

## The US Policy Agendas Project

*Rebecca Eissler and Bryan D. Jones*

### 21.1 US Policy Agendas

Frank Baumgartner and Bryan D. Jones started the US Policy Agendas Project (PAP) in 1993 as a way to systematically measure government attention within and across specific policy areas and time. As the founding project of the Comparative Agenda Project (CAP) network, Baumgartner and Jones created the initial coding scheme that has since evolved into the common CAP coding scheme. One of the major aims of the initial work was to allow the creation of reliable time-series data on the policy topics scheduled for public debate. As such, a key criterion was backward compatibility—the need to adjust past policy categories if and when new categories were added in the future. This is necessary to make policy categories comparable across time. The model here was the National Income and Products Accounts, which established such a system to assess changes in various components of national economies (see Jones, 2016).

The development of the PAP coding scheme began as a pilot project at Texas A&M University while Baumgartner and Jones were studying US congressional hearings. Hearings were selected because they met a key criterion: hearings form a record of institutional attention at an early stage of the policy process. This characteristic was vital to their goal of understanding agenda-setting dynamics. By systematically collecting and coding data related to the activities of the federal government since 1947, the PAP has led to political scientists being able to quantitatively understand dynamics that had previously only been explored via individual case studies (Eissler and Russell, 2016).

Over the past twenty years, the project has grown to examine a broader range of government activities. The project, currently located at the University of Texas at Austin's Department of Government, has expanded to include

upwards of a dozen additional datasets, such as the New York Times Index, Roll Call Votes, Executive Orders, and State of the Union Addresses, as well as hosting datasets created and maintained by scholars outside the project that utilize the coding scheme. The quantity and variety of both in-house and affiliated datasets represent the collaborative and evolutionary nature of the project, as many of the current datasets started because of student researchers' interest in using the project's structure to advance their personal research (Eissler and Russell, 2016).

## 21.2 US Political System

The American national government has long been characterized by the separation of powers between the three branches. The legislative branch is made up of two chambers, the House of Representatives and the Senate, to which members are elected from single member districts using plurality rules. This structure has led to two-party domination of the political system. The Democratic party is traditionally a center-left party, while the Republican party is traditionally a center-right party, however in the past twenty years, both parties have moved away from the center towards their right and left extremes (Theriault, 2008). The House of Representatives is made up of 435 members and the distribution of seats is determined by the population in the states; with reapportionment occurring following the census, which happens once a decade in years ending in zero. The Senate is made up of 100 members, with each state permanently maintaining two seats. A president, who is both head of state and head of government, occupies the executive branch. His selection and powers operate separately from the legislative branch. As such, there is often competition between the two main policymaking branches of the federal government. The third branch, the judiciary, is made up of a system of courts that culminate in the Supreme Court, an institution that can weigh in only on "cases or controversies" that appeal to the Court for a decision. The modern Supreme Court is made up of nine justices who sit in one panel. They are appointed by the president and confirmed by the Senate. They have complete control over which cases they chose to hear.

## 21.3 Datasets

The PAP maintains and updates thirteen datasets covering the full range of government activities. Our legislative datasets include congressional hearings, public laws, and roll call votes. Our executive branch datasets include executive

orders and State of the Union addresses. We code Supreme Court decisions to represent the activities of the judicial branch. We also code the policy content of the two main political parties via the party platforms datasets. The PAP tracks the budgetary activities of the federal government via the budget authority and budget outlays datasets. Our media datasets include the *CQ Almanac* and the New York Times Index. We track interest group activity via the Encyclopedia of Associations dataset. Finally, we track public opinion about policies using the MIP dataset (see Table 21.1).

All data is coded by hand, with each observation being assigned codes by two coders who work independently. Following coding, the two codes are reconciled to make a final determination. We strive for 90 percent agreement on major topic and 80 percent on minor topic following reconciliation. Additionally, all datasets contain the policy topic code according to the CAP coding scheme and the legacy PAP coding scheme. The legacy PAP code has been retained for replication and the continuation of research that is based solely in the US context. The differences between the two systems are minimal.

The primary differences between the CAP codes and the legacy PAP codes include: a separate subtopic within social welfare for food assistance, which gets incorporated into low-income assistance in the CAP codebook; arts and culture as a subtopic within education that gets moved to its own major topic

**Table 21.1.** PAP-maintained datasets

Dataset name	Type	No. of observations	Years	Unit of analysis
Congressional Hearings	Legislative	94,882	1947–2013	Description of hearing activity
Public Laws	Legislative	20,403	1948–2014	Title of each public law
Roll Call	Legislative	50,148	1947–2013	Description of vote
Executive Orders	Executive	4,190	1945–2015	Title of order
State of the Union	Executive	22,794	1946–2016	Quasi-sentence
Democratic Party Platform	Political parties	13,633	1948–2008	Quasi-sentence
Republican Party Platform	Political parties	19,836	1948–2016	Quasi-sentence
Supreme Court Cases	Judiciary	8,955	1945–2009	Summary of case
Budget Authority (Adjusted)	Budget	7,935	1947–2015	Dollars per OMB functions and sub-functions
<i>Congressional Quarterly Almanac</i>	Media	14,217	1948–2011	Articles in the publication
New York Times Index	Media	53,495	1946–2013	Random sample of index entries of newspaper content
Encyclopedia of Associations	Interest groups	972	1966–2001	Groups
Most important problem (MIP)	Public opinion	1,344	1947–2012	Answers to open-ended question

Source: Comparative Agendas Project—United States

in the CAP codebook; parental leave and child care is moves from labor in the PAP codebook to social welfare in the CAP codebook; a separate subtopic within health for other health benefits and procedures that gets incorporated into insurer providers in the CAP codebook; a separate subtopic for pollution and conservation in coastal and other navigable waterways within the environment that gets combined with drinking water in the CAP codebook; a separate subtopic for truck and automobile transportation and safety within transportation that gets incorporated with highways in the CAP codebook; a separate subtopic for police, fire, and weapons control within law and crime that gets moved into a single category for all law enforcement agencies in the CAP codebook; and a separate subtopic for military veterans within defense that gets combined with military personnel in the CAP codebook.

The congressional hearings dataset contains over ninety-four thousand observations, with each corresponding to a hearing held by Congress since 1947. Committees in both the House of Representatives and the Senate conduct hearings to accomplish a variety of tasks, which include investigating policy proposals and conducting oversight of bureaucratic agencies. Committees are organized according to task or policy area. These data were coded from the annual volumes published by the Congressional Information Service and, starting in the mid-2000s, from data available on the ProQuest Congressional database. In addition to basic identifying information, the dataset includes a number of variables containing additional information about the hearing purpose and the committee conducting the hearing.

The public laws dataset contains all public laws enacted since 1948. A public law is created by first being introduced as a bill into one of the two chambers of Congress, which then receives the approval of a majority in both chambers, before being signed into law by the president. The list of public laws was collected from the appendix of the annual editions of the *CQ Almanac*.

The roll call votes dataset records all the votes in the House and Senate in which the individual politicians' choices as either for or against the measure are recorded according to the CAP coding scheme since 1946. Many types of legislative action can be put to a roll call, such as votes for amendment to proposed legislation, as well as the vote for final passage of a bill. Although the data have come from multiple sources over time (more details in the data codebook), since the 107th congress, votes are obtained from Govtrack.us, an open-source, online database of congressional voting records.

The executive orders dataset contains the orders issued by a president to administrative agencies since 1945. These actions carry the force of law and do not require legislative approval. However, because presidents issue them unilaterally, they can be revoked unilaterally by a successor or over ruled by a public law. Observations are compiled from the Federal Register's Executive Orders "Disposition Table Index."

The State of the Union dataset provides information on the presidential State of the Union address since 1946. Although many of these are delivered orally in front of a joint session of Congress, occasionally these take the form of written documents. These speeches occur once a year in fulfillment of the constitutional requirement that the president “give to the Congress Information on the State of the Union, and recommend to their Consideration such Measures as he shall judge necessary and expedient.” This dataset codes the text of the speech at the quasi-sentence level, which means that observation is the text between periods, semi-colons, and other punctuation marks.

Our party platforms datasets were originally compiled by Christina Wolbrecht and are maintained by the PAP. The datasets contain the Democratic Party and Republican Party platforms, which are written every four years during the presidential nominating conventions. The platforms are obtained from The American Presidency Project, an online database of presidential materials.

The Supreme Court case dataset contains all cases granted on certiorari or on appeal and argued before the Court since 1945. This dataset is the only publicly available dataset that examines the Court from a policy perspective. The Supreme Court has complete control over which cases it hears. Most years, the Court receives upwards of seven thousand petitions and, on average, only hears eighty cases, a number that has been on the decline for many years (Liptak, 2009).

The budget authority dataset is based on the Budget of the United States Government since fiscal year 1947. Originally compiled by James True, the dataset measures budget authority and provides an inflation-adjusted view of the budgetary process. The data is organized according to the functions and sub-functions of the federal government, rather than by the CAP policy topics.

The budget outlays dataset, compiled by Bryan D. Jones, Frank R. Baumgartner, and John Lovett, provides two “synthetic” series of annual, long-term budget outlays. No single series of expenditures (outlays) exist for the US federal government dating back to the founding of the republic, but two separate data series, compiled by the Treasury Department and the Office of Budget and Management allow for the construction of two “synthetic” datasets covering from 1791 to present. Outlays are noted for domestic and defense expenditures. The dataset includes both outlays and inflation-adjusted values.

The *Congressional Quarterly (CQ) Almanac* dataset contains information from all articles in the main chapters of the *CQ Almanac*. The *CQ Almanac* is published by CQ Press, which specializes in political news. Every year, they publish the *Almanac*, which highlights the important legislative events of the year. Each article usually covers one legislative initiative.

The New York Times (NYT) Index dataset is a sample of articles from *The New York Times* since 1946. *The New York Times* is a daily publication that is considered representative of the broader media agenda. The sample is

generated by collecting information about the first entry on every odd numbered page of the NYT index volumes.

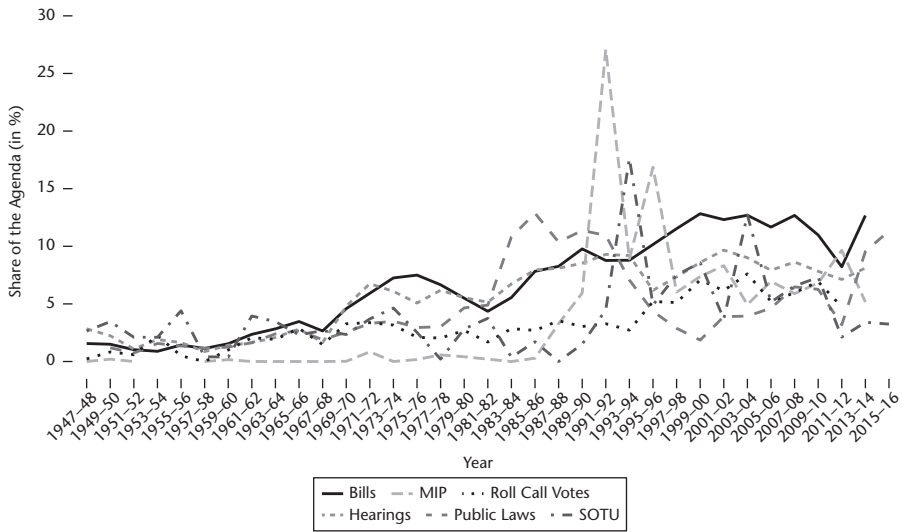
The *Encyclopedia of Associations* dataset tracks number of groups according to major topic policy areas since 1966 and is generated via the Gale Research, later Thomson/Gale, printed volume of the same name. We created a simple list of each group and coded the groups according to the PAP coding scheme. Complete data are available in five-year increments, as well as estimated annual counts for the whole time period.

The most important problem (MIP) dataset contains responses to Gallup's question "what do you think is the most important problem facing this country today?" Individual responses are aggregated at the annual level and coded according to major policy topic since 1947.

The PAP also acts as host to a number of datasets updated and maintained by other scholars, but that utilize the coding scheme. One such example is the Congressional Bills dataset, which is maintained by the Congressional Bills Project under the direction of E. Scott Adler and John Wilkerson, allows scholars to learn about the bills that are introduced into the US Congress. Another dataset the PAP hosts is a dataset assembled by Sam Kernell that examines presidential veto rhetoric offering insight into the threats that presidents make to veto legislation. On the public opinion side, the Project hosts the Policy Moods data, compiled by James A. Stimson and K. Elizabeth Coggins. This data allows users to study longitudinal measures of public opinion related to different policy issues. The PAP also hosts two media datasets. The New York Times Front Page dataset, by Amber Boydstun, examines articles that appear on the front page of *The New York Times* to measure the print media agenda, while the TV News Policy Agenda data, by Joe Uscinski, examines over sixty-five thousand TV news stories from the Vanderbilt archive. Finally, the PAP hosts the Tax Expenditure dataset, which was created and is maintained by Christopher Faricy. This data is based on the Congressional Joint Committee on Taxation's annual five-year estimates of federal tax expenditures. All of these datasets extend the efforts of the PAP to measure what government pays attention to over time.

## 21.4 A Look into the Data

Over the post-World-War-II period, health care policy has appeared repeatedly on the government agenda. The PAP datasets let us trace the attention to the issue over time and the many attempts to enact policy reforms. Figure 21.1 shows the percent of each dataset that is devoted to health. Prior to the late 1960s, healthcare was a small, but regular presence on the policy agenda across the legislative and executive branches, while barely registering on the



**Figure 21.1.** Attention to health across policy tools by Congress, 1947–2015  
 Source: Comparative Agendas Project—United States

mass public agenda. In the 1970s and 1980s, there was a steady increase in legislative activity, while the share of the president and public agenda stayed largely the same. Only in the 1990s and 2000s do we see healthcare become a significant portion of all agendas, particularly surrounding the Clinton healthcare reform attempt in the early 1990s and the successful passage of the Affordable Care Act during the Obama administration. This is merely one demonstration of how these data can be used understand attention to American policy over time and institutions.

## References

Eissler, R., and Russell, A. (2016). The Policy Agendas Project. In *American Governance*, ed. S. Schechter. New York: Macmillan, 50–1.

Jones, B. D. (2016). The Comparative Policy Agendas Projects as Measurement Systems. *Journal of Public Policy*, 36: 31–46.

Liptak, A. (2009). The Case of the Plummeting Supreme Court Docket, *New York Times*, Sept. 28, 2009.

Theriault, S. M. (2008). *Party Polarization in Congress*. New York: Cambridge University Press.