When Is There a Single Media Agenda?

John Lovett
Frank R. Baumgartner

Department of Political Science
University of North Carolina at Chapel Hill

jllovett@email.unc.edu
frankb@unc.edu

Abstract

When the *New York Times* devotes attention to a certain political issue, do other media sources follow? Given the number of scholars interested in measures of "the media agenda" it is surprising that no one has answered this question. We estimate a model that explains the conditions under which a single national media agenda is likely to be present. We create 90 different keyword searches covering the full range of topics across the Policy Agendas Project and gather counts of numbers of stories per month from 12 national and regional media sources with data going back to 1980 where possible. We estimate a factor analysis for each series to determine whether a single factor emerges. We then estimate a regression model to predict the variance explained by the first factor in each model. The results show the conditions under which it is highly likely that any national source will produce time series results consistent with any other. Key independent variables are the average number of stories, the variance in stories per month, and the presence of any "spike" in the data series. For keyword searches with high average stories per month, a single factor is the rule. However, for searches with low average numbers of stories per month, only those series with obvious spikes such as those associated with the Olympic Games will produce a single series; absent such a spike, low-frequency data are highly idiosyncratic. Our large-scale empirical assessment should provide guidance to scholars assessing the quality of time series data on media coverage of issues.

Paper presented at the annual meetings of the Midwest Political Science Association, Chicago, April 12-14, 2012.

When Is There a Single Media Agenda?

The Question

Many scholars want to know the level of media attention to different policy topics.[1] Since Schattschneider, the issue of public salience has been a stable in studies ranging from lobbying (see for example Kollman 1998, Goldstein 1999, Smith 2000), presidential behavior (see Brody 1991, Edwards, Mitchell, and Welch 1995, Kernell 1997, Edwards and Wood 1999), public opinion (Behr and Iyengar 1985, Zaller 1992, Kellstedt 2003), media agenda-setting and framing (Cohen 1963, McCombs 1972, Gans 1979, Entman 1989, Kosicki 1993, Hamilton 2004) and policy change (Baumgartner and Jones 1993, Soroka 2002, Baumgartner et al. 2008). These are, of course, just a few of the studies that have relied on counts of stories to assess how much media coverage surrounds a particular policy issue. In an important and highly critical article, John Woolley (2000) suggested that the choice of a particular media source as opposed to the use of a range of sources was highly problematic. It is common in the literature to use a single source such as *New York Times*, the *Readers' Guide to Periodical Literature*, or a single consistently published weekly such as *Time Magazine* and to make an argument that the source is reflective of media coverage more broadly. Woolley's point is, not so fast. Our goal here is to find out the conditions under which we observe high correlations among many media outlets in their coverage of a wide range of issues over time. If the correlations are high, then a single source is

[1] We thank Derek Epp and Mel Atkinson for help with data management, Jim Stimson and Elizabeth Coggins for listening and commenting, and UNC undergraduates Maria Sturdivant, Anna Melillo, Bianca Bell, Jazmine Walker, Eric Hall, Nicole Franceschini, Amber Clifford for help in data collection. Tim Jurka wrote a parsing program that was essential to the data management aspect of the project. Amber Boydstun also provided technical assistance. Some of the data used here were originally collected by Frank R. Baumgartner and Bryan D. Jones, with the support of National Science Foundation grant numbers SBR 9320922 and 0111611, and were distributed through the Department of Government at the University of Texas at Austin. Neither NSF nor the original collectors of the data bear any responsibility for the analysis reported here.

more acceptable. If the correlations in coverage are low, then scholars need be more careful in justifying one source rather than another, or to use multiple sources.

Many studies have been done that have attempted to assess the reliability of media sources. Frank Baumgartner and Bryan Jones (1993) devoted an appendix to showing the correlations between coverage of several issues in the *New York Times* and the *Readers' Guide*. Stuart Soroka (2002) compared several different Canadian newspapers in their coverage of particular issues over time. Frank Baumgartner and colleagues (2008) assessed the reliability of their use of the *New York Times* coverage of capital punishment by looking at electronic searches of death penalty coverage in multiple sources and showing a high correlation in amounts of coverage over time and a similar surge in attention to the concept of "innocence" in all newspapers they investigated, including the *Houston Chronicle*. Still, it is unclear from the literature how much difference it makes if one uses this rather than that newspaper, or if one would reach the same general or specific conclusions about the impact of media coverage on a range of dependent variables if one used an index of multiple sources or if one included one or another source in that index. In sum, many scholars rely on counts of media coverage to inform their studies in a wide variety of topic areas, but we have very little systematic evidence about the idiosyncratic v. generalizable nature of the different media sources that might be used.

Empirical Approach

Our approach to this question is very simple. We want to gather a lot of data about media coverage of a very wide range of policy topics (and indeed some non-policy topics as well). We seek to maximize variance on several dimensions: a) time coverage so that we have as long an historical period under review as possible; b) media sources so that we can create the broadest possible index and see if there are common trends across different types of media; and c) policy

topics so that we are not conducting another case study as have been common in the literature but can make some generalizations about media coverage across the full range of issues that might appear on the media agenda.

We explain our data collection process in detail in Appendix A. The essentials are this. First, we developed a set of key-word searches that covers a wide range of policy- and non-policy related topics. We did this by going systematically through the topic codebook of the Policy Agendas Project and identifying topics that could be searched successfully with electronic keywords with several examples coming from each of the major topic domains as defined in the Agendas Project classification system. Our goal was not to replicate the coding done in the Agendas Project, as this is based on reading texts and abstracts and no keyword system could successfully replicate this work. Rather, we used the codebook to guide us in covering an extremely broad range of topics. We have several searches in each of the 19 major topics of the Agendas Project codebook, ensuring that we cover all domains of US politics and international affairs. Further, we use the extended codes of the Agendas Project media database to include some "non-policy" items such as sports, weather, fires and other items unrelated to government and public policy so that we can assess if there are differences in how policy- and non-policy related topics are covered in various media sources. It could well be, for example, that sports or weather coverage is highly localized so no national trends might emerge on such topics, whereas there is greater correspondence when we look at coverage of something like the national unemployment numbers or the inflation rate.

We conducted 90 searches across the full range of topics available, as explained in more detail below. We are not aware of any published work in political science that has used as many

as ten such searches. Our goal here is to provide a broad platform for assessing the characteristics of media coverage series.

Our second goal was to include many media sources. We used Lexis-Nexis Academic Universe through our University library electronic database collection and identified 12 national and regional sources for study. Table 1 shows the television and newspapers searched and the dates they became available through Lexis-Nexis.

(Insert Table 1 about here)

We searched four national newspapers (the *New York Times,* the *Wall Street Journal*, the *Washington Post,* and *USA Today)*, three regional newspapers (the *Houston Chronicle*, the *Philadelphia Inquirer,* and *the St. Louis Post-Dispatch*), and five national television networks (ABC News, CBS News, CNN, Fox News, NBC). We began the searches on the date the source became available (June 1, 1980 for the longest running sources) through June 30, 2010. As the table shows, we have 30 years of data for four sources: ABC News, the *New York Times*, the *Washington Post*, and the *Wall Street Journal.* Four other sources, CBS News, CNN, *USA Today,* and the *St. Louis Post-Dispatch* have data going back to at least 1990. The shortest series in the dataset is Fox News, which begins in December of 1997.

We sought as our third goal to have as much time coverage as possible. As Table 1 indicates we have from 14 to 30 years of data for each of the series. Finally, as described in Appendix A, we aggregate all of our data for each keyword-source combination to get a monthly count. It is then a simple matter to compare the patterns of attention across all the sources available for each topic. The empirical question of interest then is when do the different sources show the same pattern in their coverage over time and when to they follow idiosyncratic or unrelated patterns.

Results

Table 2 shows that the number of "hits" per search varies dramatically across our list of 90 keywords. The number of articles per month (including all sources) ranges from 1.7 to 687, with a median of about 60 stories. Thirty searches lead to fewer than 23 stories per month across all the sources searched. Some of these are relatively trivial but others include such searches as "water pollution," "farm subsidies," "joblessness" and a more complicated set of terms designed to capture racial or ethnic discrimination. As we will see below, a key first distinction in the likelihood that a single national media agenda will emerge from the data is the simple count of stories per month. With very few hits even across many sources, much of what we see is noise.

(Insert Table 2 about here)

A middle group of stories has between 23 and 90 stories per month. Finally, those topics with relatively high numbers of hits per month include such searches as "Medicare" "United Nations," "Federal Reserve," and "world cup."

For each of the 90 searches listed in Table 2, we performed an identical factor analysis to assess the degree to which the various sources vary over time according to a consistent pattern. Figures 1 through 4 give a sense of some of the patterns we observed. Each figure presents the 12 data sources as gray lines and the first factor that resulted from the factor analysis with a dark black line. The factor, or index, is measured on the left scale in each figure and is standardized to have a mean of zero and standard deviation of one. The right axis lists the number of stories per month in each of the individual newspaper or television news sources.

(Insert Figures 1 – 4 about here)

Recall that our question to identify the conditions under which a single national media agenda emerges, and the figures allow us to go some distance in understanding these conditions. Figure 1 shows the results for the acronym "NATO." Indeed we see a single factor which

explains 93 percent of all the variance in the individual series. This is largely due to an extreme spike associated with military action by NATO in 1999, with multiple smaller spikes occurring later in the series. The figure also shows very high average levels of coverage: Table 2 listed this search as 85th of 90 in terms of amount of coverage per month. Figure 2 shows coverage of HMOs. Here again we see a high proportion of variance explained by the first factor, and relatively robust overall levels of coverage. While the individual series might vary somewhat, it is clear that there was much more coverage of HMOs in the 1998-2002 period, less from 2002 through 2005, and that the word had almost disappeared in the period after 2007.

Figure 3 shows a peculiar but revealing pattern: that of cyclical coverage or the importance of statistical spikes in the data series. Beach Volleyball may not be an important element of public policy, but like elections it is visible in the media only at certain periods of the calendar (in this case, associated with the Olympics), and if one wants to know whether this topic is on the agenda or not, it matters very little which particular source one consults. The first factor predicts 81 percent of all the variation in the 12 individual series.

Finally, Figure 4 shows a case where no single factor emerges: Water pollution shows about the same average coverage as beach volleyball (they were ranked 19th and 22nd respectively in Table 2), but it has no peaks or spikes in the data. Absent a galvanizing event or major spike in attention to the topic at any particular moment, the individual news outlets pay more or less attention to the topic over time in a way that is not consistent across outlets. The first factor in our factor analysis explains only 18 percent of the variation.

A Statistical Model of the Emergence of a National Media Agenda

How can we summarize the results of 90 individual factor analyses? We use the proportion of variance explained by the first factor in a series of identical factor analyses as the most

appropriate dependent variable in the analysis that follows. We compute this proportion by dividing the eigenvalue produced by the factor analysis by the total number of sources used. As Figures 1 through 4 made clear, the series differ dramatically in the degree to which they show a common factor, and the proportion explained (or eigenvalue)[2] is a good indicator of this characteristic. Table 3 gives summary statistics of three key variables in the analysis to follow.

(Insert Table 3 about here)

Across our 90 searches, the first factor explains between 12 and 93 percent of the variance, with a mean value of 40 and a median of 36 percent. Two key explanatory variables are the amount of coverage (mean articles) and a variable we refer to as "event spike." This is defined as the difference between the maximum value and the mean, divided by the mean, and the table shows that this variable has a minimum of 0.74 and a maximum of over 20. In the text below we will refer to the statistical power of the first factor as the presence or absence of a single national media agenda. Where that number is high, a single factor explains the vast bulk of variance in media coverage across all media sources over time, so it matters little which individual media source one consults. Where that number is low, no single factor emerges from the data and scholars need beware of any analysis which purports to show that "media coverage" affects some outcome variable. In these cases, media coverage is highly idiosyncratic by source.

Table 4 shows the results of a simple model explaining the presence or absence of a single national media agenda.

(Insert Table 4 about here)

[2] The eigenvalue is the proportion of variance explained multiplied by the number of sources. Since we have the same number of sources in all cases, the two numbers are perfectly correlated in a linear manner.

The first model shows that we can explain the emergence of a single national media agenda with two very simple variables: the mean number of stories and our measure of spikes in attention associated with particular events. Logging both variables because they have declining marginal effects at very high levels, the two variables explain 89 percent of the variance in the power of the first factor.

Model 2 in Table 4 is a robustness test to see if these patterns are different for those series with different levels of news coverage. We divide the series roughly into thirds based on the number of stories per month and do the same analysis with the high-coverage set as the reference category. With the same two independent variables as well as dummy variables for low and medium coverage, we see no effect of the dummies and no difference in the coefficients for the others. (We also included a measure of variance rather than mean articles and find similarly robust results. There is too much collinearity between event spike, mean articles, and variance in articles per month to include all three in the same model however.) Similarly when we run separate regressions for the low, medium, and high groups from Table 2 on the number of articles per month, we find that the event spike and mean articles have similar coefficients as in Table 4 and that the variance explained ranges from .79 to .87. We conclude from this that we have highly robust results with a two-variable model. Low numbers of stories per month and low variability in the number of articles per month will produce low probabilities of a single factor emerging whereas a high number and a high variance in the number of articles per month will lead to a high probability that a single national media agenda exists.

If we choose an arbitrary threshold of 50 percent variance explained by the first factor in order to identify a single national agenda, we can show the likelihood of this occurrence by looking at the mean number of articles per month. Table 5 shows these results.

(Insert Table 5 about here)

Table 5 shows that for the 33 cases with fewer than 30 articles per month, a single factor emerged only 3 times. Among the 31 cases with more than 90 stories per month, 16 showed a single factor.

Variation across Media Sources

Just as we can analyze our 90 keyword searches to see which cases see a single agenda and which see idiosyncratic patterns, we can look at our 12 media sources to see which ones are highly correlated with the others and which show the most idiosyncratic patterns. Figure 5 shows the average factor loadings for each of the media sources across the 90 keyword searches we conducted.

(Insert Figure 5 about here)

Across the 12 sources, CNN, the *New York Times*, and the *Washington Post* tend to be most central to a national media agenda. The *Wall Street Journal* is the most distinctive of all, with significantly lower average loadings.[3] The *Philadelphia Inquirer* and Fox News are also quite low, as the figure shows.

A Core Set of National Sources

Based on the previous section, we limit our analysis in this section to seven sources, rather than 12; these might be considered a core set of national media sources. Figure 6 shows the seven sources chosen as well as their average loadings on 90 keyword searches after eliminating the five regional or idiosyncratic sources from the previous analysis.

(Insert Figure 6 about here)

[3] Note that Lexis-Nexis makes full text available for the sources used, except for the *Wall Street Journal*, in which case only the abstracts are included in the Lexis search. We cannot determine if the *Journal* follows a different agenda with for example more business news or if the difference observed is due to the different amount of text available for the searches.

Eliminating the *Wall Street Journal*, Fox News, and the three regional newspapers (Houston, St. Louis, and Philadelphia), we see that the seven remaining sources have relatively similar factor scores, indicating that none is particularly more central to the system than any other. We can then replicate some of the analysis done in the first empirical section by limiting ourselves to these seven sources. Table 6 replicates the analysis from Table 4, showing the predictors of the presence of a single national media system.

(Insert Table 6 about here)

As in Table 4, the results of the first model in Table 6 are both impressive in variance explained (again 89 percent of the variance explained, with just two independent variables) and in their robustness to alternative specifications. High numbers of articles and high variance in this value explain the presence of a national agenda. Table 7 replicates Table 5 for the national core dataset and shows that 90 percent of those series with more than 90 stories per month exhibit a single national agenda based on a first factor explaining more than 50 percent of the variance in total attention.

(Insert Table 7 about here)

Table 8 summarizes our results relating to which sources are at the center of the national media system, when we can speak about such a thing. Rather than comparing across the keyword topics, Table 8 looks at the individual sources. We observed a single factor across the 7 sources in 41 of 90 cases. In these 41 cases, CNN showed a factor loading above .70 on this first factor 37 times; it was above .80 28 times; and above .90 on this factor 14 times. Similarly for the other sources, the table shows that CNN, the *New York Times*, and NBC were virtually identical in their high loadings on the observed statistical factor.

(Insert Table 8 about here)

Table 9 extends the analysis from Table 7 and shows the predicted value of the strength of the first factor when we manipulate the mean number of articles as well as the spike variable. Looking at the first row, if we had relatively low coverage (23 articles per month) and no particularly strong spike in coverage (maximum = 67 stories), then the first factor is predicted to explain 34 percent of the variance in media coverage. Moving to high average coverage (90 articles) with a strong peak in the data (611 articles) generates an expected value of 72 percent variance explained.

(Insert Table 9 about here)

We can say based on this that a single national media agenda is likely to exist if there is high coverage, high variance in the coverage over time, and if we look at the core data sources listed in Table 8 rather than regional newspapers, the *Wall Street Journal*, or Fox News.

Discussion

We have provided a simple but empirically the most extensive study to date of the conditions under which a disparate set of electronic and print media cover the same news stories. Using our new and expansive dataset covering 90 policy and non-policy topic areas across 12 sources over as much as 30 years, we show that the presence or absence of a single national media agenda can be explained by two main factors: how much coverage a topic receives, and whether an event producing a spike in attention occurs. Roughly 90 percent of the variation in the explanatory power of a national media factor score can be explained by these two explanatory variables.

We also find that there are variations across sources with how they load onto the primary media factor. National sources such as the *New York Times* and CNN perform particularly well on the first media factor, with others like the *Wall Street Journal* and the *Philadelphia Inquirer* showing more variation relative to the primary factor. Testing the national core set using the

models from testing the entire series, we find similar effects on the proportion of variance from the mean and event spike. Further tests show that the *New York Times*, CNN, and NBC have particularly high loadings across cases with a single media factor.

There is no right answer about whether an analyst need or need not include many media sources in measuring the media agenda. It depends on the theoretical question being investigated, of course. But rather than give a blanket statement about whether one must or must not include many sources, it is more useful to think of the question in a different way. There are clear indications that for cases with high average stories per month and with high variation in how many stories appear each month, virtually any source will show similar patterns to the others. Therefore it may be a waste of resources to construct a complicated media index in these cases. (Of course, depending on how the data are collected, the additional resources for a broader rather than a narrower data collection project may be trivial so there is no reason to limit oneself to a single source.) For cases with low average coverage, and not spikes in attention related to focusing events, on the other hand, it may be misleading even to use the concept of "media agenda." In these cases, no single factor emerges and what we see in the data appears to be completely idiosyncratic patterns across the different sources. Some of these topics are relatively surprising as one might expect that something like "water pollution" would be an important enough topic that a general media factor would be clear. In those cases with no single media agenda, we do not believe that an index is any more useful than a particular source. Rather, the data coming from any media search, with low counts, are unreliable and should not be used as an indicator of media concern with the issue. Sometimes it is important to admit that a measurement strategy has simply failed. The appropriate response in these cases is to ensure

that the keyword searches generate enough hits to rise above the random background noise that affects all such searches.

References

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

Baumgartner, Frank R., Suzanna L. De Boef and Amber E. Boydstun. 2008. *The Decline of the Death Penalty and the Discovery of Innocence.* New York: Cambridge University Press.

Behr, Roy L. and Shanto Iyengar. 1985. Television News, Real-World Cues, and Changes in the Public Agenda. *Public Opinion Quarterly.* 49(1): 38-57.

Brody, Richard A. 1991. *Assessing the President: The Media, Elite Opinion, and Polls.* Stanford: Stanford University Press.

Cohen, Bernard C. 1963. *The Press and Foreign Policy*. Princeton: Princeton University Press.

Edwards, George C., and B. Dan Wood. 1999. Who Influences Whom? The President, Congress, and the Media. *American Political Science Review*. 93(2): 327-344.

Edwards, George C., III, William Mitchell, and Reed Welch. 1995. Explaining Presidential Approval: The Significance of Issue Salience. *American Journal of Political Science.* 39:108-134.

Entman, Robert M. 1989. How the Media Affect What People Think: An Information Processing Approach. *Journal of Politics*. 51(2): 347-370.

Gans, Herbert J. 1979. *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. Chicago: Northwestern University Press.

Goldstein, Kenneth M. 1999. *Interest Groups, Lobbying, and Participation in America*. New York: Cambridge University Press.

Hamilton, James T. 2006. *All the News That's Fit to Sell: How the Market Transforms Information Into News*. Princeton: Princeton University Press.

Kellstedt, Paul M. 2003. *The Mass Media and the Dynamics of American Racial Attitudes.* New York: Cambridge University Press.

Kernell, Samuel. 1997. *Going Public: New Strategies of Presidential Leadership*. 3rd ed. Washington, DC: CQ Press.

Kollman, Ken. 1998. *Outside Lobbying: Public Opinion and Interest Group Strategies*. Princeton: Princeton University Press.

Kosicki, Gerald M. 1993. Problems and Opportunities in Agenda-Setting Research. *Journal of Communication.* 43:100-127.

McCombs, Maxwell E. and Donald L. Shaw. 1972. The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly*. 36(2): 176-187.

Smith, Mark A. 2000. *American Business and Political Power: Public Opinion, Elections, and Democracy*. Chicago: University of Chicago Press.

Soroka, Stuart N. 2002. *Agenda-Setting Dynamics in Canada.* Vancouver: University of British Columbia Press.

Woolley, John T. 2000. Using Media-Based Data in Studies of Politics. *American Journal of Political Science*. 44(1): 156-173.

Zaller, John. 1992. *The Nature and Origin of Mass Opinion.* New York: Cambridge University Press

Appendix A. Constructing our Keyword Database

We sought to create a large list of keywords that would cover the full range of topics related to public policy and some that were unrelated to government. We used the topic codebook of the policy agendas project (www.policyagendas.org) to provide a list of topics. Within each major topic we sought more particular subtopics that could reasonably be searched with keywords. However, we sought to avoid false hits more than we sought to cover the full range of issues included in that subtopic. Therefore, the keyword searches listed here should not be understood to replicate the associated agendas project subtopic; they are almost always much more narrow in scope. To ensure clean searches, we followed an iterative process. We first ran a set of simple search terms and read the results. If there were many false hits we revised the search terms in order to eliminate them while still producing a clean set of hits related to the topic of interest. We then ran the revised search and read those results. We continued this process until fewer than 15% of the articles sampled were false hits. In cases where we could not achieve this level of accuracy, we abandoned that search. So our 90 searches can all be understood has having been vetted until they produced at least 85 percent positive hits. We chose this threshold because of experience suggesting that it is sufficient to produce reliable counts over time; ultimately we are interested in how many stories appear in each month. Table 2 lists the terms and associated agendas project subtopic codes.

We hired students to conduct the searches for each of the keywords listed in Table 2 through Lexis. This generated a large number of full-text files consisting of the text of the story as well as a variety of identifying variables such as the date, source, byline, and embedded keywords. These text files were then run through a data parser to convert he text-based results into a database that we could more easily analyze. For this paper, we focus on just two pieces of

information: the date and the source. We then used Stata's "collapse" command to create counts of numbers of stories for each keyword and each source by month, or 361 observations for the longest series, as indicated in Table 1.

With a database that listed the number of hits in each of our 12 sources for each month available from 1980 through 2010, we then used Stata's "factor" command to extract the factors present for each of the 90 monthly series; note that in these analyses each source is a variable, and we ran them separately for each of the 90 keyword searches. The results of the factor analysis (e.g, the eigenvalue associated with the first extracted factor) then became a variable in another dataset with 90 observations, used for the regressions reported in this paper.

We performed a variety of robustness checks on our factor analyses designed to test potential variations due to time and sources, including testing for variations when adding in or leaving out CNN or the *Wall Street Journal*, to see whether removing or adding in the two would result in different factor loadings (CNN because of its larger article space and ability to update throughout the day, the *Wall Street Journal* because of its specific business focus and that Lexis-Nexis only provides abstracts for the articles.). The tests of variation removing and adding both to the factor analysis showed no significant or systematic differences in the pattern of the data when using the two sources in factor analysis.

We finally settled on two series of factor analysis to use for data analysis. The first is the entire dataset of 12 sources run together as one factor analysis. The second is a national core dataset of the 7 sources that loaded the most frequently and generally on the first factor. These sources are ABC News, CBS News, CNN, NBC, *The New York Times*, *The Washington Post*, and *USA Today*.

Tables and Figures

Table 1: Media Sources and Dates Available for Search.

Source Name	Start Date	Months Available
Fox News	November 26, 1997	151
NBC	January 1, 1997	162
The Philadelphia Inquirer	January 1, 1994	197
The Houston Chronicle	September 15, 1991	225
CBS	February 1, 1990	244
CNN	January 1, 1990	245
USA Today	January 3, 1989	257
The St. Louis Post-Dispatch	January 1, 1989	257
ABC	June 1, 1980	361
The Wall Street Journal	June 1, 1980	361
The New York Times	June 1, 1980	361
The Washington Post	June 1, 1980	361

Note: Searches were conducted from the first date available through June 30, 2010, which determines the number of months available listed in the last column. All searches were conducted through Lexis-Nexis.

Table 2: Search Terms and Number of Articles Per Month.

Agendas Project Code	Description	Search Term	Mean Articles	Rank
1302	Poverty and Assistance for Low-Income Families	BODY("Aid to Families with Dependent Children" OR AFDC OR "Temporary Assistance for Needy Families")	1.7	1
205	Handicap or Disease Discrimination	BODY(handicap! OR disabled OR "chronic illness" w/5 "discrimination" AND (U.S. OR United States OR America!))	2.2	2
505	Fair labor standards	BODY("minimum wage" w/5 (rate OR regulation))	2.2	3
2701	Fires	BODY("forest fire" w/ 15 (damage OR disaster OR fight))	2.6	4
1900	General State Department	BODY("US Department of State")	3	5
401	Food Inspection and Safety(inspection of domestic and imported food)	BODY(food w/5 (inspection OR regulation) w/15 (U.S. OR United States OR America!))	4.4	6
2105	U.S. Dependencies and Territorial Issues	BODY("Puerto Rico" OR Guam OR "American Somoa" w/10 (statehood OR policy!))	4.6	7
604	Vocational Education	BODY(vocational w/5 education AND (U.S. OR United States OR America!))	5.7	8
2102	Native American Affairs	BODY("Native American" OR "American Indian" w/5 (education OR court! OR health OR casino))	7.2	9
108	Industrial policy	BODY(("industry" or "industrial") w/15 "productivity"))	7.5	10
1707	Broadcast Industry Regulation (TV, Cable, Radio)	BODY("Federal Communication Commission")	8.2	11
402	Farm Subsidies	BODY((farm! OR agriculture) w/5 (subsidy! OR "price support")) w/15 (U.S. OR United States OR America!))	8.9	12
1408	Elderly and Handicapped Housing	BODY(housing w/5 (elderly OR handicapped))	9.1	13

207	Freedom of Speech & Religion	BODY("school prayer")	10.5	14
2801	Arts and Entertainment	BODY("Opera" w/5 (opening))	11.1	15
204	Age discrimination	BODY(age w/5 "discrimination" AND (U.S. OR United States OR America!))	11.2	16
802	Electricity and Hydroelectricity	BODY(hydroelectric AND (U.S. OR United States))	11.4	17
405	Animal and Crop Disease and Pest Control	BODY("pest control")	12.3	18
701	Drinking Water Safety	BODY("water pollution")	12.3	19
1205	Prisons	BODY(prison! w/5 (construction OR contracts OR overcrowding)	12.6	20
103	Unemployment Rate	BODY(("unemployment rate" OR "joblessness" OR "jobless rate") w/5 (United States OR U.S. OR America!))	12.9	21
2904	Sports and Recreation	BODY("beach volleyball")	15.7	22
201	Ethnic Group Discrimination	BODY(("race" OR "ethnicity" OR "ethnic group") w/5 "discrimination") AND (U.S. OR United States OR America!))	16	23
504	Employee Relations and Labor Unions	BODY(NLRB)	16.3	24
2100	General Interior	BODY("Department of Interior")	16.9	25
206	Voting Rights and Issues	BODY("Voting Rights Act")	21	26
508	Parental Leave and Child Care	BODY(leave w/5 (paternity OR maternity OR parental))	21	27
1915	Panama Canal Issues and Other International Canal Issues	BODY("Panama Canal")	21.2	28
501	Workplace Safety	BODY(OSHA)	22.2	29
1708	Weather Forecasting and Related Issues, NOAA, Oceanography	BODY(NOAA)	22.9	30
603	Education of	BODY(education AND ("Head Start"	23.3	31

	Underprivileged Students	OR NHSA))		
2603	Weather and Natural Disasters	BODY("drought" w/15 ("rain"))	26	32
1798	Research and Development	BODY("National Science Foundation")	27.2	33
2702	Fires	BODY("house fire")	28.4	34
1203	Illegal Drug Production, Trafficking, and Control	BODY("Drug Enforcement Administration")	32.7	35
1605	Arms Control and Nuclear Nonproliferation	BODY(nonproliferation OR "Arms Control and Disarmament Agency")	33.2	36
202	Gender and sexual orientation discrimination	BODY(("women" OR "sex" OR "sexual orientation") w/5 "discrimination") AND (U.S. OR United States OR America!))	36.5	37
704	Hazardous Waste	BODY(waste w/5 (hazardous OR toxic))	38.2	38
705	Air polution	BODY("air pollution" AND (U.S. OR United States OR America!))	38.5	39
2804	Arts and Entertainment	BODY("book review")	39.2	40
1403	Urban Economic Development and General Urban Issues	BODY(urban w/5 (revitalization OR renewal OR sprawl OR "economic development"))	41.5	41
1006	Highway Construction, Maintenance, and Safety, Auto Safety	BODY("National Highway Traffic Safety Administration" OR NHTSA)	42.4	42
1808	Exchange Rates and Related Issues	BODY("exchange rate")	43.5	43
1802	Trade Negotiations, Disputes, and Agreements	BODY(NAFTA OR GATT OR "most favored nation")	45.1	44
1301	Food Stamps, Food Assistance, and Nutrition Monitoring Programs	BODY("Food stamps" OR "WIC")	56.4	45
2802	Arts and	BODY("Brad Pitt")	61.9	46

	Entertainment			
1001	Mass Transit	BODY("mass transit")	63.7	47
343	Controlled and Illegal Drug Abuse, Treatment, and Education	BODY("drug abuse")	64.1	48
335	Prescription Drug Costs and coverage	BODY("prescription drug!" w/5 (cost! OR price! OR coverage))	66	49
709	Species and Forest Protection	BODY("endangered species")	66.3	50
1609	VA Issues	BODY("veterans affairs")	68.1	51
1807	Tariff and Import Restrictions, Import Regulation	BODY(tariff! or "import quota!")	71.7	52
1005	Railroad Transportation and Safety	BODY(AMTRAK)	77.1	53
1401	Housing and Community Development	BODY(HUD)	77.4	54
1523	Domestic Disaster Relief	BODY("Federal Emergency Management Agency")	79.6	55
1500	General Commerce	BODY("Department of Commerce" OR "National Bureau of Standards")	84.7	56
601	Higher Education	BODY(student! w/5 (loan! OR grant!))	86.4	57
1800	General Foreign Trade	BODY("Federal Trade Commission")	87	58
400	General Agriculture	BODY("Department of Agriculture" OR DOA)	88.6	59
302	HMOs	BODY("HMO")	89.6	60
806	Alternative and Renewable Energy	BODY(energy w/5 (solar OR wind OR alternative))	90.1	61
1202	White Collar Crime and Organized Crime	BODY(crime w/5 ("white collar" OR organized))	94.6	62
2903	Sports and Recreation	BODY("LPGA")	96.3	63
606	Special Education	BODY("special education")	98.7	64
1003	Airports, Airlines, Air Traffic Control	BODY("Federal Aviation Administration" OR "Civil Aeronautics Board")	102.1	65

	and Safety			
2013	Census	BODY(census w/5 (bureau or U.S.))	108.4	66
1208	Family Issues	BODY("domestic violence")	111.2	67
2601	Weather and Natural Disasters	BODY("hurricane" w/15 (damage OR disaster))	112.3	68
1201	Executive Branch Agencies Dealing With Law and Crime	BODY(DOJ OR FBI)	118.4	69
2003	Postal Service Issues (Including Mail Fraud)	BODY(“postal service’’)	119.3	70
301	Comprehensive health care reform	BODY(reform w/5 ("health care" OR "Medicare")	132.1	71
707	Recycling	BODY(recycl!)	136.2	72
2803	Arts and Entertainment	BODY("Michael Jackson")	145.1	73
1207	Child Abuse and Child Pornography	BODY("abuse" w/5 (child!))	165.3	74
1906	International Finance and Economic Development	BODY(IMF OR "world bank")	180.7	75
333	Mental Health	BODY("mental!" w/5 (health OR retardation OR ill) AND (U.S. OR United States OR America!))	192.4	76
1409	Housing Assistance for Homeless and Homeless Issues	BODY(homeless! AND (U.S. or United States))	210.6	77
2009	IRS Administration	BODY(IRS)	219.4	78
2101	National Parks, Memorials, Historic Sites, and Recreation	BODY("national park" or "national memorial")	250.2	79
2902	Sports and Recreation	BODY("world cup")	254.5	80
2602	Weather and Natural Disasters	BODY("earthquake")	257.1	81
803	Natural gas	BODY("natural gas")	278.4	82
1502	Securities and Commodities Regulation	BODY("Securities Exchange Commission")	279.6	83

1701	NASA , U.S. Government Use of Space, Space Exploration Agreements	BODY(NASA)	301.6	84
1602	U.S. and Other Defense Alliances, U.S Security Assistance	BODY(NATO)	329.1	85
1603	Military Intelligence, CIA, Espionage	BODY(CIA)	438.6	86
104	Monetary Supply	BODY("fiscal policy" OR "monetary policy" OR "Federal Reserve")	465.5	87
1303	Elderly Issues and Elderly Assistance Programs (Including Social Security)	BODY("social security" OR SSA)	468.9	88
1507	Bankruptcy	BODY(Bankruptcy)	528.0	89
1926	International Organizations other than Finance: United Nations	BODY("United Nations")	687.7	90

Note: Agendas codes were used to ensure that we covered the full range of possible topics. The keywords do not, however, cover the full range of issues associated with each sub-topic listed.

Table 3: Summary Statistics

Variables	Mean	Median	Minimum	Maximum
Factor 1 Variance Explained	0.40	0.36	0.12	0.93
Mean Articles	100.90	62.83	1.71	687.74
Event Spike	4.16	2.70	0.74	21.42

Note: Table refers to all 12 media sources. The "Event Spike" variable is defined as the difference between the maximum and the mean, divided by the mean.

Table 4: Regression on Proportion of Variance Explained by First Factor, 12 Sources

	Model 1 Coef. (SE)	Model 2 Coef. (SE)
Event Spike (logged)	0.20 (0.01)*	0.20 (0.01)*
Mean Articles (logged)	0.11 (0.01)*	0.11 (0.01)*
Low Articles	-	-0.01 (0.04)
Medium Articles	-	-0.01 (0.03)
Intercept	-0.24 (0.03)*	-0.20 (0.04)*
R^2	0.89	0.89
N	90	90
F	294.6 (0.00)	150.4 (0.00)

* = p<0.05 (Two-Tailed Test). White's Robust Standard Errors used.
Low: <30 articles per month; Medium: 30-90 articles per month; High: >90 articles per month

Table 5: Comparison of Number of Factors by Media Coverage (Single Factor = Proportion>.50)

	1 Factor N (%)	2+ Factors N (%)	Total N (%)
Low	3 (9.1%)	30 (90.9%)	33 (100%)
Medium	8 (30.8%)	18 (69.2%)	26 (100%)
High	16 (51.6%)	15 (48.4%)	31 (100%)
Total	27 (30.0%)	63 (70.0%)	90

Chi^2 = 13.77, p=0.001
Chi^2 (Medium: High) =2.52, p=0.112
Low: <30 articles per month; Medium: 30-90 articles per month; High: >90 articles per month

Table 6: Regression on Proportion of Variance Explained by First Factor, National Core Set

	Model 1 Coef (SE)	Model 2 Coef (SE)
Event Spike (logged)	0.19 (0.01)*	0.19 (0.01)*
Mean Articles (logged)	0.12 (0.00)*	0.11 (0.01)*
Low Articles	-	-0.06 (0.05)
Medium Articles	-	-0.06 (0.02)
Intercept	-0.16 (0.02)*	-0.06 (0.05)
R^2	0.89	0.90
N	90	90
F	382.5 (0.00)	217.6 (0.00)

* = p<0.05 (Two-Tailed Test). White's Robust Standard Errors used.
Low: <30 articles per month; Medium: 30-90 articles per month; High: >90 articles per month

Table 7. Comparison of Number of Factors by Media Coverage, National Core Set (Single Factor = Proportion>.50)

	1 Factor N (%)	2+ Factors N (%)	Total N (%)
Low	9 (21.4%)	33 (78.6%)	42 (100%)
Medium	15 (51.7%)	14 (48.3%)	29 (100%)
High	17 (89.5%)	2 (10.5%)	19 (100.0%)
Total	49 (54.4%)	41 (45.6%)	90

Chi^2 = 25.08, p=0.000
Chi^2 (Medium:High) =7.36, p=0.007
Low: <30 articles per month; Medium: 30-90 articles per month; High: >90 articles per month

Table 8: Number of Times Sources Load on First Factor in 1 Observed Factor Cases

	Number of loadings above this threshold:		
	.70	.80	.90
CNN	37	28	14
The New York Times	37	27	14
NBC	37	26	13
ABC	32	25	11
CBS	37	23	10
The Washington Post	31	24	7
USA Today	32	22	5

Note: A single first factor emerged in 41 cases. The table shows the number of times each source loaded above the threshold indicated. For example, the first column shows that CNN, NYT, and NBC had loadings above .7 on the first factor 37 out of 41 times. They loaded above 0.9 on this factor 13 or 14 times

Table 9. Predicted Strength of a National Media Factor

Scenario	Mean Number of Articles	Maximum articles per month	Predicted Value, Strength of First Factor
Low-low	23	67	.34
Low-medium	23	98	.45
Low-high	23	156	.56
Medium-medium	60	256	.56
High-low	90	263	.51
High-medium	90	383	.61
High-high	90	611	.72

Note: Average articles at 23, 60, and 90 represent approximately the 33rd, 50th, and 67th percentile on the observed number of articles per month. The maximum number is calculated to generate a value of our "event spike" variable representing the 25th, 50th, or 75th percentile respectively. Predicted values are calculated with Model 1 from Table 6. Recall that the mean and maximum articles per month are based on all 7 sources in the core set, so if this model were used as a guide to the number of stories in any single source the number needed would be much lower.

Figure 1: High Media Coverage, One Retained Factor (NATO)

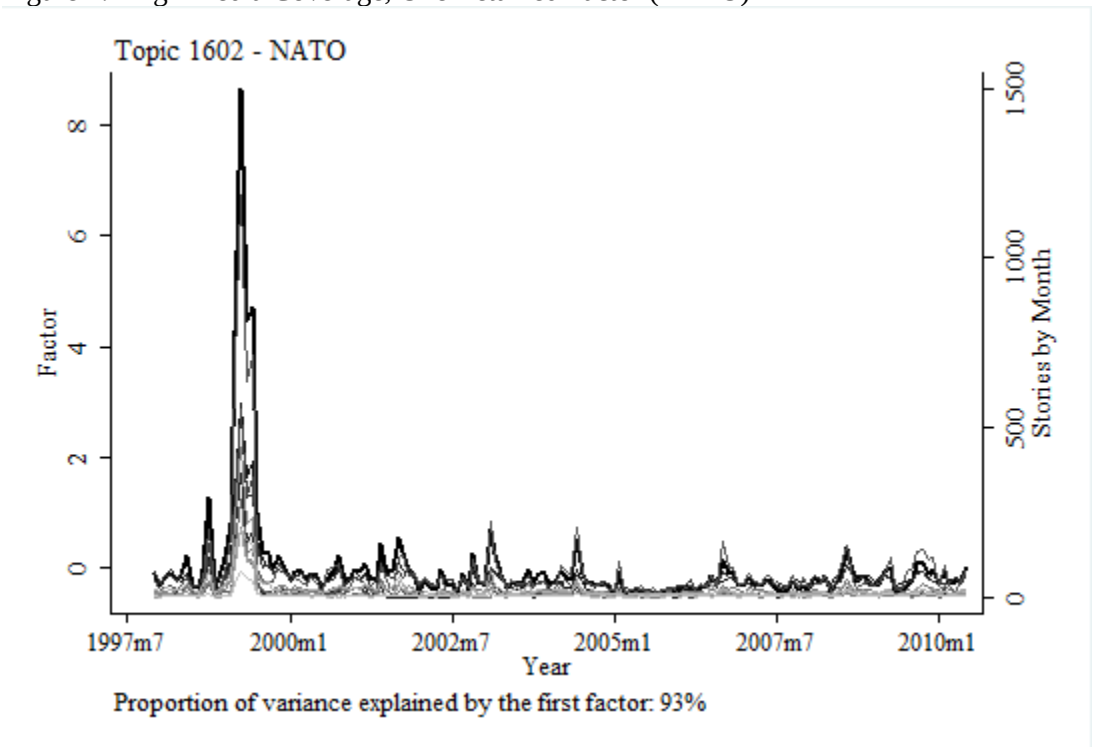


Figure 2: Medium Media Coverage, One Retained Factor (HMOs)

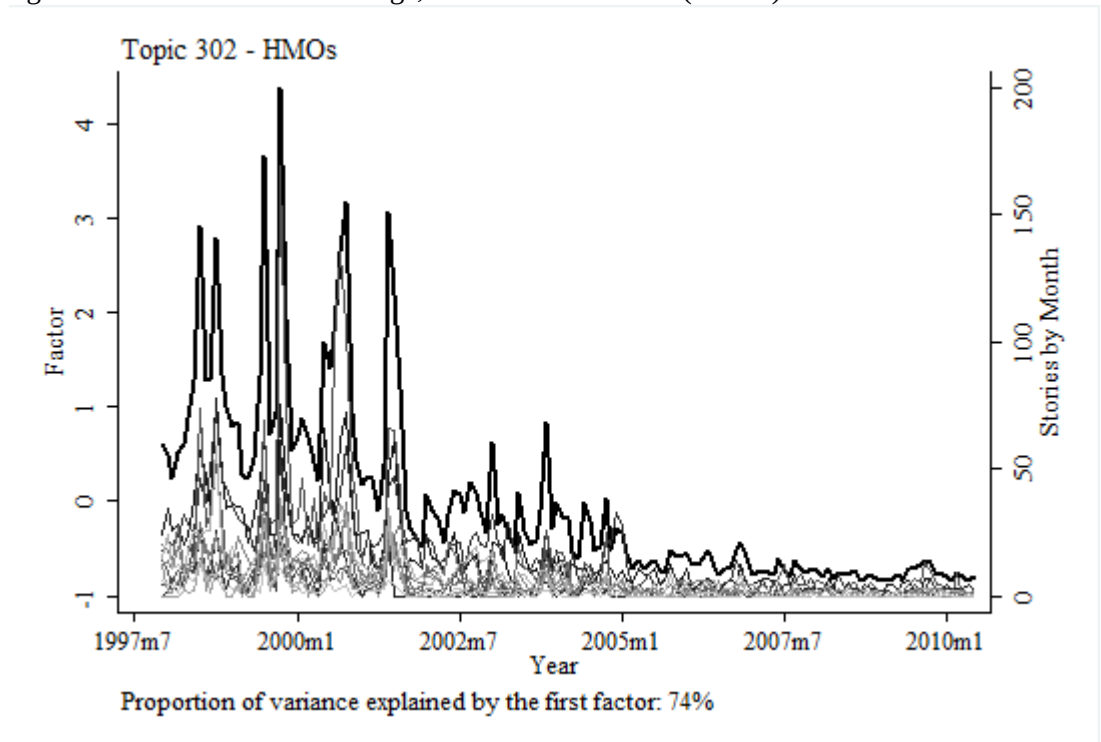


Figure 3: Low Media Coverage, One Retained Factor (Beach Volleyball)

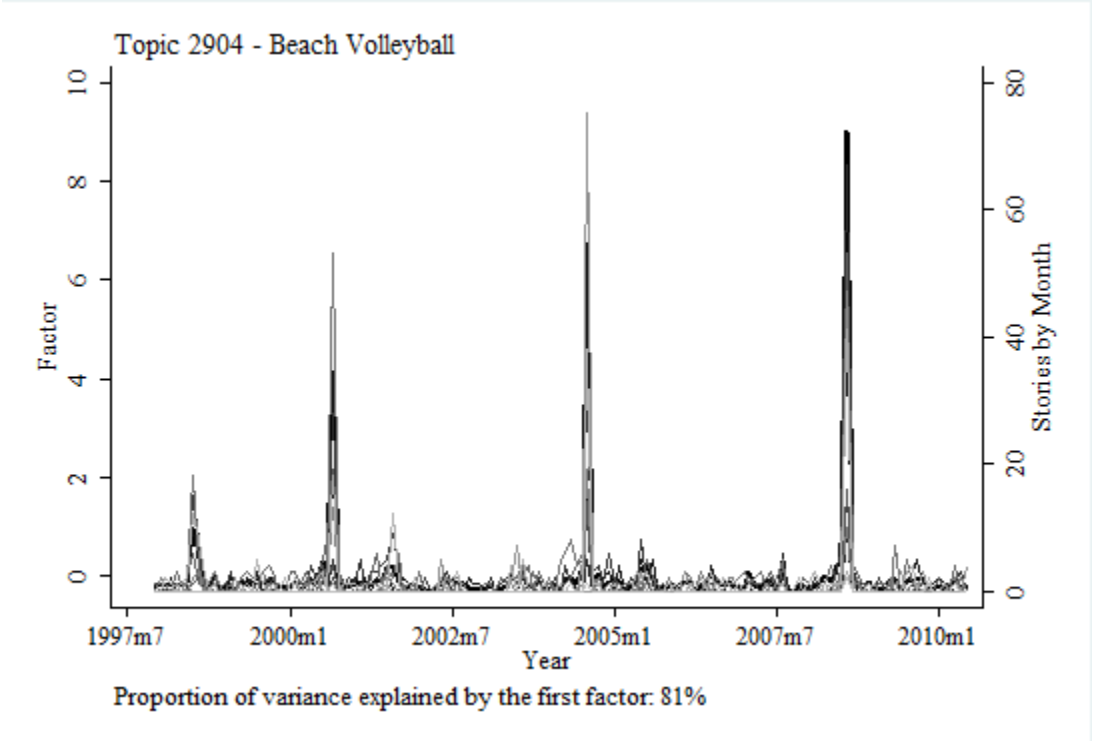


Figure 4: Multiple Retained Factors (Water Pollution)

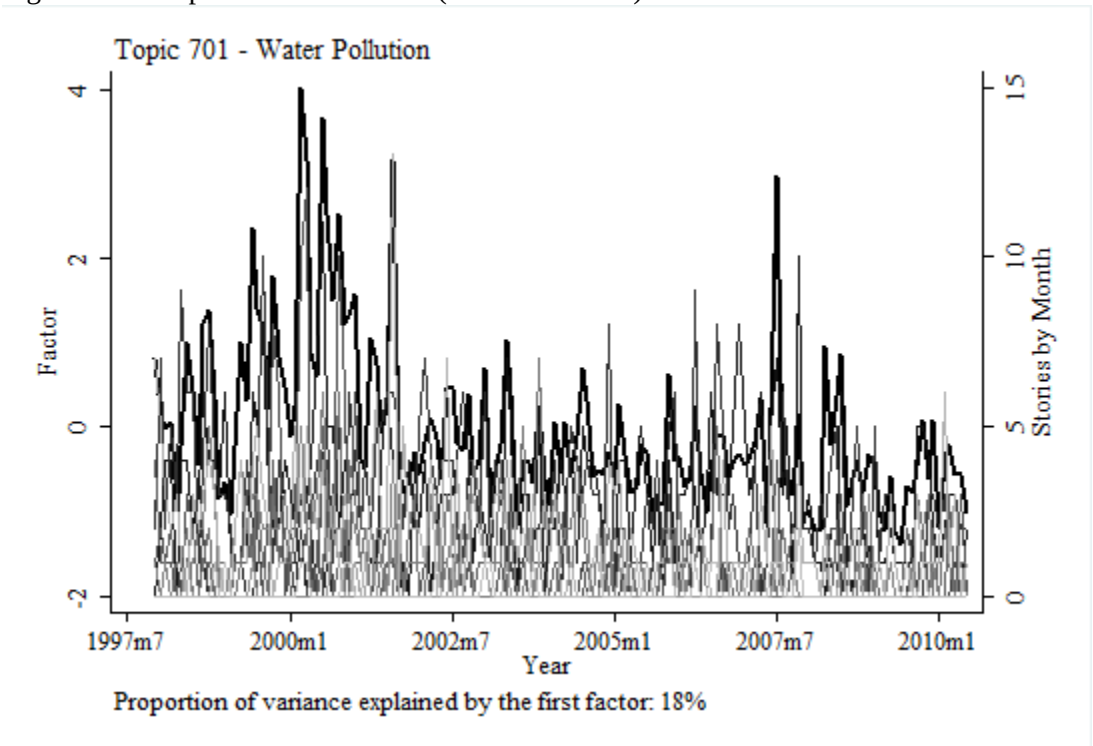
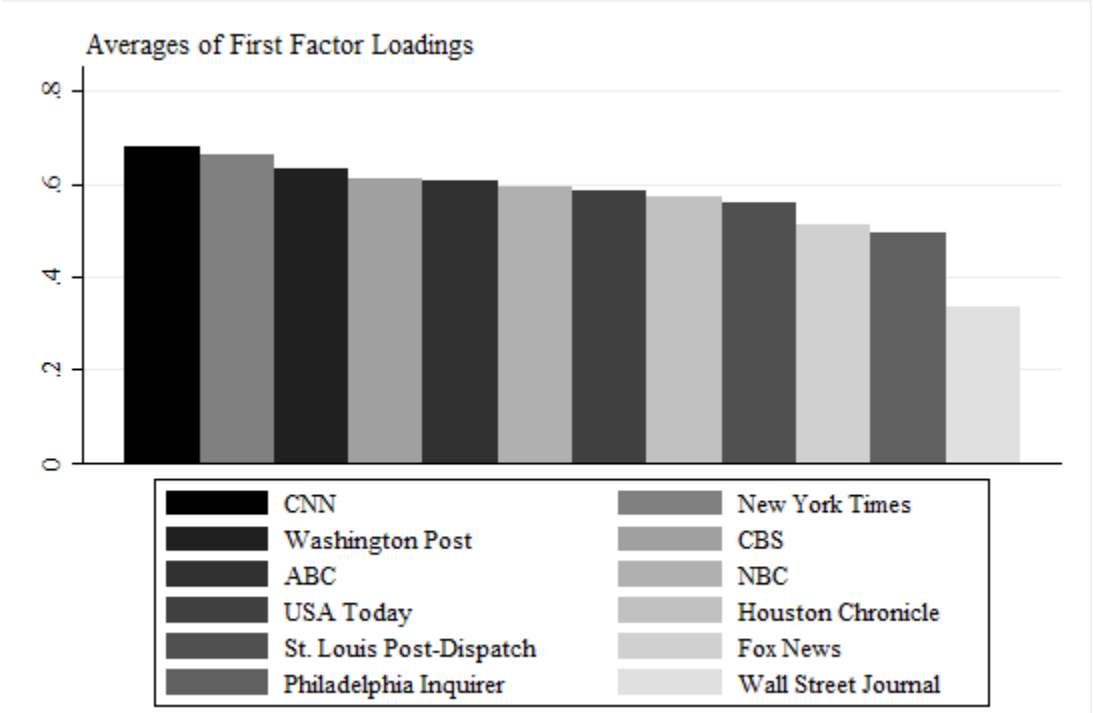


Figure 5: Mean Loadings by Source



Note: High loadings mean the source correlates highly with a national agenda across 90 keyword searches. Low loadings mean the source tended to correlate poorly with the other sources.

Figure 6: Mean Loadings by Source, National Core Set