Draft version December 15, 2018
Typeset using LaTeX default style in AASTeX62

The Fiscal Responsibility Index: A Relative Measure of Deficit Contributions for Members of Congress

Ben Christensen

Submitted to Brigham Young University - ACME Senior Project

ABSTRACT

Congress has maintained an almost continuous budget deficit since the 1970s, but Congress doesnt legislate in unanimity. In this paper I investigate which Senators and Representatives add most and least to the budget deficit. The Fiscal Responsibility Index is a score that allows constituents to compare how often their representatives vote to increase or decrease the budget deficit. The costs and revenues associated with each bill is taken from the Congressional Budget Offices published cost estimates. These estimates are published on separate web pages for each bill and the reports are displayed in idiosyncratic ways, so I scraped the data using web-crawling and multiple regular expressions to find the estimates embedded within text. These methods are subject to error, especially for the last four sessions of Congress.

1. INTRODUCTION

The United States debt is approximately $21.9 trillion, and the deficit is about $840 billion. To put that into perspective, the United States Gross Domestic Product is $19.4 trillion. In almost every year since 1970 the federal government has spent more than it has raised in revenues. The size of the debt is a significant problem of national interest, but I will not discuss it in depth here.

One of the inherent difficulties in fiscal responsibility is that beneficiaries of government spending reward members of Congress in popularity and campaign contributions while victims of increased taxation punish members of Congress. These incentives may be more pressing than the more disconnected goal of a balanced budget that requires working-often across the aisle-with other members of Congress to achieve.

Although many Democrat politicians speak of higher taxes and Republican politicians speak of lower spending, there seem to be many low-taxing Democrats and high-spending Republicans. When either party takes control of Congress and the White House the deficit would disappear if they were more committed to their budget deficit-ideologies.

The purpose of the Fiscal Responsibility Index (FRI) is to create an incentive for politicians to reduce the budget deficit. For every bill signed into law, each Yea-voting-representative loses a point for every dollar the bill is estimated to cost and gains a point for every dollar the bill is estimated to raise in revenue. Symbolically,

$$FRI_i = \sum_{j=1}^{n} (Revenue_j - Cost_j) * g(i,j)$$

where i indexes members of Congress, j indexes bills signed into law, and

$$g(i,j) = \begin{cases} 1 & \text{if Congressman } i \text{ voted for bill } j \\ 0 & \text{otherwise} \end{cases}$$

The FRI should not be treated as a deficit. It does not include the total yearly revenue the federal government receives, but rather changes in the revenue the government will receive. Similarly with costs, it does not include the total federal government spending. It only includes discretionary spending or changes in mandatory spending such as Social Security or Medicaid.

This leads to another point of clarification: members with very negative FRI are not necessarily responsible for the federal debt or even the federal deficit. About two thirds of the federal budget is for mandatory spending, much of which was committed to before the Congressional Budget Office began publishing cost estimates. The FRI is a relative

score to evaluate how members of Congress have been contributing to or taking away from the national debt since 1997.

Only 1.5% of members of Congress have a nonnegative FRI. Republicans and Democrats have almost identical average scores. Senators tend to have lower scores than House Representatives. Members of Congress who have served longer tend to have lower scores. In this paper I describe and visualize these results and discuss their limitations. Collecting the cost estimates for passed bills proved extremely difficult. This may be why a project like this hasnt been done before. In the following section I explain the web-scraping process. Then I visualize the resulting data and perform regression analysis to determine associations of FRI scores with representative characteristics. After discussing limitations, I suggest ways to improve the reliability of the FRI.

2. WEB-SCRAPING

The Congressional Budget Office (CBO) only has cost estimates publicly available from 1997 to present. This is the limiting factor for data collection because Congressional voting records are available since 1990 for both the House and the Senate. This gives 11 sessions of congress, 105th - 115th, where 115th is the current session of Congress. Each session spans two years. All of the data scraped is publicly available and the program follows outlined sleep times to not overload websites with requests. The web-scraping processes are outlined in the following subsection. Images of typical webpages scraped for data are included for the reader's convenience. Figure 1 corresponds to subsections 2.1 and 2.2, Figure 2 corresponds to subsections 2.3 and 2.4, and Table 1 and Figure 3 correspond to subsection 2.5.

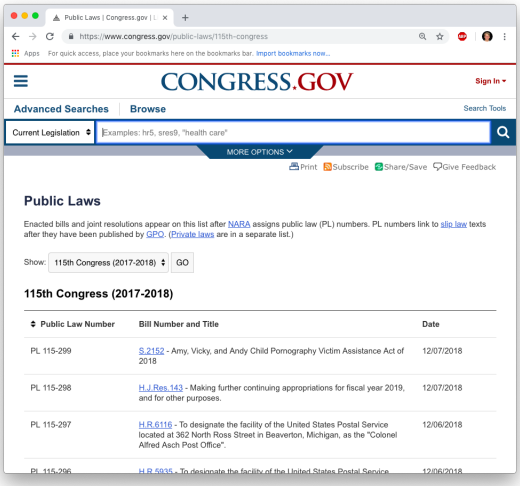
2.1. *Bills Signed into Law*

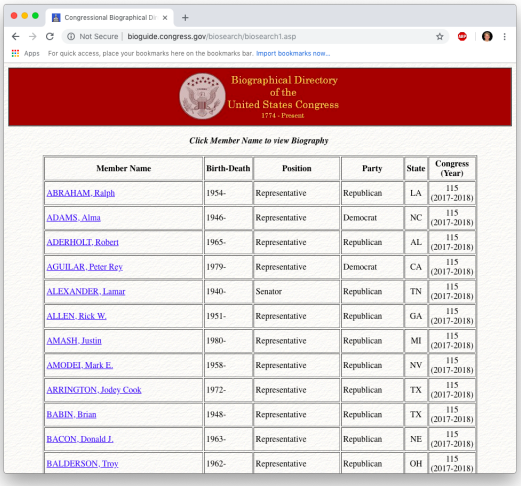
The simplest web-scraping task was to find all the bills signed into law since 1990. Each of these bills is listed by session of congress on Congress.gov. The format for a bills name is H.R.X if the bill is started in the House and S.X if the bill is started in the Senate where X is an integer greater than 1. A bill name is only unique within a session of Congress so I keep the session with the bill name (e.g. H.R.4322-113th).

2.2. *Members of Congress*

Finding the names of every member of Congress was also a simple task. Using the lookup tool on Congress.gov my code searches for all representatives by session number of Congress and collects their full name, party, state, year of birth and creates a python dict object, adding the session number to another attribute titled sessions which is a list object. I use this list to create a variable for each member called tenure which is the number of sessions served in Congress.

Figure 1. Typical Webpages for 2.1 and 2.2





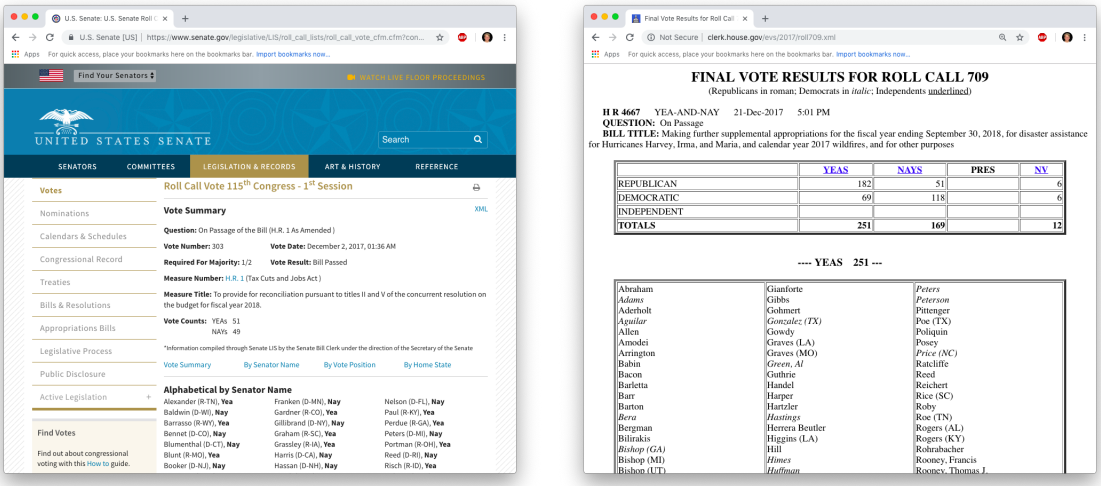
2.3. *Senator Voting Records*

Collecting the voting records for each Senator that served since 1990 was a moderately difficult task. Using BeautifulSoup in Python I parsed through the html to find the voting record for each Senator. Fortunately, the formatting was consistent across roll call vote pages and Congress sessions. I created an attribute for each Senator in my Python dict called Yeas which is a list containing every bill for which that Senator voted Yea. Since only the Senators last name is listed in the voting record I use the listed party, state information, and session number to match the voting record to the correct Senator.

2.4. *House Representative Voting Records*

Collecting voting records for House Representatives was more difficult than for Senators because the formatting was not consistent. If no one voted Nay the html code for the webpage would be formatted differently. I created several if statements to account for these differences and added the voting record information to Representatives the same way I did for Senators. For this website, some Representatives are listed with their full name and others with only their last name. Still others have last name and a state. I use this available information and session number to match the voting record to the correct Representative.

Figure 2. Typical webpages for 2.3 and 2.4



2.5. *Cost Estimates*

The CBO does not provide estimates for every bill, but for the bills it does provide estimates for, it publishes a PDF document containing the estimates. Collecting these estimates was particularly troublesome. Unfortunately, the CBO does not publish a dataset containing all of these estimates in one place so to get them requires web-scraping and web-crawling.

If the estimates were published consistently this could be as simple as finding the names of bills signed into law, but the publications are inconsistent in where and how the information is recorded.

For some bills, a summary gives the estimates in text on the same webpage on which the PDF is published. This summary is easily accessible without even opening the PDF report. The actual estimates are nested in the text of the summary. The $ symbol is unique to the estimates within the text, but whether the estimate is a cost or revenue depends on the context of the estimate. Using regular expressions, I find if a cost key-phrase comes before the estimate. If a revenue key-phrase does not come between the cost key-phrase and the estimate itself, the program codes the estimate as a cost. The revenue case is symmetric to that process.

Table 1. Key Phrases used in 2.5

Cost Key Phrase
costs provides
additional|increase|resul* AND spending|outlay
discretion* AND spending
decrease|reduc AND revenue
revenue AND lower|losses

Revenue Key Phrase
additional|increase|resul* AND sav*|revenue|collection|assessments
reduc*|decrease AND cost|spend|outlay
cost|spend|outlay AND decrease|lower
offsetting|rais* AND collect|receipts

Where * represents any character, | represents the mathematical OR, and AND means both words must be found in the string.

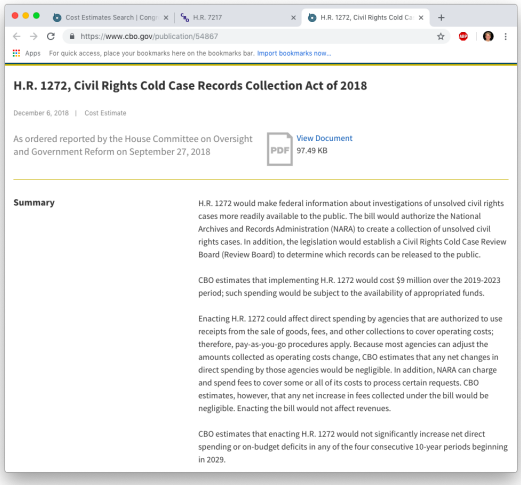
Sometimes an estimate is annual. Using the date range given, I multiply the annual estimate by the number of years. If no date range is given for the annual estimate, I multiply by the number of years that have passed since the bill was signed into law. Sometimes an estimate is provided for a certain year but then a total estimate is given for a range of years that includes the year the first estimate was provided for. Using regular expressions that capture the years I check for these cases and adjust the collected cost or revenue accordingly. Sometimes a document refers to an estimate multiple times that my code is unable to recognize. This seems to be the most common mistake my code makes.

For other bills, no summary is given on the webpage. The PDF document must be downloaded to obtain the estimates. Using a module in Python called textract, I convert the PDF into a Python string object so I can use the regular expressions I created and used as above. In these documents, it is most common for the summary to be headed with the title SUMMARY in all-caps and if another subsection follows it is also headed with an all-caps title. Using this nature of all-caps I attempt to limit the string Im searching through to only the summary itself because other subsections often repeat the estimates and if there is an estimate contained in the report, it seems to always be included in the SUMMARY section whenever there is one.

For some of these bills with no easy-grab summary, the PDF document is not in paragraph format but in table format. Although tables are usually easier to collect data from in html code, they are more difficult to find when the table must first be converted by the textract module from a PDF to a string. It is hard to predict how the estimates will appear in the resulting string. The process is even more difficult because the tables vary widely in format and in how they reference costs and revenues. These factors prevented me from finding a way to collect estimates from tables. So, every CBO estimate that has no summary in html code and only lists the estimates in table format within the pdf is omitted from the FRI. The number of CBO reports for which I find no estimate is included in the data visualization section. Omitted estimates will be included in this number, but this number will also include reports for which the CBO estimates a cost of zero and a revenue of zero.

Visually testing my code by random samples of several dozen cost estimates, it seems my code correctly finds cost and revenue estimates 66% of the time. The two major improvements to my code would be fixing the double counting that sometimes occurs when a total estimate is listed multiple times and finding a way to collect estimates from PDF tables.

Figure 3. Typical Webpages for 2.5



H.R. 1272, Civil Rights Cold Case Records Collection Act of 2018

December 6, 2018 | Cost Estimate

As ordered reported by the House Committee on Oversight and Government Reform on September 27, 2018

View Document
97.49 KB

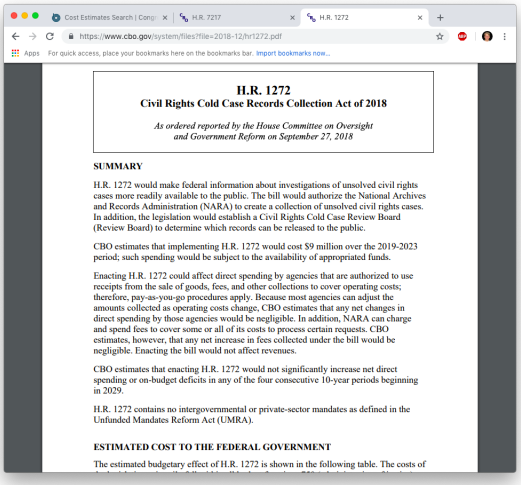
Summary

H.R. 1272 would make federal information about investigations of unsolved civil rights cases more readily available to the public. The bill would authorize the National Archives and Records Administration (NARA) to create a collection of unsolved civil rights cases. In addition, the legislation would establish a Civil Rights Cold Case Review Board (Review Board) to determine which records can be released to the public.

CBO estimates that implementing H.R. 1272 would cost $9 million over the 2019-2023 period; such spending would be subject to the availability of appropriated funds.

Enacting H.R. 1272 could affect direct spending by agencies that are authorized to use receipts from the sale of goods, fees, and other collections to cover operating costs; therefore, pay-as-you-go procedures apply. Because most agencies can adjust the amounts collected as operating costs change, CBO estimates that any net changes in direct spending by those agencies would be negligible. In addition, NARA can charge and spend fees to cover some or all of its costs to process certain requests. CBO estimates, however, that any net increase in fees collected under the bill would be negligible. Enacting the bill would not affect revenues.

CBO estimates that enacting H.R. 1272 would not significantly increase net direct spending or on-budget deficits in any of the four consecutive 10-year periods beginning in 2029.



H.R. 1272
Civil Rights Cold Case Records Collection Act of 2018

As ordered reported by the House Committee on Oversight and Government Reform on September 27, 2018

SUMMARY

H.R. 1272 would make federal information about investigations of unsolved civil rights cases more readily available to the public. The bill would authorize the National Archives and Records Administration (NARA) to create a collection of unsolved civil rights cases. In addition, the legislation would establish a Civil Rights Cold Case Review Board (Review Board) to determine which records can be released to the public.

CBO estimates that implementing H.R. 1272 would cost $9 million over the 2019-2023 period; such spending would be subject to the availability of appropriated funds.

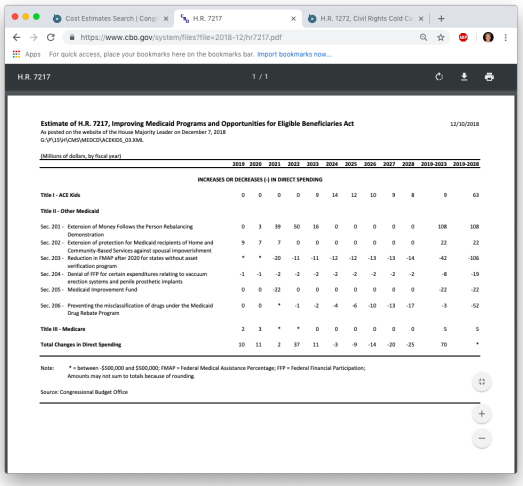
Enacting H.R. 1272 could affect direct spending by agencies that are authorized to use receipts from the sale of goods, fees, and other collections to cover operating costs; therefore, pay-as-you-go procedures apply. Because most agencies can adjust the amounts collected as operating costs change, CBO estimates that any net changes in direct spending by those agencies would be negligible. In addition, NARA can charge and spend fees to cover some or all of its costs to process certain requests. CBO estimates, however, that any net increase in fees collected under the bill would be negligible. Enacting the bill would not affect revenues.

CBO estimates that enacting H.R. 1272 would not significantly increase net direct spending or on-budget deficits in any of the four consecutive 10-year periods beginning in 2029.

H.R. 1272 contains no intergovernmental or private-sector mandates as defined in the Unfunded Mandates Reform Act (UMRA).

ESTIMATED COST TO THE FEDERAL GOVERNMENT

The estimated budgetary effect of H.R. 1272 is shown in the following table. The costs of



H.R. 7217 1 / 1

Estimate of H.R. 7217, Improving Medicaid Programs and Opportunities for Eligible Beneficiaries Act 12/10/2018

As posted on the website of the House Majority Leader on December 7, 2018
G:\PUBS\H\CMS\MEDCD\ACEKIDS_03.XML

(Millions of dollars, by fiscal year)

	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2019-2023	2019-2028
INCREASES OR DECREASES (-) IN DIRECT SPENDING												
Title I - ACE Kids	0	0	0	0	9	14	12	10	9	8	9	63
Title II - Other Medicaid												
Sec. 201 - Extension of Money Follows the Person Rebalancing Demonstration	0	3	39	50	16	0	0	0	0	0	108	108
Sec. 202 - Extension of protection for Medicaid recipients of Home and Community-Based Services against spousal impoverishment	9	7	7	0	0	0	0	0	0	0	22	22
Sec. 203 - Reduction in FMAP after 2020 for states without asset verification program	*	*	-20	-11	-11	-12	-12	-13	-13	-14	-42	-106
Sec. 204 - Denial of FFP for certain expenditures relating to vaccum erection systems and penile prosthetic implants	-1	-1	-2	-2	-2	-2	-2	-2	-2	-2	-8	-19
Sec. 205 - Medicaid Improvement Fund	0	0	-22	0	0	0	0	0	0	0	-22	-22
Sec. 206 - Preventing the misclassification of drugs under the Medicaid Drug Rebate Program	0	0	*	-1	-2	-4	-6	-10	-13	-17	-3	-52
Title III - Medicare	2	3	*	*	0	0	0	0	0	0	5	5
Total Changes in Direct Spending	10	11	2	37	11	-3	-9	-14	-20	-25	70	*

Note: * = between -$500,000 and $500,000; FMAP = Federal Medical Assistance Percentage; FFP = Federal Financial Participation;
Amounts may not sum to totals because of rounding.

Source: Congressional Budget Office

3. DATA VISUALIZATION

Note that the FRI scores are very large. Most of the following plots are on the scale of trillions of dollars. The summary statistics for FRI are listed in Table 2. First I visualize the way each cost estimates were scraped for each 2-year session of Congress to find potential errors in the code or gaps in the data, then I visualize the distributions of the FRI by relevant groups.

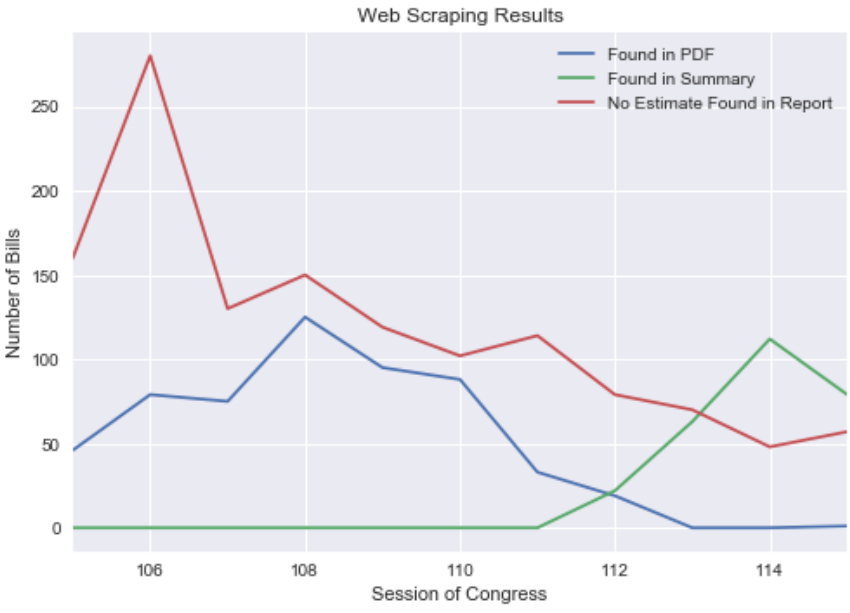
Table 2. FRI Summary Statistics

count	1305
mean	-497,145,500,000
std	627,785,200,000
min	-2,813,974,000,000
25%	-909,289,900,000
50%	-228,470,000,000
75%	-11,000,000,000
max	3,700,000,000

3.1. *Data Description*

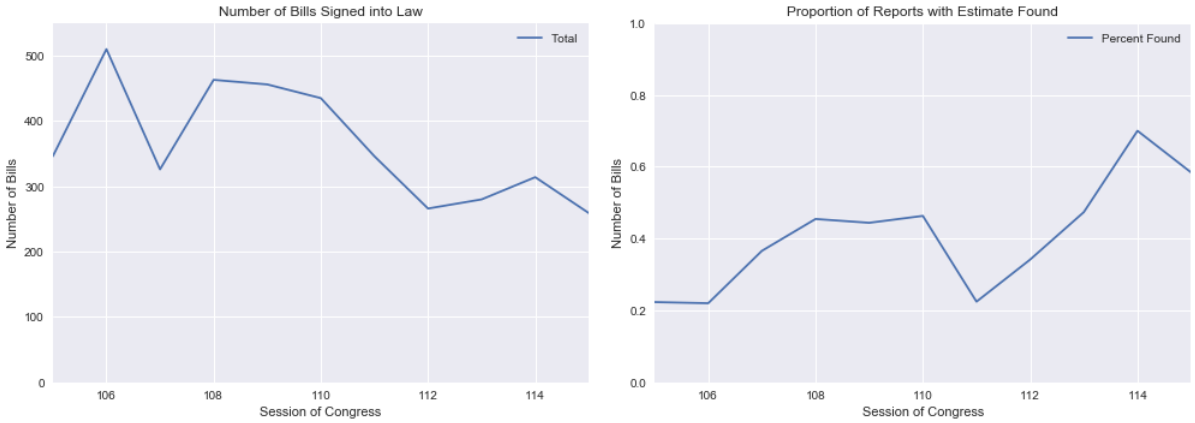
Figure 4 shows the web-scraping results for all the bills signed into law over the last 20 years. Found in PDF means that the PDF had to be downloaded to find the cost estimate, whereas Found in Summary means the cost-estimate could be found directly on the web page. No Estimate Found in Report means that the CBO published a report but my code did not find a cost estimate. It is not uncommon for the CBO to publish reports with no cost estimate (because the cost and revenues for the bill are estimated to be zero). This subsection also includes bills my code fails to find the cost estimate for such as when the cost estimate is listed in a table.

Figure 4.



As seen in Figure 5, except for the 111th session, my web-scraping code finds a cost estimate for a larger proportion of CBO reports for later years. This could be because the CBO produced fewer reports with zero-estimates or because my code was more successful on later reports.

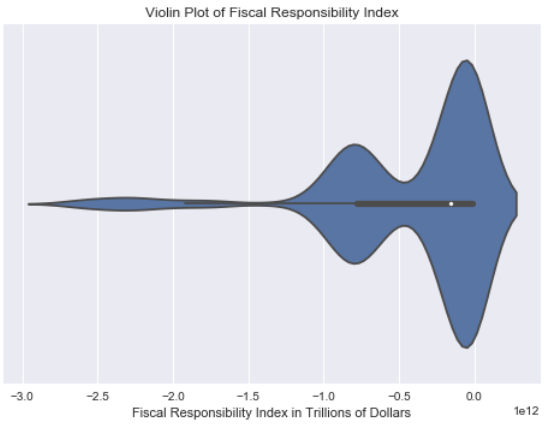
Figure 5.



3.2. *Distributions and Group Averages*

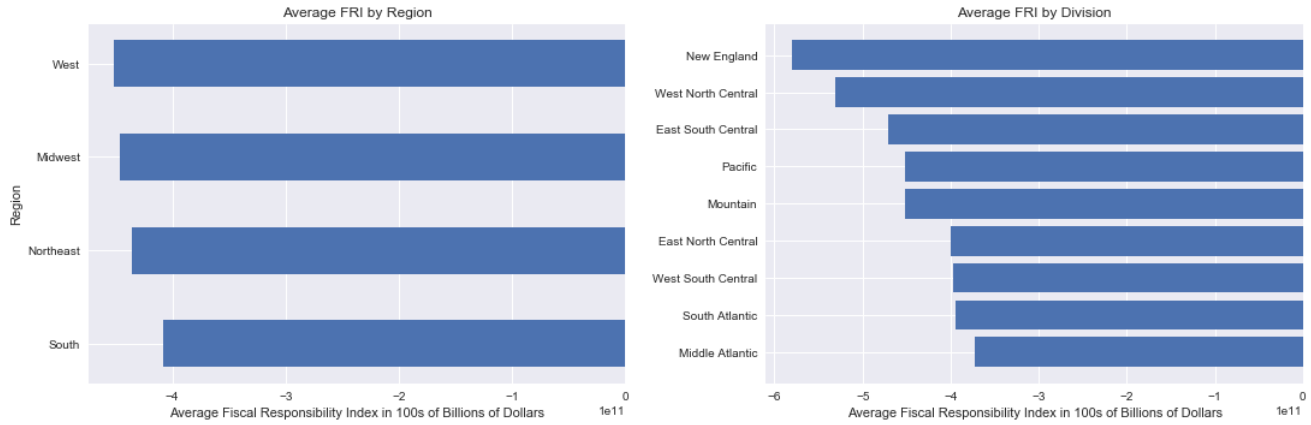
Figure 6 is a violin plot that shows the distribution of the FRI. The bimodality of the distribution is curious. The different average spending between Representatives and Senators could be a partial cause, but those groups also see bimodal distributions. Note also that the data is quite skewed to very low FRI scores.

Figure 6.



Using US Census Regions and Divisions as groups I compare average FRI scores. The spread is quite small for regions but is much large by division. For this reason, in the analysis I choose to use divisions rather than regions as features. I would not use both because divisions are partitions of the regions.

Figure 7.

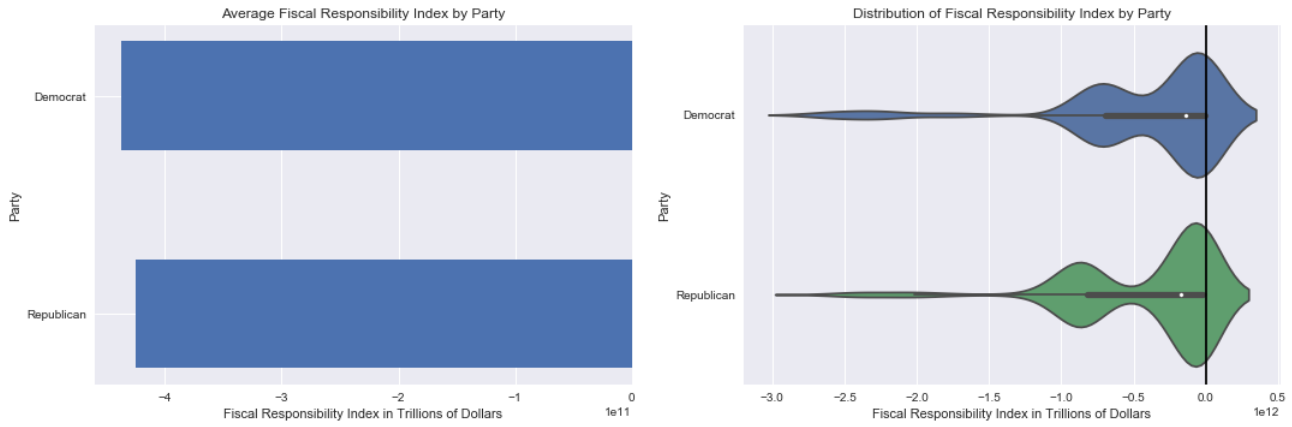


The Average FRI by Party and Position are given in Table 3. There are only three Independents in the data, so I exclude them from the plots in figure 8. Democrats seem to have slightly lower FRI scores than Republicans on average, but the difference is small enough to be considered trivial. Even the distributions are almost identical.

Table 3. Average FRI by Party and Position

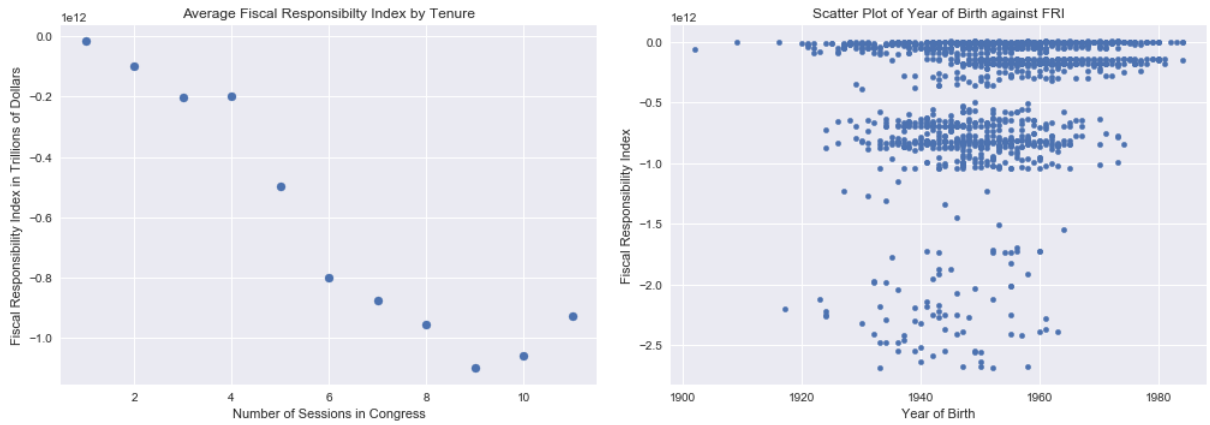
	Position	
Party	Representative	Senator
Democrat	-371,617,000,000	-1,086,663,000,000
Independent	-859,746,400,000	-853,134,000,000
Republican	-416,850,900,000	-850,020,000,000

Figure 8.



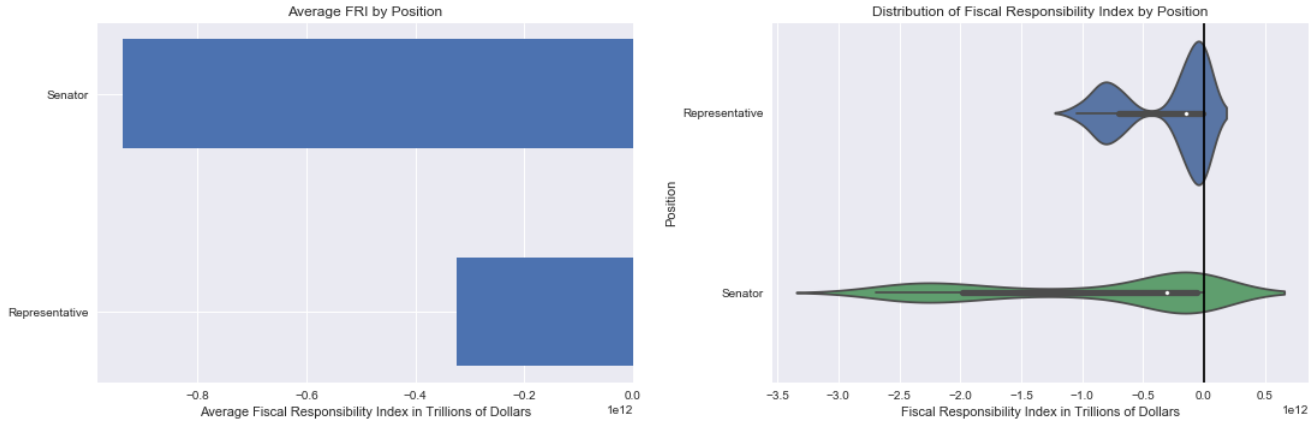
The negative correlation is strong for FRI and number of sessions in Congress, (which I will often call tenure for simplicity). Since the correlation seems to reverse for large values of tenure, I will include a tenure squared variable in the regressions in the next section. The correlation between FRI and year of birth is weakly positive. Year of birth may be correlated with tenure the associations will be clearer in the regression results.

Figure 9.



Senators far outspend Representatives on average as is seen in the following two figures. The FRI for Senators are very skewed towards lower scores. Yet both distributions seem to be bimodal. It is not clear what factor causes the bimodality.

Figure 10.



4. DATA ANALYSIS

For data analysis I perform linear and logistic regressions. I include a full-feature linear regression in Table 4 and a linear regression with features that optimizing AIC and BIC in Table 5. The logistic regression is in Table 6. The optimal features are Senator, Republican, Tenure, Tenure_sq, and Year of Birth where Senator, Republican and each division feature are binary. Tenure takes on integer values from 1-11 and Year of Birth is an integer from 1902 to 1984.

In the full feature regression, the coefficients for Senator, Republican, Tenure, Tenure_sq, and Year of Birth are statistically significant. Its not surprising then that these are the same features that optimize AIC. Each are significant at the 98% confidence level.

Table 4. Full Feature Linear Regression Results

Dep. Variable:	Score	**R-squared:**	0.536
Model:	OLS	**Adj. R-squared:**	0.531
Method:	Least Squares	**F-statistic:**	114.5
Date:	Thu, 13 Dec 2018	**Prob (F-statistic):**	4.54e-204
Time:	00:18:14	**Log-Likelihood:**	-36716.
No. Observations:	1302	**AIC:**	7.346e+04
Df Residuals:	1288	**BIC:**	7.353e+04
Df Model:	13		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	-1.571e+13	1.98e+12	-7.941	0.000	-1.96e+13	-1.18e+13
Senator	-4.85e+11	3.24e+10	-14.977	0.000	-5.49e+11	-4.21e+11
Republican	-7.471e+10	2.5e+10	-2.989	0.003	-1.24e+11	-2.57e+10
Division_East South Central	3.091e+10	5.74e+10	0.539	0.590	-8.17e+10	1.43e+11
Division_Middle Atlantic	8.02e+08	4.49e+10	0.018	0.986	-8.73e+10	8.89e+10
Division_Mountain	2.363e+10	4.97e+10	0.475	0.635	-7.39e+10	1.21e+11
Division_New England	5.321e+10	6.02e+10	0.883	0.377	-6.49e+10	1.71e+11
Division_Pacific	5.628e+10	4.54e+10	1.239	0.216	-3.28e+10	1.45e+11
Division_South Atlantic	5.329e+10	4.09e+10	1.302	0.193	-2.7e+10	1.34e+11
Division_West North Central	-6.996e+09	5.14e+10	-0.136	0.892	-1.08e+11	9.39e+10
Division_West South Central	4.845e+10	4.8e+10	1.009	0.313	-4.58e+10	1.43e+11
Tenure	-1.873e+11	1.57e+10	-11.940	0.000	-2.18e+11	-1.57e+11
Tenure_sq	5.741e+09	1.35e+09	4.238	0.000	3.08e+09	8.4e+09
Year of Birth	8.197e+09	1.01e+09	8.110	0.000	6.21e+09	1.02e+10

Because of multicollinearity I omit Year of Birth in the limited regression, but the rest of the features work without error. Being a Senator is associated with a -523 billion lower FRI. To put that in perspective, the average FRI is -497 billion. Being a Republican is associated with a -63 billion lower FRI. Each additional session of Congress served is associated with about a -150 billion lower FRI after accounting for the squared tenure coefficient which is not large enough to outweigh the linear tenure coefficient for even the maximum value of tenure. In the logistic regression I use the division features and the features from the limited linear regression. The dependent variable is 1 if FRI is nonnegative and 0 if FRI is negative. In this model, Senator, Republican, and Tenure increase the likelihood of a negative score with a confidence level of 99.9%. Tenure_sq and the Pacific division decrease the likelihood of a negative score with the same confidence level. Middle Atlantic and New England divisions also decrease the likelihood of a negative score, but with a confidence level of 90%.

Being able to control for multiple variables is useful in separating effects. For example, in the data visualization section, New England had the lowest average FRI, but in the logistic regression, it is associated with a lower likelihood for a negative score. This may because states often have similar party representations within divisions.

Table 5. Limited Feature Linear Regression Results

Dep. Variable:	FRI	**R-squared:**	0.511
Model:	OLS	**Adj. R-squared:**	0.509
Method:	Least Squares	**F-statistic:**	338.8
Date:	Thu, 13 Dec 2018	**Prob (F-statistic):**	1.16e-199
Time:	00:22:00	**Log-Likelihood:**	-36750.
No. Observations:	1302	**AIC:**	7.351e+04
Df Residuals:	1297	**BIC:**	7.354e+04
Df Model:	4		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	3.537e+11	3.92e+10	9.018	0.000	2.77e+11	4.31e+11
Senator	-5.233e+11	3.2e+10	-16.330	0.000	-5.86e+11	-4.6e+11
Republican	-6.252e+10	2.47e+10	-2.527	0.012	-1.11e+11	-1.4e+10
Tenure	-1.938e+11	1.6e+10	-12.144	0.000	-2.25e+11	-1.62e+11
Tenure_sq	5.59e+09	1.38e+09	4.053	0.000	2.88e+09	8.3e+09

Table 6. Logit Results

Dep. Variable:	Positive_FRI	**No. Observations:**	1305
Model:	Logit	**Df Residuals:**	1291
Method:	MLE	**Df Model:**	13
Date:	Wed, 12 Dec 2018	**Pseudo R-squ.:**	0.1508
Time:	23:51:03	**Log-Likelihood:**	-568.40
converged:	True	**LL-Null:**	-669.32

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	-86.6242	12.799	-6.768	0.000	-111.710	-61.538
Senator	-0.8340	0.249	-3.350	0.001	-1.322	-0.346
Republican	-0.6691	0.157	-4.260	0.000	-0.977	-0.361
Tenure	-0.5441	0.099	-5.503	0.000	-0.738	-0.350
Tenure_sq	0.0325	0.009	3.634	0.000	0.015	0.050
Year of Birth	0.0444	0.007	6.788	0.000	0.032	0.057
Division_East South Central	0.1202	0.395	0.304	0.761	-0.655	0.895
Division_Middle Atlantic	0.4763	0.278	1.713	0.087	-0.069	1.021
Division_Mountain	0.3708	0.314	1.180	0.238	-0.245	0.987
Division_New England	0.5991	0.370	1.621	0.105	-0.125	1.324
Division_Pacific	1.0175	0.277	3.678	0.000	0.475	1.560
Division_South Atlantic	0.3739	0.264	1.416	0.157	-0.144	0.892
Division_West North Central	-0.0591	0.368	-0.161	0.872	-0.781	0.663
Division_West South Central	0.3803	0.307	1.239	0.215	-0.221	0.982

The tenure variable is inaccurate because of the sample size. Many members of Congress in the dataset served in sessions before the 105th so tenure is biased down for them. In future years the tenure variable will become more accurate.

5. CONCLUSION

The largest limitation in these results is the inaccuracy in collecting the CBO cost estimates of every bill signed into law during the 105-115th sessions of Congress. As I discussed in section 2, the most helpful improvements on the web-scraping code will be to correct double counting of cost and revenue estimates and to find a way to scrape estimates from tables within PDFs. Approximately 33% of the cost and revenue estimates are inaccurate. This limitation may be related to the bimodality of the distributions of the FRI. It also casts considerable doubt on the point estimates generated in the linear regressions. But considering the lack of available data on CBO cost and revenue estimates, 66% accuracy is considerable progress towards making the FRI functional as an accountability measure. Some of the relationships are strong enough to plausibly persist with a more accurate FRI. The number of sessions a member of Congress serves seems to significantly reduce his or her FRI. And Senators appear to far outspend Representatives. This is true even when controlling for tenure although it may not be as strong with a more complete tenure variable.