2

Studying Policy Dynamics

Frank R. Baumgartner, Bryan D. Jones, and John Wilkerson

All of the chapters in this book have in common the use of a series of datasets that comprise the Policy Agendas Project (also included, with full documentation, on the attached CD). The project had its genesis in previous work in which two of us used publicly available sources such as congressional hearings and the *Readers' Guide to Periodical Literature* to trace public and media attention to policy issues over the post-World War Two period. In that project, we studied particular issues such as the civilian use of nuclear power, pesticides, smoking and tobacco, and other topics, covering nine issues overall (see Baumgartner and Jones 1993). In the Policy Agendas Project, we decided to be much broader. With funding from the Political Science Division of the National Science Foundation, we began a much more ambitious project: to trace public attention to all issues, not just a few, and to cover the entire post-1947 period. The Project includes data on all congressional hearings, all laws, all stories in the *Congressional Quarterly Almanac*, a sample of stories in the *New York Times Index*, and the entire federal budget. We put together these datasets with the hope of encouraging the systematic study of policy change over long periods of time.

Most of the chapters that follow supplement our datasets with additional information, and we think this is the most fruitful way for most people to make use of the data. In this chapter we lay out some of the basics of the Policy Agendas Project datasets. Our goal here is three-fold. First, we explain the content, construction, and logic of the datasets. Second, we show why tracing policy activity across time depends on the availability of these datasets and others like them that ensure comparability across time. We highlight a critical warning signal to would-be policy analysts: Most archived datasets are not comparable across time, even if they appear to be. In contrast with many other datasets that appear to be useful for over-time comparisons, our datasets were specifically designed to ensure consistent coding over time. We spend considerable time in this chapter explaining the difficulties of this, as well as the rarity of it. Finally, we provide information that can help other potential users—the readers of this book—to use the datasets themselves.

The critical problem with the rapidly expanding set of computerized and searchable databases covering such things as congressional hearings, media coverage, presidential papers, and other sources of historical information about various public policy issues stems from the fact that designers of these datasets have information retrieval in mind, not the creation of a consistent public record. Keyword searches, the most common form of analysis that one can do, may not reveal consistent results when done over long periods of time because vocabularies change. Indexers change their practices over the decades; key-words are multiplied so that the user can find every mention of a given word, no matter how tangential the topic may have been to the main point of the article, story, presidential paper, or congressional hearing. In sum, many hidden problem arise in the use of available searchable electronic databases—users must beware of what appears to be a consistent series, but which may in fact harbor many hidden inconsistencies. In order to construct a consistent time-series over a period as long as those we explore here, one must pay considerable attention to the details of how data sets were created in the first place, and for what purposes they were originally collected. We show in this chapter how rare it is to find comparable data sources over long periods of time, and how we solved this problem in the construction of the Policy Agendas Project. In contrast to other sources of public policy information, our datasets were specifically designed to ensure historical continuity.

The Policy Agendas Project: Five Datasets on US Public Policy

Overview

Five datasets make up our project. Each is designed with a simple logic: It should be useful in and of itself to allow analysts to trace attention and government decisions over time, and it should provide enough information about the sources of the material so that anyone who wants to find out more detail about particular issues, decisions, or periods of attention can quickly find more complete information. Therefore, our logic is to gather a minimum of information about each congressional hearing, for example, but also to gather a complete set of identification materials so that the user can get more complete information from the same source materials as we used in the first place. Here we give a simple overview of each of the datasets.

The congressional hearings dataset consists of over 67,000 records corresponding to each hearing held in Congress since 1947. These data were coded from the annual volumes published by the Congressional Information Service and available in most government documents sections of major libraries. Variables coded include several identification variables (CIS identification numbers, date, committee(s) and subcommittee(s) involved), the topic codes, a short textual summary, as well as a series of variables which indicate whether the hearing: dealt with proposed legislation or was more of a fact-finding nature; considered an administration proposal or not; considered appropriations matters or not; mentioned the creation of a new agency or not; mentioned the creation of a new program or not; the number of days the hearing lasted; and the number of sessions in the hearing.

The *CQ Stories* dataset consists of a record for each article in the *Congressional Quarterly Almanac* from 1947 onward. In the cases where very long articles are broken into substantially distinct sub-sections (as is sometimes the case, for example, of discussions of huge omnibus bills, or of the President's budget proposals), we have a separate record for each of these sections. In all, there are over 12,000 records in this dataset, which includes identification materials so that a user can find the original story; committee(s) and subcommittee(s) involved; mentions of any committee reports, bill numbers, and Public Law numbers if applicable; information concerning how far through the legislative process the bill proceeded (e.g., whether it passed the House, passed the Senate, was vetoed, was signed into law, or was attached to some other bill as part of an omnibus legislation). In addition, we note the length of the story. This dataset is likely to be useful especially for those interested in a further check on the activities and level of interest in Congress to various issues, since the editors of CQ make efforts to cover the most important issues that Congress deals with, including those that are debated but not passed into law. Further, it is especially valuable as a sophisticated index to the *CQ Almanac* itself, since it can be used to identify all stories on a given topic or with other attributes.

Our dataset on Public Laws lists the PL number, the Congress, the sponsor of the legislation, the sponsor's party, the House and Senate report numbers (if any) concerning the bill, an indication of whether there was a Conference Report, whether the bill was commemorative or substantive, and whether or not the law was previously vetoed by the President. Of course, all these datasets also include a textual summary and a full 4-digit set of topic and subtopic codes.

The *New York Times Index* dataset consists of a sample of entries in that source, with a total of approximately 36,000 records dating back to 1947. Like the other datasets, this one includes identification material, a short textual summary, and a topic code. The topics are coded only by the major topic categories rather than the 226 subtopics used in the congressional databases, however, because so many stories in the media are on questions that do not correspond exactly to what congressional hearings focus on. In addition, there are entire topic areas, such as obituaries, art and book reviews, sports, and other events that are not included in the congressional databases, as we describe in more detail below. There are a variety of filter variables in the *New York Times* dataset as well designed to allow users to include or exclude local news items, international news items, stories that have anything to do with government and public policy as opposed to those that do not, and whether the story was on Page 1. For all stories that mention public policy, we also note whether the story mentions any of a number of institutions of government: the President, Congress, federal agencies, the courts, state and local governments, campaigns and elections, and interest groups.

Our dataset on the federal budget is based on annual figures reported by OMB in the annual budgets submitted by the President, but it is adjusted so that the spending categories are consistent over time. We report spending totals (budget authorizations) for 74 narrowly categories of spending and 17 major areas, as described in more detail below. Table 2-1 provides a summary of our data files:

Table 2-1. Summary of Policy Agendas Project Datasets

Dataset	Period Covered	Source	Unit of Analysis	Number of Cases	Number of Variables
Congressional Hearings	1946–1994	CIS Abstracts	Hearing	67,291	20
US Public Laws	1948–1994	*CQ Almanac* Appendix	Public Law	16,318	17
Congressional Quarterly	1948–1994	*CQ Almanac*	Story	12,583	37
US Budget Authority	1947–1997	*Budget of the United States*	OMB sub-function	115	1
New York Times	1947–1994	*NYT Index*	Story abstracts	36,403	20

The Topic and Subtopic Coding System

Perhaps the most important element that determines the usefulness of our datasets to a large and diverse audience is the extensive set of topic codes that we have devised. In contrast with most sources of longitudinal data, including the federal government's own reports of these materials, we have worked hard to guarantee temporal consistency for our topic categories. Typically, keyword searches, published indices, and other sources of publicly available data over time suffer from a tendency to revise or "improve" the categorization system over time. Our topic categories are consistent over time. This includes our version of the OMB budget authority; we spent over two years simply reading through the footnotes to the annual federal *Budget* noting how OMB had altered their spending classifications over time, and we adjusted the figures so that they are consistent. OMB itself does not have a consistently defined longitudinal time series of the federal budget that goes back as far as the one we have created here. In the case of the budget dataset, we use the OMB classification of 74 categories of spending (though we adjust it for changing definitions over time and for inflation). For each of the other datasets, we use the following major topics:

Table 2-2. Major Topic Categories

1. Macroeconomics
2. Civil Rights
3. Health
4. Agriculture
5. Labor, Immigration, and Employment
6. Education
7. Environment
8. Energy
10. Transportation
12. Law, Crime, and Family Issues
13. Social Welfare
14. Community Development and Housing
15. Banking, Finance and Domestic Commerce
16. Defense
17. Space, Science, Technology, and Communications
18. Foreign Trade
19. International Affairs
20. Federal Government Operations
21. Public Lands and Water Management

Additional Major Topics Used for *New York Times Index* Only

24. State and Local Government Administration
26. Weather and Natural Disasters
27. Fires
28. Arts and Entertainment
29. Sports and Recreation
30. Death Notices
31. Churches and Religion
99. Other, Miscellaneous, and Human Interest

New York Times stories are coded only by the major topics indicated (and, since there are many topics, such as book reviews, sports results, or home improvement ideas, on which Congress is rarely called to legislate, there are no corresponding topics in the congressional databases for a few categories that exist in the NYT database, as indicated above). Each of the congressional databases is broken down further by subtopic. For example, the major topic of Health is broken down into the following subtopics:

Table 2-3. Health Care Subtopics

300	General (includes combinations of multiple subtopics)
301	Health Care Reform, Health Care Costs, Insurance Costs and Availability
303	Medicare and Medicaid
306	Regulation of Prescription Drugs, Medical Devices, and Medical Procedures
307	Health Facilities Construction and Regulation, Public Health Service Issues
309	Mental Illness and Mental Retardation
310	Medical Fraud, Malpractice, and Physician Licensing Requirements
311	Elderly Health Issues
312	Infants, Children, and Immunization
313	Health Manpower Needs and Training Programs
315	Military Health Care
332	Alcohol Abuse and Treatment
333	Tobacco Abuse, Treatment, and Education
334	Illegal Drug Abuse, Treatment, and Education
349	Specific Diseases
398	Research and Development
399	Other

All in all, there are 27 major topic categories and 226 subtopics in our coding system, as listed in Appendix A and on the attached CD. Users can locate all hearings, CQ stories, and public laws on a given subtopic or broad topic area with ease. For most categories, one or more of the budget categories used by OMB may also correspond. The US Government classifies spending in many ways, but the most useful system for tracking spending by topic is the functional classification system developed by OMB and reported in the annual *Budget of the US Government*. Table 2-4 presents the major topic classifications used by OMB; the complete list of 74 detailed subtopics is presented in Appendix A and on the attached CD:

Table 2-4. Major Spending Classifications Used by OMB

Code	Title
050	National Defense
150	International Affairs
250	General Science, Space, and Technology
270	Energy
300	Natural Resources and Environment
350	Agriculture
370	Commerce and Housing Credit
400	Transportation
450	Community and Regional Development
500	Education, Training, Employment, and Social Services
550	Health
570	Medicare
600	Income Security
650	Social Security
700	Veterans Benefits and Services
750	Administration of Justice
800	General Government
900	Net Interest
950	Undistributed Offsetting Receipts

Sources: See Appendix A.

OMB reports spending by 19 major categories (called "functions" in OMB parlance) and also by 74 more detailed "subfunctions." Functions 900 (Net Interest) and 950 (Undistributed Offsetting Receipts) are largely financial categories that we typically do not analyze since they do not correspond to any substantive government activities or programs. We have, therefore, 66 detailed topical categories of spending in 17 major areas of government activity, not counting the financial categories also reported by OMB. Most, though not all, of the OMB subfunctions and functions correspond closely with one or more of our subtopics and topics into which we have coded our congressional and *New York Times* materials. For most areas of spending, therefore, one can note whether spending corresponds with attention to the given topic area in the media, congressional activities, or whether it is related to legislative activities.

More important than our coding system to many users may be the fact that our datasets include a textual summary that includes a short description of each story, hearing, law, or abstract. These short summaries allow users to recode our datasets according to their own needs. The following table provides examples of records for a health subtopic from each of the datasets.

Table 2-5. Selected Textual Summaries from Four Datasets

Topic Code	Entry Summary
	Hearings
301	Federal health care spending
301	Health care reform and the role of medical technologies
301	Health maintenance organizations and hospitals providing managed health care
301	Health care access problems of disadvantaged and minority persons
301	Hospital financial practices and issues
	***CQ* Stories**
301	Minority Health: a non-controversial draft bill to authorize at least $144 million in fiscal 1994 to improve the health of minorities.
301	Alternative Health-Care Proposals: alternative plans made by Congress as opposed to the Clinton plan
301	Health Care Debate Takes Off: Congress gets up to speed on the complex economics and policies driving the US health care system
301	Health care program with included tax increase on the wealthy.
301	Health care reform bill to impose national limits on health spending and expanded access to health insurance for pregnant women, children and those who worked for small businesses.
	Public Laws
301	Amend the Public Health Service Act to provide an improvement in the health of members of minority groups
301	Provide federal assistance in establishing and expanding health maintenance organizations.
301	Revise and extend the program for the establishment and expansion of the health maintenance organizations.
301	Enact the Health Maintenance Organization Amendments of 1978.
	New York Times Index
3	Pres. Clinton's plan to save $35 billion from Medicare over next four years
3	Column article on both governmental and employers' long-term care policies and state intervention
3	Cost of health services should be distributed uniformly in all the states by financing it nationally
3	Hillary Clinton will appear before five committees of congress during hearings on Admin's health care plan
3	Letter from Western Pennsylvania Blue Cross executive officer explains how Penn. keeps percentage of people without health insurance under 10 percent

With the combination of an extensive and consistent set of topic codes in each dataset and the textual summaries available for each entry, users can combine or recode the datasets to meet their needs. All in all, the extensive topic and subtopic coding is the key to making these linked datasets useful to a broad audience in public policy. In the chapters that follow, a variety of uses of these topic categories are shown.

The major topic and subtopic content codes of the Policy Agendas dataset were developed through an iterative procedure that involved proposing an initial set of categories, coding congressional hearings to one and only one of these topic categories, and then modifying the categories until inter-coder reliabilities were achieved at the levels of at least 95 percent for the major topic and 90 percent for the minor topic codes (see Baumgartner, Jones, and MacLeod, 1998a). This topic system was then used to code all US Public Laws, *Congressional Quarterly* stories, and a sample of *New York Times* stories since 1947. Similar success in reliability was achieved for these datasets.

We began by coding congressional hearings, and continually updated and revised our coding system until we had done several years worth of coding. After we had done about 10,000 hearings, we reached a point where few changes were needed any more: each new case fit within one of our established categories, and two coders working independently from each other consistently coded the same cases identically over 90 percent of the time (and we achieved over 95 percent accuracy across the 19 major topics). We then went backwards in time, coding hearings all the way back to 1947. Subsequently, we followed the same procedures for all the Public Laws and then for all the stories in the *CQ Almanac*. Since these two sources often are broader than the congressional hearings, we more often made use of the –00 subtopic, which includes general coverage of the entire topic area or combinations of multiple subtopics. For example, an omnibus crime bill might well cover elements of sentencing, aid to local law enforcement agencies, prison spending issues, and other subtopics. This would be coded 1200 in the Public Law and CQ Stories datasets. Congressional hearings would be coded in the same manner, but Congress is more likely to hold hearings separately on each of these different elements of the bill. If they did, then the hearings would be coded according to the various relevant subtopic codes (1210 for sentencing; 1209 for police and law enforcement issues; 1205 for prisons; and 1299 for other and miscellaneous). For the *New York Times Index*, we coded only by the major topic categories and we added several categories because there are many areas of reporting that simply have no congressional counterparts (recipes, architectural reviews, sports results, obituaries, and weather reports would be some examples).

The topic categories that we developed are mutually exclusive and exhaustive. In the inevitable cases where a hearing, law, or story covered more than one topic, we coded it by the topic that was predominant. Each major topic also includes a general subtopic (always numbered –00) that includes cases where several different subtopics, all within the same major topic area, are discussed. In addition, where we noted large numbers of cross-references, we created distinct subtopic codes specifically for these cases. An example would be military health care issues: Are those defense issues or health issues? Our answer is to code them in their own category so that users can decide that question for themselves. Table 2-3 above showed that category 315 is reserved for this topic, just as category 311 is reserved for elderly health issues. By creating a series of "intersection" topics, we built into our coding system a level of flexibility that users with different interests may exploit. With over 220 topics in our system, most are quite specific. Still, users should note that we allow only for a single topic code for each item. Where the case clearly crosses boundaries, we either created a new category specifically for it (if there were many such cases), or we coded the case by the topic that predominated. In cases that were evenly balanced, our rule was to use the category that is listed first in the list of topics. The attached CD includes an extensive explanation of each topic category with examples of cases coded into each as well as "see also" references to related subtopics where similar cases may be coded.

Customizing and Supplementing Project Data

Among the greatest advantages of the Policy Agendas datasets is their reliability across time. The biggest problem with any single dataset is that it may not suit all needs. While our coding system is reliable, a student of a particular policy area may find that it does not match exactly the aims of his or her study. We suggest three strategies: creating one's own customized set of sub-topics; supplementing our data with further analysis; and searching and recoding our data based on the textual summaries. Many users will find that the datasets included here are sufficient for their needs, for example to compare attention and spending on defense issues to domestic policies such as education, health care, or transportation. A much broader community of users will be served, however, by some combination of our data and some others.

The first possibility for customizing our datasets is quite straightforward. Appendix A shows the full set of topic and subtopic codes we use. While we combined various subtopics into more inclusive major topic categories, these aggregations can easily be re-done, either to make our topic categories broader or narrower. Those interested in all defense-related issues, for example, might choose a set of subtopic codes that is centered within our major topic of 16 (Defense), but which also includes subtopic 315 (Military Health Care). Finding all foreign policy-related topics would include a combination of topics 16 (Defense), 18 (Foreign Trade), 19 (International Affairs) and possibly a few subtopics from other areas such as 530 (Immigration and Refugee Issues), which we code as part of the major topic of Labor, Immigration, and

Employment. The simple point is that users can easily recombine our 226 categories into customized topic areas that suit their needs. The more subtle point is that users should note that the ways in which we aggregated our 226 subtopics into 19 major topic categories might not suit all needs.

Second, our data can easily be supplemented. Each record contains the information needed to find the original source material. So, for example, one can identify every congressional hearing or *CQ* story focusing on immigration and refugee issues by searching on topic number 530. With this list, it is straightforward to then go to the CIS Abstracts, to the *CQ Almanac*, or to another source to gather information about who testified at the hearings, which legislation was considered, and what arguments were discussed, or what types of refugee issues were being debated. Several of the chapters that follow use our datasets as the base and supplement them with additional coding from the original sources.

The third way to customize our datasets is to make use of the short textual summary included in each record, illustrated in Table 2-5 above. These can be searched to identify mentions of key-words in combination with our topic coding system, or instead of using our topic coding system. This may be used in two different ways: to create entire new topics (searching for all mentions of words related to the elderly, for example, as one dissertation student we supervised successfully did in order to identify all hearings dealing with issues of concern to that demographic group), or in combination with a subtopic selection in order to narrow down one or more of our subtopics to an even more precise definition. In the example above of topic code 530, immigration and refugee issues, one can read through the summary to identify one or another of those more precise topics, or to find only those cases dealing with refugees from Asia, for example.

With the combination of an extensive and consistent set of topic codes in each dataset and the textual summaries available for each entry, users can combine or recode the datasets to meet their needs. All in all, the extensive topic and subtopic coding is the key to making these linked datasets useful to a broad audience in public policy. The attached CD includes a full set of codebooks explaining each variable in each of the datasets, a full description of the topic codes, including examples and "see also" references, and the datasets themselves, as well as some simple annual counts from each of the datasets. Our web site (http://depts.washington.edu/ampol/agendasproject.html) includes this information as well as various updates, bibliographic information, and other useful items.

Information Systems in the Study of Policy Processes

Recent years have seen the development of computerized search techniques and large publicly available databases of many types. These have created great new research opportunities, but also some new problems. Ironically, we often suffer from too much information, or more precisely from information that appears to be reliable at first glance, but which on deeper inspection proves to suffer from massive reliability problems. Most importantly, many large and historical datasets are designed for information retrieval, but are almost useless for the types of trend and pattern recognition that we have designed into our datasets. This is mostly because of three problems: backwards compatibility (that is, no effort is typically made to ensure that topic categories are reliable over time); over-categorization (that is, multiple keywords are used to index each item, but the keywords are not consistent and each item may be coded many times); and uniqueness (that is, each dataset, collected by a different institution, agency, corporation, or scholar, uses a different set of key-words and subject categories than the others).

Three Problems

The major problem today in quantitative policy studies involves moving from information systems based on retrieval to ones capable of trend recognition. The availability of many impressive and useful retrieval systems does not mean that they can be used for trend recognition. To do this, three critical problems must be addressed.

Backward Compatibility. Analysts maintaining existing databases that have been used to monitor policymaking—budgets, legislative activity, press coverage, etc.—tend to add and subtract categories over time. Normally no thought is given to making sure that previous uses of categories are consistent with present ones. This means that a category system applied in 1970 can be a quite different entity by 1995, even though it purports to assess the same material. Existing indexing systems do not value temporal consistency, but consistent categories are essential to studying policy change over time. Indexing systems must be continuously adjusted if temporal consistency is to be maintained, with all relevant material re-classified any time new categories are added. In the case of indices to hearings and media coverage that are published annually, however, there is no opportunity to go back and recode the previous years once the analyst decides that a new keyword or new subject category must be added. In our datasets, we did exactly this: we coded information covering the entire post-war period with the same topic categories in mind.

Over-Categorization. Over-categorization is the propensity to place items in multiple non-exclusive categories. A single congressional proposal applying civil rights legislation to providers of home health care, for example, might be classified as a commerce bill, a labor bill, a health care bill, an elderly bill, *and* a civil rights bill. An important example is Legislative Indexing Vocabulary (LIV), developed by the Library of Congress. This indexing system was developed to enable congressional staff and other researchers to identify legislative actions that are relevant to their interests. This search tool is available to the public via THOMAS, the Library of Congress web site. The LIV currently includes more than 7,000 subject terms, and a given bill can be coded as relevant to several dozen of these terms. While such an approach is desirable for information retrieval, it is practically useless for studying policy trends. The main purpose of the bill cannot be deciphered from the government's (or any other) indexing system. Since the purpose is to allow retrieval of all items even tangentially or incidentally related to the topic, many datasets typically are marked by extensive over-categorization. Similar problems come with the use of full text searches for key words. A search for all bills that mention the word "cancer" would find thousands of bills, but many of these would not be primarily about that topic, but would have mentioned it only in passing, their main thrust being elsewhere. For some users, this all-inclusive search process is exactly what is needed; for others, less so. The important point is that the user should realize that all indexing systems are not equally useful for all purposes. Typically, information retrieval systems are designed to err in the direction of including too much. Most importantly, they typically do not distinguish between those cases where the keyword refers to the main topic of the item found and those where the keyword is merely one of a laundry list of topics that may have been mentioned. Of course, to find every case where a given topic was mentioned, this is exactly what some users want, so these retrieval systems play an important role and their design is not a flaw; rather, it simply needs to be understood.

Uniqueness. Even if indexing systems overcome the backward compatibility and over-categorization problems, there is an additional issue: comparisons across datasets. One indexing system often cannot be compared to another. As a consequence, causal relations among arenas cannot be examined. For example, one might wonder if media coverage leads or lags governmental interest and activity in a policy area. This would not be possible even if both governmental activity and media coverage indexes were consistently categorized and were backward compatible unless both arenas were coded according to the same indexing system. While policy scholars emphasize "process," in fact we lack the tools needed to study the evolution of a policy idea quantitatively.

It is sometimes possible to construct parallel datasets using multiple data sources with different, but similar, categories and subject headings. For example, two of us previously constructed a series of comparable datasets tracing congressional and media attention to certain issues: pesticides, nuclear power, child abuse, and other topics (Baumgartner and Jones 1993). To do this, however, we were careful to construct separate lists of keywords and subject headings for each different data source (see Woolley 2000 for a discussion of the importance of making these comparisons with care). While this was possible for us to do in a small number of cases, it is not feasible to make such disparate coding systems match up across the board. While some areas of fit can often be found, there is no general solution. We solved this problem simply by applying the same coding system to all four of our related datasets. (We were forced to admit defeat in our efforts to make the OMB datasets completely compatible with our other ones, though it does correspond in most categories.)

Information Retrieval v. Pattern Recognition

Providers and immediate users of most data sources are generally interested in *retrieval*, whereas students and scholars are interested in *patterns and trends*. The critical component for assessing trends and patterns is the reliability of the measuring instrument. If the relationship between the indicator and the measurement object change over time, then the instrument cannot be used to follow trends. The critical component in retrieval is to make sure that the users of information find all relevant material. The expert indexer wants to make certain that a researcher interested in a topic will find and be able to recover a particular document. Librarians want an indexing system that allows them to find all relevant documents and would prefer to have this system err in the direction of providing more rather than fewer citations.

If, as is often the case, the indexed material evolves in content, then the indexer has no compunction about adding new key terms to aid the retriever of information in finding just what he or she wants. But it is extremely rare for the provider of information to go back in time and make sure the indexing categories are consistent. Any student of trends can make large-scale mistakes by assuming that a category this year contains the same content as last year's. Examples of this process would include something like the subject headings in the *New York Times Index* or the *Readers' Guide to Periodical Literature*. Considering a topic such as racial integration, new subject headings are often added whenever important new topics arise in the real world: so "busing" would become a relevant and widely used subject heading in the 1960s, just as "Brown v. Board of Education" would refer to many important stories on integration after, but obviously not before, 1954. There is certainly no reason why an indexer would not add new and important subject categories to allow readers to find the stories they are looking for. Our simple point is that users looking to trace changes over time must be especially aware of these changes in coding and category definitions (see also Woolley 2000).

There are three types of information systems, each of which is designed to achieve a different goal. It is useful to be aware of the goals of each type of information system. By far the most commonly used information systems are *retrieval systems*, designed to allow a user to use his or her knowledge to find a document, use multiple keywords to characterize each document in the system. An example is THOMAS, Congress' information system for citizens. Each bill, hearing, law, etc. is characterized by multiple categories—the Library of Congress' Legislative Indexing Vocabulary, which currently contains thousands of key terms. Because there are so many terms used to characterize each item, one has no way of tracing changes in policy categories across time.

Pattern recognition systems operate to find patterns empirically in a body of data. Designers develop computer programs embodying cluster or scaling routines that empirically search the data, and report patterns. An example is Poole and Rosenthal's (1997) comprehensive study of roll-call voting patterns in Congress since 1789, a research tool used regularly by congressional scholars. Their NOMINATE system basically recognizes patterns of voting in each Congress. Pattern recognition systems, however, cannot be linked to other databases, since it is the result of a scaling procedure using roll call votes, but with no effort to categorize by topic.

Trend recognition systems rely on the highest level of designer knowledge. Expert coding systems are established, generally based on some explicit or implicit theory of the subject matter. Categories in the coding schema are exhaustive and mutually exclusive, and the expert coders must make key decisions about where items are to fall. An important example is the tabulation of a country's economic output by sector. Economists must maintain consistent coding categories that are exhaustive and mutually exclusive. When the economy develops in a manner that requires new categories, experts must ensure compatibility by re-coding previous information according to the newly designed categories. Otherwise it will seem as if there is an explosion of new economic activity when the categories are added. A great variety of economic statistics are designed to allow comparisons over time, and the designers of these systems, such as the US Bureau of the Census or the Bureau of Labor Statistics, are loathe to revise their categories.

Since they are careful to revise them only when necessary (and to conduct careful studies to note the measurement-induced changes in the trends they seek to trace), users can safely use these data series to analyze trends over time.

Parallel trend recognition systems are trend recognition systems in which the same indexing system is used for several different arenas—as, for example, when media coverage and congressional hearings are coded according to the same policy content system. While it is common in full text databases to be able to search across databases using the same or similar keywords, we have discussed above how these full text searches can often err on the side of including cases that are only tangentially related to the topic searched for. In our databases, we used the same coding process to code congressional hearings, statutes, *CQ Almanac* stories, and entries in the *New York Times Index*. Therefore, common analyses can be done comparing these disparate data sources. Table 2-6 summarizes the differences in various types of information systems.

Table 2-6. Types of Information Systems

Purpose of the System	Substantive Base of the Initial Indexing System	Example
Information Retrieval	Low	THOMAS, other key-word based systems
Pattern Recognition	Medium	Nominate
Trend Recognition	High, in one issue-area	CPI, other economic indicators
Parallel Trend Recognition	High, in many issue-areas	Policy Agendas Project

Information systems designed for one purpose are not normally adaptable for other purposes. That is, a retrieval system is not easily adaptable to follow trends. Similarly, trend systems are not always the best devices for retrieving information. We believe that the datasets that comprise the Policy Agendas Project are the best available for the purposes that we use them for in this book. They are specifically designed to allow for parallel trend recognition, and the textual summaries that are included as parts of each record also allow users to conduct key-word searches. Of course, for some uses, a full text keyword search is what is needed, and for some uses our particular topic coding system will not be appropriate.

Conclusion

The Policy Agendas Project consists of five linked and comparable datasets covering the entire Post-World War Two period. These datasets allow the systematic comparison of a variety of issues and the exploration of a number of questions that have not before been subject to systematic quantitative analysis. We believe that the ability to make comparisons across many issues and to compare the policy process more generally will lead to a variety of new understandings about the American policy process. For example, we can systematically compare issues on the dimension of their agenda-status, and when we do so we observe that very few issue-areas remain out of the realm of public discussion over the entire period. Similarly, when we look at budgetary incrementalism, we find that almost all areas exhibit periods of stasis as well as periods of dramatic change. Full understandings of the policy process will come with the kinds of broad based comparisons that our new datasets allow, especially as these are compared and integrated with previous studies based on smaller sets of comparisons.

Many previous works in public policy have focused on certain patterns of behavior, such as institutionally created equilibrium behaviors leading to steady, routine, bureaucratic outputs. Others have contrasted with this literature in their studies of particular policy areas that have exhibited periods of dramatic change, reversal, or macro-political intervention. Because of the lack of a common set of indicators, however, scholars have not been able to characterize any given pattern of behavior as particularly common, uncommon, or so rare as to be a complete anomaly. While we do not expect that the creation of these datasets will put an end to various disputes in the literature, we do hope that they will encourage broad studies of the policy process, studies that note simultaneously the importance of inertia, incrementalism, and institutionally driven patterns of stability at the same time as they note that these characteristics of many policies coexist with their polar opposites: dramatic policy changes and the creation of new institutional structures that occur from time to time in all areas of politics and public policy. As two of us wrote in a previous book, we believe that these seeming opposites are actually part of a single underlying process (Baumgartner and Jones 1993).

The chapters that follow make use of our datasets in order to explore a great variety of theoretical and applied issues in American public policy. The first several chapters trace changes in the institutional settings of policymaking: Congressional processes have been greatly affected by the changing nature of the public agenda over time, as has the Supreme Court. Contributors note these changing patterns of behavior, stressing how institutions of government have been forced to change as a result of the changing nature of the issues before them. Of course, government helps to create public policy issues, so the relation between issues and institutional structures is a tightly woven one; the next few chapters illustrate the reasons why one should not treat one in isolation from the other. In the last section of the book, our contributors focus on the analysis of several different policy areas: defense spending, telecommunications, and other areas. In each of the chapters that follow, data from the Policy Agendas Project are combined with additional information gathered especially for that chapter. Thus, the chapters that follow provide important demonstrations of our theoretical points of the importance of punctuated equilibrium models to any understanding of public policy over the long-term, but they also provide examples of how to use the datasets included in the CD attached to this volume. We encourage readers to take these chapters only as the starting point and to explore further issues in their own analyses, using the following chapters as examples of the types of work that the Policy Agendas Project allows.