# Sticky Legacies: Persistence of State Constitutional Provisions 

Nancy Martorano Miller
University of Dayton, nmiller1@udayton.edu
Maria Aroca
Rice University, mp.arocav2@gmail.com
Keith E.Hamm
Rice University
Ronald D. Hedlund
Northeastern University

Follow this and additional works at: http://ecommons.udayton.edu/pol_fac_pub
Part of the Political Theory Commons, and the Public Affairs, Public Policy and Public Administration Commons

## eCommons Citation

Miller, Nancy Martorano; Aroca, Maria; Hamm, Keith E.; and Hedlund, Ronald D., "Sticky Legacies: Persistence of State Constitutional Provisions" (2016). Political Science Faculty Publications. 108.
http://ecommons.udayton.edu/pol_fac_pub/108

# Sticky Legacies: <br> Persistence of State Constitutional Provisions 

Nancy Martorano Miller<br>University of Dayton<br>Nmiller1@udayton.edu

Maria Aroca<br>Rice University<br>mp.arocav2@gmail.com

Keith E. Hamm
Rice University
hamm@rice.edu

Ronald D. Hedlund<br>Northeastern University<br>r.hedlund@neu.edu


#### Abstract

In this paper, we assess the evolution of 32 state constitutions and the U.S. Constitution over a $100+$ year time period (1776-1907). We construct an original sectionlevel dataset containing the text of every section within a constitution for every year between the adoption of the state's first constitution and 1907. We classify each section by topic and compare the content of each new constitution as well as the impact of amendments. With a subset of these data, we analyze the extent to which sections were added, deleted, modified and remained the same over time using a novel approach that relies on an edit distance measure to quantify the similarity between sections of two constitutional documents. We are also able to empirically evaluate the level of similarity of modified sections as new constitutions were adopted or alterations were made to an existing constitution. Finally, we determine which topic areas were subjected to the largest amount of change. We demonstrate that it is possible to systematically assess a large corpus of constitutional documents to test theories of institutional change, provide empirical support to existing qualitative accounts, and create operationalizations of concepts such as "stickiness" that are comparable across states and over time.


Paper prepared for presentation at the Annual Meeting of the American Political Science Association, Philadelphia, PA, September 1-5, 2016.

# Sticky Legacies: Persistence of State Constitutional Provisions 

"Thus, the states remain significant determinants of the quality of life of the American people. The way in which each state frames and allocates powers through its constitution reflects certain conceptions of government and understandings of the two faces of politics - power and justice. That is, state constitutions are important determinants of who gets what, when and how in America because they are conceptual and at times, very specific statements of who should get what, when and how (Elazar 1982: 17)."

## Why Study State Constitutions?

The activities of state governments have a profound impact on the quality of the lives of those living within their borders. In addition to establishing the basic structure of governance, these documents also set the rules with regards to citizen participation in the governing process; guarantee the rights and liberties of citizens; outline the functions and powers of local governments; and in many cases, direct public officials to enact/implement specific public policies as well as provide specific services to citizens. Thus, the impact of these documents on our everyday lives is far-reaching.

Historically, state governments have been front and center in the governing process. In the United States' earliest days, the sovereignty of the ability of the states to govern as they saw fit was a core tenet of the nation's founding. Today, the U.S. states are often characterized as "laboratories of democracy" for their ability to innovate and to formulate public policies aimed at solving some of the most pressing public policy issues. Elected officials in the states have often had both the public's support and willingness to experiment with innovative solutions to policy that is absent at the federal level. The policy areas of same-sex marriage, marijuana legalization, and immigration are all areas where states have adopted policies in the face of inaction by the federal government. Of course, the policy adopted in each area varies significantly across states with the policies adopted reflecting the political culture and values of the states' populations.

A society's constitution represents a fundamental statement of that society's core cultural, political and social values at the time of its adoption, and we should expect that these documents have significant impact on how policy is adopted as well as on the content of the policy adopted. Donald Lutz has studied state constitutions extensively, and he believes that a written constitution is a "political technology" or "the very embodiment of the technology for achieving the good life (Lutz 1982). Further, Lutz argues eloquently for the need to study the content and effect of these documents:
"Constitutional variables do not matter, we are told, because political variables do not seem to matter. However, studies of political variables usually ignore constitutional variables. For instance, the presence of a oneparty system as opposed to a two-party system is a political variable, not a constitutional one. The presence of a unicameral versus a bicameral
legislature, on the other hand is a constitutional variable. Herein lies part of the problem. Many important constitutional provisions do not vary much from state to state, and can hardly be used to explain variance in public policy. On the other hand, the tendency for certain constitutional features to be similar over most states must itself be subject to investigation, just as these similarities must be considered as possible reasons for similarities or lack of diversity in certain types of public policy. Put another way, these constitutional similarities may so structure the political variables that public policy will tend to respond to environmental conditions in a differential manner simply because constitutional variables have already had their effect by limiting the range of possible policy outcomes to a relatively narrow set a set of possibilities within which environmental factors are determinative (Lutz 1982: 28)."

A handful of other political scientists and legal scholars writing at the same time as Lutz argued for more systematic study of state constitutions (Elazar 1982; Kincaid 1988, Friedman 1988). For the most part, this call has gone largely unheeded by political scientists, and legal scholars have chosen to focus more on interpretations and analysis of isolated provisions found in a specific state constitution or on deep constitutional histories in single states. ${ }^{1}$

We believe that this inattention is driven by a perception that state constitutions are similar to the U.S. Constitution as well as the challenge of determining how best to systematically evaluate these documents. In previous research, we convincingly illustrate that state constitutions are in fact very different from the federal document (Martorano Miller, Hamm and Hedlund 2009; 2010; 2011; 2014a; 2014b, 2015, 2016), and in this paper we provide a systematic analysis of pre-20 ${ }^{\text {th }}$ Century constitutions. Current events also provides us with an example of why we should care about the political "playing fields" that state constitutional documents create. A recent session of the Texas Legislature provides evidence to this claim. Ross Ramsey, of the Texas Tribune published a story on May 27, 2015 titled, "Analysis: Votes Count, but the Rules Can Count More," regarding the passage of legislation in the Texas legislative assembly. The Texas Constitution contains some very specific rules for the consideration of bills and a well-schooled member or minority of members can use that knowledge to stop legislation from passage. This in fact occurred with bills and amendments regarding abortion policy, same-sex marriage policy and child welfare that session (Ramsey 2015a, 2015b).

## The Scope of This Study

[^0]In this paper, we are presenting our first attempts to leverage the power of machine reading and analysis of state constitutions in order to better describe and understand how these documents have evolved over time. We have chosen to focus on the development of the first state constitutions or the time period of $1776-1907$. For this roughly 100 year time period we have collected and processed each constitution and constitutional amendment for thirty-two states as well as the U.S. Constitution. ${ }^{2}$ First, we will use these data to illustrate some basic descriptions about the evolution of state constitutions. Second, we will use the cases of Alabama, Colorado, Illinois, Minnesota, New Jersey, Wyoming and the U.S. Constitution to illustrate the power of approaching the analysis of texts in this manner.

The data for this project comes from mainly two sources: 1) the NBER/Maryland State Constitutions project (NBER) ${ }^{3}$ and 2) Francis Newton Thorpe's, The Federal and State Constitutions, Colonial Charters, and the Organic Laws of the State, Territories, and Colonies; Now or heretofore Forming the United States of America (Thorpe) ${ }^{4}$. The NBER project is a federally funded project that created a searchable web-based database of state constitutions and their amendments. The Thorpe resource is a seven-volume compendium of important state documents that was commissioned by the U.S. Congress and published in 1909. NBER relied heavily on the Thorpe work and two of the co-authors of this study crosschecked both sources for accuracy as we processed the data. Appendix A contains a detailed discussion of how we used the data found in the NBER database ${ }^{5}$.

We also adopted the coding scheme of NBER. In this research our unit of analysis is the section level of a constitution. Most state constitutions are organized by Article, then Section and subsection. ${ }^{6}$ The NBER treated each section of a constitution separately and coded each as belonging to one of 91 topic areas ${ }^{7}$. We then collapsed those categories into

[^1]seven broad topic areas found in most state constitutions: 1) Bill of Rights/Minorities, 2) Constitutional Amendment Process, 3) Suffrage and Elections 4) Finance and Taxation, 5) Policy, 6) Structure of Government and 7) Federal Government. We excluded from our analysis the following types of sections: preambles, schedules, certificates and ordinances as these sections have no direct impact on the processes of governance of the governing structures themselves.

Using $R$, a software environment for statistical computing, we transformed the text files of these constitutions into a section-level dataset where each unit is a section within a constitution. A dataset was created for every year since the adoption of the first constitution of the state up until 1907. The datasets include the topic and full text of each section. If amendments were made or a new constitution adopted, we replaced the text of each section with the text that resulted from any modification. If sections were deleted, they were dropped for the dataset of that year and if they were added, they were included in the dataset of that year. Our final datasets are then a reflection of the way each constitution looked as of that particular year. This data structure allows us to make section level comparisons of a constitution in a given point in time, to that same constitution at a later or previous point in time, as well as to empirically track the evolution of each state's constitutional document(s). With these data we are able to assess change in the aggregate across many states overtime as well as explore that change in some detail within a state. Moreover, we are able to compare the text of every section of a state's constitution at point $t$ to the text at point $t_{\mathrm{n}}$. To assess the similarity of two sections and after properly processing and cleaning the text, we used a modified normalized version of the Levenshtein edit distance, which calculates the minimum number of insertions, deletions and substitutions of words that would be required to transform the text in any given section at time $t$, to the text of that same section at time $t_{\mathrm{n}}$. The score takes values between 0 and 1 , where 1 indicates maximum similarity, and 0 indicates null similarity ${ }^{8}$.

Using this approach, we have traced the evolution of each section of a sub-sample of states over time. We will first present data from our thirty-two state constitutions plus the U.S. Constitution sample that provides a basic overview of these documents. We will then proceed to use our smaller six state constitutions plus U.S. Constitution sample to illustrate the power of utilizing these techniques to leverage these text sources.

## What do We Know About State Constitutions?

Regarding constitutional politics, the U.S. states differ significantly from the national government. Collectively, the states have adopted 145 constitutions since the nation's founding ${ }^{9}$, and the current state constitutions have been amended over 10,000 times (Tarr

[^2]2014). During the time period of this study (1776-1907), the thirty-two states and the United States adopted 80 constitutions and 614 amendments (See Figure 1). In contrast, the United States has had a single federal constitution that has been amended only twentyseven times in 229 years with twenty-one of those amendments occurring before 1908. Figure 1 illustrates that during the first 100 years, the states were much more able and willing to modify their constitutional documents with some states like California, Minnesota and Michigan choosing a path that relied more on the amendment process while most Southern states adopted five or more constitutions during this period. ${ }^{10}$

## [Figure 1 about here]

Further, state constitutions on average tend to be longer than the U.S. Constitution and are more likely to contain explicit provisions outlining the structures of government and directing the state government to adopt specific policies or provide specific services to citizens (see recent work by the authors: Martorano Miller, Hamm and Hedlund 2009; 2010; 2011; 2014a; 2014b, 2015, 2016).

Additionally, scholars such as Lutz (1982) and Kincaid (1988) have argued that while the scope of influence of the national government has expanded significantly since the Nation's founding, the true basis of governance in the United States most certainly exists in the individual state governments. These scholars assert that the Framers intentionally drafted the U.S. Constitution as an incomplete document - purposely leaving to the states all residual powers not enumerated in the federal document. Thus, it was left to the states to "define and implement many provisions of the U.S. Constitution and structure the potentially vast domestic powers reserved to the states and to the people by the U.S. Constitution (Kincaid 1988: 12)" and to "develop the description of a way of life (Lutz 1982: 41)" in the new formed nation.

Figure 1 also provides additional empirical support for John Dinan's (2009) extensive research on the state constitutional convention proceedings. He convincingly argues and provides evidence that reinforces the contention of earlier scholars that the founders intended the states to "flesh out" governance in their respective constitutional documents. In his book, Dinan (2009) makes a strong case for the belief that the states have been better at revisiting their constitutions and revising their institutions and governing principles based on past experiences or fundamental shifts in culture, etc. He asserts that constitutional revision and amendment processes at the state level are easier relative to the process of revising or amending the U.S. Constitution, and thus have allowed state governments to evolve in ways that make them more responsive to modern problems.

State constitutions, while containing many of the provisions found in the federal document, also differ markedly from our federal constitution in the manner in which government structures and relative power was originally perceived. For example, the U.S. Constitution clearly delineates the Congress' legislative powers in Article I, Section 1 and

[^3]limits the Congress to only those "legislative Powers herein granted." The Tenth Amendment further limits the federal government by granting the states the power not delegated to the federal government or prohibited to the states. Given the relative brevity of the national document, this granted substantial, non-delineated power to the state governments. The constitutions adopted by the states reflected the vast nature of what was left unwritten in the federal Constitution and largely conceived of state government power as plenary and granted significant general powers to the legislative branch in particular (Tarr 1998; Elazar 1982). Tarr explains,
"...state governments have historically been understood to possess plenary legislative powers - that is, those residual legislative powers not ceded to the national government or prohibited to them by the federal Constitution. As the Kansas Supreme Court has observed: 'When the constitutionality of a statute is involved, the question presented is, therefore, not whether the act is authorized by the constitution, but whether it is prohibited thereby (Tarr 1998: 7).'"

Elazar also writing about the plenary nature of state governments explains that in comparison to the U.S. Constitution, the state constitutions need to be more comprehensive and explicit about limiting and defining the scope of governmental powers to prevent their growth and expansion (Elazar 1982). Thus, these descriptions of state constitutions by Tarr (1998) and Elazar (1982) echo and support Lutz's perspective on the U.S. Constitution as rather "incomplete" in its "dos and don'ts" of the governance of the United States and its citizens. As a result, the individual states really have had no choice but to draft and adopt more detailed directives regarding how society would determine who gets what, when and how as well as establishing the balance of power and justice within that society.

State constitutional scholars largely agree that a constitution represents a peoples' belief of how their society should operate at a distinct point in time. Therefore, we can anticipate that the content of these documents may be heavily influenced by the historic eras within which they have been adopted. Albert Sturm (1982) identifies five periods of state constitutional development:

1. The First State Constitutions (1776-1780)

- Brief documents
- Focused on outlining basic government structures and protections for citizens (e.g. Bill of Rights type items)
- Marked by establishing of strong legislatures with significant plenary powers.

2. Early 19th Century Developments (1800-1860)

- This period marked the rise of Jacksonian Democracy.
- Increase in suffrage rights, but no major provisions of a Bill of Rights nature
- More limitations on legislative power
- Clarifications/further delineation of judicial power
- Increase in executive power
- Provisions for constitutional amendment and revision

3. Civil War, Reconstruction and Its Aftermath (1860-1900)

- Significant impact on Southern states
- Addition of Jim Crow provisions in Southern constitutions
- Provisions that authorized the government to control and regulate economic activity
- Increase in the establishment of agencies not controlled by the executive
- More detailed directives for the legislature, judiciary and executive

4. Beginnings of Reform (1900-1950)

- Pressure to reform state government after massive corruption uncovered
- Extension of popular control of government - rise of initiative and referenda
- Influence of the National Municipal League's Model State Constitution
- Additional limitations on legislatures

5. Constitutional Modernization (1950 forward)

- Revisions to state government needed to meet the expansion of state government functions

Figure 2 empirically illustrates Sturm's evolutionary path of state constitutions. The figure presents a simple count of the number of sections devoted to the seven topics we identified in the previous section. Given the timeframe of our data we will focus on the first three eras identified by Sturm. Sturm describes an evolutionary path in which the earliest state constitutional documents were brief documents that outlined only the basic government structures. Over time, these documents become more complex further delineating the structures of government as well as better defining the power relationships between them. Furthermore, as the need arose more sections were added giving the state government the power to regulate finance and taxation as well provide clearer instructions regarding public policy responsibilities. Figure 2 clearly shows that between 1776 and 1908 the number of sections found in state constitutions increased at a steep rate with increasing number of sections in the areas of finance and taxation as well as public policy post-1850.

## [Figure 2 about here]

Figure 2 only reveals an aggregate picture of this evolution. Could it be the case that the increase in sections is simply an artifact of additional states entering the Union over time? Figure 3 offers two bubble charts. The first is a bubble illustration of the 1819 Alabama Constitution. In this constitution there were just four topics, with all but two sections falling into either the structure of government topic ( 76 sections) or the bill of rights and minorities topic ( 39 sections). One section concerned amending the constitution and the other policy. The sections of the Alabama Constitution of 1901 are spread more broadly across six topics, with a significant increase in the sections concerning policy (from

1 section to 58 sections), suffrage and elections (from 0 sections to 26 sections) and structure of government (from 76 to 138 sections).
[Figure 3 about here]

## Stickiness and Similarity in Constitutions

Students of institutions have often considered the stickiness of institutions - that is once a structure, rule, procedure or informal norm has become adopted it is notoriously difficult to repeal, change, etc. In this section, we will explore stickiness as well as the notion of similarity. For our purposes, we consider a constitutional section as sticky if it endures unchanged over time since its adoption. We are also interested in the similarity of each a section over time and will employ Levenshtein distance scores to determine the extent to which the adoption of new constitutions and amendments modified existing sections.

Figure 4 provides a series of Venn diagrams for the U.S. Constitution and 6 states. The circle to the left represents the text of the initial constitutional document and the circle to the right represents the text of the constitutional document as it existed in 1907.11 The Venn diagrams simply compare to see if a section either: 1) exists only in the initial document (deleted), 2) in both documents (remained) or 3) only in 1907 (added). The Venn diagrams do not take into account any modifications made to sections that endured at both time periods. The point of Figure 4 is merely to illustrate that there appears to be significant variation in the stickiness of sections across states. As expected the U.S. Constitution appears very sticky with $99 \%$ of the sections in 1787 remaining in $1907{ }^{12}$. Minnesota is similar with $96 \%$ of the sections existing in 1857 carrying over to 1907. In contrast, only $47 \%$ of the sections in the 1819 Alabama Constitution still existed in the

[^4]1907 document. In Colorado and Wyoming there were very few deletions or additions if any. Of course, given the time period of this study, this result is not too surprising since Colorado achieved statehood in 1876 and Wyoming in 1889 leaving little time to engage in major constitutional alterations.

## [Figure 4 about here]

Using the Levenshtein distance technique discussed in Appendix B, we can easily compare the text by section of a state's constitution at any two points in time. ${ }^{13}$ In essence, this measure provides us with a similarity score for each section between two time-points. In this analysis, we are comparing the sections of the first constitution with the constitution in 1907. A score of one indicates that the section is identical in 1907 as it was in the first constitution. A score of zero indicates that the section was added or deleted sometime between the adoption of the first constitution and 1907. The range of scores in between represent a relative similarity in text between first constitution and 1907. What these scores do not reveal is the substance of these sections - do they grant a government power or do they mandate or restrict government power? Our hope is that in the future, we will able to combine this analytical technique with our previous work on the substance of these documents to more efficiently and accurately empirically analyze the substance and impact of these changes.

The similarity scores between the first constitution and 1907 in our sub-sample reveals that there is quite a bit of variation in the extent to which the written text of these documents endure over time across states. Table 1 provides a descriptive summary of the percentage of sections that were unchanged, modified, added and deleted in our subsample. In some cases (United States, Colorado, Minnesota and Wyoming), a large percentage of sections endured unchanged from the first constitution though 1907. In others (Alabama, Illinois and New Jersey) very few sections went unchanged. Similar patterns emerge for modifications, additions and deletions. The point is that there is tremendous variation across cases in these behaviors.

## [Table 1 about here]

Figure 5 presents the distribution of similarity scores for six states and the U.S. Constitution. Each dot on the plot represents a section. At the very top of each box-plot is a dark line of dots (in blue) - these are the sections that possess perfect similarity between the first document and 1907. For all but one case, there are quite a few cases where a section endures virtually unchanged over time. The one exception is New Jersey, where not a single section remained intact between 1776 and 1907. This is not too surprising since the 1776 New Jersey Constitution is more similar to a colonial charter in its language than the constitutions of most states upon joining the United States.
[Figure 5 about here]

[^5]Moving to the other extreme in Figure 5, are the lines of dots (in green and yellow) running along the bottom of each box-plot. These dots represent the sections of the constitution that have been added or deleted since the first constitution was adopted. Across our cases, it is not surprising to find that in most instances quite a few sections have been added or deleted to the text as it evolved over time. The one exception is Wyoming, where no new sections were added nor any sections deleted between 1889 and 1907.

Great change in a constitutional text occurs through the modification of existing sections. Over time, structures become more delineated and citizens may decide that the government needs to be mandated to act on its behalf or limited in its action. The middle section of dots (in red) in each box-plot represents the sections that have been modified between the first constitution and 1907. As is evident, there is quite a wide range of modification occurring in these sections. In some cases, the modification to the section is relatively minor with similarity score greater than .75. In others the modification is extreme with a similarity score less than .25. In Alabama, Colorado, Illinois, and Minnesota there is extensive variation in similarity scores across modified sections. The dark red line within the box plot is the median similarity score. The fact that there is a great amount of distance between the median and the upper and lower quartiles as well as fairly long whiskers to the minimum and maximum similarity scores indicate there is substantial variation ripe for future analysis.

A different pattern of modification emerges in each of the three remaining cases. In New Jersey, the box-plot reveals that modified sections of its constitution in 1907 exhibit little similarity to the text of the sections adopted in 1776 with a median similarity score of less than .25 and a very squat quartile box with tiny whiskers. Only 3 sections of the U.S. Constitution were modified with two of the modified sections possessing similarity score of less than .5 and one with a score in excess of .75 . Finally, in Wyoming there was not a single section modified between 1889 and 1907.

We have also produced this same set of box-plot breaking the sections out by constitutional topic. These are presented in Figure 6. What this figure shows is that the sections that were modified dealt largely with government structure and bill of rights and minorities sections. While there were some modifications to sections dealing with the other topics, the change between 1776 and 1907 in areas such as finance and taxation, policy and suffrage and elections was due more to the addition of new sections (see lower left box of Figure 6) than changes in existing sections. The patterns of section addition, deletion and modification exhibited in the table further support Sturm's (1982) account of the eras of state constitutional development.

## [Figure 6 about here]

In this section we have effectively illustrated that there seems to be some evidence that constitutional provisions are somewhat sticky - once included in a document, it is likely that they will remain. However, we have also shown, especially in the states, that
modifications of these provisions will occur. What we cannot illustrate in this paper is whether or not these modifications resulted in significant substantive changes in these sections. Tackling that question will be the focus of a future paper.

## Constitutional Change Over Time and The Choice to Amend or Adopt New Documents

Many scholars have asserted that the manner in which political institutions and processes evolve are by their nature path dependent. That is the initial choices and decisions made regarding those institutions and process start a trajectory in which those early choices influence all future decision-making and changes in those institutions and processes. Thus, any attempt to account for institutional evolution must take this path dependency into account (Pierson 2000a, 2000b; Jervis 2000; Thelen 2000; Bridges 2000). Comparative politics scholars studying democratization and the adoption of other governing institutions as well as scholars studying American political development have long acknowledged the importance of path dependency in their work (e.g., Lipset and Rokkan 1967; North 1990; Schickler 2001; Skocpol 1992; Collier and Collier 1991; Ertman 1996; Hacker 1998). These same authors assert that events - both small and large - that occur as the path is travelled create "forks" in the road that may lead to different institutional changes and evolution depending on the turn taken.

In this section we will discuss how initial procedures for changing a constitutional document may influence how a document is altered over time. Figure 7 is a box-plot of U.S. Constitutional change over the time period of study. The U.S. has only ever had a single constitution and as is evident from the plot during its first roughly 100 years was modified only minimally mostly through the addition of a handful of amendments. Since the initial constitutional convention, there have been no successful attempts to call a convention to propose amendments to the constitution nor have most of the over 10,000 proposed amendments successfully navigated the amendment process.

## [Figure 7 about here]

The state of Alabama provides a stark contrast to our national constitutional experience. Between 1819 and 1901, Alabama adopted six separate constitutions. Figure 8 displays a box-plot of the similarity scores tracking constitutional change over time. The box plots in the gray (shaded) areas represent the distribution of similarity scores when we compare the text of a new a constitution to the text of that constitution in the year just prior to the adoption of a new constitution. The box-plots in the unshaded areas represent a comparison of the text of the old constitution in its last year of existence to the text of the new constitution that succeeded it. We wanted to make this distinction to separate the changes that occur during the amendment process that takes place between the adoption of two constitutions, from those that take place as a result of the adoption of a new constitution. For example, Alabama adopted its first constitution in 1819 and its second in 1861 when it seceded from the United States at the start of the Civil War. The first box-plot
in the gray area conveys the similarity of the sections of the 1819 constitution as adopted, to what the text read in 1860 after it had been amended several times over 41 years. The second box-plot in the white area compares the sections of the 1819 Constitution as it read in 1860 to the text of the new Constitution adopted in 1861.

## [Figure 8 about here]

In Alabama, despite provisions in each of the six constitutions that provided for the amending of the existing document, it is clear major constitutional changes were reserved for when new documents were adopted. In fact, prior to 1907, the only Alabama constitution that was ever amended was the 1819 Constitution. Recall the earlier discussion of path dependency and the impact of small vs. large events. Alabama's decision to rely on the adoption of new documents to alter its constitution is likely driven by its location in the deep-South.

The experience of the Southern states during the Civil War and Reconstruction era illustrates this point beautifully. Prior to their secession from the Union each of these states had adopted constitutions that were similar to those found in the Northern states save provisions that furthered and protected the institution of slavery. Somewhat interestingly, these states did little to alter their constitution as they created the Confederate States of America (CSA) - most of the states simply just amended their existing documents by changing United States of America to Confederate States of America and adopted provisions that further entrenched and protected the institution of slavery. The significant changes did not occur until after the war was lost and the CSA was no more. The price of reentering the Union was the drafting and establishment of a new state constitution that met the specific conditions set by the federal government for reentry. Once the federal government returned governance to the defeated Southerners, they adopted another set of constitutions that institutionalized Jim Crow laws aimed diminishing the civil and political rights of African-Americans. Figure 1 clearly showed that the Southern states in our study had adopted significantly more new constitutions versus states in other regions. This is due in large part to the Civil War experience. Three distinct constitutions were adopted by Southern states during this period in history: 1) Confederate States of America (Civil War) constitution; 2) Reconstruction constitution (needed for readmittance to the United States) and 3) redeemer constitutions (reestablishment of the old order once reconstruction ended). Thus, in this case a large event - the Civil War - created a "fork in the road" that caused these states to choose the patch of adopting new constitutions versus amending an existing one.

Additionally, the events of the Civil War created conditions that led to substantial change in these documents. Let's return to our Alabama example. Figure 9 uses median similarity scores to illustrate the impact of the adoption of a new constitution. In this case we are plotting the median similarity score in the constitution for the time period of 18191907. We are comparing the sections of a document in a given year, to the sections of the document the following year. The line remains flat at one (perfect matching of sections) for most years. The exceptions are 1865 and 1868 (Reconstruction constitutions), 1875 (redeemer constitution) and 1901 where the median similarity score was less than .5
indicating that the sections that were modified exhibited a low degree of similarity from their previous version. In fact, the median similarity score for the 1875 redeemer constitution is zero indicating that many sections were deleted entirely and/or completely new sections were added. The only time a new constitution was adopted in Alabama and the median similarity score remained at one was when the 1861 Civil War constitution was adopted. Histories of the Alabama constitution indicate that the state made minimal changes to the document upon secession (Winkle 2014).

## [Figure 9 about here]

This notion that a large event can lead to significant change also echoes the work Baumgartner and Jones (1993) who developed a theory of punctuated equilibrium to explain policy change. In their work they show that policy does not change incrementally. Rather, policy is stable for a long period of time and then an event occurs that causes the public and elected officials to call for policy change that leads to significant policy alteration, which then leads into another extended period of stability. This brief analysis of the case of Alabama provides some evidence that punctuated equilibrium may have some ability to account for constitutional change under certain historical conditions.

## Discussion

The description of the historical evolution of state constitutions in this paper represents in many ways the view from 50,000 feet. The most general conclusion we can draw is that at the state level there was tremendous change in constitutions between 1776 and 1907. Specifically, we have also shown that number of sections increased and the topics addressed in a constitution increased over time and are correlated with the time period of the adoption of the constitution. In general, as time passes, constitutions become more complex - the document contains more provisions over a wider range of topics.

We also, in very general terms, were able to address the notions of stickiness and similarity. Once a section is included in a constitution, it is likely that it will remain (stickiness). However, there appears to be a wide range in variance regarding the similarity of the text of these sections over time in some states, and at this time we are unable to determine the true impact of the modifications revealed by the similarity scores. In future research we intend to dig deeper into the tension between stickiness and similarity within sections of a constitution over time.

The other contribution of this paper is the technique that we have used to analyze and compare the constitutions in our sample. The techniques employed here have provided leverage over the data we had been unable to previously achieve. Human reading and coding of these texts is time consuming. The use of machine-reading and automated text analytics have allowed us to be able to compare over 350,000 sections of constitutions over a 100-year period. Our hope is that we will be able to further develop these techniques to tackle the coding of constitutional content in the future. At a minimum, the similarity score makes the analysis of change more efficient and less resource demanding
by helping us quickly identify which sections have changed so that we can later assess the depth and meaning of that change.

In a future analysis of state constitutions, we plan to tackle the question of content. What do these documents say specifically regarding governance in a state? Past research we have conducted reveals that the provisions found in state constitutions restrict, exclude, mandate or empower structures such as the legislature. What is the balance between these types of provisions within states overtime as well as across states in different eras? What does this mean for policy-making and the content of that policy?

Additionally, to what extent do states learn and/or borrow from one another? In this analysis we clearly show that states that entered the United States later in the period of study began statehood with more complex constitutions - more provisions in more areas. How much of what exists in one state's constitution can be found in the constitutions of other states? It is likely that as states entered the Union or chose to engage in whole scale revision of an existing constitution, that they consulted the constitutions of other states in an attempt to identify the best practices of the day. Squire (2012), in his historical study of the evolution of American legislatures provides extensive evidence that states modeled their legislatures on already existing legislatures. Therefore, it is likely that particular provisions experience a diffusion into the states over time. With more elaborated measures of text similarity that allow us to compare, not just the literal, but the semantic similarity between two texts, which are currently being developed in the field of Natural Language Processing, we can begin to tackle questions of diffusion on larger scales than ever before.

Finally, the approach and methods we have used in this paper can easily be used to study other phenomena in the states. For example, our method of machine reading and comparing a state's constitution over time as well as constitutions across states can easily be translated to studies of public policies, the development of legislation or a state's statutory code.

## APPENDIX A

To conduct a quantitative analysis of state constitutions the texts of the constitutions must be in a readable format (.txt) for processing, and we must be able to easily search and capture fragments of the text (such as articles or sections) to conduct targeted comparisons. The NBER/Maryland State Constitutions project website has a public repository of constitutions saved in .txt format. Additionally, the coding rules used by the NBER project allow us to retrieve items within a constitution in an automated way. Each of the constitutions in the repository has been coded using the same scheme. Our initial analysis was conducted using 80 constitutions between 1776 and 1907. Out of these 80, NBER had complete information on 55 constitutions, and had recorded amendments for 45 of these. The missing constitutions include: California 1849, California 1879, Georgia 1777, Georgia 1789, Georgia 1798, Georgia 1861, Georgia 1865, Georgia 1868, Georgia 1877, Iowa 1844, Nebraska 1866, Nebraska 1875, New Jersey 1776, New Jersey 1844, Oregon 1857, Tennessee 1796, Tennessee 1835, Tennessee 1870, Virginia 1776, Virginia 1830, Virginia 1851, Virginia 1869, Virginia 1902, Wisconsin 1848, United States 1787.

These missing constitutions were coded by the authors. Amendments were also coded by the authors for these constitutions as well as for the constitutions of Alabama and Delaware, which were not completed by the NBER. These missing constitutions and amendments were mostly retrieved from Francis Newton Thorpe's, The Federal and State Constitutions, Colonial Charters, and the Organic Laws of the State, Territories, and Colonies; Now or heretofore Forming the United States of America (Thorpe). The scans from the volumes were converted into .txt format using ABBYY, an OCR and text scanning software. After the constitutions were converted into .txt format they were coded following NBER coding rules. It is important to note that NBER texts were all revised and cross checked by the authors also using Thorpe. In some cases, complimentary sources were used. These include the Delaware Journal of Corporate Law ${ }^{14}$, the GeorgiaInfo online almanac ${ }^{15}$, Acts and Resolutions of the General Assembly of the State of Georgia for various years ${ }^{16}$, the Code of the State of Georgia by Howard Van Epps 189517, documents from the TNGenWeb Project ${ }^{18}$, documents from the For Virginians: Government Matters website

[^6]owned by The Center for History and New Media (CHNM) at George Mason University ${ }^{19}$ and information from the constitutionsus website ${ }^{20}$.

Once these text were in readable .txt format we used R to transform the text files of these constitutions into a section-level dataset where each unit is a section within a constitution. A dataset was created for every year since the adoption of the first constitution of the state up until 1907. The datasets include the topic and full text of each section. If amendments were made or a new constitution adopted, we replaced the text of each section with the text that resulted from any modification. If sections were deleted, they were dropped for the dataset of that year and if they were added, they were included in the dataset of that year. Our final datasets consist of the sections of each constitution for the 32 states and the U.S. as of each year between the first constitution and 1907.

[^7]
## APPENDIX B

Before we were able to compare sections of two constitutions the text of sections needed to be pre-processed, cleaned and decoded to get an accurate similarity assessment. First we decoded the data so that all of it would be in a standard encoding format. In this case UTF-8, which is widely accepted and recommended format. Second, we removed stopwords, or commonly occurring words in the English language using the tm package ${ }^{21}$. We did not remove stop words that, although common in the English language we considered to be reflective of substantive change such as "may" and "shall". Third, we removed punctuation marks and other non-character symbols. Finally, we standardized words, replacing words that used the British spelling (i.e. "Defence"), with the American spelling so that words that had the same meaning but a different spelling would not drive down the similarity score. We used a variety of sources on the web to get a list of such words ${ }^{22}$.

After the sections texts had been processed we proceeded to assess the similarity between sections of constitutions. For this purpose, we used a modified version of the Levenshtein edit distance. The original Levenshtein algorithm calculates the least number of edit operations (deletions, insertions and substitutions) that are necessary to modify one string to obtain another string. The algorithm assigns a value of 1 to each character insertion, deletion and substitution. The normalized version of the algorithm then divides this minimum number of insertions, deletions and substitutions by the length of the longer string and subtracts the resulting value from 1 . The Levenshtein measure was originally created to compare characters in words, instead of sentences, where a small number of differences is to be expected, and is typically used by spell checker and correction systems. Since we were interested in comparing words within sentences, instead of characters within words we adapted the algorithm to use words instead of characters to calculate the distance and used the standard dynamic programming approach for computing the Levenshtein distance with these modifications. We then divided that distance by the length (amount of words) of the longer sentence and subtracted the resulting value form 1 . The values of the modified Levenshtein distance range from 0 to 1 , with 1 denoting perfect similarity and 0 denoting the lowest degree of similarity.

In the case of amendments knowing which sections to compare with another was straightforward, since they shared a common article, section and subsection (when applicable) number, so sections with the same identifiers (article, section and subsection numbers) were compared. In the case of the adoption of new constitutions this is harder to assess since sections might no longer share the same identifiers as constitutions get renumbered and reorganized. To solve this problem, we compared every possible section of a given topic in the first document with every possible section of that same topic in the

[^8]second document. The result is a matrix of similarity scores between the sections of the same topic. We then calculated the maximum similarity score for each section, working under the assumption that this would increase the probability that a section in the first document would be matched with its true equivalent in the second document. We proceeded to do a visual check of matches, to ensure that each section was matched with an equivalent section in the second constitution. We found that the algorithm was producing a correct match whenever the similarity score was above .25 . Typically, scores of the Levenshtein rarely fall to zero since some common words between sentences, for example, the word "legislator" all always present when you are comparing sections that share topic. What we want is for added or deleted sections to be compared to a blank space so that the Levenshtein distance will take a value of zero. We made sure to revise every case, ensuring that sections that were deleted from the first document, or added from the second document, were not matched with another section, and that each section was matched with its true equivalent. After this process we recalculated the similarity score.

One of our concerns when using the Levenshtein measure to calculate similarity section by section is that sections vary a great deal in the amount of words they contain, and constitutions vary in the average number of words they include in each section. If the distance is sensitive to the number of words, meaning sections with a lower number of words will have lower similarity scores, comparisons across sections and states would not be accurate. To assess whether such bias exists we plotted the similarity scores per pair of section over the total number of words in each section for all of our constitution combinations and there appears to be no bias. Similarity scores are distributed with no discernable pattern across different values of the logged number of words per section.

Finally, it is important to note some of the limitations of this approach to compare similarity of sections. The methods we have discussed to classify and measure the similarity of texts are part of what is known as a statistical approach to Information Retrieval. From a statistical approach, documents can be described by a representative set of keywords (BaezaYates and Ribeiro-Neto 1999) ${ }^{23}$. A problem of this approach, versus what is known as a natural language approach, is word sense ambiguity, or "The Vocabulary Problem" (Furnas et al. 1987) ${ }^{24}$, which refers to situations in which reformers of one constitution do not use exactly the same words as reformers of another constitution, even though the intention of both groups of reformers is the same. They might use synonyms or equivalent phrases. Another side of this problem is when two constitutions use the same terms, so both documents include the same search terms, but they are used with different intentions, or in another sense.

[^9][^10]Natural language texts contain sentences that have a grammatical structure and use vocabulary rich in synonyms. These elements contribute to a textís meaning, which is not considered in statistical information retrieval. Automated information retrieval with a statistical approach will always face these types of limitations, which is why they might not achieve human levels of performance, and are likely to identify some degree of erroneous information, or ignore relevant information.

Because we are analyzing constitutions within the same state and within a limited time period we considered and were able to determine through visual checks that the vocabulary problem was not giving bias to our results. However, moving forward into comparisons within longer time frames and comparison of the texts of different states this problem needs to be addressed. One way of proceeding is to look into natural language approaches that already exist and use methods such as domain specific thesaurally derived lexical chains and text case based reasoning to assess similarity of documents that do not share the same words but have the same underlying meaning.

## References

Baumgartner, Frank R. and Bryan D. Jones. 1993. Agendas and Instability in American Politics. Chicago, IL: University of Chicago Press.

Bridges, Amy. 2000. "Path Dependence, Sequence, History, Theory," Studies in American Political Development. 14: 109-112.

Cayton, Adam. 2015. "Why Are Some Institutions Replaced while Others Persist? Evidence from State Constitutions," State Politics and Policy Quarterly. 16: 1-23.

Collier, Ruth Berins and David Collier. 1991. Shaping the Political Arena: Critical Junctures, the Labor Movement and Regime Dynamics in Latin America. Princeton, NJ: Princeton University Press.

Dinan, John. 2009. The American State Constitutional Tradition. Lawrence, KS: University of Kansas Press.

Elazar, Daniel J. 1982. "The Principles and Traditions Underlying State Constitutions," Publius. 12: 11-25.

Ertman, Thomas. 1996. Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe. Cambridge: Cambridge University Press.

Friedman, Lawrence M. 1988. "State Constitutions in Historical Perspective," Annals of the American Academy of Political and Social Science. 496: 33-42.

Gardner, James A. 2014. "Autonomy and Isomorphism: The Unfulfilled Promise of Structural Autonomy in American State Constitutions," Wayne Law Review. 60: 31-67.

Hacker, Jacob. 1998. "The Historical Logic of National Health Insurance: Structure and Sequence in the Development of British, Canadian, and U.S. Medical Policy." Studies in American Political Development. 12: 57-130.

Hammons, Christopher W. 1999. "Was James Madison Wrong? Rethinking the American Preference for Short, Framework-Oriented Constitutions," American Political Science Review. 93: 837-49.

Jervis, Robert. 2000. "Timing and Interaction in Politics: A Comment on Pierson," Studies in American Political Development. 14: 93-100.

Kincaid, John. 1988. "State Constitutions and the Federal System," Annals of the American Academy of Political and Social Science. 496: 12-22.

Lipset, Seymour Martin and Stein Rokkan. 1967 "Cleavage Structures, Party Systems, and Voter Alignments: An Introduction." In Party Systems and Voter Alignments, ed. Seymour Martin Lipset and Stein Rokkan. New York: Free Press. Pp. 1-64.

Lutz, Donald. 1982. "The Purpose of American Constitutions," Publius: The Journal of Federalism. 12: 27-44.

Martorano, Nancy, Keith E. Hamm, Ronald D. Hedlund. 2009. "Paths, Events and Strategic Choices: The Evolution of American State Legislatures," Paper presented at the annual meeting of the American Political Science Association, Toronto, Ontario, Canada.
. 2010. "State Legislative Evolution: Path Dependency, Organization Theory and Bounded Rational Choice," Paper presented at the 10th annual State Politics and Policy Conference, Springfield, Illinois.
$\qquad$ . 2011. "Limiting the Evolutionary Path: Explaining Constitutional and Statutory Entrenchment in State Legislatures," Paper presented at the annual meeting of the American Political Science Association, Seattle, Washington.
$\qquad$ . 2014a. "With One Hand Tied Behind Their Backs: Constitutional Entrenchment in State Legislatures," Paper presented at the annual meeting of the Midwest Political Science Association, Chicago Illinois and State Political and Policy Conference Indiana University.
_ 2014b. "Constrained Behavior: Understanding the Entrenchment of Legislative Procedure in American State Constitutional Law," Paper presented at the annual meeting of the American Political Science Association, Washington, DC.
. 2015. "Working in a Cage: The Evolution of Constitutional Restrictiveness in U.S. State Legislatures," Paper presented at the annual State Politics and Policy Conference, Sacramento, California.
. 2016. "Powers, Mandates, Restrictions and Exclusions, Oh My...Exploring the Influence of State Constitutions on the Legislative Branch," Paper presented at the annual Midwestern Political Science Association conference, Chicago, Illinois.

Pierson, Paul. 2000a. "Increasing Returns, Path Dependence, and the Study of Politics," American Political Science Review. 94: 251-267.
. 2000b. "Not Just What, but When: Timing and Sequence in Political Processes," Studies in American Political Development. 14: 72-92.

North, Douglass C. 1990. Institutions, Institutional Change and Economic Performance. Cambridge: Cambridge University Press.

Ramsey, Ross. 2015a. "Analysis: Votes Count, but the Rules Can Count More," May 27, 2015. The Texas Tribune.
$\qquad$ . 2015b. "Lawmakers: Bill Slaughter Not a Partisan Deal," May 27, 2015. The Texas Tribune.

Schickler, Eric. 2001. Disjointed Pluralism: Institutional Innovation and the Development of the U.S. Congress. Princeton, NJ: Princeton University Press.

Skocpol, Theda. 1992. Protecting Soldiers and Mothers: The Political Origin of Social Policies in the United States. Cambridge, MA: Belknap Press of Harvard.

Squire, Peverill. 2012. The Evolution of American Legislatures: Colonies, Territories, and States, 1619-2009. Ann Arbor, MI: University of Michigan Press.

Sturm, Albert L. 1982. "The Development of American State Constitutions," Publius. 12: 5798.

Tarr, G. Allan. 1998. Understanding State Constitutions. Princeton, NJ: Princeton University Press.
. 2014. "Explaining State Constitutional Change," The Wayne Law Review. 60 (Spring 2014).

Thelen, Kathleen. 2000. "Timing and Temporality in the Analysis of Institutional Evolution and Change," Studies in American Political Development. 14: 101-108.

Thorpe, Francis Newton. 1909. The Federal and State Constitutions, Colonial Charters, and the Organic Laws of the State, Territories, and Colonies; Now or heretofore Forming the United States of America. Washington, DC: U.S. Congress.

Winkle, John W. 2014. The Mississippi State Constitution, second edition. The Oxford Commentaries on the State Constitutions of the United States, G. Alan Tarr, series editor. Oxford, UK: Oxford University Press.

Table 1. Similarity of Constitutional Sections, First Constitution-1907

| State | $\begin{array}{c}\text { Not Modified } \\ \text { [Similarity=1] }\end{array}$ | $\begin{array}{c}\text { Modified } \\ \text { [Similarity<1 \& >0 ] }\end{array}$ | $\begin{array}{c}\text { Added } \\ \text { [Similarity=0] }\end{array}$ | $\begin{array}{c}\text { Deleted } \\ \text { [Similarity=0] }\end{array}$ |
| :---: | :---: | :---: | :---: | :---: |
| Alabama | $4.07 \%$ | $11.92 \%$ | $65.99 \%$ | $18.02 \%$ |
|  | $[14]$ | $[41]$ | $[227]$ | $[62]$ |$]$| Colorado |
| :---: |

[^11]Figure 1. New Constitutions and Total Amendments per State 1776-1907


Figure 2. Evolution of Constitutional Topics, 1776-1907

FIGURE 3.1 SECTIONS BY TOPIC, ALABAMA 1819-1907


Figure 4. What Remains of First Constitutions?


Figure 5. Similarity of Constitutional Sections, First Constitution - 1907


Figure 6. Similarity of Constitutional Sections by Topic, First Constitution - 1907


## Topic

|  | Bill of Rights and Minorities |
| :---: | :---: |
| ¢ | Constitutional Amendments |
|  | Federal Government |
|  | Finance and Taxation |
|  | Local |
|  | Policy |
|  | Structure of Government |
|  | Suffrage and Elections |

*Includes similarity scores between the first constitution and 1907 of
first constitution and 1907 of
Alabama, Colorado, llinois,
Minnesota, New Jersey, the United
States and Wyoming.

Figure 7. United States Constitutional Changes Over Time

Amendment Type
白 added
deleted白 modified No changes

Figure 8. Alabama Constitutional Changes Over Time


Figure 9. Year by year similarity, Alabama 1819-1907



[^0]:    ${ }^{1}$ Exceptions are Hammons (1999), Gardner (2014) and Cayton (2015) who provide some very abstract quantitative analysis of state constitutions. In the Oxford Commentaries on the State Constitutions of the United States series, each state's constitutional history is given an in-depth book-length treatment.

[^1]:    ${ }^{2}$ These states are Alabama, California, Colorado, Delaware, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Mississippi, Montana, Nebraska, Nevada, New Jersey, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Utah, Virginia, Washington, West Virginia, Wisconsin and Wyoming.
    ${ }^{3}$ http://www.stateconstitutions.umd.edu/index.aspx
    ${ }^{4}$ See Appendix A for information on complimentary sources used.
    ${ }^{5}$ Note that the NBER had complete information for 55 constitutions, and complete amendments for 45 constitutions. The remaining constitutions and amendments were coded by the authors using the NBER coding scheme. For more details, see Appendix A.
    ${ }^{6}$ For states that were not organized using the Article, Section format, we treated each item as distinct section. Sections correspond to the smallest unit of analysis that divides a constitution. In some cases, clauses and subsections were used as sections.
    ${ }^{7}$ Note that the NBER had a complete topic classification for 55 constitutions. The rest of the information was coded by the authors using the NBER topic coding scheme. For more details, see Appendix A.

[^2]:    ${ }^{8}$ See Appendix B for information on the calculation of text similarity.
    ${ }^{9}$ This number is actually slightly higher if the constitutions of southern states forming the Confederacy during the Civil War are included in the tally.

[^3]:    ${ }^{10}$ The secession of these states from the Union accounts for some of these differences in number of constitution adopted that exists between Southern and non-Southern states.

[^4]:    ${ }^{11}$ For example, for the U.S. Constitution, the left circle represents the 1787 constitution as drafted and ratified. The right circle is the text of the U.S. Constitution with all amendments through 1907 included. For Alabama, the left circle represents the 1819 Constitution and the right circle, the text of the 1901 Alabama Constitution with any amendments through 1907. The Venn diagram does not take into account the $1861,1865,1868$ or 1875 constitutions and any amendments to them. The Venn diagrams simply compare the constitution at entry into the United States with the constitution at 1907.
    ${ }^{12}$ Note that for the U.S constitution clauses were treated as sections. Between 1787 and 1907 Amendments 1 through 10 added 10 sections on the topic of Bill of Rights, Amendment 11 added a section on the topic of Judicial politics, and modified an existing section (Article 3 Section 2 Clause 1) on that same topic. Amendment 12 modified an existing section (Article 2 Section 1 Clause 3) on the topic of the executive. Amendment 13 added a section on slavery (which we treated as belonging to the topic of Bill of Rights and minorities). Amendment 14 added 3 sections on citizenship (treated as belonging to the topic of Bill of Rights and minorities), modified a section (Article 1 Section 2 Clause 3), and an additional section was superseded by this amendment (Article 2 Section 1 Clause 3), which was treated as a deletion. Finally, Amendment 15 added two sections on citizenship and suffrage respectively.

[^5]:    ${ }^{13}$ We can also compare the text by section of the constitutions of two different states as well, but that is a topic and analysis that we will tackle in a future paper.

[^6]:    ${ }^{14}$ http://www.djcl.org/wp-content/uploads/2014/07/A-History-of-Delaware-CorporationLaw1.pdf
    15 http://georgiainfo.galileo.usg.edu/topics/government/related_article/constitutions/georgia-constitution-of-1861
    16
    https://play.google.com/store/books/details/Acts_and_Resolutions_of_the_General_Assembly_of_t h?id=deYXAAAAYAAI , https://books.google.com/books?id=-
    RI4AAAAIAAJ\&pg=PA19\&dq=georgia+acts+and+resolutions+1890+1891\&hl=en\&sa=X\&ved=0ahU
    KEwjVgtfJzKHOAhVISCYKHSawBigQ6AEIRzAH\#v=onepage\&q=georgia\%20acts\%20and\%20resolu tions\%201890\%201891\&f=false
    17 http://digitalcommons.law.uga.edu/ga_code/25/
    18 http://tngenweb.org/law/

[^7]:    19 http://vagovernmentmatters.org/primary-sources/518
    20 http://constitutionus.com/

[^8]:    ${ }^{21}$ See the documentation of the package for a full list of stop words https://cran.rproject.org/web/packages/tm/tm.pdf
    ${ }^{22}$ See for example http://www.tysto.com/uk-us-spelling-list.html

[^9]:    ${ }^{23}$ Baeza-Yates and Ribeiro-Neto. 1999. Modern Information Retrieval.
    http://people.ischool.berkeley.edu/~hearst/irbook/print/chap10.pdf

[^10]:    ${ }^{24}$ Furnas, G. W., Landauer, T. K., Gomez, L. M., \& Dumais, S. T. 1987. The vocabulary problem in human-system communications. Communications of the ACM, 30.s

[^11]:    *Number of sections in between brackets

