Joanna Tung
MACS30200
Data/Methods Section Draft
May 17, 2017

DATA
Sources
Annual antigovernment group (a.k.a. "patriot" group) data were obtained from the Southern Poverty Law Center's (SPLC) quarterly publication, *Intelligence Report*. The SPLC is a non-profit organization that publishes an annual record of identified antigovernment groups, characterized only by state and city or county. No membership records are provided.[1] According to the SPLC website, these records are "compiled from field reports, group publications, the Internet, law enforcement sources and news reports."[2] *Intelligence Report* records were accessible online from 2005 to 2016.[3] Links to each of the relevant articles can be found in Appendix I.

County-level demographic and household economic data are taken from American Community Survey (ACS) 1-year estimates, which is collected and maintained by the U.S. Census Bureau. These can be downloaded via FTP, using available APIs, or using the U.S. Census Bureau's American FactFinder tool.[4] Since all U.S. Census Bureau data is identified with a FIPS (Federal Information Processing Standard) county-code, we can use this as a key to create a county-level database that links observed antigovernment group types with demographic and household economic data. Many scholars have argued that county-level assessments are most appropriate for these types of mobilization studies because they permit researchers increased demographic and economic data granularity and increase the number of available observations (when compared to state-level analyses), while simultaneously acknowledging the fact that better modes of transportation have made it increasingly probable for individuals to seek group memberships outside of their immediate city.

Three different databases were used in two passes to label each SPLC antigovernment group record with the corresponding FIPS county-code. The first pass used a single database obtained through the U.S. Small Business Administration API.[5] This database contained city, county, state, and FIPS codes. According to the API webpage, data come from USGS Geographic Names Information System. The second pass used a combination of two datasets. The first contained

---

[1] It is not uncommon for antigovernment groups to obfuscate membership records for privacy and security reasons, certainly, but also to obscure the magnitude and reach of their activities.

[2] For further discussion about the complications of using SPLC data, please see the Literature Review section of the paper. Southern Poverty Law Center. *Active Patriot Groups in the US in 2016.* Retrieved from https://www.splcenter.org/fighting-hate/intelligence-report/2017/active-patriot-groups-us-2016 . Accessed April 23, 2017.

[3] Attempts to obtain digital records for this data spanning back through 1995 or digital records directly from the SPLC were unsuccessful. Our contact at the SPLC provided SPLC's digital records for hate groups from 2000-2015, but indicated that the same was not available for militia groups.

[4] Options for downloading data can be found through the ACS "Data Tables & Tools" webpage. https://www.census.gov/acs/www/data/data-tables-and-tools/index.php.

[5] Code for this is provided in the "Untitled 1.ipynb" document in the repository. U.S. Small Business Administration. "U.S. City & County Web Data API". Retrieved from http://api.sba.gov/doc/geodata.html.

city, county and state data, and was obtained from the Gaslapmedia webpage.[6] The second contained county and FIPS codes, and was obtained from the U.S. Census bureau.[7]

Data, As-Found
SPLC reported antigovernment group data in slightly different formats from year to year, but always in list format. All years, groups were reported by state. Most years, organization name was reported in bold, with relevant chapter location(s) within the state reported in the next line in regular font. Militia status was often indicated by an asterisk by the organization name or the chapter location, but in some years, militancy was indicated within the organization name field instead. SPLC provided the state and annual tally of groups based on the number of ***chapter locations*** observed, rather than by number of unique organization names.

SPLC data were manually copied from each online article and reformatted into a Pandas dataframe using a combination of Excel and Python. A total of 7,836 chapter location records were compiled for the 2007-2016 years.

Data Cleanup
The variation in SPLC's reporting format from year to year and observed inconsistencies (errors) within years immediately signaled concerns about data quality. For most years, however, these mistakes were readily identified and corrected. The 2012 year presented additional difficulties: organization name could not be distinguished by format (bold, font size, capitalizations, etc.) from the chapter location(s). As a result, each record from the 2012 report was manually classified as either organization or chapter location using a combination of records from other years and Google searches, where necessary.

After correcting for obvious errors, however, the total number of antigovernment groups reported annually by the SPLC and those I compiled from the *Intelligence Report* differed for some years. There was an especially high number of unexplained number of discrepancies found in the 2006 year. Time constraints prevented further validation of records with the SPLC. Lacking a logical explanation for the large discrepancy observed, it was determined that both 2005 and 2006 year data should be removed from the future analysis.

For the 2011 and 2015 years, SPLC's *Intelligence Report* provided a larger tally of groups for than were found listed in the *Intelligence Report* records, even after initial cleanup. Fortunately, this discrepancy could be reasonably explained by the number organization names that were erroneously misclassified as chapter locations – recall, SPLC tallies groups by the number of individual chapters, instead of by organization name. In both the 2011 and 2015 years, the number of misclassified organization names corresponded exactly to the "missing" number of group records by year.[8] Based on available evidence, our correction should thus produce a group

---

[6] Gaslap Media. "DOWNLOAD: ZIP CODE LATITUDE LONGITUDE CITY STATE COUNTY CSV." Retrieved from https://www.gaslampmedia.com/download-zip-code-latitude-longitude-city-state-county-csv/.
[7] U.S. Census Bureau. Retrieved from https://www2.census.gov/geo/docs/reference/codes/files/national_county.txt.
[8] For the 2011 year, both the Constitution Party and Freedom Yell Party in Alabama were counted as chapter locations and erroneously included in the state and annual tally. For the 2015 year, the organizations that were erroneously reported as chapter locations were: New Jersey, R.V. Bey Publications; Michigan, We the People, Washington, Washington State Militia.

number count for the 2007-2016 years that is more accurate than the annual tally provided by the SPLC.

Table 1: Comparison of Group Numbers by Years, from the *Intelligence Report*: SPLC Tally versus Number Compiled by Author

Year	Group Number: Compiled by author	Group Number: Tallied by SPLC
2016	623	623
2015	**995**	**998**
2014	874	874
2013	1096	1096
2012	1360	1360
2011	**1272**	**1274**
2010	824	824
2009	512	512
2008	149	149
2007	131	131
2006	**134**	**147**
2005	**131**	**132**

[1]RED = observed discrepancy between SPLC tally as reported in text of article and the full number of groups listed at the end of article. All data come from the *Intelligence Report.*
[2]GRAYED OUT rows = years removed from analysis

SPLC Data: FIPS County-Code Labeling
As discussed above, to label each SPLC antigovernment group record with the correct FIPS county code, I attempted two passes using different datasets. 5,483 of the SPLC records were identically matched to a County and FIPS county code between the two passes. 1,649 records were unassigned due to vague location assignments by the SPLC: "Statewide" (1625 counts) or unlisted (24 counts). The remaining 704 records are currently being assessed. Initial assessment of the remaining 704 records has so far shown that county and FIPS codes were unsuccessfully assigned either because the listed location does not belong to a county, most commonly belonging instead to a borough, or because the listed location summarizes a general region (ex. "Northeast") instead of a specific city or county. At present, the plan is to assign these 704 records to county and FIPS codes using a combination of Google searches and city-borough level datasets.

There is a large proportion of records with location fields containing the term "Statewide." This presents a particularly thorny problem. First and foremost, the label itself indicates that there is no detectable physical locus of activity.[9] Because this study is focusing also on militia activities, which require in-person, physical organizing to facilitate efficient weapon depot access and the practice of military maneuvers, the absence of detectable physical foci of activity reasonably also

---

[9] I conclude this given the fact that where there are multiple chapters of an organization within a state, the SPLC has consistently listed each location, every year in their *Intelligence Report.*

indicates an absence of strong militia activity. Thus, those records with the "Statewide" location assignment are not practically very meaningful for any analysis that attempts to solely classify counties based on a binary of militant antigovernment groups present or not present.

Secondly, because the term is so vague, it is practically difficult to translate the term "Statewide" into a meaningful analysis. An exploration of the SPLC data showed that in five of the years, all 50 states had at least one antigovernment group whose location was listed as "Statewide" by SPLC. This result again calls into question whether these records represent evidence of meaningful – that is to say, provide a significant "signal" – antigovernment or militant group organization.

Thirdly, the variability with which "Statewide" was used from year to year in the SPLC database raises serious questions about the consistency in the definition used. As the Table 2 shows, the term was not used until the 2009 year, and saw a rare drop in the 2014 year. Unless we are able to obtain information from the SPLC clarifying what the usage of this term is intended to imply, this inconsistency of usage seems to suggest that the term itself is rather meaningless to the SPLC as a whole.[10]

Table 2: Comparison of states that contain records with location listed as "Statewide", by Year

Year	Count of unique states with "Statewide" location field	Count of records with "Statewide" location field
2016	50	155
2015	48	102
2014	17	21
2013	50	351
2012	50	330
2011	50	349
2010	50	182
2009	49	135
2008	0	0
2007	0	0

---

[10] The meaningfulness of the term "Statewide" is highly questionable, especially given the fact that the SPLC reported trends showing a huge spike in antigovernment groups in 2008. Indeed, the SPLC and others have made much hay of this fact. Was this a response to the recent recession? Or perhaps was it related to the social conditions that would bring Barack Obama to the presidency in 2008? Because the SPLC is notoriously non-transparent with its data reporting and collecting processes, citing the safety and confidentiality of informants, it is difficult for us to interrogate what all the parameters of the published data truly are. However, this presents an opportunity for leverage: any analysis that chooses to compare the inclusion and exclusion of records with the term "Statewide" might identify how outcomes can vary drastically depending on whether these "Statewide" records are omitted from the analysis or not.

Ultimately, because we are interested in county-level demographic and economic factors and their effects on antigovernment group mobilization, our analysis is probably best suited to observe only those groups that have some semblance of physical rootedness within the physical community, those that are more likely to be affected by its existence within the specific county in question. Broadly speaking, it is not unreasonable to assume that the ability to detect physical foci of activity indicates a greater level of "establishment" for an organization within a community. The establishment of a physical presence reflects the willingness of an organization to be seen by its community, and in turn, mirrors the relative willingness of the community to accommodate the organization's activities. Given these reasons, at present, I have concluded that the analysis would be best served if the "Statewide" or unlisted location records are omitted from the final SPLC antigovernment dataset. Naturally, this alters the parameters of the analysis slightly. Models will account for only those antigovernment groups who have distinguishable physical foci of activities.

Lastly, we must also extend this discussion to those records which list a general region (ex. "Northeast") instead of a specific city or county. The added specificity of these regional location assignments seems to indicate a strong physical regionality of the groups; however, without more information, it is still impossible to design a reasonable and consistent way to assign each record to an applicable county or counties. Present options lead us to believe it would be best to simply omit these records from the final analysis as well. However, a final decision on this will not be made until a more thorough investigation of these organizations (using Google search) is completed.

METHOD
Difficulties with the data collection and cleaning process have delayed any methodological exploration of the SPLC data in conjunction with the U.S. Census Bureau's demographic and economic data. However, given current knowledge about SPLC data quality, I have determined that the analysis may generally proceed along two paths.

One option is to use classification models. I will attempt to use a logistic regression model, decision trees and machine learning approaches to see which method offers to most robust and meaningful classification. As planned, the records with the term "Statewide" will be omitted from the classification, all counties in the United States will be classified as having A) no antigovernment group(s), B) having non-militant antigovernment group(s), and C) having militant antigovernment group(s).[11] There is an alternate classification model that will also be explored; this will use all SPLC records (including those with "Statewide" in the location field), but will instead classify counties as have either A) militant antigovernment groups or B) having no militant antigovernment group.

The second options is to use regression models. In this approach, we will revert to the state-level analysis, allowing us to reintroduce the records with the "Statewide" label as an indicator of "disperse" mobilization activity. In this formulation, we predict the number of militant antigovernment groups within a state, giving records with a "Statewide" location field only half

---

[11] As discussed above, antigovernment groups in this model would only be those which specifically have a distinguishable physical foci of activity, as evidenced by the SPLC's ability to pinpoint a city or county location for the organization.

the count weight (0.5) of those records that list a specific city or county. This option is not preferred, given small number of available observations by year.

The final set of demographic and economic variables from the American Community Survey have not yet been selected for the final analysis. Preliminary analysis will look at the following variables:

- Veteran status (exploring theories that Veterans from the Gulf War – the first fully "volunteer army" war – were more dissatisfied by their service experience and possess the training to organize militant groups)

- Household income, Percent Rural (exploring theories about economic power and status and their impact on antigovernment mobilization)

- Population age, Race, Number of Foreign/Native/Immigrant Status (engaging theories about the social pressures caused by demographic transformations)

Appendix I

Intelligence Report articles containing antigovernment group data, years 2005-2016

2016
Southern Poverty Law Center. *Active Patriot Groups in the US in 2016.* Retrieved from
https://www.splcenter.org/fighting-hate/intelligence-report/2017/active-patriot-groups-us-2016.

2015
Southern Poverty Law Center. *Active Antigovernment Groups in the United States.* Retrieved
from https://www.splcenter.org/active-antigovernment-groups-united-states.

2014
Southern Poverty Law Center. *Active Antigovernment Groups in the United States Since 2014.*
Retrieved from https://www.splcenter.org/sites/default/files/antigovernment-groups-
2014_splc.pdf.

2013
Southern Poverty Law Center. *Active 'Patriot' Groups in the United States in 2013.* Retrieved
from https://www.splcenter.org/sites/default/files/d6_legacy_files/active_patriot_groups.pdf.

2012
Southern Poverty Law Center. *Active 'Patriot' Groups in the United States in 2012.* Retrieved
from https://www.splcenter.org/fighting-hate/intelligence-report/2013/active-
%E2%80%98patriot%E2%80%99-groups-united-states-2012.

2011
Southern Poverty Law Center. *Active 'Patriot' Groups in the United States in 2011.* Retrieved
from https://www.splcenter.org/fighting-hate/intelligence-report/2012/active-
%E2%80%98patriot%E2%80%99-groups-united-states-2011.

2010
Southern Poverty Law Center. *Active 'Patriot' Groups in the United States in 2010.* Retrieved
from https://www.splcenter.org/fighting-hate/intelligence-report/2011/active-
%E2%80%98patriot%E2%80%99-groups-united-states-2010.

2009
Southern Poverty Law Center. *Active 'Patriot' Groups in the United States in 2009.* Retrieved
from https://www.splcenter.org/fighting-hate/intelligence-report/2010/active-patriot-groups-
united-states-2009.

2008
Southern Poverty Law Center. *'Patriot' Groups.* Retrieved from
https://www.splcenter.org/fighting-hate/intelligence-report/2009/patriot-groups.

2007

Southern Poverty Law Center. *'Patriot' Groups.* Retrieved from
https://www.splcenter.org/fighting-hate/intelligence-report/2008/patriot-groups.

2006
Southern Poverty Law Center. *'Patriot' Groups.* Retrieved from
https://www.splcenter.org/fighting-hate/intelligence-report/2007/patriot-groups.

2005
Southern Poverty Law Center. *Active Patriot Groups in 2005.* Retrieved from
https://www.splcenter.org/fighting-hate/intelligence-report/2015/active-patriot-groups-2005.