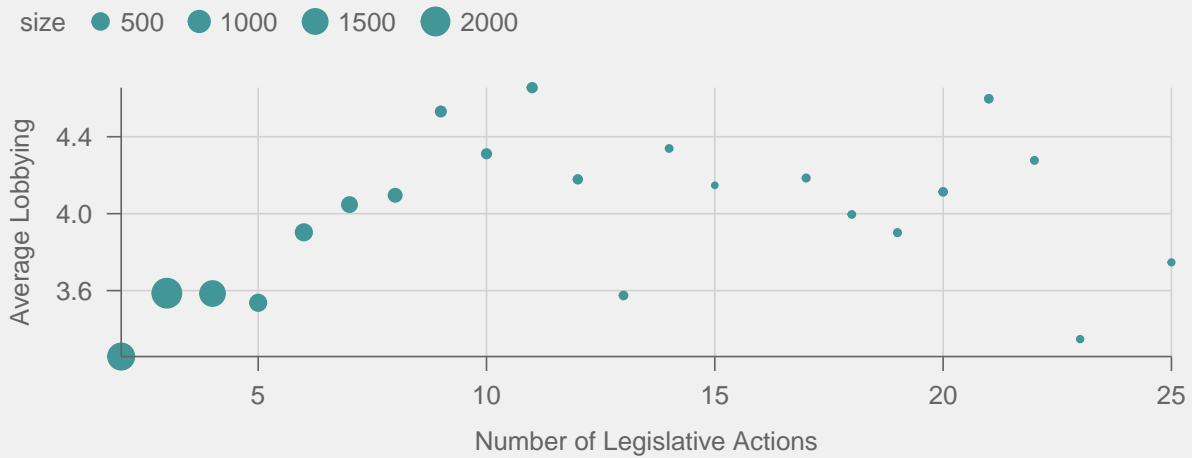


The relationship between lobbying and the number of associated legislative subjects is similarly seemingly positive, but like with the count of related bills a log transform may be appropriate. The number of legislative subjects might be thought of as a proxy for legislative complexity given that more complex laws will touch more subjects and less complex laws will touch less subjects. That said, some subjects may be more similar to each other and be more likely to appear alongside each other in the same bill so caution is warranted in applying this proxy rigidly.

Log Lobbying by Number of Legislative Actions

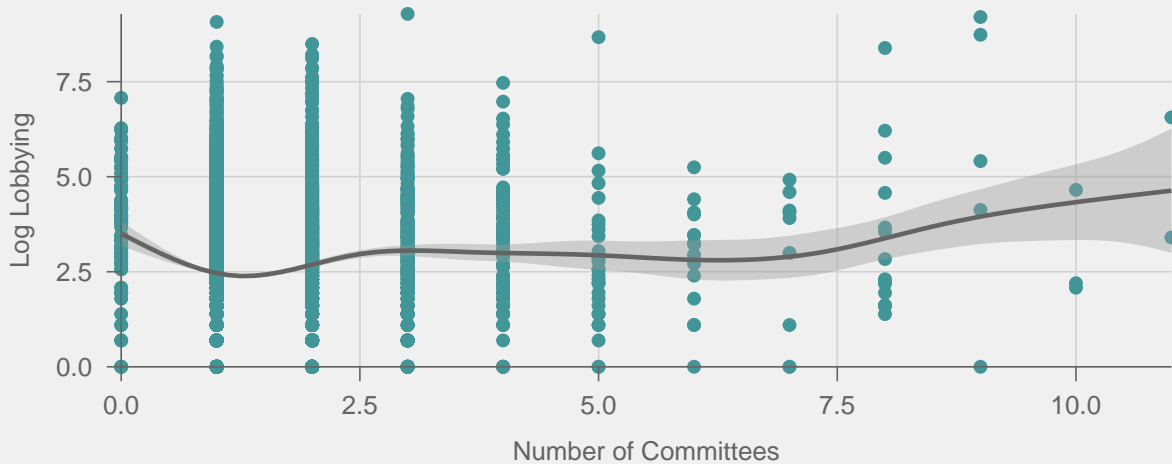


Lobbying by Binned Number of Legislative Actions



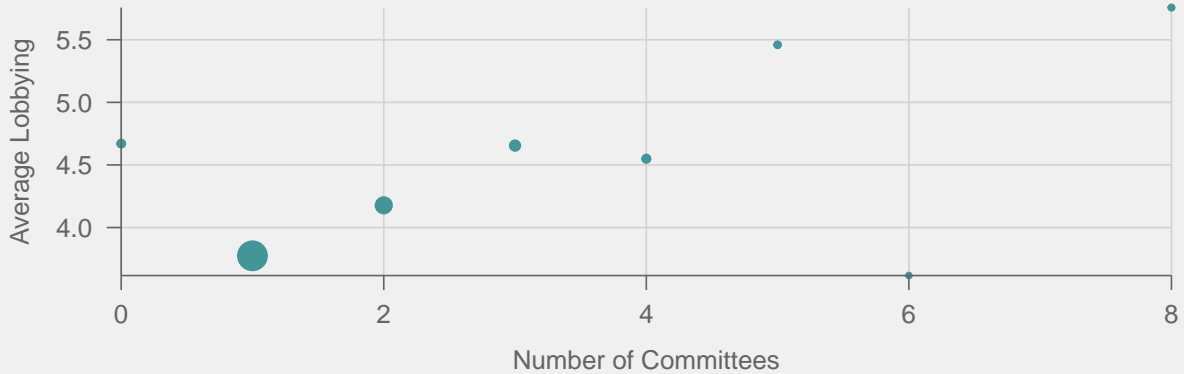
A legislative action is when any official step is taken to advance the bill including scheduling a hearing, holding a hearing, or holding a vote. Taking more legislative actions is a proxy for moving a law closer to passage, but some bills may require more actions to reach the same stage. This variable shows a similar pattern to the other two count variables. This could, however, just be a restatement of the variable that measures the furthest that the bill reaches in Congress.

Log Lobbying by Number of Committees



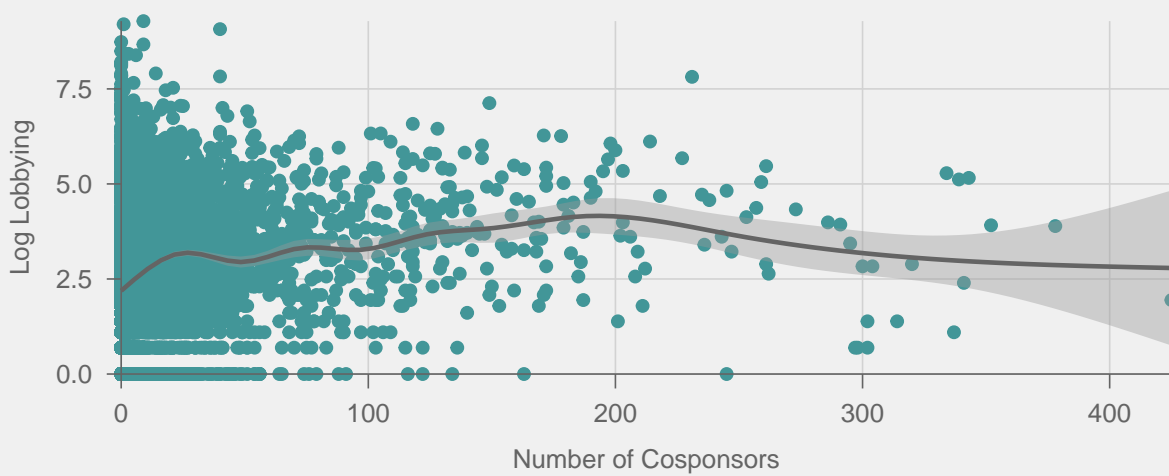
Lobbying by Binned Number of Committees

size ● 2000 ● 4000 ● 6000

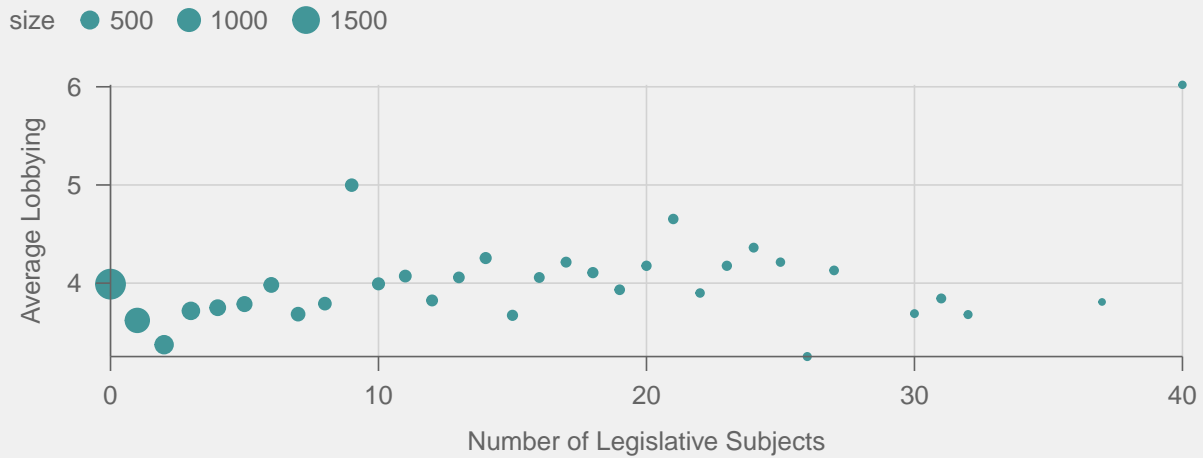


While it is rare, some bills are assigned to multiple committees if it lies in an area of significant overlap and it can also be thought of as a proxy for complexity. That said, there doesn't appear to be as significant a relationship between this variable and the number of times a bill was lobbied.

Log Lobbying by Number of Cosponsors

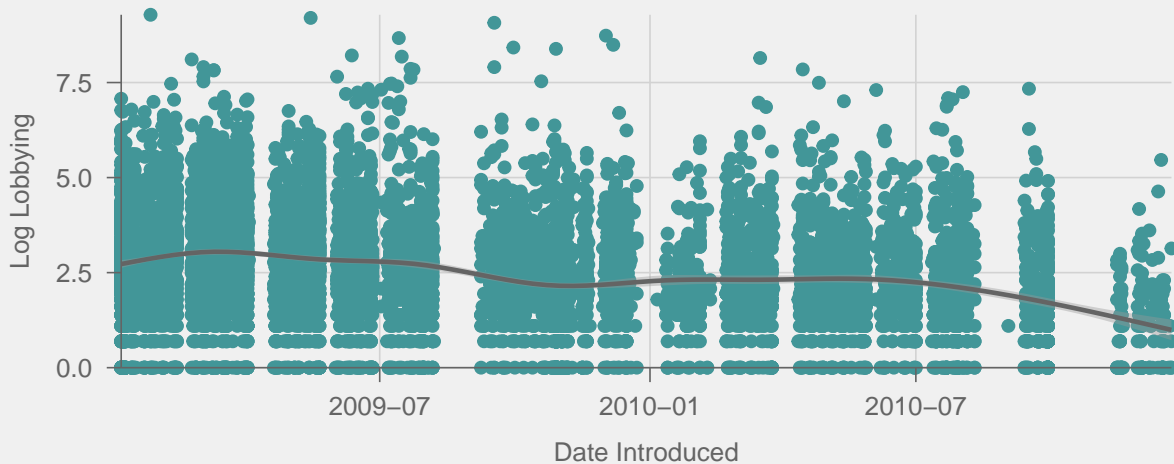


Lobbying by Binned Number of Cosponsors



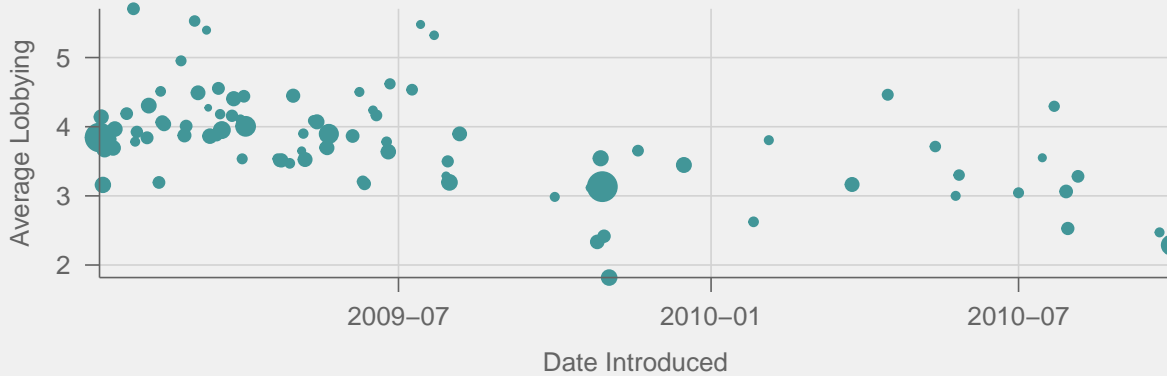
A cosponsor is a member of Congress who signs onto a bill that they did not sponsor but would like to be passed. Sometimes, cosponsors are just as involved as sponsors in trying to pass the bill (The American Legion, Part 4: Sponsors and co-sponsors). Other times, members only cosponsor to signal support but do not become involved in its passage. High-profile bills tend to draw more cosponsorships, so it is interesting to see that bills with more cosponsors are do not appear to generate more interest from lobbyists form this bivariate relationship.

Log Lobbying by Date Introduced



Lobbying by Binned Date Introduced

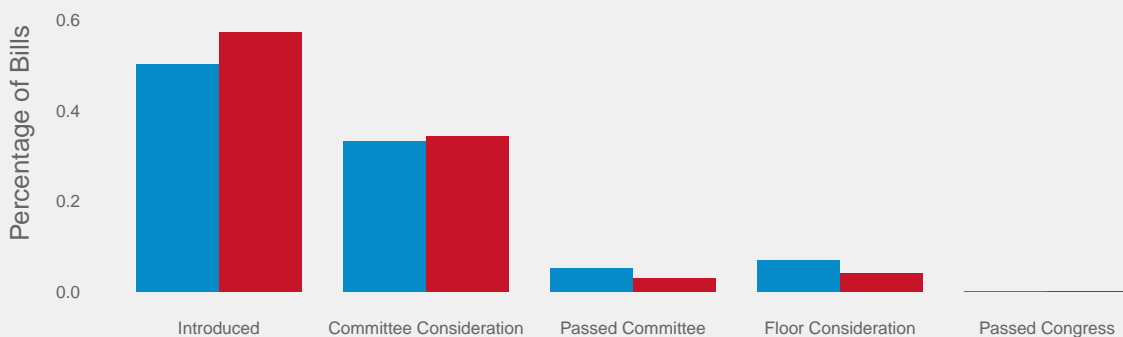
size ● 100 ● 200 ● 300



Bills are introduced throughout the two-year Congressional session. Often, the highest priorities are introduced first, but this is not always true. Bills that are introduced earlier are also able to be lobbied about for longer. As such, it is not surprising to see higher lobbying activity on bills that were introduced earlier in the term.

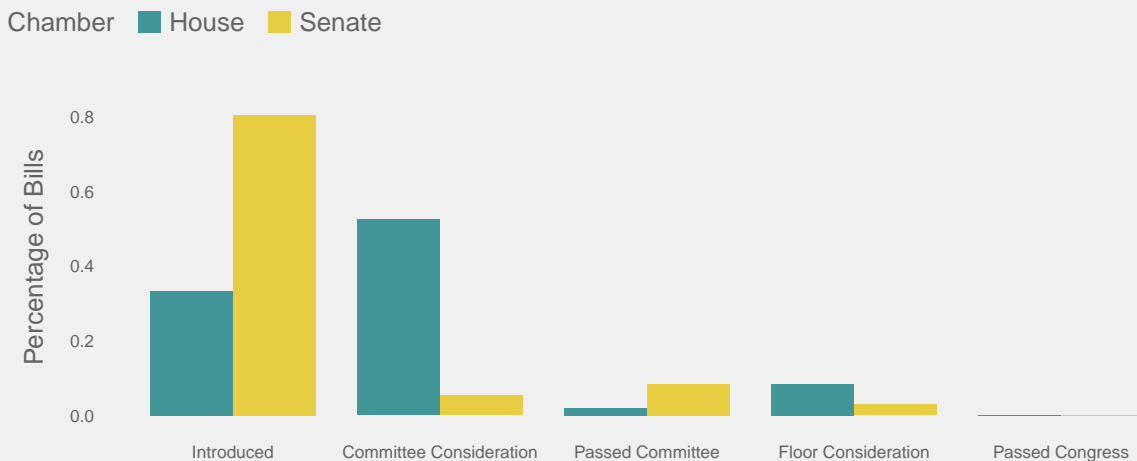
Final State of Bills by Sponsor Party

Sponsor's Party ■ Democrat ■ Republican



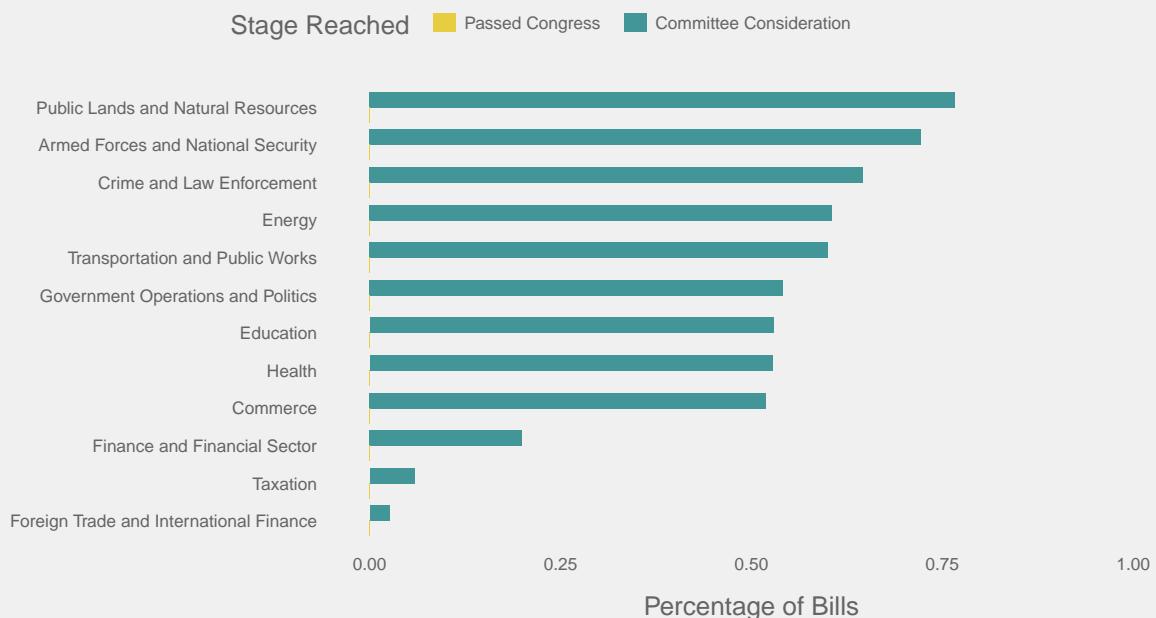
This graph shows the difference in the legislative success of bills introduced by Democrats and Republicans. Bills introduced by Democrats are more likely to pass committee and to reach stages beyond Committee Stage, which reflects Democrats controlling both chambers.

Final State of Bills by Originating Chamber



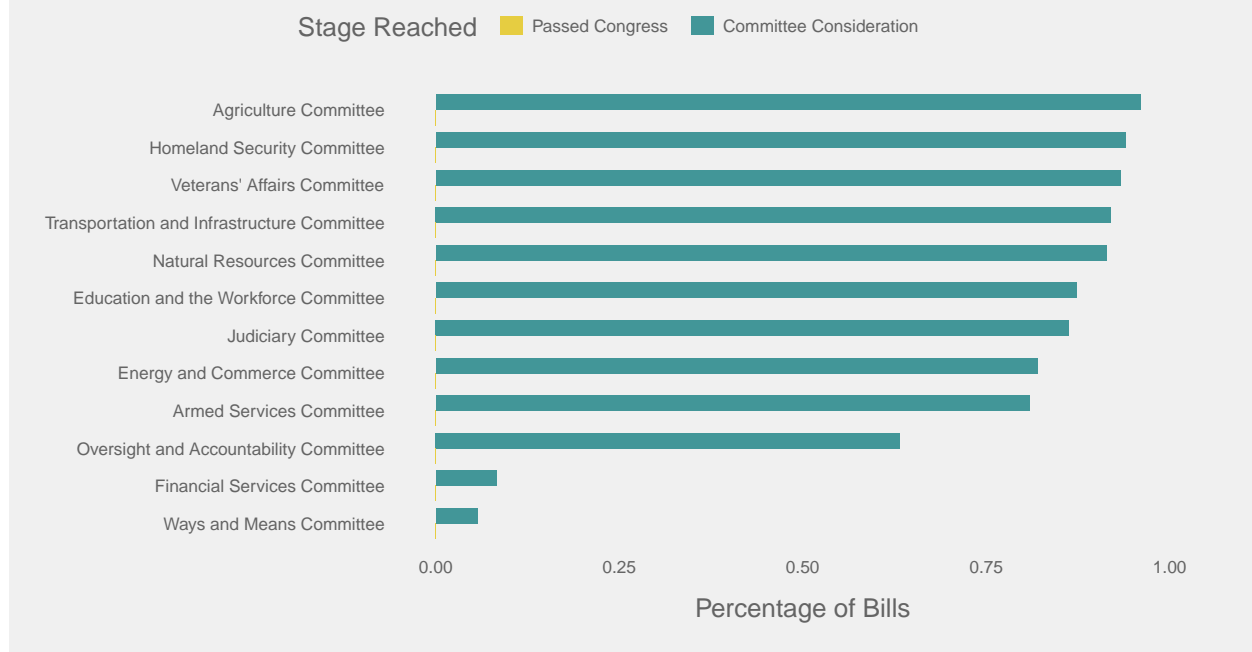
This chart compares the legislative path of bills introduced in the Senate and the House. Very few Senate bills receive any committee consideration which reflects its nature as a slower-moving body than the house. That said, it is interesting to see the chamber of origin influencing the legislative path of a bill more significantly than the party of its sponsor.

Minimum State Reached by Topic Area



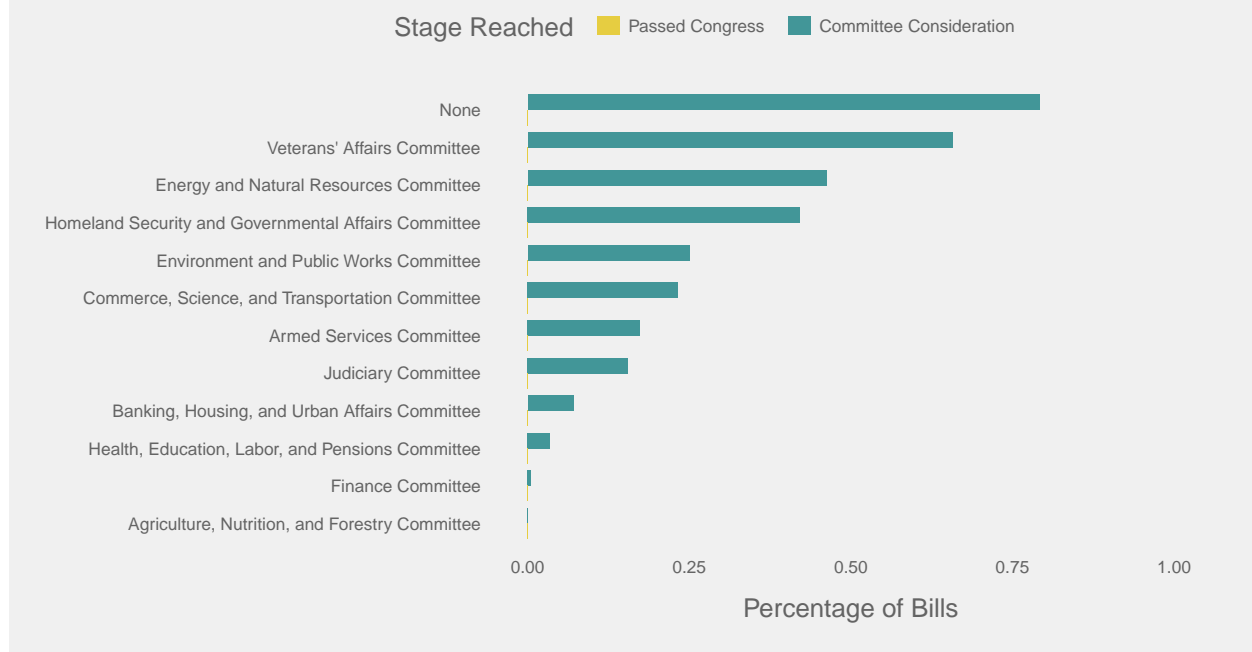
This comparison of the legislative paths of bills about different subjects shows that financial bills do not typically advance far in Congress but environmental and security bills often receive committee consideration. The bills that pass the most are those related to Government operations and similar subjects.

Minimum Stage Reached by House Committee



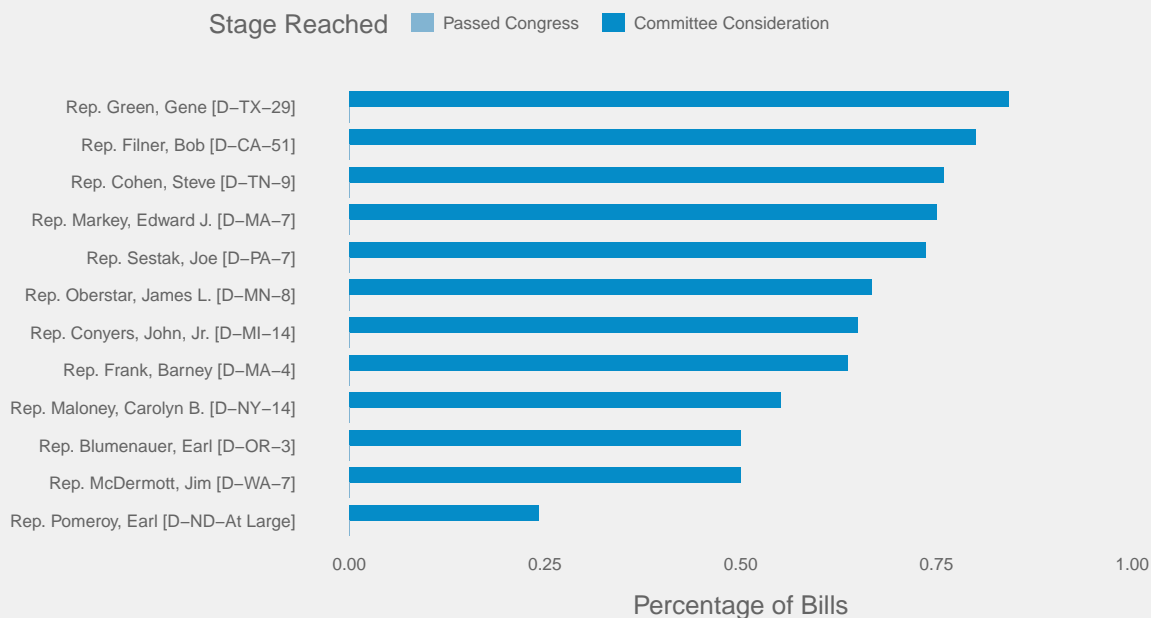
This visualization shows a similar pattern to the policy areas. In most committees, it is very easy for bills to get a committee hearing, but very hard to pass a law. The exceptions are the Financial Services Committee, and the Ways and Means Committee, which controls tax and government revenue policy, where very few introduced laws received a hearing.

Minimum Stage Reached by Senate Committee

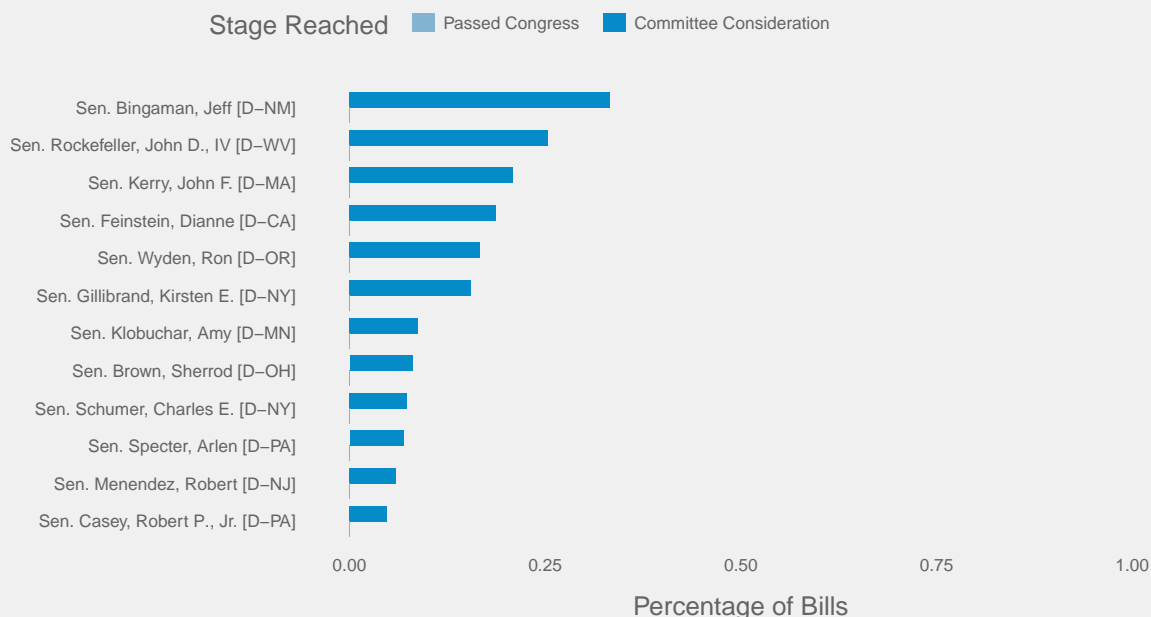


The legislative paths from different Senate committees is likewise varied. Bills without a committee automatically move to floor consideration and seem to be highly likely to pass. Like the other graphs, this one shows that financial bills often do not receive committee consideration whereas energy and environment bills do.

Bills Sponsored by Dem House Members

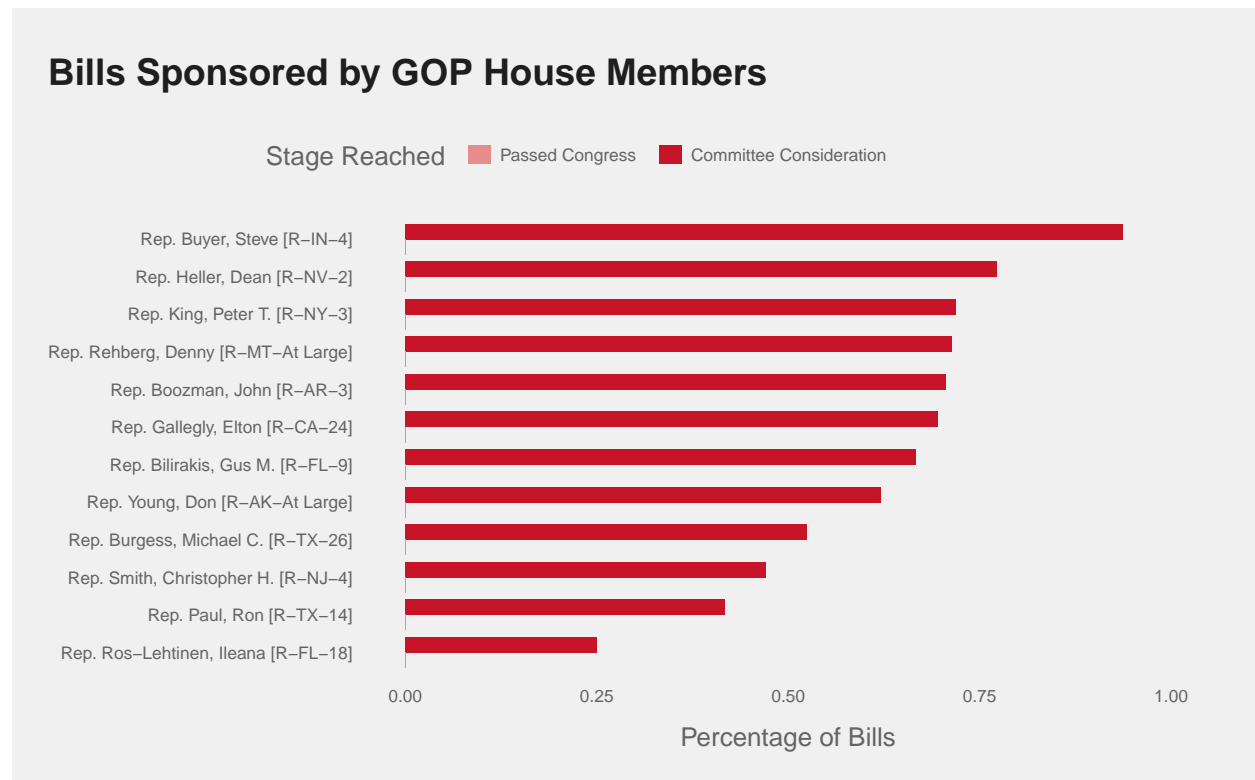


Bills Sponsored by Dem Senators

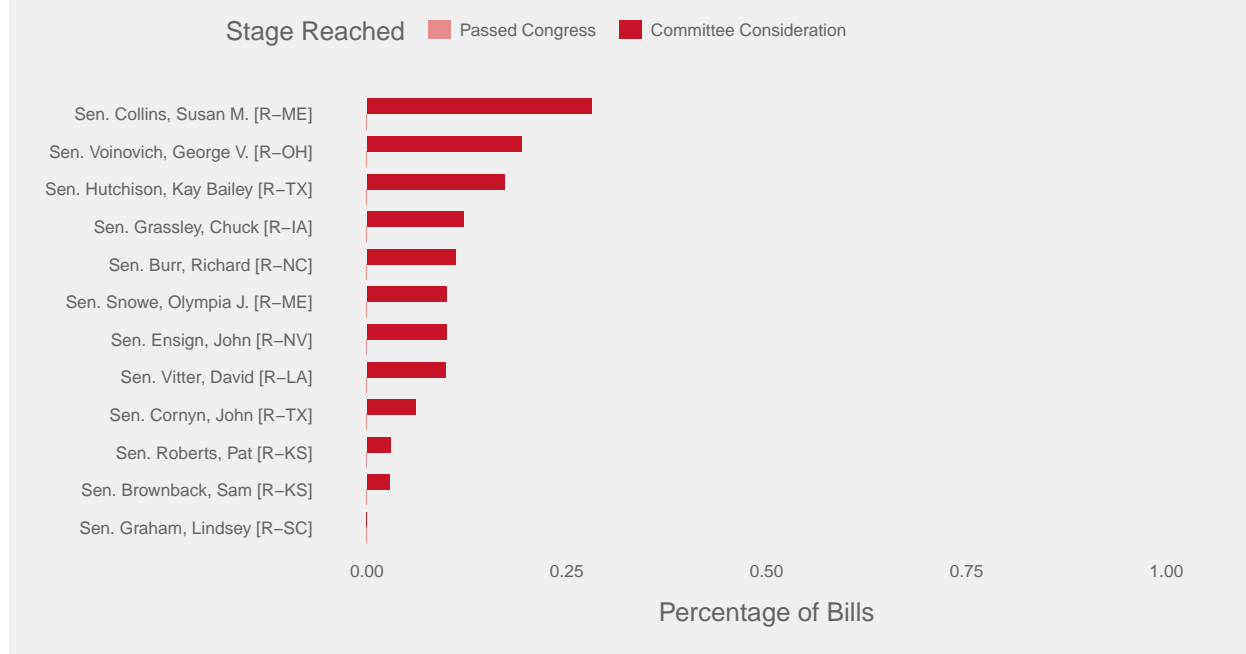


This chart shows the variety of legislative success that Democrats in Congress have in passing legislation. The bills sponsored by some members of the House appear to always receive committee consideration while others struggle to get committees to consider their bill. Senators face particularly strong challenges to their

bills receiving committee consideration. James Oberstar seems to have had a uniquely productive year; nine of the 30 bills that he sponsored became law. Upon further investigation, one of these was to rename a building, four of them were related bills about airline safety, and two of them were related to the Coast Guard Authorization Act. That leaves two remaining substantive standalone bills, one on predisaster hazard mitigation and one on tax relief and unemployment insurance. This remains a highly productive legislative session for Representative Oberstar, but perhaps not quite as much of an outlier as the visualization makes it seem.



Bills Sponsored by GOP Senators



These visualizations demonstrate the challenges facing Republicans in the Democratic Controlled House. None of the 12 most prolific bill sponsors in either chamber successfully passed more than one law that they sponsored. GOP Senators faced particularly stiff challenges as very few of their bills received committee consideration. In particular, Lindsey Graham, despite sponsoring 47 bills in the 111th Congress, was not able to receive consideration for a single one of them in a committee.

Section 3: Modelling Lobbyism

The first quantity modeled is the number of times that a bill is lobbied. There are two ways of modelling this variable. One of these is to use a model that is designed for count data. While poisson is the typical method for count data, it assumes that the mean and variance are the same; however the mean of the lobbying counts is 53.64352 and the variance is 80222.47, which are nowhere near close to each other. As such, a negative binomial distribution should be used instead, which is appropriate for count data with overdispersion.

An alternative would be to apply a transformation to the lobbying counts and create a linear model. It is theoretically preferable to use a negative binomial model, but a linear model may be more interpretable. As demonstrated in section 2, the lobbying data is highly skewed right, but appears nearly normal when log transformed. That said, the transformation did not make a perfect normal distribution so it is worth conducting a box-cox transformation, which also suggests that a log transformation is appropriate (see Appendix A).

For both models the observed variable is the number of times lobbied, the observation is the Congress.gov information about each bill, and the potential predictors are:

- The date introduced
- The party of the sponsor
- The main committee the bill is assigned to (34 levels)
- The number of committees the bill is assigned to

- The Policy Area (34 levels)
- The number of legislative subjects (log transformed)
- The number of cosponsors (log transformed)
- The number of related bills (log transformed)
- The number of actions taken on the bill (log transformed)
- The final state of the bill (7 levels: Introduced, Committee Consideration, Discharge Petition, Calendar Entry, Floor Consideration, Resolving Differences, and Passed Congress)

Chamber of Origin is also a potential predictor, but since there are no committees that are in both chambers it would be redundant to include both in a model. It is therefore only included as a potential bivariate relationship and is excluded from the multivariate models.

A negative binomial model assumes that data comes from a negative binomial distribution that can model count data. It can be thought of as flipping a weighted coin a number of times and counting how many heads appear before a set number of tails appear. (Roback & Legler, *4.9 Overdispersion* 2021)

Sampling Y from a binomial distribution with parameters r and p is equivalent to the following:

$Y|\lambda \sim \text{Poisson}(\lambda)$ and $\lambda \sim \text{Gamma}(r, \frac{1-p}{p})$ (Roback & Legler, *4.9 Overdispersion* 2021)

In this case, $E(Y) = \frac{rp}{1-p} =: \mu$ and $Var(Y) = \frac{rp}{(1-p)^2} = \mu + \frac{\mu^2}{r}$. The difference between the mean and the variance is thus determined by r , which means r can be estimated by observing the overdispersion of the response variable (Roback & Legler, *4.9 Overdispersion* 2021).

It is a generalized linear model with $\log(1-p)$ as its canonical link function.

The assumptions of a negative binomial distribution are:

- The response variable is generated from a negative binomial distribution
- The responses are independent from each other
- $\log(1-p)$ is linear with respect to the variables of X (Roback & Legler, 2021)

The fact that this is overdispersed count data gives strong reasons to believe the first assumption and the visualizations demonstrate a strong likelihood of meeting the third assumption. It is unlikely that the lobbying counts of each bill are independent given the complex dynamics of the legislative and political system, but that might not have bearing on the effectiveness of the model.

The linear model on the other hand assumes that Y is sampled from a normal distribution with a mean equal to $X\beta$, where β is a vector of coefficients and a constant error. It can therefore be simply written as:

$Y = X\beta + \varepsilon$ where ε is a vector with the same length as Y that has values sampled from $N(0, \sigma^2)$

The assumptions of a linear model are:

- There is a linear relationship between Y and X
- The errors are independent
- The values of Y are normally distributed at all levels of X
- The variance of Y is constant for all values of X (Roback & Legler, *1.3 Assumptions for Linear Least Squares Regression* 2021)

While there is a similar issue with meeting the independence of errors assumption, the linear relationship and normal distribution of Y can be created with the log transformation. The assumption of normally distributed errors with constant variance can be checked after fitting a model. The transformation, however, makes the log linear model a less neat fit than the negative binomial model.

Two types of models were fit for each class of model. First, a series of bivariate relationships. Second, a multivariate model with predictors chosen by the LASSO. To create this model, a LASSO regression on a

model with every possible predictor was fit after normalizing each variable. The predictors whose coefficients were 0 were removed and then a best fit model was fit.

The following tables summarize the bivariate relationships in both the log linear model and the negative binomial model.

Log Linear Model of Lobbying
Bivariate Relationships Table

Variable	Linear Coefficient	R Squared	Adj R Squared	Linear AIC	Linear Model P Value
Subjects Count	0.834	0.112041744	0.111929387	27836.58	<0.001***
Max Type	See Appendix B	0.048006595	0.047162737	28399.03	<0.001***
Actions Count	0.549	0.048706512	0.048586141	28381.21	<0.001***
Policy Area	See Appendix B	0.100915736	0.097146230	27999.01	<0.001***
Related Bills Count	0.703	0.061772607	0.061653889	28271.89	<0.001***
First Committee	See Appendix B	0.060404013	0.056464657	28347.41	<0.001***
Sponsors Party	-0.445	0.017636549	0.017512246	28635.27	<0.001***
Committees Count	0.161	0.007660063	0.007534499	28715.15	<0.001***
Introduced Date	-0.002	0.057122811	0.057003505	28310.97	<0.001***
Cosponsors Count	0.296	0.076145046	0.076028147	28149.86	<0.001***
Chamber	0.117	0.001465973	0.001339625	28764.34	<0.001***

Negative Binomial Model of Lobbying
Bivariate Relationships Table

Variable	Negative Binomial Coefficient	Negative Binomial AIC	Negative Binomial Model P Value
Subjects Count	1.033	71306.33	<0.001***
Max Type	See Appendix B	72219.09	<0.001***
Actions Count	0.837	72338.00	<0.001***
Policy Area	See Appendix B	72423.25	<0.001***
Related Bills Count	1.05	72505.24	<0.001***
First Committee	See Appendix B	73038.04	<0.001***
Sponsors Party	-1.099	73628.05	<0.001***
Committees Count	0.321	73865.27	<0.001***
Introduced Date	-0.002	73986.51	<0.001***
Cosponsors Count	0.118	74226.16	<0.001***
Chamber	-0.06	74333.93	0.09

Coefficients and AIC values should not be compared for different types of models. The coefficients in different models refer to different relationships. The AIC for the negative binomial model is also significantly higher because it measures penalized deviance from the full lobbying values while the linear model AIC measures penalized deviance from the log transformed values. According to the AIC of the negative binomial model, the strongest predictor is the number of legislative subjects. In general, the predictors related to the topic of the bill and how far it gets are strongest. The weakest predictor is the originating chamber, which is not going to be included in the full model anyway.

Next, we use LASSO to perform model selection as described above. The selected regularization parameter For Linear LASSO was the highest regularization parameter within 1 standard error of the lambda with the minimum cross validated mean square error.

After performing the LASSO on the log linear model and selecting λ , 51 predictors of the original 80 predictors remained. The variables that had become zero were: 9 policy area predictors (Animals, Economics and Public Finance, Emergency Management, Foreign Trade and International Finance, Housing and Community

Development, Labor and Employment, Law, None, and Private Legislation), 16 committees of origin (Banking, Housing, and Urban Affairs; Budget; Education and the Workforce; Energy and Natural Resources; Environment and Public Works; Foreign Relations; Homeland Security and Governmental Affairs; Homeland Security; Select Intelligence; Judiciary; Natural Resources; Rules and Administration; Transportation and Infrastructure; Veterans' Affairs; and Ways and Means), and 4 steps of the legislative process (Introduced, Discharge Petition, Floor Consideration, and Passed Congress). Interestingly, of the three steps that remained: Committee Consideration, Resolving Differences, and Calendar Entry, two of them are consecutive in the legislative process: Committee Consideration and Calendar Entry. This suggests that perhaps getting out of the committee stage is what lobbyists consider the crucial step in passing a law.

The full table of coefficients for the selected linear model can be found in Appendix C.

Turning first to the predictors other than the 3 categorical variables:

- There was a highly significant positive effect of the number of legislative subjects, indicating that more complex laws draw more interest from lobbyists
- Bills sponsored by Republicans received less interest from lobbyists in this democratic controlled Congress
- Bills with more cosponsors were lobbied more
- Bills with more actions on them were more likely to be lobbied.
- Bills introduced later in the session were lobbied less
- Bills with more related bills were lobbied more

Many of these relationships were visible in the visualizations.

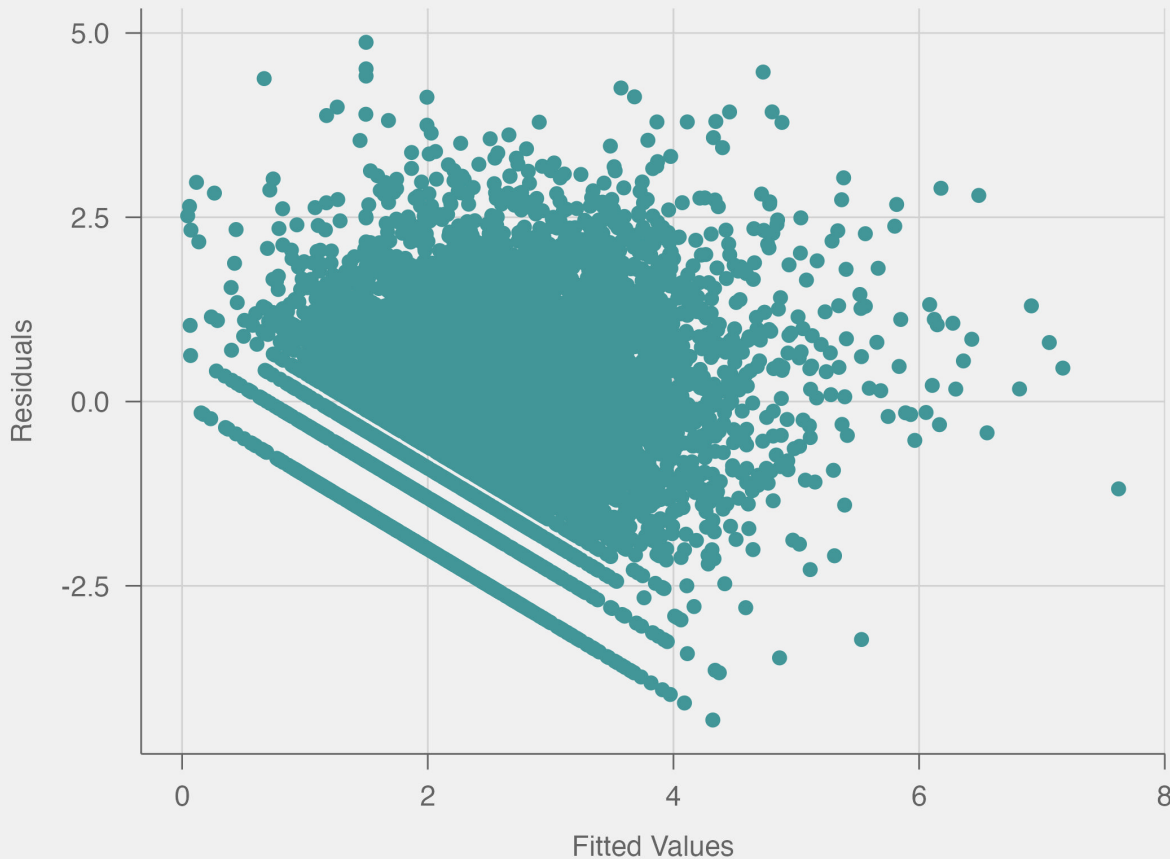
Bills were lobbied more at the Calendar Entry and Discharge Petition and lobbied less at the Committees stage. This supports the theory that the transition between committee and the floor draws significant lobbyist attention, which is visible in the sharp jump in average lobbying between the Committee Consideration Stage and the Calendar Entry Stage.

As was also evident in the visualizations, bills from the appropriations committee were heavily lobbied, as were bills from the Finance Committee. The bills from the Agriculture, Nutrition, And Forestry Committee were also heavily lobbied, which was not as evident in the visualizations. The small business committees in both chambers had the strongest negative coefficient, which perhaps reflects the fact that small businesses do not have the resources for major lobbying operations.

The issues that were lobbied the most heavily were Science, Technology, Communications; Finance; and energy. Finance and energy were two of the most hotly debated issues in this legislative session due to the financial crisis and the proposed cap-and-trade system, so these are not surprising (Cohen, 2010). It would be interesting to investigate why the Science bills were highly lobbied more because it does not immediately appear to be a major issue or to be connected with a major special interest. On the other side of the spectrum, bills about Sports, Congress itself, Social Science/History, and the Arts had the lowest coefficients. These tend to be low-temperature issues without strong special interests in play so it makes sense that they would not draw significant amounts of lobbying.

A plot of the residuals against the fitted values can indicate how well a linear model meets the assumptions. While most of the plot shows that the model meets the assumptions, the diagonal lines in the bottom left hand corner indicate a problem with results with low fitted values. This is reflective of this model's poor theoretical fit with the count data and counts against selecting it as a model for lobbying if the negative binomial model can perform similarly well.

Log Linear Model Fitted Values vs Residuals



The R function for LASSO regression on a negative binomial model from the “mpath” package does not have the same cross validation function as the function for LASSO regression in linear models from the “glmnet” package, but it provides the AIC and BIC for the models at each value of λ . The selected value of λ was whichever of model with the optimal AIC or BIC that had fewer predictors. In this case, that was the λ with the lowest BIC, which produced a model with 65 predictors selected from the original set of 80 potential predictors. The 15 predictors removed were all categorical like with the linear model. It removed four stages of the legislative process (Introduced, DischargePetition, FloorConsideration, and PassedCongress), two policy areas (Economics and Public Finance, Health), and nine committees (Energy and Commerce, Foreign Relations, Homeland Security and Governmental Affairs, Select Intelligence, Natural Resources, Rules and Administration, Veterans’ Affaris, and Ways and Means). This is a similar set of predictors that were removed in the linear model, including exactly the same steps of the legislative process as the linear model.

The full table of Coefficients can be found in Appendix D.

In general, the predictors of the negative binomial model are similar to those of the log linear model.

The direction of the coefficients for the three legislative steps is the same as the direction in the log linear model and have similar values.

The coefficients of the committees are in a similar order as in the linear model. Appropriations has the strongest positive coefficient, followed by Budget, Agriculture, and Finance. Just as in the linear model, the strongest negative coefficients were the small business committees.

As with the linear model, economic, energy/environmental, and health subjects had the highest coefficients and the lowest were low-salience subjects like Congress, Social Sciences, and the Arts. Civil rights was also among the policy areas with the lowest coefficients. There are no industries that would highly lobby on this subject, but it is higher salience than other subjects with low coefficients so it is somewhat surprising that it has such a low coefficient.

The Root Mean Square Error of the log linear model after performing 5-means cross-validation and exponentiating the predictions was 268.2252. The Root Mean Square Error of the Negative binomial model after performing 5-means cross-validation was 261.7179. This means that the Negative Binomial has a stronger theoretical relationship to the data, stronger case to meet the model assumptions, and a stronger fit to the data, though it is possible that with a different iteration of the 5-means cross-validation the linear model would have performed better. For both models, even given the scale of the data, this is a relatively high error, which reflects the high variance in the lobbying counts. The high RMSE suggests that this model may be more useful to understand the relationships within the lobbying data than it is for predicting how much a bill will be lobbied.

Often, linear models are more interpretable than other models, but this advantage is weak when a transformation is applied to the response. As such, it seems that the negative binomial model is likely superior to the log linear model.

Section 4: Modelling Final Bill State

Another type of model is that can be built with this data is a model for how far a bill reaches in the legislative process. This treats the final legislative state of the bill as y , the response variable, and the log number of times lobbied as one of the columns in X . The other predictors used in the modelling in Section 3 can also be used as predictors in this section with the exception of the number of actions because that variable is essentially a proxy for how far a bill gets so it is not an interesting relationship to analyze. Including it could also dilute the effects of other predictors and interfere with potential insights.

The path of a law in Congress can be complex, but in general it follows a series of steps beginning with its introduction, followed by consideration in committee, then consideration on the floor of Congress. As a result, the bill state should be considered an ordinal categorical variable. For the purposes of modelling the many legislative steps were simplified into 5:

1. Introduced into Congress
2. Considered by a Congressional Committee: This can involve any type of hearing, markup, or vote on the bill by a committee or subcommittee
3. Passed by a Congressional Committee: After a bill is passed in committee, it is placed on the legislative calendar. Some bills can bypass the committee process with a discharge petition, which is a rare process used to consider a bill on the Floor of Congress even if it hasn't passed committee. Bills whose highest state was either "Calendar Entry" or "Discharge Petition" were
4. Floor Consideration: Not all bills that pass committee receive consideration by the entire chamber. Bills in this state were debated or voted on in the floor of Congress, but did not pass both chambers. If a bill passed both chambers but the chambers were not able to agree on the same version of the law, it was also classified in this state.
5. Passed Congress: Bills that passed both chambers and went to the president were included in this category. All but one of these bills became law, the other was vetoed.

(Congresswoman Eleanor Holmes Norton, How a bill becomes a law)

Because the response variable is ordinal, the appropriate model is an ordinal logistic regression model. This models the log odds of the cumulative probability of being up to a given state in the model as a linear

function of its predictors, whereas traditional multinational regression models models the log odds of the probability of being at one state.

This is represented as:

$\log\left(\frac{P(Y \leq l)}{P(Y > l)}\right) = \zeta_l - \eta_1 X_1 - \eta_2 X_2 \dots - \eta_p X_p$ for a model with p predictors where l represents one of the states in the model (Nahhas, 6.22 *Ordinal Logistic Regression*).

Each ζ term is analogous to an intercept and is the log odds of being anywhere up to state l when all predictors are at 0 or their reference level. Each η term is analogous to a β value and returns the log odds ratio of the cumulative probabilities. A positive value of η corresponds to a higher cumulative odds of being at higher levels (Nahhas, 6.22 *Ordinal Logistic Regression*). These coefficients can be more challenging to interpret than those in a linear model, but when they are exponentiated, they are odds ratios which is easier to interpret.

The assumptions of Ordinal Logistic Regression are:

- The dependent variable is measured on an ordinal level
- One or more of the independent variables are either continuous, categorical or ordinal
- No Multi-collinearity
- Proportional Odds - each independent variable has an identical effect at each cumulative split of the ordinal dependent variable

As discussed, the first assumption is met because the stages of the legislative process follow each other in a set order. The second assumption is met as all predictors are either continuous or categorical. There may be multi-collinearity in the variables, but not significant enough to affect the validity of the results. Finally, there is no evidence to suggest that the Proportionality Odds assumption is either met or not met. It could be true, for example, that lobbying is highly associated with bills moving from the introduced stage to being considered in committee but not associated with any other steps. If the model performs poorly and unevenly, a violation of this assumption could be one reason why.

The following table summarizes the bivariate ordinal models between the bill state and all of the potential predictors.

Ordinal Model of Bill Progress
Bivariate Relationships Table

Predictor	Coefficient	SE	t.value	p.value	AIC
First Committee	See Appendix E	See Appendix E	See Appendix E	See Appendix E	14209.33
Policy Area	See Appendix E	See Appendix E	See Appendix E	See Appendix E	15709.53
Chamber	-1.795	0.053	-33.754	<0.001***	16613.79
Committees Count	0.742	0.031	23.644	<0.001***	17214.34
Subjects Count	0.981	0.038	26.015	<0.001***	17253.95
Cosponsors Count	0.284	0.015	18.324	<0.001***	17603.13
Lobbying	0.227	0.015	15.235	<0.001***	17711.33
Cosponsors Count	0.009	0.001	13.625	<0.001***	17756.53
Related Bills Count	0.583	0.043	13.61	<0.001***	17760.76
Introduced Date	-0.184	0.022	-8.322	<0.001***	17874.05
Sponsors Party	-0.362	0.049	-7.404	<0.001***	17888.56

According to the AIC, the strongest predictors by a significant margin were the categorical predictors of the committee, the policy area, and the Chamber of the bill. That suggests that there are some subjects where bills are more likely to be passed and some committees that move bills more quickly. Lobbying had a moderately strong and significant effect on bills moving through Congress, but it was not as individually

impactful as the first committee, the policy area, or even predictors like the number of committees, the number of subjects, or the number of cosponsors.

Just as with the log linear model and the negative binomial model for predicting lobbying, variable selection can be performed with LASSO. A function to perform LASSO model selection with cross validation can be performed on ordinal models is contained in the ordinalNet package. The function performs 5-fold cross-validation on the data and recommends a λ regularization parameter for each fold. The selected one for the whole model is whichever of these recommended values is the highest. In this case, the model recommended two different λ values depending on the fold, but the one that was chosen is inconsequential because the predictors that are zero are identical in each group. The regularization recommended the removal of nine predictors: 7 policy areas (Commerce, Environmental Protection, Families, Housing and Community Development, Private Legislation, Social Sciences and History, and Transportation and Public Works) and 2 committees (Oversight and Accountability and Business and Entrepreneurship).

The full table of coefficients is in Appendix F. The p-values are calculated using a Likelihood Ratio Test performed with the Anova() function in the car package, which is recommended for large ordinal models (Nahhas, 6.22 *Ordinal Logistic Regression*).

The variable of highest interest is lobbying, which has a positive, statistically significant coefficient. The number of legislative subjects, the number cosponsors, and the number of committees all had positive coefficients, indicating that higher values are associated with bills moving further in Congress. Bills introduced later in the session and bills introduced by Republicans did not go as far as bills introduced earlier in the session.

One interesting detail is that some of the predictors that had the lowest coefficients in the model of lobbying have very high coefficients in this model and vice versa. The Small business committee, arts bills, sports bills, and civil rights bills, for example all have high coefficients. On the other hand several economic and energy issue areas (Finance, Energy, Economics, Foreign trade, and taxation) and committees (Finance) have very low coefficients. There are other issues and committees that have strong coefficients in the same direction in the models with both response variables such as Appropriations. This is a confusing phenomenon because bills that move further in the legislative process are lobbied more. This strong difference perhaps suggests that there are more complex dynamics in play than simply bills that are lobbied more being more likely to move further. Future work could perhaps see if there are some issue areas where lobbying helps bills move further and some where lobbying typically prevents them from going further.

Using 5-means cross validation, the classification error rate was 30.33%. It was more successful at classifying bills that were either in the “Introduced” or “Committee Consideration” stage and more poorly at later stages in the legislative process, as the table below shows. This suggests that the model, while perhaps a useful tool for understanding relationships in the data, is not a strong predictive tool and further work is necessary. The log-loss was 0.8419733.

Classification Table
Actual States vs. Predicted State

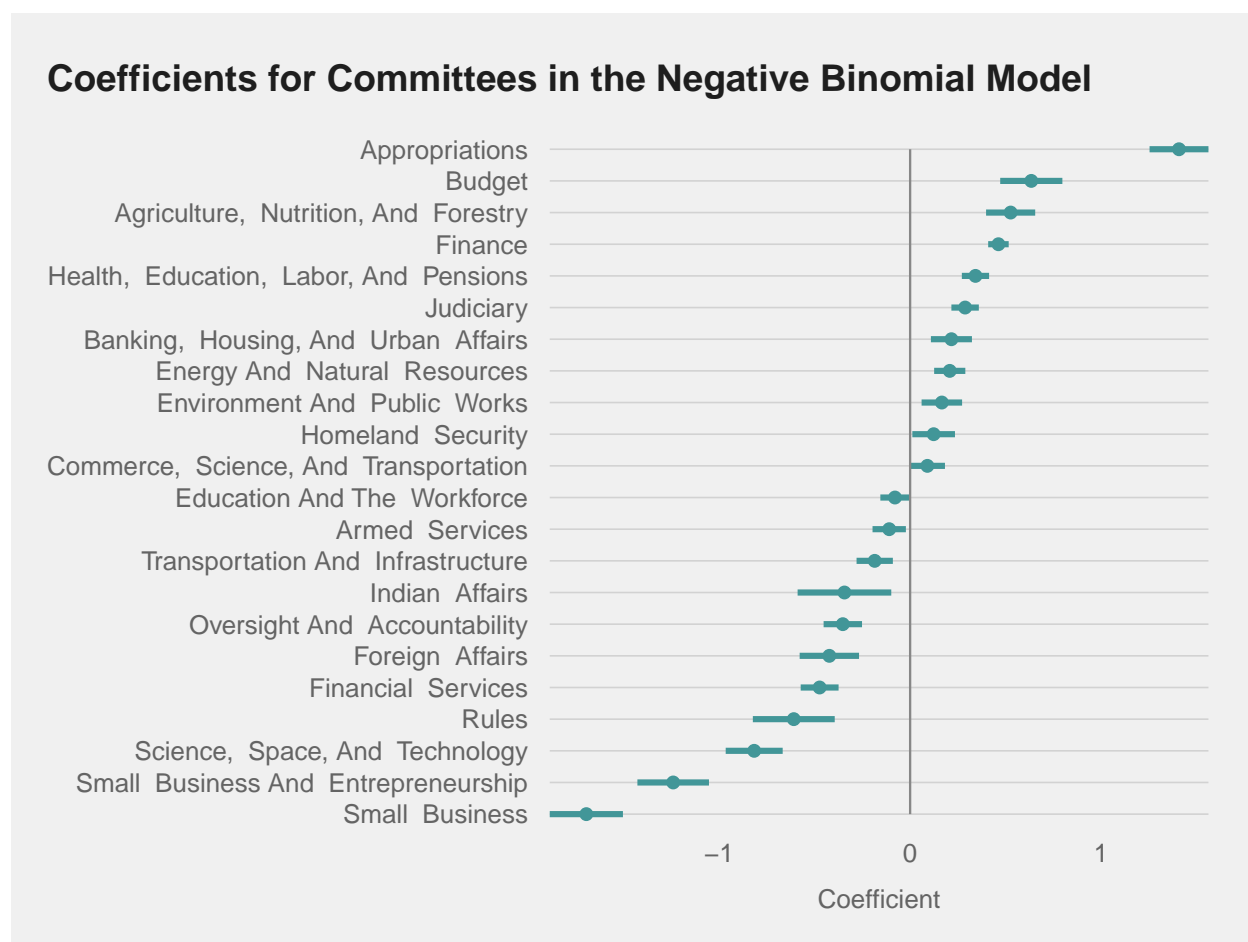
Actual State	Prediction				
	Introduced	Committee Consid.	Passed Comm.	Floor Consid.	Passed Congress
Introduced	3329	795	0	0	1
Committee Consideration	372	2227	0	12	48
Passed Committee	56	266	0	10	26
Floor Consideration	98	364	0	5	27
Passed Congress	67	171	0	5	26

One possible reason that the model performs better at classifying bills that are earlier in the legislative process than those that are later is that the model violates the proportional odds assumption. Future work could potentially replace the ordinal logistic model with a series of four binary logistic regressions where the

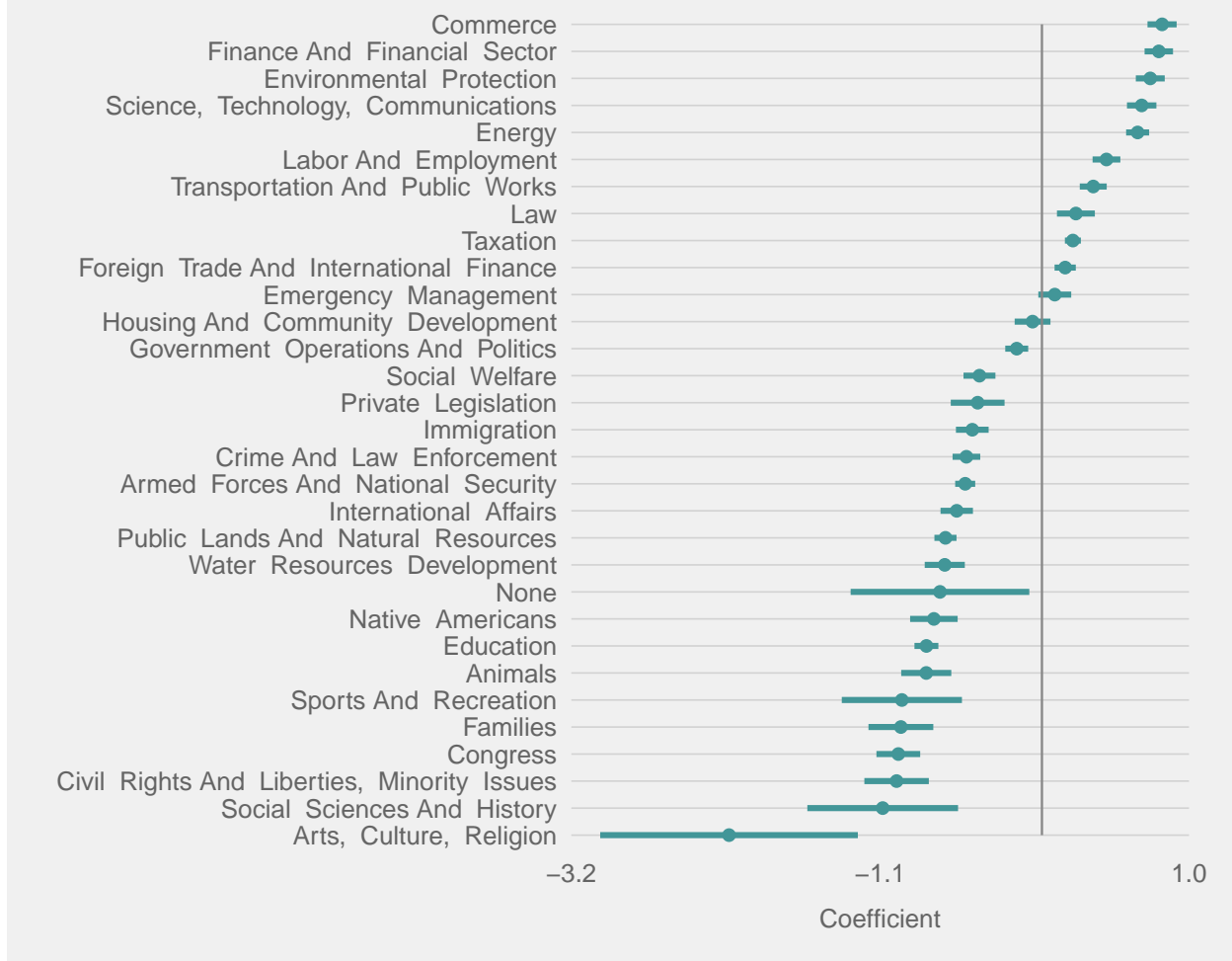
states are reaching at minimum each of the four states past Bill Introduction and infer probabilities of being in each state based on the results of those models. This would allow the predictor coefficients to vary at each state and might yield more accurate predictions.

Section 5: Visualization and interpretation of the results

The predictors of lobbying that are among the most relevant to improving government transparency are the Policy Area and the Committee. Understanding which policy areas are most lobbied helps understand which issues special interests are most invested in. Understanding which committees generate the most lobbying activity could likewise shed light into where lobbyists' points of entry into Congress tend to be.



Coefficients for Policy Areas in the Negative Binomial Model



The Committees that had the highest coefficients:

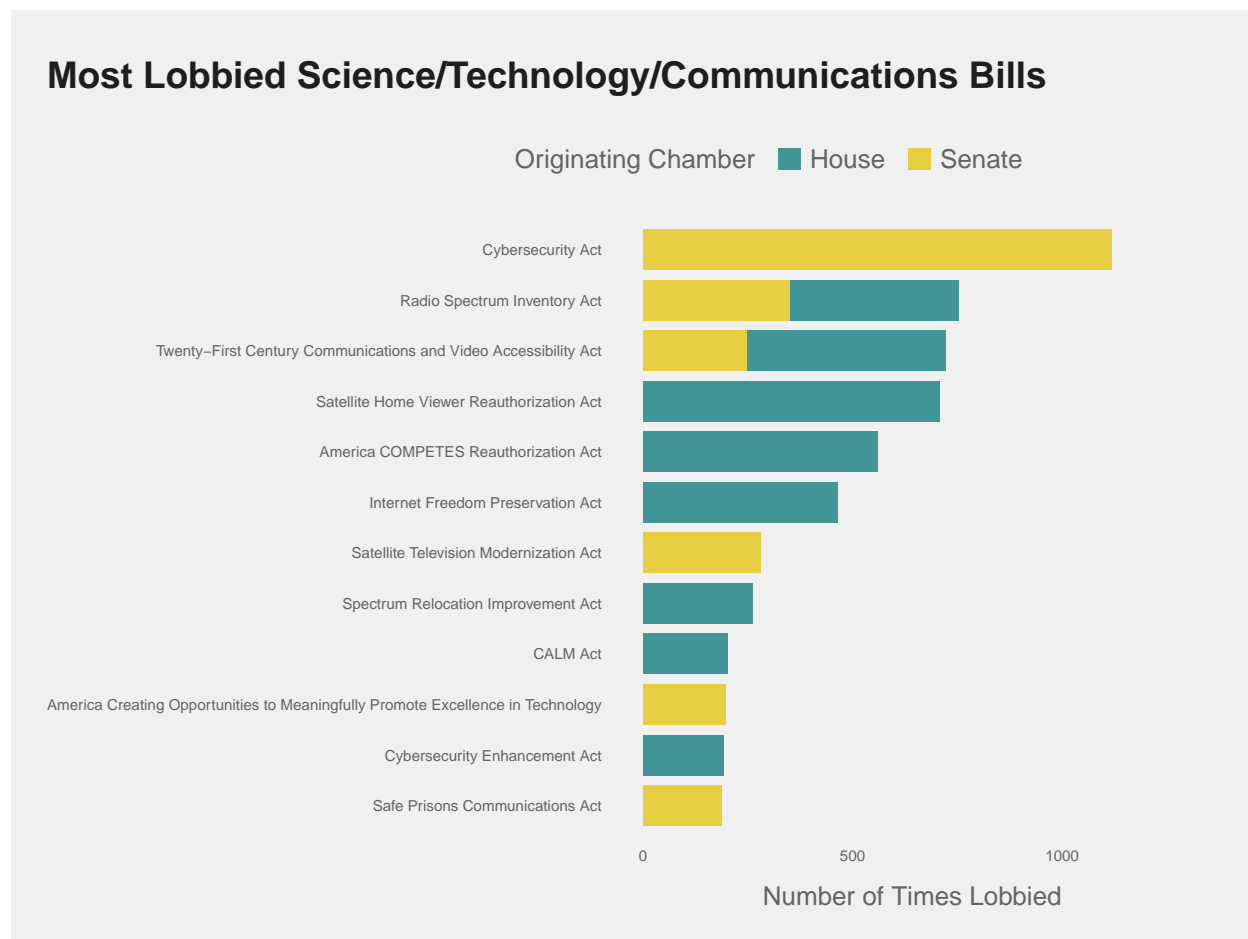
- Appropriations
- Budget
- Agriculture, Nutrition and Forestry
- Finance
- Health, Education, Labor and Pensions

The Policy Areas that had the highest coefficients were:

- Commerce
- Finance and Financial Sector
- Environmental Protection
- Science, Technology, and Communications
- Energy

Some of these committees are highly predictable such as anything related to health, finance, or the energy and environment as those were the major legislative issues of the 111th Congress (Cohen, 2010). Science,

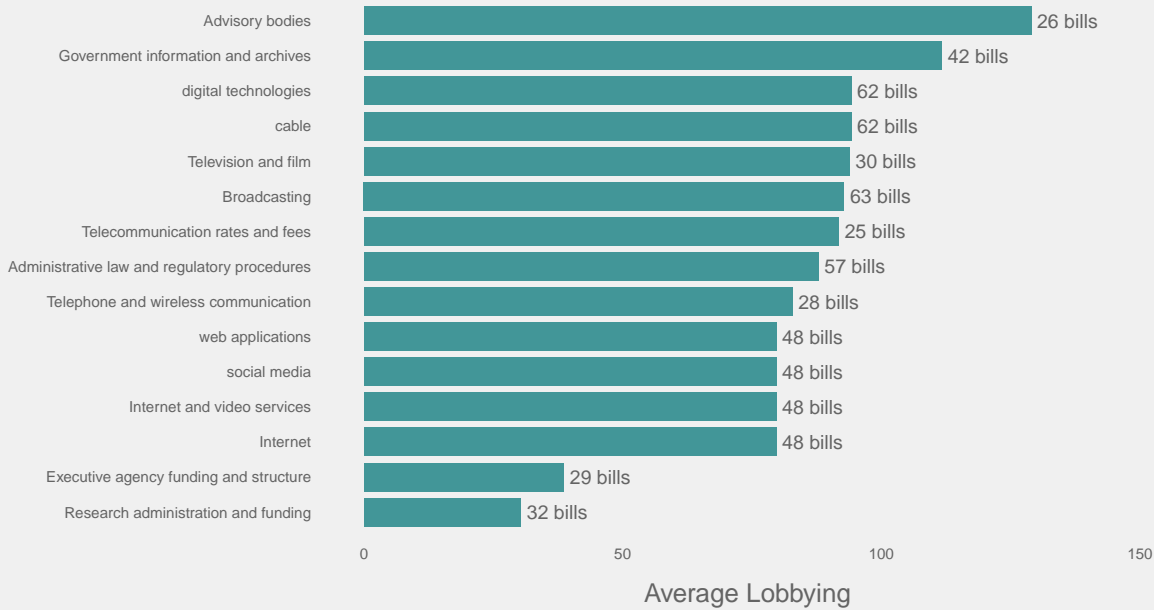
Technology, and Communication is more surprising. Commerce (the subject) and Appropriations are also both very vague and broad subjects, and it would be interesting to investigate in more detail.



The main topics discussed in the most lobbied Science, Technology, and Communications bills were primarily covering cybersecurity and telecommunications related.

In addition to the 32 category policy grouping, the Library of Congress also sets a much more detailed policy are with 1,004 groups and a bill can fit into multiple of these categories (Library of Congress, Find bills by subject and Policy Area). The following chart shows which of these sub-subjects were the ones that were most lobbied within Science, Technology, and Communications bills.

Most Lobbied Science and Technology Subjects



This plot shows the 15 legislative subjects that had the highest number of Science, Technology, and Communications Bills. Somewhat surprisingly, the most highly lobbied subject among the top 12 subjects was Advisory bodies followed by Government information and archives. Perhaps these bills are highly relevant to government contracts and government contractors are lobbying for them to be passed. Subjects like digital technologies, cable, television and film, broadcasting, and the internet reinforce that much of the lobbying was driven by telecommunication utilities bills.

Another rich source of data about bills is their summaries produced by Congressional Research Services (CRS). By exploring the difference in the summaries of the most lobbied bills, it is possible to learn on an even more granular level what subjects attract the attention of lobbyists.

First, it is useful to understand the nature of the summaries before considering lobbying. This is accomplished by creating a document term matrix for the CRS summaries weighted by term frequency-inverse document frequency (TF-IDF). A document term matrix is a matrix where every row corresponds to a document, in this case the CRS summaries, and every column refers to a word that appears in any of the documents. The entries are then calculated as follows:

$DTM_{ij} = tf_{ij} * \log(\frac{N}{df_j})$ Where tf_{ij} is the number of times that word j appears in document i , N is the number of documents in the corpus, and df_j is the number of documents that word j appears in. This statistic is designed to give high values for words that are not only frequent in document i , but are also relatively unique to document i . Entries are then normalized so that each row has the same sum to prevent longer documents from dominating the results.

The DTM was prepared using summaries of all topics, so the TF-IDF statistic highlights the words that are unique to the Science, Technology, and Communication Category and deemphasizes words that appear in many CRS summaries across all topics. The data was prepared by grouping similar words together and eliminating simple words that are common in all English speech such as “we”, “a”, and “such.”

This shows the words corresponding to the columns with the highest sum of TF-IDF values in the rows corresponding to Science, Technology, and Communication bills.

Bills about Science and Technology

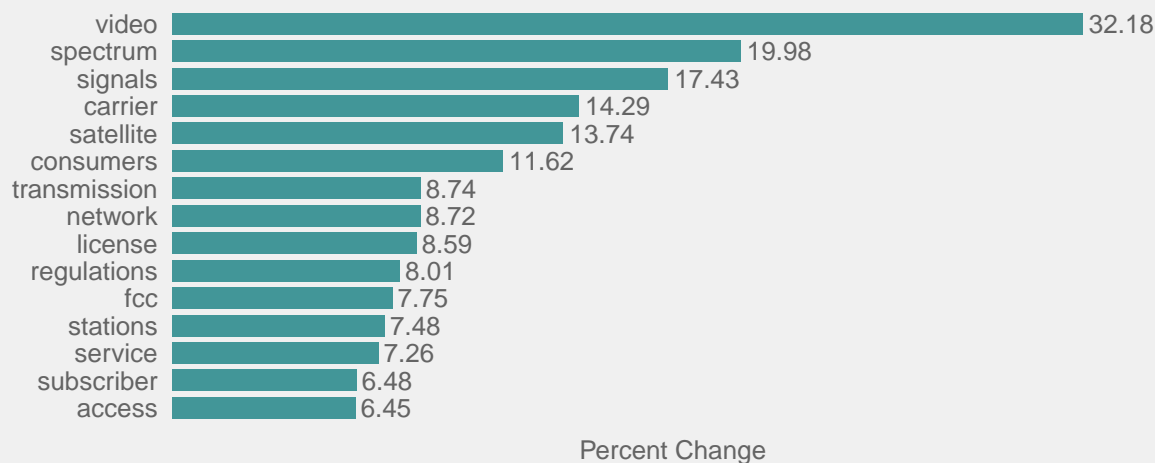


Words like broadcasting, broadband, and television appear commonly, underscoring the importance of telecommunications utilities in this category. The FCC, the agency that regulates telecommunications is also a common word (FCC, *What we do*).

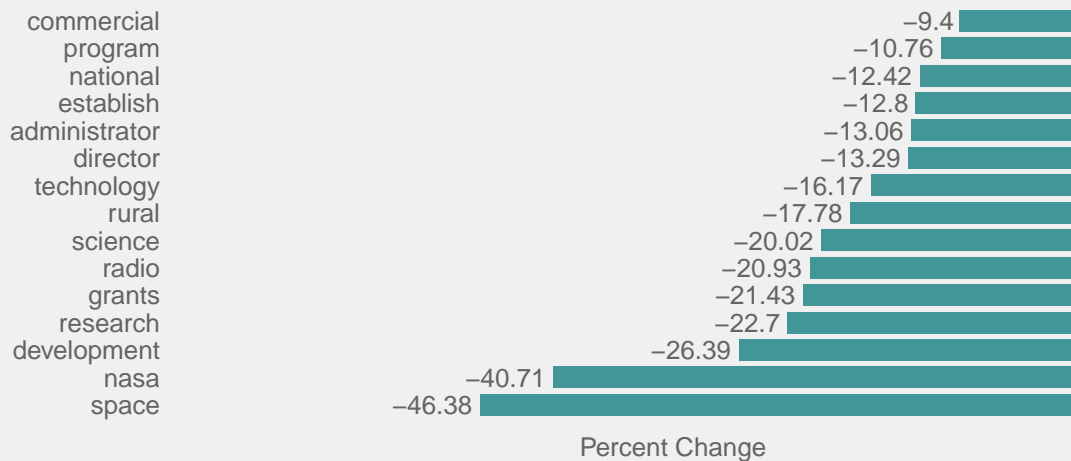
To understand which words appear in the most lobbied bills, the TF-IDF scores can be weighted by the log number of times lobbied. Each row in the Document Term Matrix was multiplied by the log number of times that the bill the summary was written for was lobbied. The column sums of these weighted values were then taken. The percent change in the z-score from the unweighted value to the value weighted by log lobbying was then taken. Words with high percent change in the z-score appear more frequently in bills that are lobbied more while words with lower percent change in the z-score appear more frequently in bills that are lobbied less.

The following graph shows the 15 words with the highest percentage change and with the lowest percentage change among the 50 words with the highest scores prior to weighing by lobbying.

Top words by Percent Change



Lowest Words by Percent Change



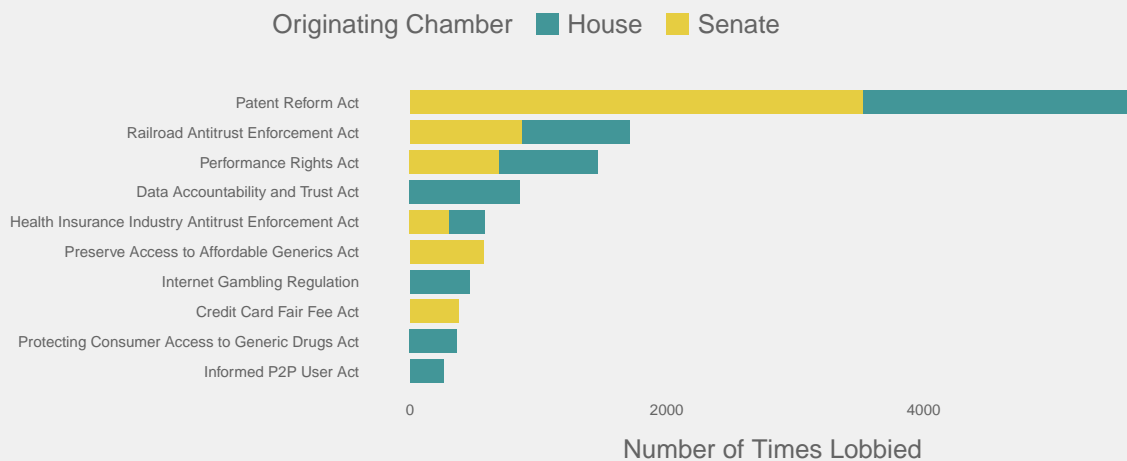
The top words included a lot of words that indicated potential for new spending such as “establish” and “grants.” It also included “federal” and “national” which suggests that nationwide projects are of more interest to lobbyists. The lowest words on the list somewhat surprisingly included broadband and satellite, both of which are related to telecommunications utilities. This suggests that these newer technologies are not as salient to lobbyists. Science and research topics like “science,” “nasa,” and “space” were also included in this list.

Based on the above analysis one can conclude:

- Telecommunication utility bills are of high interest to lobbyists, although older technologies may be more salient than newer ones like broadband and satellite
- Lobbyists are especially interested in scientific projects when there is potential for federal funding
- Cybersecurity is of high interest to lobbyists
- Science and Research is not of high interest to lobbyists

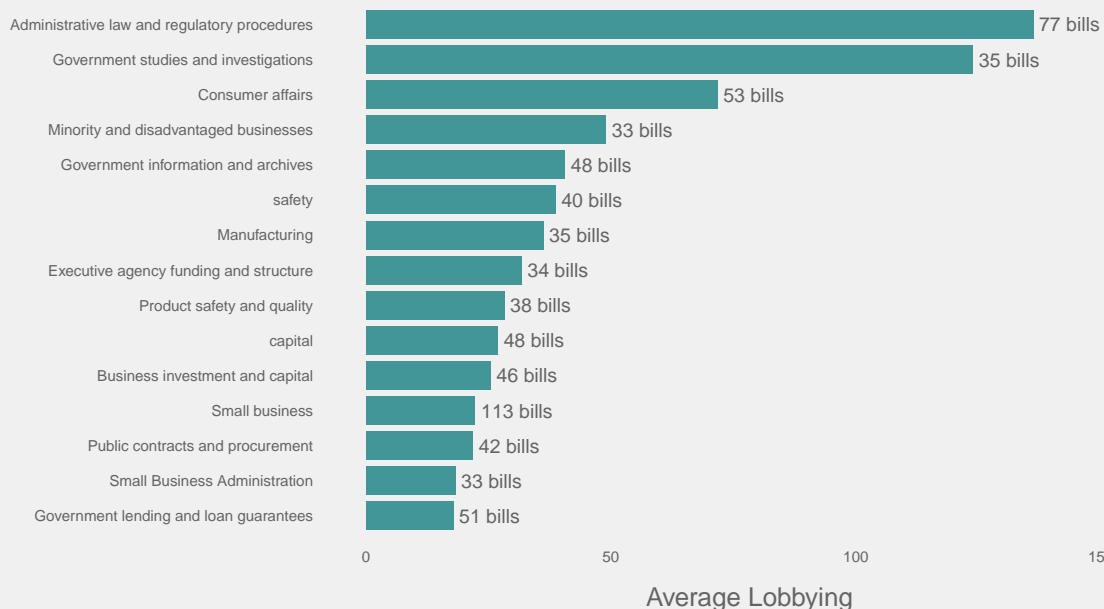
The next subject that generated high amounts of lobbying for initially unclear reasons was Commerce.

Most Lobbied Commerce Bills



This chart immediately offers clues about why commerce bills drew heavy lobbying because the Patent Reform Act generated an unusually high number of interested lobbying clients. The next highest bill, railroad antitrust act may signal that other antitrust bills are drawing heavy interest from lobbyists.

Most Lobbied Commerce Subjects



The chart of the lobbying of the 15 most common legislative subjects shows that the lobbyist interest in Commerce bills primarily comes from government regulation and government oversight. The bills that drew less lobbying interest included Small Business bills, mirroring a finding from modelling.

The following wordcloud was prepared for Commerce bills in the same manner as the one for Science, Technology, and Communications:

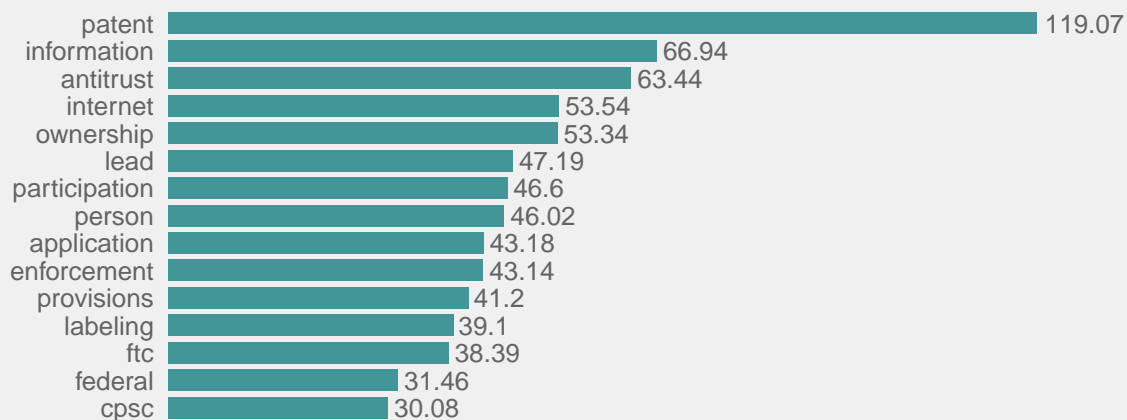
Bills about Commerce



The two most prominent words by far are “small” and “business.” Given that the SBA, the Small Business Administration also appears it is likely that “small” and “business” appear frequently together. Other words include loan, trademark, and antitrust, each of which relates to the regulatory interest in the previous chart (USAGov, Small Business Administration (SBA)).

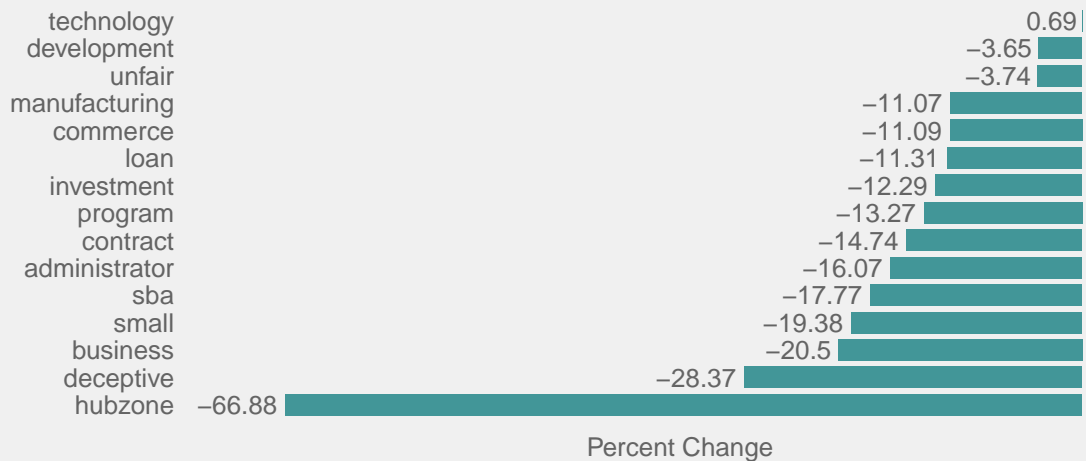
The following barcharts on the percent change in word use between unweighted and weighted TF-IDF scores was prepared with the same method as for Science, Technology, and Communications bills:

Top words by Percent Change



Percent Change

Lowest Words by Percent Change



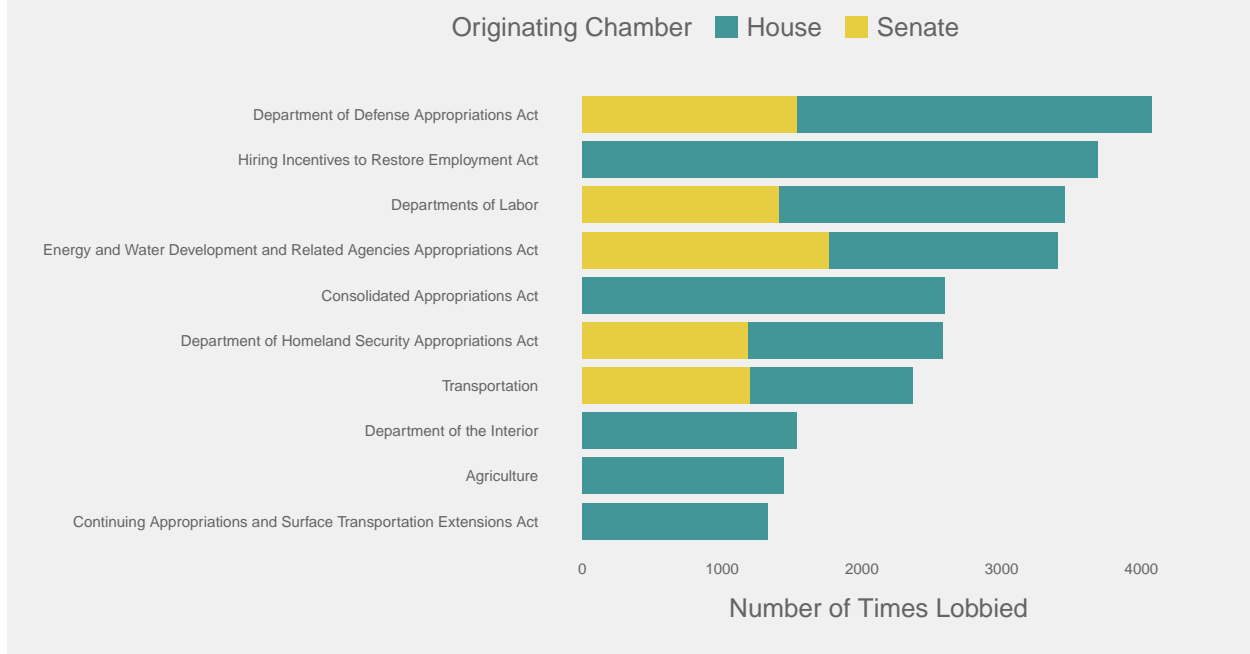
These charts seem to confirm the findings of the analysis of legislative subjects and analysis of the highest lobbied bills. By far, the word with the highest percent change was patent, and the most lobbied bill was about patents. Other words in the top words by percent change include regulatory language such as “antitrust,” “application,” “enforcement,” “provisions,” and “ftc.” On the other hand, the least lobbied commerce bills were about small businesses and seemingly about local investment with words like “hubzone,” “investment,” “investment,” and “loan.”

Based on the above analysis one can conclude:

- Patents are a highly salient issue to lobbyists
- The most salient commerce issues tend to be regulatory
- Small business issues do not draw significant attention from lobbyists

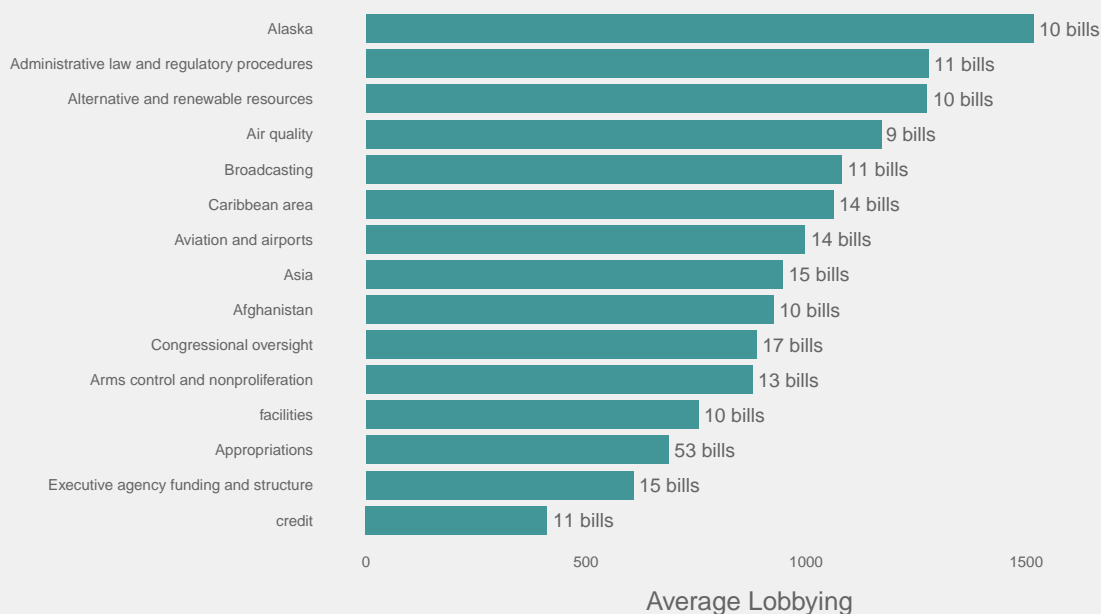
Lastly, the Appropriations Committees of both chambers are among the most powerful and are responsible for creating the federal budget. Given this and the fact that it was a strong predictor of both lobbying and bills advancing, it would be interesting to understand which appropriations subjects are of greatest interest to lobbyists.

Most Lobbied Appropriations Bills



The appropriations committee is responsible for passing several large bills that determine the budget of a portion of the government, and that is reflected in this chart which has the bill to determine how much money several executive agencies will receive (Saturno & Lynch, 2023). These appear to be generating significant interest from lobbyists.

Most Lobbied Appropriations Subjects



Appropriations is a unique subject because it appears that most legislative subjects were only in a few bills. That said, from this chart it seems there is a strong lobbying for government funding on Alaska, alternative energy, aviation, and air quality. This chart also underscores the interest from lobbyists on broadcasting and government regulation.

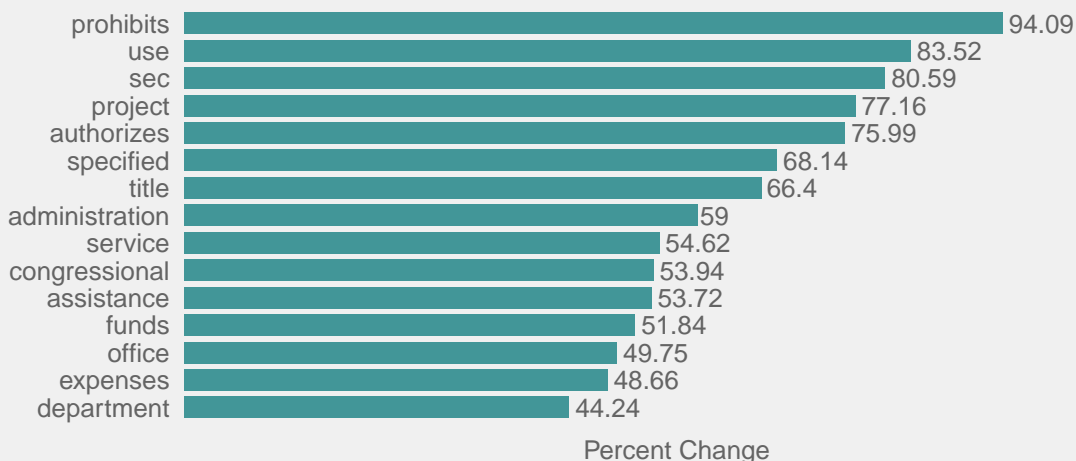
This wordcloud was prepared with the same method as the wordclouds for Science/Technology bills and Commerce bills.

Bills about Appropriations

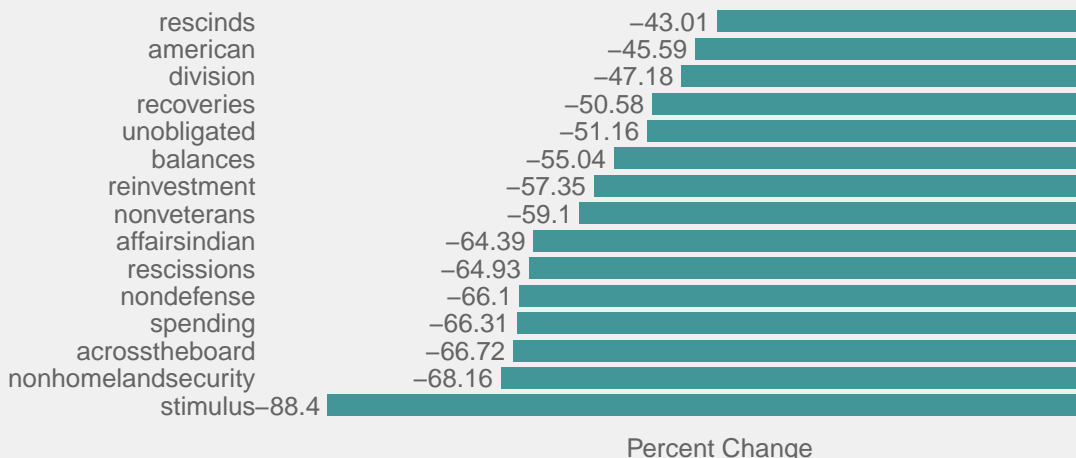


This wordcloud highlights the nature of these bills as spending and appropriations bills, but does not reveal much more about their topics.

Top words by Percent Change



Lowest Words by Percent Change



The words that appear in bills with more lobbying appear to be fairly typical language of spending and budgeting to different agencies and offices. This does not reveal much about what projects are lobbied the most heavily, but it underscores the heavy lobbying on appropriations bills in general. The words with the lowest percent change on the other hand reveal some information about what subjects are lobbied the least. Words like “nonveterans,” “nondefense,” and “nonhomelandsecurity” highlight that defense and military appropriations is lobbied more than other types of appropriations.

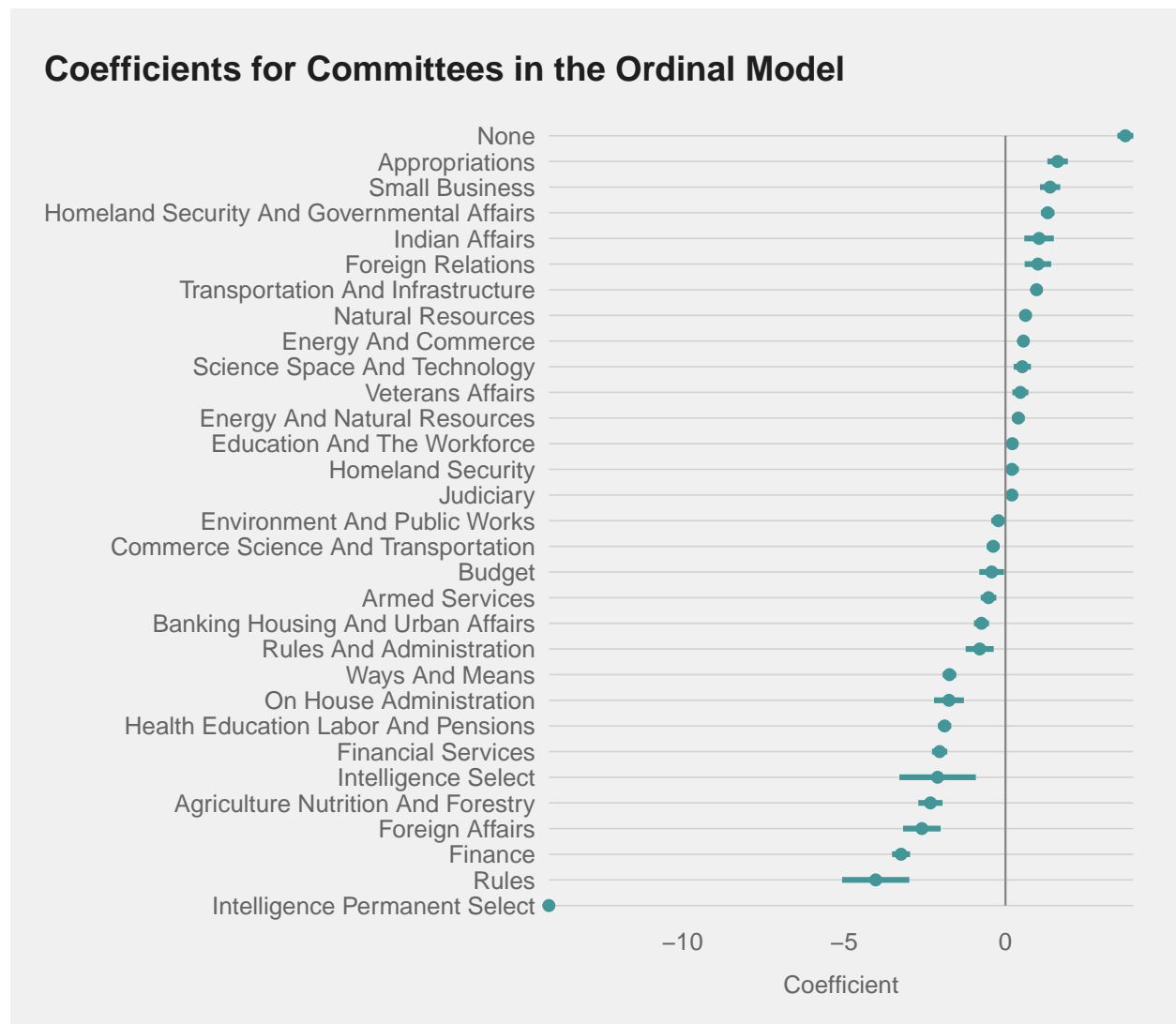
Based on the above analysis one can conclude:

- Appropriations regardless of subject tend to be highly lobbied

- Defense appropriations are among the most lobbied subjects within appropriations
- There is some evidence for strong lobbies on Alaska, alternative energy, aviation, and air quality, though there is no information about which direction that lobbying supports (i.e. whether they support more or less spending on those subjects).

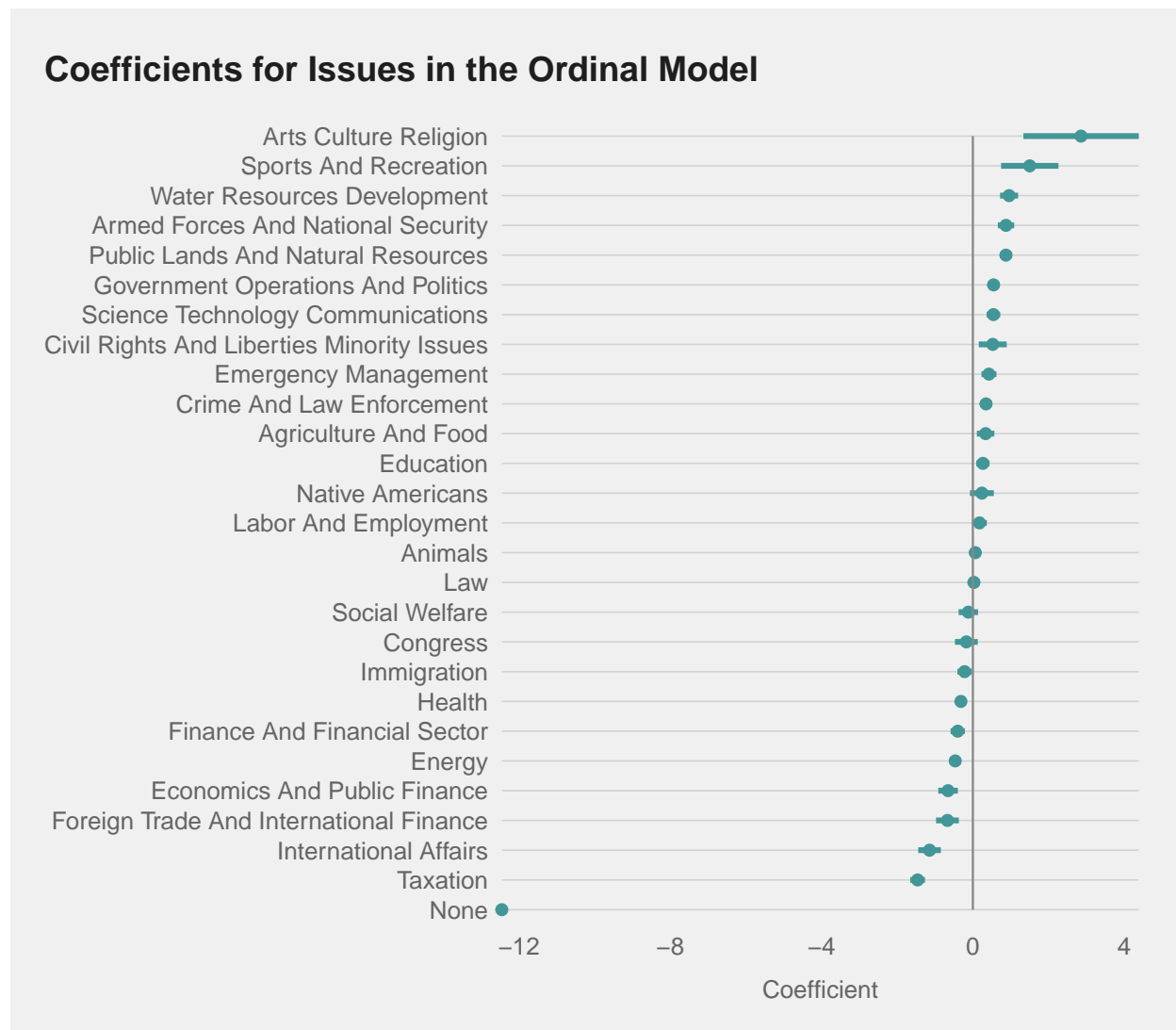
The examples above were chosen based on probability of finding a surprising result, but this type of analysis would be useful for other issues and committees and could perhaps be explored in future work.

It is also relevant to understand the bills from which issues are most likely to advance in Congress. Members of Congress propose bills all the time and understanding which bills are most likely to advance can provide information about which proposed bills are most serious. Similarly, understanding which committees' bills move the furthest could also help political observers understand which committees were the most productive and which are moving the slowest.



It is striking how there are several Committees with strong negative values including Ways and Means, House Administration, Health Education Labor and Pensions, Financial Services, Intelligence Select, Agriculture, Foreign Affairs, Finance, Rules, and Intelligence Permanent Select. These committees might be overworked with too many bill proposals or they might be uniquely gridlocked. The committees with the strongest

positive values tend to be on issues where bipartisanship is known to be common such as Appropriations and Homeland Security or issues with low salience with the public such as Small Business and Indian Affairs.



One striking feature of this model is that many of the coefficients are near zero, which suggests that perhaps more variables could be removed to further guard the model against over fitting. Similar to the committees coefficients, several high salience issues like Finance, Energy, Economics, Foreign Trade, International Affairs, and taxation have very low coefficients. One possible explanation for this is that members of Congress often propose bills with little chance of passing to demonstrate their commitment to an issue, which is more common for bills on issues of high salience. On the other hand, low-salience issues like the arts, recreation, water resources, and public lands have high coefficients, perhaps because with less public attention there were more opportunities for cooperation without as much scrutiny.

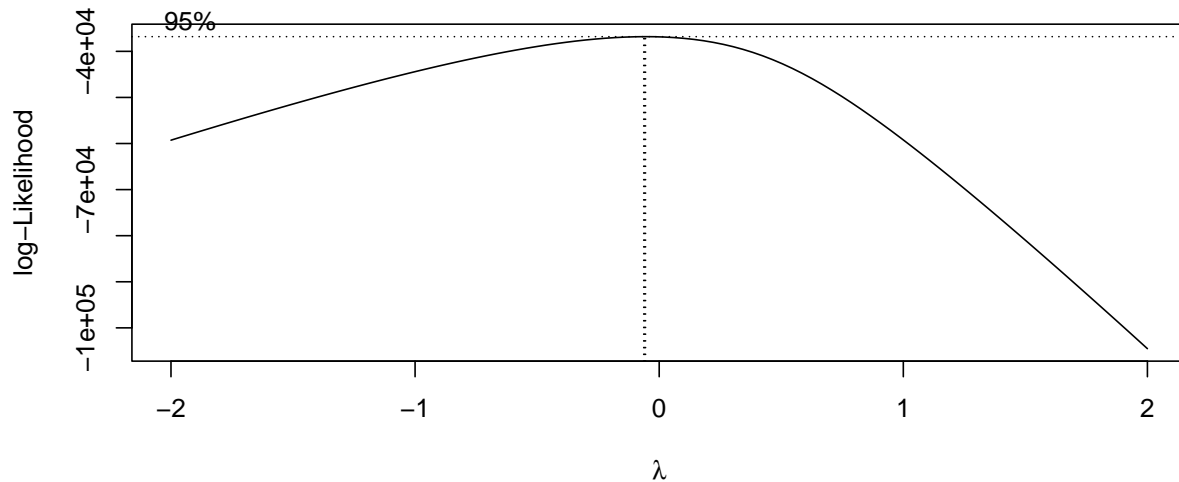
Section 6: Conclusions and Recommendations

The analysis in this paper is broad and covers many topics, but there are some key takeaways. First, bills that were lobbied more advanced further in Congress, though there was no mechanism to determine whether this was a causal relationship in either direction. Issues of high salience tended to be among the most lobbied

in the 111th Congress, but bills on these issues tended to not go as far in Congress. Conversely, bills on issues of low salience were not lobbied heavily but were more likely to advance in Congress. There were some subjects and issues that were consistently heavily lobbied and more likely to pass, such as appropriations bills and other bills related to federal spending. Bills related to regulation and telecommunications were also heavily lobbied. The complexity of laws, as measured by the number of legislative subjects, made bills more likely to be lobbied and more likely to advance through Congress. In addition to these broad findings, there are additional factors that affect how often a bill is lobbied and how likely it is to advance in Congress.

There is a significant amount of future work that could build on this paper. First, due to the rate limits on the Congress.gov API, it was not possible to generate more data than the 7905 bills in this data set. This analysis could be improved by including the other bills in the 111th Congress and assuming that those that did not appear in the OpenSecrets data were lobbied 0 times. Another way that this could be extended is to include more than the 111th Congress, which would transform this analysis from an investigation into a unique period of legislative history into something more generalizable. This would come at the expense of the ability to point to major legislative battles like the Affordable Care Act, Wall Street Regulation, and the Cap and Trade bill as explanations for phenomena, but this might be a worthy tradeoff for generalizability. There are also more aspects of lobbying to analyze, in addition to the raw number of appearances of a bill in LDA reports. Incorporating an element of money could help distinguish between different types of lobbyists such as businesses and activist groups. A more rigorous form of Natural Language Processing on the CRS summaries of bills could also yield additional insights about which types of bills are heavily lobbied, as would conducting the analysis of summaries performed in this paper for more subsets of the data. Lastly, there was no mechanism to see if additional lobbying had a causal relationship with bills moving further in Congress, so it would be interesting to generate a mechanism for causal inference in this case.

Appendix A: Box-Cox Transformation of Lobbying Data



This is a likelihood plot of the boxcox transformations for a model with every possible predictor. Given that $\lambda = 0$ is close to the maximum log-likelihood, it is most appropriate to use a log transformation in linear modelling.

Appendix B: Remaining tables of Coefficients in the bivariate models of lobbying

Log Linear Model of Lobbying
Bivariate Relationships for Final Bill State

Predictor	Estimate	SE	t.value	p.value
Intercept	2.362	0.023	104.084	<0.001***
Committee Consideration	0.142	0.036	3.913	<0.001***
Discharge Petition	0.431	0.365	1.182	0.24
Calendar Entry	1.175	0.082	14.327	<0.001***
Floor Consideration	0.775	0.070	11.078	<0.001***
Resolving Differences	3.151	0.486	6.479	<0.001***
Passed Congress	0.840	0.092	9.162	<0.001***

Log Linear Model of Lobbying
Bivariate Relationships for Committee

Predictor	Estimate	SE	t.value	p.value
Intercept	2.572	0.144	17.821	<0.001***
Agriculture, Nutrition, and Forestry Committee	0.066	0.213	0.308	0.76
Appropriations Committee	2.180	0.238	9.172	<0.001***
Armed Services Committee	-0.468	0.171	-2.740	0.006**
Banking, Housing, and Urban Affairs Committee	0.428	0.178	2.400	0.02*
Budget Committee	0.430	0.252	1.704	0.09
Commerce, Science, and Transportation Committee	0.543	0.174	3.126	0.002**
Committee on House Administration	-0.875	0.276	-3.170	0.002**
Education and the Workforce Committee	-0.118	0.161	-0.729	0.47
Energy and Commerce Committee	0.371	0.154	2.414	0.02*
Energy and Natural Resources Committee	-0.011	0.165	-0.064	0.95
Environment and Public Works Committee	0.319	0.183	1.747	0.08
Finance Committee	-0.197	0.151	-1.311	0.19
Financial Services Committee	-0.116	0.167	-0.698	0.49
Foreign Affairs Committee	-0.597	0.215	-2.777	0.006**
Foreign Relations Committee	-0.534	0.235	-2.272	0.02*
Health, Education, Labor, and Pensions Committee	0.234	0.161	1.453	0.15
Homeland Security and Governmental Affairs Committee	-0.268	0.186	-1.439	0.15
Homeland Security Committee	0.396	0.193	2.053	0.04*
Indian Affairs Committee	-0.316	0.288	-1.100	0.27
Intelligence Permanent Select Committee	-0.345	0.505	-0.683	0.49
Intelligence Select Committee	-0.225	0.664	-0.338	0.74
Judiciary Committee	-0.043	0.153	-0.278	0.78
Natural Resources Committee	-0.433	0.166	-2.607	0.009**
None	1.096	0.223	4.921	<0.001***
Oversight and Accountability Committee	-0.525	0.181	-2.899	0.004**
Rules and Administration Committee	-0.569	0.288	-1.980	0.05*
Rules Committee	-0.502	0.306	-1.644	0.1
Science, Space, and Technology Committee	0.103	0.234	0.439	0.66
Small Business and Entrepreneurship Committee	-0.355	0.254	-1.398	0.16
Small Business Committee	-0.652	0.260	-2.506	0.01*
Transportation and Infrastructure Committee	0.046	0.175	0.264	0.79
Veterans' Affairs Committee	-0.436	0.166	-2.626	0.009**
Ways and Means Committee	-0.134	0.153	-0.878	0.38

Log Linear Model of Lobbying
Bivariate Relationships for Policy Area

Predictor	Estimate	SE	t.value	p.value
Intercept	2.710	0.129	21.009	<0.001***
Animals	0.204	0.242	0.841	0.4
Armed Forces and National Security	-0.558	0.142	-3.935	<0.001***
Arts, Culture, Religion	-1.905	1.011	-1.883	0.06
Civil Rights and Liberties, Minority Issues	-0.569	0.293	-1.941	0.05
Commerce	0.005	0.158	0.030	0.98
Congress	-0.994	0.211	-4.708	<0.001***
Crime and Law Enforcement	-0.301	0.151	-1.992	0.05*
Economics and Public Finance	0.670	0.182	3.673	<0.001***
Education	-0.500	0.151	-3.314	<0.001***
Emergency Management	-0.151	0.179	-0.846	0.4
Energy	0.532	0.151	3.518	<0.001***
Environmental Protection	0.395	0.164	2.411	0.02*
Families	-0.760	0.298	-2.554	0.01*
Finance and Financial Sector	0.423	0.150	2.824	0.005**
Foreign Trade and International Finance	-0.916	0.140	-6.522	<0.001***
Government Operations and Politics	-0.577	0.148	-3.884	<0.001***
Health	0.194	0.136	1.421	0.16
Housing and Community Development	-0.357	0.174	-2.058	0.04*
Immigration	-0.206	0.172	-1.193	0.23
International Affairs	-0.592	0.166	-3.568	<0.001***
Labor and Employment	0.168	0.167	1.004	0.32
Law	0.100	0.187	0.533	0.59
Native Americans	-0.582	0.202	-2.878	0.004**
None	-1.712	0.721	-2.375	0.02*
Private Legislation	-1.387	0.237	-5.858	<0.001***
Public Lands and Natural Resources	-0.719	0.147	-4.903	<0.001***
Science, Technology, Communications	0.792	0.170	4.655	<0.001***
Social Sciences and History	-1.009	0.647	-1.558	0.12
Social Welfare	-0.354	0.183	-1.935	0.05
Sports and Recreation	-0.714	0.518	-1.379	0.17
Taxation	-0.026	0.137	-0.187	0.85
Transportation and Public Works	0.329	0.154	2.146	0.03*
Water Resources Development	-0.754	0.204	-3.690	<0.001***

Negative Binomial Model of Lobbying
Bivariate Relationships for Final Bill State

Predictor	Estimate	SE	z.value	p.value
Intercept	3.391	0.021	158.333	<0.001***
Committee Consideration	0.206	0.034	6.042	<0.001***
Discharge Petition	0.304	0.344	0.885	0.38
Calendar Entry	1.775	0.077	23.089	<0.001***
Floor Consideration	1.375	0.066	20.956	<0.001***
Resolving Differences	3.200	0.455	7.032	<0.001***
Passed Congress	2.310	0.086	26.887	<0.001***

Negative Binomial Model of Lobbying
Bivariate Relationships for Committee

Predictor	Estimate	SE	z.value	p.value
Intercept	3.788	0.141	26.809	<0.001***
Agriculture, Nutrition, and Forestry Committee	-0.232	0.209	-1.109	0.27
Appropriations Committee	2.671	0.232	11.519	<0.001***
Armed Services Committee	-0.171	0.167	-1.022	0.31
Banking, Housing, and Urban Affairs Committee	0.819	0.174	4.705	<0.001***
Budget Committee	1.904	0.246	7.728	<0.001***
Commerce, Science, and Transportation Committee	0.324	0.170	1.906	0.06
Committee on House Administration	-1.320	0.273	-4.833	<0.001***
Education and the Workforce Committee	-0.165	0.158	-1.043	0.3
Energy and Commerce Committee	0.251	0.150	1.673	0.09
Energy and Natural Resources Committee	0.142	0.162	0.877	0.38
Environment and Public Works Committee	0.612	0.179	3.425	<0.001***
Finance Committee	-0.134	0.147	-0.906	0.36
Financial Services Committee	-0.167	0.163	-1.022	0.31
Foreign Affairs Committee	-0.911	0.211	-4.313	<0.001***
Foreign Relations Committee	-0.516	0.231	-2.237	0.03*
Health, Education, Labor, and Pensions Committee	0.230	0.157	1.463	0.14
Homeland Security and Governmental Affairs Committee	-0.222	0.182	-1.220	0.22
Homeland Security Committee	0.788	0.189	4.177	<0.001***
Indian Affairs Committee	-0.749	0.283	-2.647	0.008**
Intelligence Permanent Select Committee	-1.293	0.501	-2.584	0.01**
Intelligence Select Committee	-1.238	0.659	-1.879	0.06
Judiciary Committee	0.156	0.150	1.040	0.3
Natural Resources Committee	-0.851	0.163	-5.221	<0.001***
None	0.883	0.218	4.055	<0.001***
Oversight and Accountability Committee	-0.911	0.178	-5.124	<0.001***
Rules and Administration Committee	-0.897	0.283	-3.168	0.002**
Rules Committee	-0.244	0.299	-0.815	0.42
Science, Space, and Technology Committee	-0.285	0.229	-1.241	0.21
Small Business and Entrepreneurship Committee	-0.629	0.250	-2.518	0.01*
Small Business Committee	-1.483	0.258	-5.751	<0.001***
Transportation and Infrastructure Committee	0.791	0.171	4.626	<0.001***
Veterans' Affairs Committee	-1.000	0.163	-6.138	<0.001***
Ways and Means Committee	0.190	0.150	1.271	0.2

Negative Binomial Model of Lobbying
Bivariate Relationships for Policy Area

Predictor	Estimate	SE	z.value	p.value
Intercept	4.508	0.125	36.009	<0.001***
Animals	-1.269	0.236	-5.375	<0.001***
Armed Forces and National Security	-0.866	0.138	-6.291	<0.001***
Arts, Culture, Religion	-3.409	1.061	-3.214	0.001**
Civil Rights and Liberties, Minority Issues	-1.331	0.287	-4.646	<0.001***
Commerce	-0.127	0.153	-0.830	0.41
Congress	-1.690	0.207	-8.178	<0.001***
Crime and Law Enforcement	-1.118	0.147	-7.620	<0.001***
Economics and Public Finance	1.435	0.177	8.115	<0.001***

Education	-1.437	0.147	-9.783	<0.001***
Emergency Management	-0.942	0.174	-5.422	<0.001***
Energy	-0.120	0.147	-0.815	0.41
Environmental Protection	0.441	0.159	2.773	0.006**
Families	-1.849	0.292	-6.322	<0.001***
Finance and Financial Sector	0.110	0.145	0.756	0.45
Foreign Trade and International Finance	-1.800	0.137	-13.178	<0.001***
Government Operations and Politics	-1.344	0.144	-9.307	<0.001***
Health	-0.283	0.132	-2.143	0.03*
Housing and Community Development	-1.044	0.169	-6.181	<0.001***
Immigration	-1.074	0.168	-6.405	<0.001***
International Affairs	-1.215	0.162	-7.515	<0.001***
Labor and Employment	-0.133	0.162	-0.820	0.41
Law	-0.493	0.182	-2.709	0.007**
Native Americans	-1.640	0.198	-8.296	<0.001***
None	-3.061	0.739	-4.143	<0.001***
Private Legislation	-2.632	0.236	-11.156	<0.001***
Public Lands and Natural Resources	-1.705	0.143	-11.933	<0.001***
Science, Technology, Communications	-0.181	0.165	-1.097	0.27
Social Sciences and History	-2.311	0.644	-3.587	<0.001***
Social Welfare	-1.480	0.178	-8.300	<0.001***
Sports and Recreation	-1.475	0.507	-2.908	0.004**
Taxation	-0.698	0.133	-5.247	<0.001***
Transportation and Public Works	-0.298	0.149	-2.001	0.05*
Water Resources Development	-1.767	0.200	-8.830	<0.001***

Appendix C: Table of Coefficients for Log Linear Model of Lobbying

Log Linear Model of Lobbying
Coefficients Table

Variable	Coefficient	SE	t.value	p.value
Intercept	26.460	1.054	25.094	<0.001***
Committee Consideration	-0.040	0.038	-1.054	0.29
Calendar Entry	0.536	0.074	7.209	<0.001***
Resolving Differences	1.226	0.411	2.983	0.003**
Actions	0.280	0.031	8.896	<0.001***
Policy Area: Armed Forces And National Security	-0.424	0.072	-5.869	<0.001***
Policy Area: Arts, Culture, Religion	-1.699	0.859	-1.978	0.05*
Policy Area: Civil Rights And Liberties, Minority Issues	-0.547	0.229	-2.393	0.02*
Policy Area: Commerce	0.478	0.106	4.499	<0.001***
Policy Area: Congress	-0.933	0.156	-5.996	<0.001***
Policy Area: Crime And Law Enforcement	-0.305	0.077	-3.936	<0.001***
Policy Area: Education	-0.627	0.080	-7.848	<0.001***
Policy Area: Energy	0.654	0.079	8.293	<0.001***
Policy Area: Environmental Protection	0.339	0.096	3.550	<0.001***
Policy Area: Families	-0.737	0.232	-3.173	0.002**
Policy Area: Finance And Financial Sector	0.720	0.081	8.929	<0.001***
Policy Area: Government Operations And Politics	-0.414	0.082	-5.067	<0.001***
Policy Area: Health	-0.001	0.059	-0.018	0.99
Policy Area: Immigration	-0.208	0.106	-1.967	0.05*
Policy Area: International Affairs	-0.742	0.118	-6.270	<0.001***

Policy Area: Native Americans	-0.663	0.171	-3.871	<0.001***
Policy Area: Public Lands And Natural Resources	-0.633	0.072	-8.786	<0.001***
Policy Area: Science, Technology, Communications	0.793	0.110	7.193	<0.001***
Policy Area: Social Sciences And History	-1.188	0.544	-2.182	0.03*
Policy Area: Social Welfare	-0.344	0.116	-2.968	0.003**
Policy Area: Sports And Recreation	-0.860	0.431	-1.996	0.05*
Policy Area: Taxation	0.123	0.053	2.338	0.02*
Policy Area: Transportation And Public Works	0.281	0.085	3.317	<0.001***
Policy Area: Water Resources Development	-0.682	0.141	-4.829	<0.001***
Subjects	0.614	0.028	21.769	<0.001***
Cosponsors	0.220	0.011	19.974	<0.001***
Introduced Date	-0.002	0.000	-24.772	<0.001***
Related Bills	0.322	0.028	11.310	<0.001***
Sponsorsparty R	-0.249	0.032	-7.877	<0.001***
Agriculture, Nutrition, And Forestry Committee	0.539	0.136	3.953	<0.001***
Appropriations Committee	1.093	0.169	6.477	<0.001***
Armed Services Committee	-0.160	0.093	-1.722	0.09
Commerce, Science, And Transportation Committee	0.160	0.095	1.688	0.09
Committee On House Administration	-0.431	0.209	-2.061	0.04*
Energy And Commerce Committee	0.023	0.060	0.375	0.71
Finance Committee	0.427	0.050	8.519	<0.001***
Financial Services Committee	-0.440	0.083	-5.277	<0.001***
Foreign Affairs Committee	-0.330	0.167	-1.982	0.05*
Health, Education, Labor, And Pensions Committee	0.293	0.074	3.960	<0.001***
Indian Affairs Committee	-0.260	0.262	-0.994	0.32
None	0.333	0.153	2.172	0.03*
Oversight And Acability Committee	-0.175	0.107	-1.647	0.1
Rules Committee	-0.613	0.229	-2.679	0.007**
Science, Space, And Technology Committee	-0.650	0.162	-4.023	<0.001***
Small Business And Entrepreneurship Committee	-0.865	0.201	-4.307	<0.001***
Small Business Committee	-1.051	0.204	-5.155	<0.001***
Committees	-0.109	0.022	-4.972	<0.001***

Appendix D: Table of Coefficients for Negative Binomial Model of Lobbying

Negative Binomial Model of Lobbying
Coefficients Table

Variable	Coefficient	SE	z.value	p.value
Intercept	28.831	0.990	29.128	<0.001***
Committee Consideration	-0.061	0.038	-1.631	0.1
Calendar Entry	0.521	0.068	7.635	<0.001***
Resolving Differences	1.384	0.372	3.721	<0.001***
Actions	0.325	0.029	11.236	<0.001***
Policy Area: Animals	-0.787	0.170	-4.618	<0.001***
Policy Area: Armed Forces And National Security	-0.522	0.068	-7.669	<0.001***
Policy Area: Arts, Culture, Religion	-2.128	0.876	-2.429	0.02*
Policy Area: Civil Rights And Liberties, Minority Issues	-0.988	0.219	-4.508	<0.001***
Policy Area: Commerce	0.817	0.099	8.240	<0.001***
Policy Area: Congress	-0.977	0.149	-6.566	<0.001***
Policy Area: Crime And Law Enforcement	-0.514	0.094	-5.459	<0.001***
Policy Area: Education	-0.785	0.081	-9.663	<0.001***

Policy Area: Emergency Management	0.087	0.112	0.780	0.44
Policy Area: Energy	0.651	0.078	8.316	<0.001***
Policy Area: Environmental Protection	0.737	0.098	7.495	<0.001***
Policy Area: Families	-0.959	0.220	-4.352	<0.001***
Policy Area: Finance And Financial Sector	0.795	0.097	8.225	<0.001***
Policy Area: Foreign Trade And International Finance	0.158	0.072	2.192	0.03*
Policy Area: Government Operations And Politics	-0.171	0.078	-2.197	0.03*
Policy Area: Housing And Community Development	-0.064	0.122	-0.524	0.6
Policy Area: Immigration	-0.474	0.111	-4.259	<0.001***
Policy Area: International Affairs	-0.579	0.110	-5.258	<0.001***
Policy Area: Labor And Employment	0.439	0.094	4.649	<0.001***
Policy Area: Law	0.231	0.129	1.793	0.07
Policy Area: Native Americans	-0.735	0.162	-4.546	<0.001***
Policy Area: None	-0.693	0.607	-1.141	0.25
Policy Area: Private Legislation	-0.437	0.183	-2.390	0.02*
Policy Area: Public Lands And Natural Resources	-0.656	0.075	-8.730	<0.001***
Policy Area: Science, Technology, Communications	0.678	0.100	6.779	<0.001***
Policy Area: Social Sciences And History	-1.083	0.512	-2.113	0.03*
Policy Area: Social Welfare	-0.425	0.109	-3.915	<0.001***
Policy Area: Sports And Recreation	-0.952	0.409	-2.331	0.02*
Policy Area: Taxation	0.210	0.054	3.871	<0.001***
Policy Area: Transportation And Public Works	0.349	0.092	3.807	<0.001***
Policy Area: Water Resources Development	-0.661	0.136	-4.849	<0.001***
Subjects	0.704	0.027	25.675	<0.001***
Cosponsors	0.221	0.010	21.568	<0.001***
Introduced Date	-0.002	0.000	-28.455	<0.001***
Related Bills	0.277	0.026	10.476	<0.001***
Sponsorsparty R	-0.243	0.029	-8.266	<0.001***
Agriculture, Nutrition, And Forestry Committee	0.526	0.129	4.091	<0.001***
Appropriations Committee	1.408	0.154	9.133	<0.001***
Armed Services Committee	-0.110	0.087	-1.255	0.21
Banking, Housing, And Urban Affairs Committee	0.216	0.107	2.012	0.04*
Budget Committee	0.634	0.163	3.899	<0.001***
Commerce, Science, And Transportation Committee	0.090	0.092	0.981	0.33
Committee On House Administration	-0.542	0.199	-2.728	0.006**
Education And The Workforce Committee	-0.079	0.077	-1.036	0.3
Energy And Natural Resources Committee	0.207	0.082	2.541	0.01*
Environment And Public Works Committee	0.166	0.106	1.565	0.12
Finance Committee	0.462	0.054	8.614	<0.001***
Financial Services Committee	-0.474	0.099	-4.775	<0.001***
Foreign Affairs Committee	-0.423	0.155	-2.720	0.007**
Health, Education, Labor, And Pensions Committee	0.342	0.071	4.791	<0.001***
Homeland Security Committee	0.123	0.112	1.103	0.27
Indian Affairs Committee	-0.344	0.245	-1.403	0.16
Judiciary Committee	0.288	0.072	4.008	<0.001***
None	0.373	0.141	2.644	0.008**
Oversight And Acability Committee	-0.352	0.101	-3.506	<0.001***
Rules Committee	-0.609	0.214	-2.842	0.004**
Science, Space, And Technology Committee	-0.816	0.149	-5.472	<0.001***
Small Business And Entrepreneurship Committee	-1.240	0.187	-6.623	<0.001***
Small Business Committee	-1.695	0.192	-8.846	<0.001***
Transportation And Infrastructure Committee	-0.185	0.095	-1.953	0.05
Committees	-0.067	0.020	-3.299	<0.001***

Appendix E: Bivariate Relationships in Ordinal Model

Ordinal Model of Bill Progress
Bivariate Relationships for Committee

Predictor	Estimate	SE	t.value	p.value
Agriculture, Nutrition, and Forestry Committee	-3.022	0.379	-7.983	<0.001***
Appropriations Committee	1.056	0.303	3.483	1
Armed Services Committee	-0.604	0.204	-2.956	0.002**
Banking, Housing, and Urban Affairs Committee	-1.796	0.240	-7.493	<0.001***
Budget Committee	-1.443	0.376	-3.842	<0.001***
Commerce, Science, and Transportation Committee	-1.034	0.225	-4.596	<0.001***
Committee on House Administration	-2.241	0.439	-5.102	<0.001***
Education and the Workforce Committee	-0.169	0.188	-0.900	0.18
Energy and Commerce Committee	-0.268	0.179	-1.499	0.07
Energy and Natural Resources Committee	-0.257	0.203	-1.271	0.1
Environment and Public Works Committee	-0.877	0.241	-3.643	<0.001***
Finance Committee	-5.040	0.273	-18.443	<0.001***
Financial Services Committee	-3.098	0.247	-12.562	<0.001***
Foreign Affairs Committee	-4.073	0.540	-7.543	<0.001***
Foreign Relations Committee	-1.036	0.331	-3.132	<0.001***
Health, Education, Labor, and Pensions Committee	-2.699	0.217	-12.432	<0.001***
Homeland Security and Governmental Affairs Committee	0.906	0.240	3.777	1
Homeland Security Committee	0.192	0.232	0.829	0.8
Indian Affairs Committee	0.380	0.364	1.043	0.85
Intelligence (Permanent Select) Committee	-13.378	0.000	-1865198.099	<0.001***
Intelligence (Select) Committee	-2.271	1.157	-1.962	0.02*
Judiciary Committee	-0.481	0.180	-2.665	0.004**
Natural Resources Committee	0.093	0.194	0.480	0.68
None	1.993	0.251	7.951	1
Oversight and Accountability Committee	-0.559	0.218	-2.564	0.005**
Rules and Administration Committee	-1.565	0.434	-3.606	<0.001***
Rules Committee	-4.473	1.031	-4.337	<0.001***
Science, Space, and Technology Committee	0.002	0.272	0.007	0.5
Small Business and Entrepreneurship Committee	-0.839	0.369	-2.274	0.01*
Small Business Committee	0.197	0.317	0.623	0.73
Transportation and Infrastructure Committee	0.200	0.204	0.979	0.84
Veterans' Affairs Committee	0.173	0.195	0.883	0.81
Ways and Means Committee	-3.517	0.211	-16.696	<0.001***
Introduced CommitteeConsideration	-1.137	0.168	-6.755	<0.001***
CommitteeConsideration PassedCommittee	1.154	0.168	6.859	1
PassedCommittee FloorConsideration	1.629	0.170	9.589	1
FloorConsideration PassedCongress	2.785	0.177	15.696	1

Ordinal Model of Bill Progress
Bivariate Relationships for Policy Area

Predictor	Estimate	SE	t.value	p.value
Animals	1.275	0.322	3.960	<0.001***
Armed Forces and National Security	1.031	0.190	5.415	<0.001***
Arts, Culture, Religion	3.085	1.463	2.109	0.02*

Civil Rights and Liberties, Minority Issues	0.946	0.390	2.424	0.008**
Commerce	0.686	0.215	3.191	<0.001***
Congress	0.141	0.305	0.463	0.32
Crime and Law Enforcement	0.845	0.202	4.173	<0.001***
Economics and Public Finance	0.361	0.263	1.374	0.08
Education	0.188	0.203	0.928	0.18
Emergency Management	1.124	0.238	4.729	<0.001***
Energy	0.464	0.202	2.291	0.01*
Environmental Protection	0.727	0.219	3.317	<0.001***
Families	-0.450	0.446	-1.010	0.84
Finance and Financial Sector	-0.866	0.216	-4.003	1
Foreign Trade and International Finance	-3.251	0.283	-11.503	1
Government Operations and Politics	1.101	0.202	5.441	<0.001***
Health	0.170	0.184	0.923	0.18
Housing and Community Development	-0.696	0.256	-2.714	1
Immigration	0.635	0.227	2.797	0.003**
International Affairs	-0.560	0.240	-2.336	0.99
Labor and Employment	0.389	0.222	1.751	0.04*
Law	0.636	0.254	2.503	0.006**
Native Americans	1.159	0.273	4.240	<0.001***
None	-13.159	0.000	-10510109.267	1
Private Legislation	-0.205	0.318	-0.644	0.74
Public Lands and Natural Resources	1.289	0.196	6.565	<0.001***
Science, Technology, Communications	1.202	0.229	5.238	<0.001***
Social Sciences and History	-0.102	0.925	-0.110	0.54
Social Welfare	-0.404	0.259	-1.561	0.94
Sports and Recreation	1.273	0.757	1.681	0.05*
Taxation	-2.432	0.217	-11.223	1
Transportation and Public Works	0.835	0.207	4.034	<0.001***
Water Resources Development	1.457	0.265	5.506	<0.001***
Introduced CommitteeConsideration	0.148	0.175	0.845	0.2
CommitteeConsideration PassedCommittee	2.171	0.178	12.228	<0.001***
PassedCommittee FloorConsideration	2.634	0.179	14.728	<0.001***
FloorConsideration PassedCongress	3.774	0.186	20.338	<0.001***

Appendix F: Table of Coefficients for Ordinal Model of Legislative State

Ordinal Model of Bill Progress
Coefficients Table

Variable	Coefficient	SE	t.value	p.value
Lobbying	0.154	0.021	7.340	<0.001***
Policy Area: Agriculture And Food	0.337	0.231	1.461	0.15
Policy Area: Animals	0.064	0.023	2.752	0.007**
Policy Area: Armed Forces And National Security	0.876	0.216	4.046	<0.001***
Policy Area: Arts Culture Religion	2.860	1.525	1.875	0.06
Policy Area: Civil Rights And Liberties Minority Issues	0.525	0.370	1.420	0.16
Policy Area: Congress	-0.172	0.302	-0.570	0.56
Policy Area: Crime And Law Enforcement	0.348	0.161	2.165	0.03*
Policy Area: Economics And Public Finance	-0.656	0.258	-2.541	0.01*
Policy Area: Education	0.265	0.176	1.508	0.13
Policy Area: Emergency Management	0.424	0.198	2.144	0.03*

Policy Area: Energy	-0.467	0.153	-3.056	0.002**
Policy Area: Finance And Financial Sector	-0.400	0.189	-2.117	0.03*
Policy Area: Foreign Trade And International Finance	-0.672	0.301	-2.233	0.02*
Policy Area: Government Operations And Politics	0.547	0.157	3.476	<0.001***
Policy Area: Health	-0.316	0.125	-2.525	0.01*
Policy Area: Immigration	-0.218	0.191	-1.142	0.25
Policy Area: International Affairs	-1.145	0.299	-3.832	<0.001***
Policy Area: Labor And Employment	0.175	0.194	0.905	0.37
Policy Area: Law	0.028	0.027	1.043	0.3
Policy Area: Native Americans	0.238	0.316	0.755	0.45
Policy Area: None	-12.455	0.000	-67016207.811	0.05
Policy Area: Public Lands And Natural Resources	0.875	0.163	5.381	<0.001***
Policy Area: Science Technology Communications	0.544	0.179	3.039	0.002**
Policy Area: Social Welfare	-0.120	0.257	-0.468	0.64
Policy Area: Sports And Recreation	1.504	0.759	1.982	0.06
Policy Area: Taxation	-1.461	0.195	-7.495	<0.001***
Policy Area: Water Resources Development	0.958	0.239	4.008	<0.001***
Subjects Count	0.281	0.054	5.166	<0.001***
Cosponsors Count	0.003	0.001	2.377	0.02*
Introduced Date	-0.150	0.027	-5.487	<0.001***
Related Bills Count	0.310	0.053	5.805	<0.001***
Sponsors Party R	-0.374	0.061	-6.130	<0.001***
Agriculture Nutrition And Forestry Committee	-2.323	0.375	-6.195	<0.001***
Appropriations Committee	1.616	0.317	5.100	<0.001***
Armed Services Committee	-0.524	0.241	-2.172	0.03*
Banking Housing And Urban Affairs Committee	-0.743	0.233	-3.191	0.001**
Budget Committee	-0.428	0.382	-1.120	0.26
Commerce Science And Transportation Committee	-0.380	0.201	-1.891	0.06
Committee On House Administration	-1.749	0.462	-3.785	<0.001***
Education And The Workforce Committee	0.212	0.177	1.197	0.23
Energy And Commerce Committee	0.556	0.156	3.554	<0.001***
Energy And Natural Resources Committee	0.398	0.195	2.037	0.04*
Environment And Public Works Committee	-0.219	0.219	-1.001	0.32
Finance Committee	-3.232	0.278	-11.627	<0.001***
Financial Services Committee	-2.039	0.237	-8.592	<0.001***
Foreign Affairs Committee	-2.588	0.582	-4.450	<0.001***
Foreign Relations Committee	1.006	0.410	2.452	0.02*
Health Education Labor And Pensions Committee	-1.885	0.206	-9.146	<0.001***
Homeland Security And Governmental Affairs Committee	1.308	0.210	6.228	<0.001***
Homeland Security Committee	0.200	0.215	0.929	0.35
Indian Affairs Committee	1.041	0.457	2.280	0.02*
Intelligence Permanent Select Committee	-14.140	0.000	-8106157.018	<0.001***
Intelligence Select Committee	-2.101	1.184	-1.775	0.05*
Judiciary Committee	0.199	0.160	1.241	0.21
Natural Resources Committee	0.622	0.192	3.235	0.001**
None	3.712	0.247	15.049	<0.001***
Rules And Administration Committee	-0.798	0.435	-1.836	0.06
Rules Committee	-4.019	1.040	-3.864	<0.001***
Science Space And Technology Committee	0.522	0.268	1.950	0.05
Small Business Committee	1.383	0.313	4.420	<0.001***
Transportation And Infrastructure Committee	0.964	0.180	5.353	<0.001***
Veterans Affairs Committee	0.463	0.248	1.863	0.06
Ways And Means Committee	-1.734	0.210	-8.274	<0.001***
Committees Count	0.615	0.037	16.819	<0.001***

Cospensors Count	0.085	0.028	3.063	0.002**
Introduced Committee Consideration	1.611	0.182	8.836	<0.001***
Committee Consideration Passed Committee	4.216	0.191	22.105	<0.001***
Passed Committee Floor Consideration	4.746	0.193	24.569	<0.001***
Floor Consideration Passed Congress	6.004	0.202	29.699	<0.001***

Appendix G: References

The American Legion. (n.d.). Part 4: Sponsors and co-sponsors. The American Legion. <https://www.legion.org/legislative/thomas/8974/part-4-sponsors-and-co-sponsors>

Ballotpedia. (n.d.). 111th United States Congress. Ballotpedia. https://ballotpedia.org/111th_United_States_Congress

Broder, J. M. (2010, March 25). “Cap and trade” loses its standing as energy policy of Choice. The New York Times. <https://www.nytimes.com/2010/03/26/science/earth/26climate.html>

Citizens Against Government Waste. (2023, April 28). 2023 congressional pig book. Citizens Against Government Waste. <https://www.cagw.org/reporting/pig-book>

Cohen, R. E. (2010, October 1). The legacy of the 111th congress. Politico. <https://www.politico.com/story/2010/10/the-legacy-of-the-111th-congress-042983>

Ferraro, Thomas (2010, December 23). Moderate Republican seeks Biden’s old U.S. Senate seat. Reuters. <https://www.reuters.com/article/us-usa-politics-castle-idUKTRE5BM0LT20091223/>

Fisher, M. (2010, February 24). Why Obama dropped the public option. The Atlantic. <https://www.theatlantic.com/politics/archive/2010/02/why-obama-dropped-the-public-option/346546/>

How a bill becomes a law. Congresswoman Eleanor Holmes Norton. (n.d.). <https://norton.house.gov/how-a-bill-becomes-a-law>

Library of Congress. (n.d.). API Documentation Bill Endpoints. GitHub. <https://github.com/LibraryOfCongress/api.congress.gov/blob/main/Documentation/BillEndpoint.md>

Library of Congress. (n.d.-a). Find bills by subject and Policy Area . Congress.gov. <https://www.congress.gov/help/find-bills-by-subject>

Library of Congress. (n.d.-b). John D. Rockefeller IV. Congress.gov. <https://www.congress.gov/member/jay-rockefeller/1424>

Library of Congress. (n.d.-c). John F. Kerry | Congress.gov | Library of Congress. Congress.gov. <https://www.congress.gov/member/john-kerry/K000148>

Library of Congress. (n.d.-d). Patrick J. Leahy | congress.gov | library of Congress. Congress.gov. <https://www.congress.gov/member/patrick-leahy/L000174>

Library of Congress. (n.d.-e). Public Laws of the 111th Congress. Congress.gov. <https://www.congress.gov/public-laws/111th-congress>

Library of Congress. (n.d.-f). Senator Arlen Specter. Congress.gov. <https://www.congress.gov/member/arlen-specter/S000709>

Library of Congress. (n.d.-g). Senator Chuck Grassley. Congress.gov. <https://www.congress.gov/member/charles-grassley/G000386?r=63>

Library of Congress. (n.d.-h). Senator Jeff Bingaman. Congress.gov. <https://www.congress.gov/member/jesse-bingaman/>

Library of Congress. (n.d.-i). Senator Richard J. Durbin. Congress.gov. <https://www.congress.gov/member/richard-durbin/D000563>

Library of Congress. (n.d.-j). Senator Saxby Chambliss. Congress.gov. <https://www.congress.gov/member/saxby-chambliss/C000286>

Nahhas, R. W. (n.d.). 6.22 Ordinal Logistic Regression. In *Introduction to Regression Methods for Public Health Using R*. essay. Retrieved December 17, 2023, from <https://www.bookdown.org/rwnahhas/RMPH/blr-ordinal.html#blr-olr-adjusted>.

National Conference of State Legislatures. (2021, September 3). Summary how states define lobbying and lobbyist. National Conference of State Legislatures. <https://www.ncsl.org/ethics/how-states-define-lobbying-and-lobbyist>

OpenSecrets. (n.d.-a). Lobbying Industries. OpenSecrets. <https://www.opensecrets.org/federal-lobbying/industries?cycle=a>

Our vision and mission: Inform, empower & advocate. OpenSecrets. (n.d.-b). <https://www.opensecrets.org/about>

Public Citizen. (n.d.). Lobbying Disclosure Act: A Brief Synopsis of Key Components. Public Citizen. <https://www.citizen.org/wp-content/uploads/brief-synopsis-of-lda.pdf>

Ray, M. (2023, March 27). Barney Frank. Encyclopædia Britannica. <https://www.britannica.com/biography/Barney-Frank>

Roback, P., & Legler, J. M. (2021). 4.9 Overdispersion. In *Beyond multiple linear regression: Applied generalized linear models and multilevel models in R*. essay, CRC Press. Retrieved December 17, 2023, from <https://bookdown.org/robback/bookdown-BeyondMLR/>.

Saturno, J. V., & Lynch, M. S. (2023, May 17). The appropriations process: A brief overview. Congressional Research Service Reports. <https://crsreports.congress.gov/product/pdf/R/R47106>

Small Business Administration (SBA). USAGov. (n.d.). <https://www.usa.gov/agencies/small-business-administration>

Wallach, P. A. (2015, April 13). America's lobbying addiction. Brookings. <https://www.brookings.edu/articles/americas-lobbying-addiction/>

What we do. Federal Communications Commission. (n.d.). <https://www.fcc.gov/about-fcc/what-we-do>

Zeballos-Roig, J. (2020, June 1). Congress' response to the 2008 financial crisis sparked a decade of inequality and resentment. the US is now repeating that mistake. Business Insider. <https://www.businessinsider.com/congress-repeating-critical-mistake-2008-financial-crisis-inequality-2020-5>