Congressional Research Service Service Reports Codebook

Policy Agendas Project
2019

An Overview of the Data Set

This dataset includes reports and revisions to reports from the Congressional Research Service (CRS) between 1997 and 2017. CRS provides expert information to members of Congress across a broad range of policy issues. Members of Congress may privately request a report from CRS on a given policy topic.

Until recently, these reports were not available to the public, and not subject to Freedom of Information Act Requests. Beginning in 2017, the Congress began requiring that CRS reports be released on Congress.gov. However, this new dataset provided insufficient metadata to collect these data. Instead, we collected our data from EveryCRSReports.com, which maintains their own comprehensive database of CRS reports, along with metadata.

The Problem of Revisions

Each row in the dataset represents either a new report, identified by the unique identifier in the 'Source' column, or a revision to an existing report. For some reports, EveryCRSReport.com was able to identify the amount of text that was changed. These amounts can vary from just a few words to a majority of the report. We do not include the EveryCRSReport.com estimates of revision amounts in these data.

Figures 1 shows the frequency of new CRS reports. Figure 2 shows the frequency of both new and revised CRS reports in the dataset. Figure 3 shows the percentage of reports coded in each policy area by year for both. New reports are correlated with report revisions, but with some exceptions. Most notably, foreign affairs revisions are independent of new reports. Thus, we urge researchers to consider whether or not including revisions is appropriate for their individual projects.
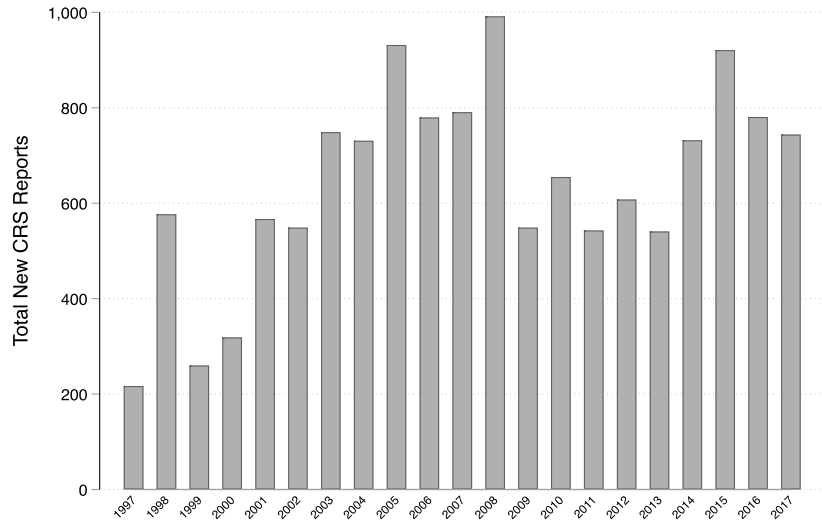
Figure 1: Frequency of New CRS Reports, 1997-2017

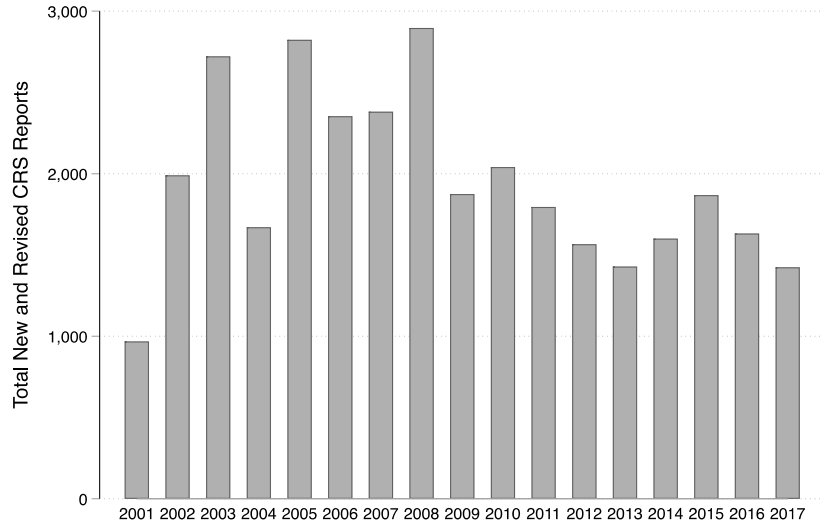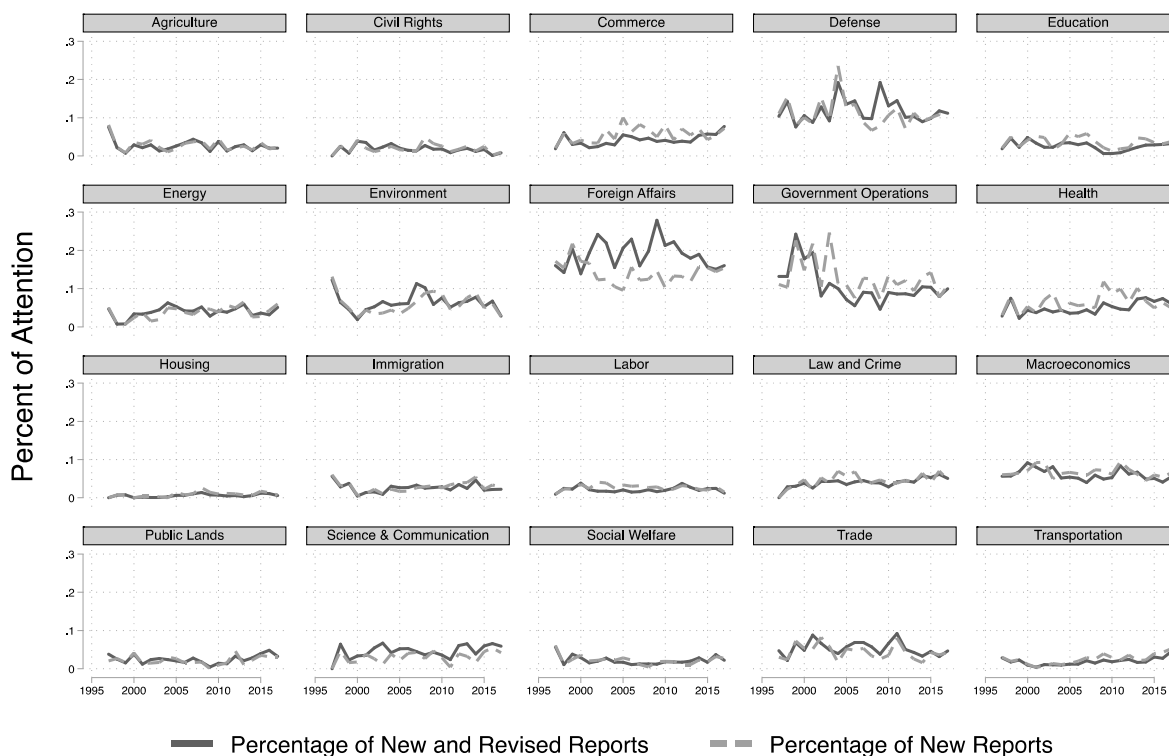Figure 2: Frequency of New and Revised CRS Reports, 1997-2017

Figure 3: Percentage of Attention by CRS Reports to Issue Areas, 1997-2017

Percentage of New and Revised Reports   Percentage of New Reports

Content Coding

We have coded each entry by policy content, along with other variables of interest. We employed the standard Policy Agendas Project topic scheme in our policy content coding, and each entry was assigned one and only one content code. This allows researchers to trace activity on a particular topic across the period. The full description of our content categories may be found in the major topics codebook at http://www.comparativeagendas.net/us.

Two trained coders assigned each observation to a single subtopic and major topic code to each observation based off their reading of the title and summary. If they could not make a decision based on these columns, they looked up the full report using the url column. When they disagreed, a third coder broke the tie.

The spreadsheet was sorted by the "crs_category" column for coders in order to facilitate faster coding, but was not used to assign a topic code to the observation.

All revisions to a report were assign to the same policy topics as the original unique report.

Variable Names and Descriptions

id

This column records the unique identifier assigned by the U.S. Policy Agendas Project for each observation. It has no substantive application.

source

The source number is a unique identifier used in by the Congressional research service. Each new unique report is assigned an identifier. Revisions of that report continue to use the same identifier.

version_no

This column records the revision number of the report. A new report is coded as version_no=1, the first revision to the new report is coded as version_no=2, etc. If researchers choose to only analyze new reports, they should drop all reports where version_no>1.

first_version

This column reports the date of the first version of this report. If the report was not revised, first_version should be identical to last_version and version_date.

last_version

This column reports the date of the final revision to this unique report. If the report was not revised, last_version should be identical to first_version. If this revision is the final revision to the report, last_version should be identical to version_date.

version_date

This columns reports the date of the new report or revision in the row.

day

The day column records the day of the month in version_date.

month

The month column records the numerical month in in version_date.

congress

This column denotes the Congress during version_date.

year

The year column records the year in version_date.

url

This column records the URL of the full report on EveryCRSReports.com.

crs_category
This column records the category code assigned by CRS to the report. The spreadsheet was sorted by this column for coders in order to facilitate faster coding, but was not used to assign a topic code to the observation.

summary
This column contains any available summary text of the report.

description
This column contains the title of the report.

pap_majortopic
This column records the Policy Agendas Project's major topic code that corresponds to the information found in the Description column. The codes and examples for this column are found in the Major Topics Codebook on the Policy Agendas Project website.

pap_subtopic
This column denotes the Policy Agendas Project's subtopic code that corresponds to the information found in the Description column. The codes and examples for this column are found in the Major Topics Codebook on the Policy Agendas Project website.

majortopic
This column records the Comparative Agendas Project's major topic code that corresponds to the existing Policy Agendas Project topic code. Please visit the Comparative Agendas Project website (www.comparativeagendas.net) for more information.

subtopic
This column records the Comparative Agendas Project's subtopic topic code that corresponds to the existing Policy Agendas Project topic code. Please visit the Comparative Agendas Project website (www.comparativeagendas.net) for more information.

Filter_Procedural

This column is coded=1 if the report is dedicated solely to Congressional operations. Examples include reports on how to vote, how to hire staff, how to request information from the Library of Congress, or how Senate points of order work.

Filter_NewReport

This column is coded=1 if version_no=1.