June 2012

Developing a Georgia Policy Database: a Research Proposal

Paul E. Rutledge
University of West Georgia, prutledg@westga.edu

Follow this and additional works at: https://digitalcommons.kennesaw.edu/gjpp

Part of the American Politics Commons, Comparative Politics Commons, Models and Methods Commons, and the Political Theory Commons

Recommended Citation
Available at: https://digitalcommons.kennesaw.edu/gjpp/vol2/iss1/2

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in Georgia Journal of Public Policy by an authorized editor of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.
Developing a Georgia Policy Database: a Research Proposal

Paul E. Rutledge
Assistant Professor
Department of Political Science and Planning
University of West Georgia
prutledg@westga.edu

Scholars of state and local politics have long faced the problem of data availability. The crux of the problem is consistent and reliable measures that are amenable analysis over time or across states. The problem with data on the state or local level is one of information retrieval. Such tasks are especially laborious, and are necessarily focused on a small part of a much broader system of policy dynamics. The lack of a systematic framework for data collection or analysis makes projects that focus on a time span of longer than a few years or more than a handful of issues difficult to achieve without substantial resources.

Early scholars of public policy faced similar problems with data availability. Scholars have produced impactful analyses of a limited number of policy issues over a short period of time despite those disadvantages. Such studies have given particular relevance to many of the prevalent theories of the policy process. Research on state politics and public policy, whether focused on a limited number of issues over a short time, or seeking to understand more longitudinal or cross-sectional policy change would benefit from the establishment of more systematic coding schemes and improvements in data availability.

The Comparative Agendas Project affords scholars with a framework for data collection and analysis largely free of the burdens discussed above. Such an undertaking requires a collaborative effort of many scholars. However, the benefit of adopting this framework is not limited to the improved ability to conduct longitudinal or cross sectional analyses. In addition to these benefits, scholars who prefer to focus intense attention to particular policies in a shorter time frame would also benefit from the availability of data that provides an effective starting point. Further, scholars would benefit from having data that is coded similarly across different governments. Comparing public policies across states or nations becomes less burdensome by having a consistent framework within which to conduct analyses.

The advantages discussed make the comparative agendas framework attractive for scholars of public policy on the national and state level. As such, the purpose of this article is partially a proposal, with the pitch being focused on the establishment of a Georgia Policy Database. Such a database would adopt the coding scheme developed by Baumgartner and Jones (1993) in the Policy Agendas Project (www.policyagendas.org), which has since been used as a springboard to launch comparative projects in fourteen countries, the European Union, and the Commonwealth of Pennsylvania. While the project has only been adopted in one of the fifty states thus far, the volume of research that has resulted from the Comparative Agendas Projects demonstrate clearly the potential that such a project has to increase and improve scholarly activity on Georgia politics and policy. Further, the study of state politics in general would be improved by more states adopting the systematic framework for comparative analysis. Finally, this essay will

Georgia Journal of Public Policy, Spring 2012
demonstrate that the benefits of adopting the comparative agendas framework in Georgia would not be limited to scholars. Instead, because of the user friendly nature of the data, practitioners, politicians, and citizens are also likely to benefit tremendously from the project.

This essay will proceed in the following manner. First, attention will be focused on the benefits of the comparative agendas framework with some attention focused on the scholarly productivity associated with the projects. Secondly, a specific discussion of the framework as it has been applied to the Commonwealth of Pennsylvania will illustrate how a framework that was developed on the national level can be adapted to the states. Finally, the essay will conclude by discussing the possibility of such a project in Georgia, with a focus on what is necessary for such an undertaking. In the end, the hope is that this article will be a first step in the formation of a scholarly network focused on the study of public policy in Georgia that through collaborative effort can establish a Georgia Policy Database.

**The Comparative Agendas Project**

The Comparative Agendas Project ([www.comparativeagendas.org](http://www.comparativeagendas.org)) is a project that seeks to bring together scholars for the development of systematic and comparable indicators of government activity across political systems. Scholars work from within the settings of their political system to develop a data coding scheme that allows for systematic analysis on a national or comparative focus. Each of the projects are based initially off of the coding scheme developed by Baumgartner and Jones (1993) in the Policy Agendas Project ([www.policyagendas.org](http://www.policyagendas.org)). The coding scheme is applied to policy related events specific to a political system. Examples of the data available from the U.S. Policy Agendas Project (also listed in table 1 below) include State of the Union Speeches, Executive Orders, budget allocations, Congressional hearings, public laws, Supreme Court cases, *New York Times* stories, Gallup’s most important problem public opinion series, and other “spin-off” projects that have coded Congressional bills ([www.congressionalbills.org](http://www.congressionalbills.org)). Clearly, this list of policy activity includes appropriate indicators of policy priorities for each of the national government institutions, the media, federal spending, and the public. A classification scheme that can be applied to each of these actors and institutions opens a wealth of measurement tools for answering several important research questions that have been the focus of policy research for decades.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Civil Rights, Minority Issues, and Civil Liberties</td>
<td>12. Housing and Community Development</td>
</tr>
<tr>
<td>3. Health</td>
<td>13. Banking, Finance, and Domestic Commerce</td>
</tr>
<tr>
<td>5. Labor, Employment, and Immigration</td>
<td>15. Space, Science, Technology, and Communications</td>
</tr>
<tr>
<td>6. Education</td>
<td>16. Foreign Trade</td>
</tr>
<tr>
<td>7. Environment</td>
<td>17. International Affairs and Foreign Aid</td>
</tr>
</tbody>
</table>
The U.S. Policy Agendas Project has resulted in some of the most important theoretical advances governing the agenda-setting and public policy literature. In their seminal work resulting from the project, Baumgartner and Jones (1993) traced political attention across the nineteen major topic codes, listed in table 1, and two hundred twenty-five subtopic codes in an investigation of the shifting of political priorities on the national level. Each record is coded according to a two digit major topic code and a two digit subtopic code. Therefore, each record will have a four digit code, with the first two digits reflecting the major topic and the last two reflecting the subtopic. For example, records relating to the training of medical personnel are coded as 0325, with the “03” indicating that it is a health care issue and the “25” indicating that the record is considered health and manpower training.

The records coded in the U.S. Policy Agendas Project, and then emulated in other political systems, have included bill introductions, legislative hearings, executive orders, executive speeches such as the annual State of the Union Address, media stories, Supreme Court cases, public views about policy priorities such as Gallup’s most important problem survey, interest group populations, public laws, and budgetary allocations. Each record for the above measures is double blind coded, typically by students, and then a random sample of the records are coded a third time by project coordinators to ensure coding reliability. Each record is assigned one only one major and subtopic code, with “ties” determined by the main focus of the policy or who benefits most from the policy (where applicable). For example, if one were coding a law that expanded the ability of students enrolled full time in a university to remain covered by a parent’s health insurance, the primary focus of the law is the expansion of health care coverage. While it impacts students enrolled in higher education, the impact of the law is focused more fully on expanded health care coverage. Therefore, the law would be coded in major code 3, health care, and subtopic code 302 insurance reform, availability, and cost.1

Baumgartner and Jones (1993) in their initial work resulting from the project introduced the theory of punctuated equilibrium to the study of public policy, providing evidence that incremental theories of policy change that were so dominant in the literature until that time were complimented by periods of rapid and expansive change that shifts the future equilibrium point. Further, Jones and Baumgartner (2005) posited further that both incremental policy change and policy punctuations could be attributed to positive and negative feedback producing institutional friction; the resistance level associated with changes in different stages of the policy process. In general, the later in the policy process one proceeds, the higher the costs associated with shifting priorities. When the associated costs are overcome in stages with higher friction, however, the rate and scope of change tends to increase, leading to more explosive shifts in the policy area as a result of positive feedback overcoming a set threshold for change.

Since the Policy Agendas Project data has become publicly available, it has been used in a wide variety of studies. Many of the bigger theoretical questions scholars of American politics encounter can be investigated effectively using the data supplied by the agendas project. Some examples include the legislative impact of divided government (Jones, True, and Baumgartner, 1997) representation (Jones and Baumgartner 2004; Jones, Larsen-Price, and Wilkerson 2009), the politics of budget allocations (Jones, Baumgartner, and True, 1998), major changes in public policy (Jones, Baumgartner, and Talbert 1993; Talbert, Jones, and Baumgartner 1995), and the

---

1 For a full list of subtopic codes, please see [www.policyagendas.org](http://www.policyagendas.org).
relationship between institutions and interest groups (Leech et al 2005; Baumgartner et al 2011). Similarly, scholars who study individual issue areas have made effective use of the data to study policy dynamics within their area of study. While this list provides a broad overview of the theoretical advances the Policy Agendas Project has made possible in the study of American politics, it certainly does not do sufficient justice to the wealth of publications that have resulted from the project.

Recently, scholars have recognized the potential of adopting the coding framework first introduced in the U.S. Policy Agendas Project to other nations for comparative research on public policy. Comparative Agendas Projects have been initiated in Australia, Belgium, Canada, Denmark, the European Union, France, Germany, Hong Kong, Israel, Italy, the Netherlands, Portugal Scotland, Spain, Switzerland, the United Kingdom, and finally in the Commonwealth of Pennsylvania. Despite the comparative aspect of the agendas project being in its nascent stages, a tremendous amount of scholarship has already resulted. Baumgartner et al (2009) have found evidence of punctuated equilibrium as a model of policy change in Belgium, Denmark, and the United States. Further, punctuated equilibrium has been found to produce a power law functional distribution in budgetary allocations across six nations (Jones et al, 2009). Policy studies have also been focused on nation-specific agenda setting dynamics in Denmark (Green-Pederson, 2006) and France (Baumgartner, Foucault, and Francois 2006; 2009) among others. Finally, dynamic studies of singular policy areas including health care (Green-Pedersen and Wilkerson 2006), immigration (Green-Pedersen and Krogstrup 2008), and food safety and pharmaceuticals (Chagues and Palau 2009), among numerous others have been published making use of the comparative agendas project data2.

Studying policy change comparatively using the Policy Agendas framework has led to tremendous theoretical advances in the study of public policy in numerous contexts, as briefly demonstrated by the discussion above. The potential for comparative studies of public policy has already been demonstrated in the volume of work arising from the Comparative Agendas Project while still in their respective nascent stages. Taken together, the research that has been produced from the agendas projects both in the U.S. and comparatively demonstrates the potential that such a project has on the state level. In the next section, I will discuss the development of the Pennsylvania Policy Database, and provide an example of the research made possible by the site’s policy analysis tool. Finally, I will conclude by discussing the benefits and requirements for creating a Georgia policy database.

The Pennsylvania Policy Database

The dilemmas of conducting longitudinal studies of public policy discussed above are especially pronounced in the context of the U.S. states. Information retrieval is a major issue because most states lack a consistent or comprehensive record of public policy. This makes studying policies

---

2 The studies referenced here are merely a sample of a much larger volume of research that has been produced using the U.S. and Comparative Policy Agendas Projects. For space limitations, I have been far too brief to give the studies due justice. For a broader list of studies, please refer to the Comparative Agendas Project website at www.comparativeagendas.org, and click on publications.
over time or across venues very difficult, and leads scholars to research designs that are limited in nature simply by resource availability. Further, the lack of a standardized coding framework leads to a literature that is tremendously difficult to advance in a linear fashion, with studies directly building upon one other. It is precisely these problems in the study of state politics that led researchers at Temple University in Philadelphia, Pennsylvania to construct a Pennsylvania Policy Database (www.temple.edu/papolicy). Researchers hope that adopting the framework first introduced on the national level in the Policy Agendas Project described above will lead to numerous gains on the state level. While the project was initially designed for a single state, McLaughlin et al (2010) suggest that the benefits of the project are not exclusive to the Commonwealth of Pennsylvania.

The benefits of the Policy Agendas Project as a model for the study of state politics are effectively described in McLaughlin et al (2010). First, the framework provides a systematic and comprehensive model to reconstruct policy history in multiple venues. As such, the longitudinal problem that has plagued state policy research (and policy research more generally) is eliminated with the coding scheme. In many states, records are decentralized and difficult to assemble. The Pennsylvania Policy Database has as one of its main purposes the creation of a centralized location for policy records to make research on policies or processes a simpler undertaking. Many of the most important theoretical advances have been made possible through such large scale projects as the National Election Studies, the Party Manifestos Project, the Correlates of War Project, the Militarized Interstate Dispute Dataset, and of course the Policy Agendas Project and Comparative Agendas Projects discussed above. The creation of a publicly available dataset has led to an abundance of scholarship in each of these examples, and I expect that a further expansion of the comparative agendas project framework to additional states, including Georgia, would lead to a similar profusion of scholarship both within and across states.

Second, because state governments deal with many of the same policy issues the national government faces, the codebook can be adopted to the state level with only minor changes. The Pennsylvania Policy Database did require some minor modifications to the codebook. The major topic codes associated with the Pennsylvania Project are provided in table 2 below (table 1 above contains the major codes associated with the national level project). A comparison of the Pennsylvania project’s major topic codes in table 2 and the national level topic codes in table 1 above demonstrate that the changes are minimal. On the major topic level, only major topic code 24 has changed. In the U.S. project, major code 24 is reserved for state and local government administration because of the relationship between the national and state governments. The equivalent functional relationship on the state level is that between state and local governments. As such, the Pennsylvania database contains State and Local Government Administration as major topic code 24.

There are a few other changes on the subtopic level. Each major topic code contains a number of subtopic codes to allow researchers to focus their inquiries on more specific policy issues than would be made possible by the major topic codes only. In total, the national project codebook includes 225 subtopic codes, whereas the Pennsylvania codebook contains 249 subtopic codes3. Several subtopic codes specific to the state level were necessarily added to and amended from

---

3 A complete list of the subtopic codes used for the Pennsylvania Project database can be accessed at www.temple.edu/papolicy.
the national level codebook to avoid missing important areas of state policy. For example, subtopic code 345 in the Pennsylvania codebook deals with the Provision and Regulation of Ambulance services, which is specific to the state level.

| TABLE 2 |
|-----------------|-----------------|-----------------|
| 2. Civil Rights, Minority Issues, and Civil Liberties | 12. Housing and Community Development |
| 5. Labor, Employment, and Immigration | 15. Space, Science, Technology, and Communications |
| 6. Education | 16. Foreign Trade |
| 7. Environment | 17. International Affairs and Foreign Aid |
| 10. Law, Crime, and Family | 20. Local Government and Governance |

While minor changes were required to adapt the codebook from the national to state level, the Pennsylvania project codebook maintains a nearly perfect similarity with the national codebook. This allows for the possibility for systematic studies of policy diffusions between state and nation to take place in the federalism context. Finally, and most important for the purposes of this article, McLaughlin et al (2010) note that since the codebook has been adapted to the state level already in the Pennsylvania project, it is easily adaptable with even fewer changes for use by scholars in other states. That said, McLaughlin and his colleagues admit that the construction of the codebook is an iterative process rather than a finished product. New issues arise which require researchers to make tough decisions. However, such issues are the exception rather than the rule.

| TABLE 3 |
|-----------------|-----------------|-----------------|
| Dataset Description | Policy Agendas Data | Pennsylvania Database Data |
| Legislative Hearings | Congressional Hearings | State Legislative Hearings |
| Legislation | U.S. Public Laws | Acts, Bills, and Resolutions |
| Executive Orders | Executive Orders Issued by the President | Executive Orders Issued by the Governor |
| Executive Speeches/Messages | State of the Union Addresses | Governor’s Budget Messages |
| Supreme Court Decisions | U.S. Supreme Court Decisions | Pennsylvania Supreme Court Decisions |
| Budget | U.S. Budget Authorizations | Pennsylvania Budget Expenditures |
| Media Coverage | New York Times Index Sample | Samples of State Capital News Digest from State Press Office |
| Public Opinion | Gallup’s Most Important Problem Series | State Public Opinion Polls (where available) |
| Most Important Policy Issues | Congressional Quarterly | Governing Magazine |

Table 3 compares the records available in the national and Pennsylvania projects. This table demonstrates the volume of policy activity made available by the projects, covering the institutions of government, the media, the public, and capitol publications covering the issues deemed to be salient at each level. An examination of the table illustrates that policy activity at each level has a functional institutional equivalent in its counterpart project. For example,
legislative hearings and bills occur on each level and are appropriate measures of policy activity in the legislatures. The State of the Union speech and the State of the State speech have the functional equivalence of executives reporting on the major issues facing the state or nation, and proposing an agenda for the legislature to address those issues. Supreme Court cases in the national and state levels are comparable indicators of activity by the judicial branch. Measuring media and public prioritizations of issues on the state level become a bit more problematic because many states do not have the functional equivalent of the New York Times, used as a proxy for media coverage on the national level, or Gallup’s Most Important Problem survey which gauges the issue citizens feel is most in need of government attention. Instead, systematic measures of public opinion that have been repeated over time are limited, and not every state has a “statewide” newspaper. McLaughlin et al (2010) note that a sample of stories collected by state press offices may offer the best hope of measuring coverage of statewide coverage of political issues. These are state specific issues that must be worked out prior to undertaking a database project by scholars in their states.

Aside from the benefits to be derived from scholars discussed above, there are also benefits to citizens and political actors in each state provided by the database. For citizens, the Pennsylvania project website is very user friendly, allowing citizens to track political activity on issues of importance to them. Political actors can access the website for an objective account of previous policy activity in the policy areas of interest to them (McLaughlin et al, 2010). Teachers can use the website for examinations of state policy activity in the classroom. McLaughlin et al (2010) use the policy analysis tool found on the Pennsylvania Project website to examine the spending and attention tradeoffs between Medicare and funding for education that have become part of conventional wisdom. Such examinations are easily undertaken through the user friendly tool in a matter of minutes. This drastically improves upon the decentralized and onerous nature of existing state level data, and opens the processes to examination by numerous individuals even if they lack the statistical and research skill possessed by many academically trained scholars of state politics. In the state of Georgia, this dataset could be used to trace governmental attention to education spending which has been increasingly salient given recent budget constraints, attention to the increasingly important issue of illegal immigration that has been addressed in many states, and the continuing problems surrounding water policy. Citizens, academics, policymakers, and lobbyists would find this dataset effective for persuading Georgia policymakers that insufficient attention has been devoted to a given issue, or legislative activity that has been undertaken has not been sufficient. It is a valuable way to gain objective evidence to inform policy debates and calls for change.

Thus far, the essay has focused on the advantages that the project offers to scholars of state politics. In conclusion, it is necessary to discuss some of the complications similar to that of measuring media and public attention, and further what is required to embark on such a major undertaking.
Developing a Georgia Policy Database

The previous sections detail the volume of studies and the benefits derived from the Policy Agendas framework in the national, comparative, and state politics contexts. While the benefits are obvious in the volume of research produced by scholars, the possible use by citizens, politicians, and educators, and the easily adaptable codebook, the complications are not as obvious. This section will focus on the complications and requirements with starting such a project in Georgia, with the admitted hope that this essay will serve as a springboard for scholars who are interested in collaborative effort to overcome them.

The biggest challenge lies in the area of funding and resources. The Pennsylvania Database project was funded by the state legislature. Students working on the project at six Pennsylvania Universities collected and double-coded more than 157,000 records covering the period from 1979-2010 at an initial cost of $488,000, with several smaller (<$25,000) appropriations in subsequent years. Recent innovations in the use of Text Tools for coding textual records have increased efficiency and reduced the cost associated with human coding of the entire dataset. Such text classification software is capable of adopting coding behavior according to records previously coded by human coders to code more than half of the entire dataset at a small fraction of the cost. As such, a Georgia Policy Database would be considerably cheaper to create than previous iterations of the Comparative Agendas Project, including the Pennsylvania Policy Database. The amount of records required to implement the project in Georgia and the decentralized nature of the records make collaborative efforts among scholars at multiple universities in the USG the most likely avenue for success, similar to the Pennsylvania database. In addition to the decentralized nature of the records making collaboration among a team of scholars more likely, collaboration also offers the advantages of broadened support and added expertise for the project. Once a team of interested scholars statewide have assembled, determinations can be made on which scholars have the best access to the records required and where different component parts of the database project should be housed. For example, researchers at the University of Pennsylvania had the best access to budget data for the Pennsylvania Project, therefore the budget data was collected and coded there. The volume of newspaper articles required work from researchers at two site locations, Penn State University and Temple University. External funding is absolutely vital to the project, and opportunities for external funding should be an early focus of interested scholars in Georgia.

Another challenge arises in locating the data. Records are generally housed in multiple locations across the state, with some data sources more easily accessed than others. Again this problem lends itself to a collaborative enterprise. An important first step is for scholars interested in the project to determine which records are most accessible at their university, and whether the data is will be available and easily accessible for student coding. The availability of data and the ease with which it can be coded by students is an important consideration in estimating a) whether the project is feasible, b) how much funding will be required to complete the project, and c) where the funding has to be allocated most. For example, in the construction of the Pennsylvania Policy Database collecting and coding newspaper records proved to be the most expensive and

---

4 For more information on Text Tools, please see the publicly available information for downloading and using Text Tools at [www.comparativeagendas.org](http://www.comparativeagendas.org).
difficult aspect of the project. Additionally, researchers have to determine whether the data is available over a sufficient time period to provide the longitudinal benefit associated with the other projects that have adopted the Policy Agendas framework.

In spite of this list of significant challenges, there are some factors which make the construction of a Georgia Policy Database far less burdensome. The codebook adapted to the state level for the Pennsylvania Policy Database can be simply adopted by other states, including Georgia. The codebook is comprehensive and detailed, making it very unlikely that a large number of coding issues will fall outside of the scheme as it currently exists. In this way, once scholars of Georgia politics have located the required records and secured funding, the system for coding decisions is already in place. Thanks to the hard work of scholars who created the Pennsylvania Policy Database, a large portion of the initial startup costs have already been borne out in a template.

Further, while resources and funding seem to be the biggest challenge, recent developments in automated text classification make it possible for a smaller portion of records to be coded by student researchers, with the rest being coded following the human coding samples with machine logarithms. These developments cut the funding that would be required for a Georgia Policy Database significantly. In comparing double human coding to automated coding, Hillard, Purpa, and Wilkerson (2007) found that automated coding was 90% accurate at the major topic code level, and 80% accurate at the minor topic level. This matches or exceeds inter-coder reliability among student coders, which is far more expensive. With the amount of funding required being significantly lower as a result of automated text classification, locating funding becomes far less of a challenge than it has been for previous projects of a similar magnitude.

Finally, scholars associated with the existing database projects have made it clear on their websites and in their research that they are readily available and willing to assist in the commencement of new projects. McLaughlin et al (2010) specifically volunteer researchers from Temple University, the main site of the Pennsylvania project, to assist researchers in adopting the Pennsylvania project’s framework for use in other states. Similarly, each year a conference of scholars engaged in the construction of databases adopting the Policy Agendas framework meets to discuss issues with adopting the framework to their nations and present the progress they have made. This network of scholars provides useful feedback in a broader collaborative effort to expand a standard classification scheme for public policy to promote comparative policy research.

The availability of a broad scholarly network and recent developments that have decreased the costs of such a project has made the construction of a Georgia Policy Database more feasible than it has ever been. The volume of scholarly output produced as a result of the projects that have been undertaken make the benefits of the project plentiful for scholars of Georgia politics. The user-friendly nature of the policy analysis tool originally developed in the U.S. agendas project which has been adopted by subsequent projects make the project valuable to practitioners and citizens alike. Scholars within the Peach State have an opportunity to be among the first to adopt this innovative framework to state policy studies. My hope is that this essay will interest scholars of Georgia politics in coming together to pursue this exciting opportunity. I encourage anyone interested in the project to use the contact information provided to contact the author. If this possibility generates sufficient interest from a network of scholars, my hope is that we will
hold a mini-conference to discuss the possibility of constructing the database. This project is a unique and exciting opportunity, and I hope this sales pitch has convinced enough scholars to work collaboratively to make it happen for Georgia.
References


