

The Diversity of Internet Media: Utopia or Dystopia?

Bryan J. Dworak
Law Student
Duke University School of Law
bryan.dworak@lawnet.duke.edu

John Lovett
Ph.D. Candidate in Political Science
UNC-Chapel Hill
jlovett@email.unc.edu

Frank R. Baumgartner
Richard J. Richardson Distinguished Professor of Political Science
UNC-Chapel Hill
Frankb@unc.edu

Abstract

At the dawn of the Internet age, some believed that the medium would be the great equalizer for information. With unlimited space, unlimited journalistic independence from beat writing, and a deliberative relationship with their audience, the media would expand their discussion of issues into an unlimited number of topics. However, others argue that Internet media continues to act similarly to traditional media, continuing to mimic stories from other news sources and showing powerful status-quo tendencies. The Internet has unquestionably removed many barriers from traditional journalism: the news hole is infinite, for one. But has it transformed the collection and dissemination of information as many hoped and expected? Using datasets on traditional media from seven sources in five countries, and Internet media data from the blogosphere and Twitter, we test whether Internet media can be better described in the terms of the utopian view or the dystopian view. We focus particularly on the diversity, volatility, and amount of friction in coverage. We find that evidence points in favor of the dystopian argument: the media agendas of traditional and new media is similarly diverse, and attention in Internet media exhibits the friction also seen in traditional media.

Keywords: media and politics, Internet media, media agenda, public policy, agenda-setting

Paper prepared for presentation at the meetings of the Midwest Political Science Association,
April 3-6, 2014.

Internet Utopias and Internet Dystopias

At the dawn of the Internet age, there was a great hope that the interconnectedness of information created by the new medium would lead to a more diverse agenda of discussion. It would be an ideal resource for discussion and dissemination—providing completely free and instant information to everyone. The traditional constraints placed on journalists would be eliminated, leading to an unlimited supply of information across a broad spectrum of topics. Journalists, unchained from the restraints of traditional media such as coverage beats and the profit motives of their employers, would look outside the parameters of normal news coverage for policy niches on which to focus their energies. It would be a new, democratizing force where information on all topics would be readily available, regardless of importance or elite pressure. The one-size-fits-all metropolitan newspaper, news magazine, or nightly television news broadcast, with its tight constraints on format, limits in its size, and traditional focus on certain elements of “high politics” or other elements deemed to be worthy of coverage would be replaced by a free-wheeling, unconstrained, and limitless source of coverage for all that one could want to know.

In this utopian vision of the Internet, new media would not simply be a replication of traditional media online. The Internet, unlike traditional news media, is not bounded by traditional limits of scarcity in news coverage. There is theoretically no need for an Internet journalist to simply take on and stick with one topic area over the course of a career. Rather, journalists could focus on any number of topics, from health care to education to foreign affairs. Further, journalists could now focus more on their own interests without having to worry about the priorities of the national media. Internet journalists, free of the traditional sense of beats and “selling news”, could now focus on their areas of interest.

The result of the Internet utopia's new journalistic freedom would be that new media news coverage would be available on topics that receive little to no attention from *The New York Times* or other traditional outlets. Online media would offer discussion on all types of topics, regardless of how much coverage the topic might receive within mainstream media sources. The result would be the democratization of ideas through a true marketplace, where individuals could find updated news and discussion on topics as varied as agriculture, transportation, and foreign policy, without the constraints of traditional media's limited page space and profit agenda.

No matter what dreams may have animated some visions of the Internet at its inception, the Internet we use each day does not appear to reflect this view. Rather, what we see may be closer to what could be called the Internet dystopia. Under the Internet dystopia, journalists still follow the practices of traditional media. Journalists do not focus on a broad range of topics in writing, and they continue to mimic one another on stories. Journalists attempt to maximize their readership and profits by writing about topics that are similar to the most popular websites, hoping to gain new followers from those seeking information on specific, popular topics. Furthermore, unlike the traditional media, the level of originality in stories decreases: essentially one source or group of sources (usually the most popular sources) creates the original article, and others share or tweak the original story. The result is a system of media that depends more on mimicry and less on originality, and one where the attention moves with high levels of friction based on events. To add insult to injury, the rise of the new media destroyed in many ways the financial model of subscriptions on which the old media system was based. With that collapse came the closure or shrinkage of many newsrooms. Thus, rather than ushering in a new era of information abundance, the Internet may have had a significant negative effect on the news.

In this article we examine both traditional and new media sources to attempt to answer the question of to what extent Internet media has realized either its utopian or dystopian vision. We focus particularly on the diversity, volatility, and amount of friction in coverage in the traditional and new media, measures of the range of coverage and of the ease with which the media shifts attention from topic to topic. As we explained before, these are basic and measurable characteristics that can be compared across any media source, and which reflect the underlying practices of news gathering. We look at data from American and international newspapers for traditional media, and political blogs and Twitter for new Internet media. We find support for dystopian media argument: there is little evidence that the media agenda of the Internet is significantly more expansive than the media agenda of traditional media. The mimicking and profit-driven processes in traditional media rear their heads in the world of online media, leading to similar patterns of media coverage on the Internet.

The Measurement and Development of Media Agendas

Generally, we know that the traditional media is relatively diverse in the topics that it covers, though far from perfectly so, and that it moves in fitful bursts. Further, we know that several media agendas generally exist at one time (Boydston, Moody, and Thomas, 2010). However, there is significant mimicking between these agendas, as a uniform media agenda can emerge in response to significant events (Atkinson, Lovett, and Baumgartner, 2014). These are called media storms—a phenomenon wherein bursts in coverage occur in the media in response to significant events of interest to the public, but fade away more slowly than they come about (Boydston, Hardy, and Walgrave, N.d.).

Media agendas in general are not smooth. Baumgartner and Jones ([1993] 2009) and Jones and Baumgartner (2005) find that media agendas are sticky, and exhibit high levels of

friction. They argue that neither the media nor policy actors focus long on one issue, a result of disproportionate information processing, wherein attention is redirected toward more pressing issues as punctuations occur (Jones and Baumgartner 2005). In their view, the media are tied by their own institutional rules to the status quo, just as government institutions are. They are less “sticky” than some government institutions, but they are nonetheless affected by organizational cultures, beats, and reader expectations that cause them to cover similar issues today as they did yesterday. Shifting attention to issues as they become more important happens, of course, but it does not proceed in proportion to the importance of the issue. Rather, attention-shifts are affected by friction: hard to bring about, and bursty when they do occur.

There are limitations to diversity, though, especially due to the limitations of traditional media sources. Boydston (2013) analyzes front-page stories from *The New York Times* and finds that issue diversity is limited: the paper tends to focus on certain topics much more than others, and the agenda in the *Times* moves in the fitful bursts, the result of an extremely limited amount of space on the front page. Baumgartner and Chaqués (2012) find similar dynamics (minimal diversity, high friction) in two Spanish newspapers. While the media can be diverse, the very construction of traditional media makes maximum diversity of sources highly problematic.

The Economy of Journalism

The logic behind news choices goes beyond agenda dynamics. There are also profit motives in play, as the media attempts to maximize readership. Hamilton (2004) argues that the news media, as well as the journalists covering stories for the media, can be understood as an economic good. Hamilton argues that in the Internet media, the economic understanding of the news is changing (190). Journalists are often viewed as economic goods which can easily be fired as excess resources.

The profit motive also helps explain the decline in traditional media. As media has become less cost effective, the need for journalists and papers has decreased, leading to major decreases. Grabowicz (2014) notes several major US newspapers, including the *Seattle Post Intelligencer* have either switched to an online-only format, or shut down entirely. Others have gone through reorganizations as the result of bankruptcy proceedings or rounds of layoffs and pay freezes. Some papers have cut down on the number of days they print or deliver each week. Ahlers (2006) further finds a substantial decline in both the circulation of newspapers and in network television news viewership from the 1980s to early 2000s. These shifts have had dramatic if not catastrophic consequences for many news organizations in the newsroom: many news-gathering journalists have lost their jobs.

The Internet and Media Agendas

Recently, scholars have focused more specifically on examining the content found in blogs, but have not found the utopia of increased general attention. Instead, researchers find that there is significant “cannibalization” of stories in the new media, particularly on blogs (Phillips 2010, 96-97). Others have found that bloggers typically rely on the traditional media for their posts and rarely do original reporting (Wallsten 2007, 568). Further, mimicry is still present, facilitated by the new forms of technology (Boczkowski, 2009).

On the other hand, others see a different relationship, where bloggers are actually the leaders, and exert influence over traditional media (Meraz, 2011; Woodyly, 2008). Further, over time the public agenda as represented by media coverage has become more diverse (McCombs and Zhu, 1995). These trends extend beyond blogs: studies have shown that Twitter users are more likely to be indirectly exposed to media through the site, increasing the diversity of news they read (An et al. 2011). Pole (2010) makes the case that the blogosphere is actually more

diverse because bloggers do not have to answer to the same sorts of constraints as traditional media journalists. She also points out the extremely low costs of entry into blogging facilitates this (128). Nahon et al. (2011) argues that information in the blogosphere is highly viral and that the lifecycle of information there is very short, leading to lower friction in the media.

One potential factor that could work in either direction (utopia or dystopia) is the networked nature of blogs and Twitter. Blogs, unlike the traditional media, frequently incorporate hyperlinks to similar resources (although often from a different point of view) on other blogs or in the old media (Farrell and Drezner 2008). The networked nature carries over to Twitter, where users are able to engage in conversation with essentially anyone else on the network, making the diversity of discussion on the network even broader (Lasora et al. 2012).

Expectations for Utopia and Dystopia

Our hypotheses about how coverage will differ in the traditional and new media are based on the notions of Internet utopia and Internet dystopia. The first hypotheses come from the very notion of what was expected within the Internet utopia. The expectation is that the expansive discursive space created by the Internet should lead to spaces where all topics receive attention. Sources themselves are generally linked to one another, sharing back and forth among one another to maximize readership collectively. The linked nature of the sources should create a greater diversity in attention because of the plurality of sources.

On the other hand, the Internet dystopia assumes that Internet media is driven by the same forces that drive traditional media: profit and economic goods. Individuals who run these sites want to maximize readership and gain a foothold as a major player, whether for reasons of profit, advocacy, or importance. Therefore, these individuals will focus on stories that will gain the most readership, to offer their own specific and unique view to the situation. As a result,

mimicking occurs: if story A is popular, then other Internet media sites will jump on story A in an attempt to maximize their own connection to story A. As a result:

Hypothesis 1a (utopia): The new media show a greater diversity of attention than the traditional media.

Hypothesis 1b (dystopia): The new media show a similar level of diversity of attention to the traditional media.

The second set of hypotheses focus on the type of coverage seen in media, namely policy-relevant coverage. If the utopia exists, we should expect that not only will there be more coverage of all issues, but more coverage of policy-relevant issues as well. Purveyors of media will expand their coverage into policy-relevant topics, knowing that they can target information to multiple populations. To that end, we offer Hypotheses 2a and 2b:

Hypothesis 2a (utopia): The new media show a higher level of policy-relevant coverage than in the traditional media.

Hypothesis 2b (dystopia): The new media show a lower level of policy-relevant coverage than in the traditional media.

In the traditional media, coverage over time is highly volatile, and can be characterized by punctuated equilibrium (Baumgartner and Jones, [1993] 2009). In punctuated equilibrium theory, change (in this case, in media attention) for the most part is incremental, except for rapid,

sudden and intense change that completely changes attention, leading to a totally new equilibrium for the subject. These shifts are quick and sudden.

In the utopian vision of the new media, we would expect these shifts to occur more gradually. Newspapers and newscasts can generally only change once per day. That means that day-to-day, their attention to certain issues might vary more than the new media. The Internet allows for sustained conversation about small issues that would never make it into a newspaper or onto a newscast even if some significant event involving the issue occurred. Twitter lends itself particularly well to this tendency. Users can instantly transmit information to other users as events unfold in front of them. It is not like the front-page of *The New York Times* where there is a very limited amount of space (only about 8 stories per day). On the Internet, “space” is cheaper and thus conversation can be more sustained and the shifts seen will appear less dramatic. If the new media are more similar to the traditional media in this respect than the vision of the Internet utopia would anticipate, then the shifts will be equally marked.

The status quo effect would dictate that the topics most likely to be covered in the media tomorrow are those issues covered in the media today. Various studies have shown this to be true in various traditional media outlets; however, in the Internet utopia, we would envision that topics covered would be determined independently based on the day’s events. Boydston (2013) finds this to be particularly true in *The New York Times*, where she observes a high friction environment. Again, the utopian and dystopian views of the Internet lead to different expectations on how attention shifts over time.

Assuming the traditional media exhibit high levels of friction, that is, it is hard for coverage to move from one issue to another over time and when it does it is characterized by marked shifts in attention, since attention changes more gradually in the new media, it would

also have less friction. In the perfect world, there would be no friction—any issue could receive a lot of attention because there are no constraints such as space on a front page, time in a newscast, or a publisher overly concerned with advertising revenues. Keeping this in mind, we offer Hypotheses 3a and 3b:

Hypothesis 3a (utopia): The new media shows lower levels of friction than the traditional media, as attention shifts more smoothly over time from topic to topic.

Hypothesis 3b (dystopia): The new media shows similar levels of friction to the traditional media, as attention lurches over time from topic to topic not in a steady manner but only after resistance.

It is unrealistic, but reflective of the Internet utopia vision, to predict that there could be a perfect representation of all topics in media. On the other hand, the Internet utopia theory supposes that unlike traditional media, the new media will distribute attention more proportionately based on the wishes of the audience. Therefore, there will be no need for sudden drastic change: the media should move gradually to cover issues as the need arises.

In sum, we lay out three very simple ways in which to compare traditional and new media. The hypotheses simply refer to the diversity of topics of coverage, the degree of policy-relevant coverage, and the ways in which attention moves from topic to topic over time. Measuring these three simple factors can allow us to assess whether the Internet is closer to the utopian visions of its dreamers, the dystopian visions of some detractors, or somewhere in between. Less dramatically, we can assess whether it is an improvement over the traditional media, worse, or more of the same. We now turn to the data.

Data

Traditional Media

The traditional media analyzed in this paper was previously collected and coded by a variety of researchers who are part of either the Policy Agendas Project (United States)¹ or the Comparative Agendas Project (international)². All data was downloaded from individual country's project websites or made available to the authors directly from the researcher.

For American data, we use Boydston's (2013) *New York Times* front-page dataset.³ The dataset includes every front-page story in *The New York Times* between January 1st, 1996 and December 31st, 2006. It is comprised of 31,034 stories, all of which were coded by Boydston and her research team. It is the primary dataset used for comparing to the new media. We also use and refer to the Policy Agendas Project's *New York Times* Index⁴ originally created by Baumgartner and Jones, which is a coded collection of stories from throughout the *Times*, based on the annual index to that newspaper.

We also utilize a variety of non-American sources. For international datasets, we use data from four countries: Spain, the United Kingdom, Switzerland, and Denmark. We look at international datasets of *El Pais* and *El Mundo*, Spanish newspapers part of a dataset created by Frank Baumgartner and Laura Chaqués⁵, a dataset of *The Times of London*, a newspaper in the United Kingdom created by Shaun Bevan⁶, a dataset of *Neue Zürcher Zeitung*, a newspaper in Switzerland created by Anke Tresch⁷, and a dataset of Danish Radio News created by Christoffer

¹ www.policyagendas.org

² www.comparativeagendas.info

³ As previously explained, made available directly from Boydston.

⁴ This data 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.

⁵ <http://www.ub.edu/spanishpolicyagendas/datasetinstruments/>

⁶ <http://www.policyagendas.org.uk/>

⁷ Not available on web; made available directly from Tresch.

Green-Pederson⁸. Full details of these datasets are available in Appendix 1. All data from the US were coded according to the Policy Agendas codebook originally created by Frank Baumgartner and Bryan Jones and adapted by Amber Boydston and all international data was coded using a similar scheme but specific to different countries. All codebooks are available from the authors or the original research team.

New Media

All new media data was collected by the authors between May 2012 and August 2012. To select blogs to include in the sample dataset, we employed a hybrid selection approach loosely modeled on Wallsten (2007). We randomly sampled thirty blogs from a larger list of blogs compiled from three lists of influential blogs. The first list was the Top 100 US Political blogs (based on data current May 10-13, 2012) from Technorati⁹, a blog aggregator and search engine that ranks and indexes blogs on different topics. The second list came from the top 50 Political Blogs according to the Blog Authority Index¹⁰, also known as the BAI (Karpf 2012a). This data was also current as of May 2012. The final list was of 30 other blogs based on a list of influential blogs created by Professor Justin Gross at the University of North Carolina-Chapel Hill. Combining these lists resulted in a total of 180 blogs from which thirty were randomly selected for sampling.¹¹ See Appendix 2 for a full list of these blogs.

We used the program OutwitHub Pro to configure HTML scrapers to collect the date, author, post title, and post text or abstract where available. The goal of these scrapes was to

⁸ <http://www.agendasetting.dk/start/page.asp?page=4>

⁹ www.technorati.com

¹⁰ <http://www.blogosphereauthorityindex.com/>

¹¹ In one case, technical difficulties immediately prohibited the use of the blog The Inquisitr as part of our dataset. That blog was removed from the dataset. With the remaining 150 blogs not selected in the original sample, we randomly sampled without replacement for one additional blog. As a result, the Daily Kos was added to the dataset. During data collection, another blog, The Political and Financial Markets Commentator, was eliminated from analysis due to technical difficulties. It was not replaced by another blog because data collection had been ongoing for over a month and replacing it would not have been feasible, as its replacement would not have been adequately represented in the full blog dataset.

represent a typical “front page” of a blog, akin to that of a newspaper for the best possible comparison of traditional media data relying on front-page coverage in newspapers. In some cases, a blog may have only been displaying the six or ten most recent posts. In the case where there were fewer than fifteen posts on the homepage, all posts were scraped¹². According to Karpf (2012b), the front page of a blog can be “the most-trafficked real estate in the political blogosphere” (64). This data was coded using the full Policy Agendas scheme as adapted by Boydston. See Appendix 3 for a full list of these two-digit codes.

For the Twitter data, we similarly started with a larger list of influential Twitter accounts. We relied on Daniel Romero et al.’s (2011) list of influential news Twitter accounts. Romero and his coauthors use a complex algorithm to calculate and rank influential users on Twitter. Their methodology goes beyond followers and number of Tweets and utilizes further data such as retweets and passivity on the network. It seeks to capture the amount of forward influence of any individual user. We started with a list of one hundred influential news Twitter accounts. To see the full list of Twitter accounts, see Appendix 4. As with the blogs, we sampled without replacement for a set comprised of thirty influential news Twitter accounts.¹³ We then scraped the twenty most-recent Tweets from each account, akin to what a Twitter “front page” would look like. This data was coded in the same manner. Data on individual blogs and Twitter accounts was also merged into larger datasets to analyze the networked nature of this sampled blogosphere and Twitter community. Blog data was collected from May 19, 2012 to August 14,

¹² In the event that a blog or Twitter user posts relatively infrequently, and posts show up in scrapes over consecutive days, that data was not eliminated from the dataset. In that case, a post may be present two, three, four, and so on times in the dataset. This decision was based on the idea that if a user were to visit that blog or Twitter timeline day after day, they would continually be exposed to that information. Attention and space is continually granted to that issue or topic until it is displaced by newer content. In a way, this is like an enhanced status quo effect.

¹³ The original Twitter dataset was to be comprised of thirty Twitter accounts. Two were eliminated from analysis because they no longer regularly post or have a different owner than they did when deemed influential by Romero et al. (2011). These are @seedmag (Seed Magazine) and @themoment (The Moment).

2012. Twitter data was collected from June 12, 2012 to August 16, 2012. Because these scrapes had to be done in real time, rather than taken from an archived source, a small number of days were missed. However, the underlying databases have posts from each of the 88 days of data collection for the blogs and 66 days for Twitter.

Results

The overall diversity of attention in a source's media agenda shows the range of topics of news available in that specific source. There are a variety ways to measure that diversity. Two that we use here in analysis are the percentage of total coverage devoted to different topics, and a more complex measure in the form of entropy. Entropy is a measure used to study disorder within political and social systems (Shannon 1948 and Coleman 1975 further explain the application to social systems). The measurement of entropy is varied, including Shannon's H Information Entropy, the Herfhdendal Index, and the normalized Herfhdendal Index. We use Shannon's H for our measure of entropy based on the success of other studies of media agendas in using the measure to describe the overall diversity of attention (Baumgartner and Jones 2009, Boydston 2013, Boydston, Bevan, and Thomas N.d., Wolfe, Boydston, and Baumgartner 2009, and Baumgartner and Chaqués 2012)¹⁴. Shannon's H is measured on a scale of 0-1, with zero representing a highly homogenous system, or a media agenda with very low (no) diversity, and one representing a highly heterogeneous system, or a media agenda with near perfect diversity. Perfect diversity in this case would mean that no single topic gains more attention than any other topic. The other end of the continuum would be no diversity of attention, in which all stories are on the same topic. The formula for entropy is as follows:

¹⁴ As a further check on the selection of Shannon's H, we also calculated the Herfhdendal Index and normalized Herfhdendal Index and ran a correlation between those and Shannon's H. They were negatively correlated with $r^2=.96$. Given this uniformity, Shannon's H is a valid measure.

Formula 1: Shannon's H Information Entropy

$$\text{Shannon's } H = - \sum_{i=1}^n p(x_i) * \ln p(x_i)$$

where:

x_i represents an issue

$p(x_i)$ is the proportion of total attention the issue receives

$\ln(x_i)$ is the natural log of the proportion of attention the dimension receives¹⁵

We first turned to a qualitative analysis of the issues represented in the different media agendas, focusing on our combined blog and Twitter datasets and the *New York Times*' front page dataset. In Table 1, we provide the percent of coverage devoted to the twenty-seven Policy Agendas topics in the three types of media outlets. There are notable differentials in some categories including Defense, International Affairs, and Arts and Entertainment. Twitter and the blogs also have significantly higher percentages of posts coded as Miscellaneous. Most of this is characteristic of the different types of media, since it is much cheaper to post human interest stories in online media versus the highly crowded front page of the *Times*. Many posts were also uncategorizable because they included several stories or were "round-up" posts with many links and stories.

(Insert Table 1 about here)

Using entropy as our measure, we undertook a more technical analysis of the different media sources in our study. First, we look at each blog and Twitter account individually. The mean entropy value for blogs was 0.59 and a median value of 0.66. The least diverse blog using this measure was *The Tory Diary* with an entropy value of 0.05. This blog focused nearly expressly on British politics (coded under Topic 19-International Affairs), despite being classified by Technorati as one of the top 100 influential political blogs in the United States. The

¹⁵ Formula from Boydston, Bevan, Thomas N.d., 14-15

most diverse individual blog was *Lawyers, Guns, and Money* with an entropy value of 0.80. There were several others near that range, while no other blog approached the extreme lack of diversity observed on *The Tory Diary*. See Appendix 5 for a display of entropy values for all blogs in our dataset.

On Twitter, there are similar results. The average entropy for individual Twitter accounts is 0.55 and the median entropy is 0.59. The least diverse Twitter account included in our dataset is @tvguide, with an entropy value of 0.09. This lack of diversity is explainable by the intense focus on arts and entertainment, which we discuss in greater detail below. The most diverse Twitter account as measured here is @nprnews, which has an entropy value of 0.83. See Appendix 6 for a full display of entropy values for Twitter accounts included in our dataset.

Thus far, our analysis has been confined to the diversity of individual new media sources. Internet media's networked nature may be creating higher levels of diversity versus traditional media. Unlike newspapers, where an ordinary individual would only read one newspaper per day, most Twitter users are reading Tweets from multiple (if not tens or hundreds of) accounts each day (and often multiple times per day) and the linked nature of the blogosphere and features like blog rolls contribute to the idea that regular readers of blogs are often reading posts on multiple blogs. This is simply the nature of the new media.

Due to the networked characteristic of both new media types in our datasets, we calculated another entropy value for each. As previously detailed, we appended all Twitter and blog datasets into one, large dataset for each with 15,856 blog posts and 18,903 Tweets. We call these our "combined" datasets for each type of new media. These values are also reflected in Appendices 5 and 6. This is a calculation of the total entropy for the blogosphere or Twitter based on our samples. Table 2 below details the results for the total entropy of blogs and Twitter,

with a variety of American and international news total source entropy values for comparison. The combined entropy for the blogs is 0.77 and the combined entropy for Twitter is also 0.77, showing a highly similar level of diversity.

(Insert Table 2 about here)

As the point of this analysis is not merely to describe the new media in qualitative and quantitative terms, but also to compare these new media sources to traditional media sources and assess the implications of that, we have compared the diversity of our new media sources to the mainstay traditional media source in the United States, *The New York Times*. Using Boydston's (2013) front-page dataset, we calculated an entropy value of 0.83. Therefore, the traditional media, at least as represented by the *Times*, is just slightly more diverse than new Internet media sources. This difference is not necessarily substantial. The diversity of the traditional and new media, at least measured numerically, is similar.¹⁶

For a more longitudinal analysis of traditional media data, we also ran entropy tests on the Policy Agendas Project's *New York Times* index of all stories between 1946 and 2008 originally created by Baumgartner and Jones. Rather than coding every front-page story in the *Times*, Baumgartner and Jones coded the first story on every odd-numbered page of the *Times* index to get an overall idea of the type of coverage displayed throughout the newspaper. Although we did our best to replicate a "front page" in the blogs and on Twitter by retrieving the most recent updates each day, this is not a perfect method, as it may not necessarily capture stories that appear between collection times if there is sufficient movement on the blog or feed in

¹⁶ To alleviate concerns about comparing approximately three months of new media data to many years of traditional media data, we ran separate entropy tests on all consecutive three-month (quarterly) periods in Boydston's front-page dataset. Taking the average of all of these entropies, the result is an overall value of 0.79. This entropy value is even closer to the values calculated for the new media. Over time, the *Times* is slightly more diverse than the new media, but barely so within three-month periods. This further supports our argument that the diversity of the new and traditional media in the United States is not significantly different.

question. The technique makes our datasets (by analogy) somewhere in between the front-page dataset and the index. The entropy value for the *Times* index is a bit higher than the front-page score, at 0.88. This demonstrates that fewer issues win the competition for the highly finite amount of space on the front page of the *Times* each day, whereas a greater number of topics can be represented throughout the paper. However, this dataset spans a much greater period of time than either Boydston's dataset or the dataset we have created. For the purposes of this analysis however, these are all "high" entropy scores and graphically do not appear much different. Coverage is distributed differently inside of the paper than on the front page. As seen in the charts, International Affairs is the most common topic on the front page, while Banking and Domestic Commerce is the most common topic overall as represented by the index.

We also compare our Twitter and blog data to our international traditional media sources. Whereas Twitter and blogs, due to their being on the Internet are easily accessible around the world, the consumption of traditional media sources is largely confined to a single country. In general, we observe that the international traditional media are also not particularly more or less diverse than the traditional or new media in the United States. The entropy values for international media range from about 0.74 for Danish News Radio to 0.86 for *El Pais*. At the very least, it is clear through the comparison between all of these traditional media sources, both American and international, and the new media sources in our dataset, that their overall entropy values are very near each other. Attention is concentrated similarly. Consequently, we can conclude that the overall level of diversity of information available is roughly the same now as it has been in the last couple of decades.

Boydston (2013) conducted similar analysis that revealed that within the traditional media, there is largely a single media agenda within television news and newspapers in the

United States, though this single agenda primarily exists among topics that have the most general coverage (Atkinson, Lovett, and Baumgartner, 2014). Using the data, we can conclude that overall, the new and traditional media display about the same level of attention diversity, at least statistically. The overall degree of attention is similar. That last point is important to note because there may be differences in the types of topics that receive different levels of attention. We will further discuss where that finite amount of attention in each medium falls, looking at possible discrepancies by topic in the traditional and new media, focusing on the question of policy coverage versus non-policy coverage.

This review of diversity of attention, then, clearly refutes the utopian hypothesis (1a) while supporting the dystopian (or at least disappointing) version: 1b, reflecting little difference in range of coverage in the new and old media.

Policy vs. Non-Policy Coverage

All of the data was coded using the full Policy Agendas Topic codes, which cover all sorts of non-policy topics including religion, sports, and entertainment. However, in looking at the political implications of media coverage, it is also relevant to look at how coverage is distributed among policy-relevant topics only, removing non-policy coverage from consideration. There are some differences in how coverage is distributed measuring entropy by only analyzing the coverage distributed over policy-relevant topics. By isolating only the policy-relevant coded posts and stories, we were able to calculate the percent of coverage in each medium devoted only to policy-relevant topics and also calculated the entropy among only those policy relevant stories. The observed results can be seen in Table 2, grouped by source.

Overall, the traditional media exhibited higher levels of policy-relevant news coverage than the new media. On the *Times* front page, 88.68% of coverage was devoted to policy topics.

In the blogosphere, this value was lower, with 82.78% of coverage being distributed across policy relevant topics. The starkest difference, however, is on Twitter. On Twitter, less than half of coverage was distributed to policy-relevant topic areas. Only 48.75% of coverage was policy-relevant. This massive difference is possibly attributable to various reasons. The ease and zero-cost nature of posting on Twitter and the large amount of information available on the network promotes a considerable number of human-interest stories to be posted. Individuals taking the time to write extended blog posts appear to be more dedicated to substantive policy-relevant discussion. Where space is expensive and highly finite, like in a newspaper, very high levels of coverage are devoted to policy-relevant topics, as it is the information that the electorate most needs (but does not necessarily most desire). This is especially true on the front page. On Twitter and blogs, there is less of a difference in terms of cost, both monetary and spatial, that constrains the coverage in those areas. This is revealed by the difference in policy-relevant coverage in the *Times* index versus the front-page dataset. In the index, only 81.27% of coverage is on policy-relevant topics. This is lower than the distribution in the blogs.

The international traditional media sources largely mirror this distribution in policy versus non-policy coverage. On average, traditional media outlets devote 87.13% of their coverage to policy-relevant topics. The international sources range from approximately 82% to 99% of coverage devoted to policy-relevant topics. Therefore, we conclude that the traditional media, especially on the front page, display high-levels of policy only topic coverage. Aside from the percentage of coverage spent on policy and non-policy relevant topics, we also calculated the entropy value across the coverage of policy-relevant topics only. In the *Times*, the entropy is actually lower across the policy-relevant topics than across all issues, though negligibly so. The entropy value is 0.83. The combined blog value follows a similar trend with a

slightly lower entropy value with a score of 0.75. The same is true across the international sources as well. This trend changes on Twitter, though. The combined entropy on Twitter actually increases when looking only at policy-relevant topics. The combined entropy of Twitter is 0.80. This means that although a smaller percentage of coverage on Twitter is devoted to policy-relevant areas, among that coverage, a greater variety of policy topics are represented. In assessing Policy-relevance, then, we can clearly refute hypothesis 2a and confirm 2b.

Friction, Skew, and Explosiveness in Media Agendas

We now turn to an analysis of how attention changes over time. Since most readers do not consume many days of news coverage in one sitting, but rather consume it on a daily basis, how attention to different topics changes over time is important to analyze. The amount of friction, or how easily attention changes over time, can be measured by kurtosis. Kurtosis refers to the peakedness of a distribution. A distribution with a large peak and fat tails has a high kurtosis and can be seen as explosive: most change is small and incremental, except for those brief moments of explosiveness in line with punctuations. Baumgartner et al. (2009) focused study on kurtosis, noting “We can assess these dynamics easily by looking at the left- and right-hand tails of the distributions” (608). We will use the same method. Kurtosis has been used to study other distributions, including both budgets (Jones et al. 2009) as well as the explosiveness of the agenda of the *New York Times* over time (Boydston 2013). In particular, Boydston (2013) finds that l-kurtosis is the superior choice of statistic to use when studying agendas based on its use in other similar studies (220).

Using l-kurtosis, we can assess the level of explosiveness of the each agenda. According to Boydston (2013), a normal distribution has an l-kurtosis of 0.123 and anything higher is considered to be a leptokurtic distribution (220). Most agendas are leptokurtic, meaning their

changes over time are characterized by a strong status-quo bias (leading to many cases in the central peak, very similar to their lagged values), at the same time as they have many cases far out in the tails of the distribution, reflecting explosive shifts (Jones and Baumgartner 2005, 110-112). In Table 3, the number of weekly-topic observations (in essence, the number of observations that can actually be used in the analysis and excluding zeroes), the kurtosis, and l-kurtosis value are displayed for all major sources in the dataset. In Appendices 5 and 6, this data is available for individual blogs and Twitter accounts, respectively.

(Insert Table 3 about here)

The combined blog data, based on 303 weekly-topic observations, has an l-kurtosis value of 0.36, well above the 0.123 threshold value for a leptokurtic distribution. All of the individual blogs are also above the threshold, though one barely so, with an l-kurtosis value of 0.13. Uppity Wisconsin, an issue-based, blog is the most leptokurtic with an l-kurtosis value of 0.70. Most, however, are around the combined value for the entire blogosphere. Regardless, we can say that the blogosphere has an explosive agenda.

Figure 1 shows a histogram of the frequency of all of the weekly percent-changes in coverage in the combined blog dataset, truncated at 300. A normal curve has been overlaid on the distribution to compare the histogram to the normal distribution. As we can see, this is a fat tailed distribution. As Boydston (2013, 217-221) observed for *The New York Times* at the monthly level, the blog media agenda is highly explosive.

(Insert Figure 1 about here)

Although we used Boydston's dataset in our analysis, she did her analysis of coverage over time at the monthly level. This is arguably a more appropriate level to analyze at given the expanse of the period her data covers. However, in seeking to provide a uniform analysis with

our new media data, we have also analyzed her data at the weekly level. The distribution can be seen in Figure 2. In this case, we truncated the distribution at 200. Although Boydston's monthly data had a clear cluster around zero, this data is similarly clustered, though the relationship at the weekly level is not as suggestive as it is at the monthly level. However, we can still see that the *Times* front page has a skewed and explosive agenda, as seen at both the monthly and weekly level based on both Boydston's and our analysis. At the weekly level, the *Times* front page data has an l-kurtosis value of 0.25 based on 6,430 observations. This value is slightly lower than what Boydston (2013, 220) calculated at the monthly level, which was an l-kurtosis of 0.344. Thus, at the weekly level, the *Times* front page is a less explosive agenda. Nonetheless, whether we aggregate these daily observations to the week or the month, we see an explosive agenda either way.

(Insert Figure 2 about here)

The *Times* index of all stories that we have also previously analyzed in this article has a slightly higher l-kurtosis value of 0.26. While this difference is inconsequential, we expected to observe a greater difference in explosiveness and skew between the front page and the index given that conversations on more topics can be more easily sustained elsewhere in the paper than on the highly competitive front page. Therefore, it appears that the *Times* as a whole is a relatively uniformly explosive agenda.

With respect to the international media, l-kurtosis values on a large number of observations for all media outlets range from 0.24 to 0.31. Therefore, none of the international media outlets are particularly more or less skewed or explosive than one another or than their American counterpart. Just as with the US sources, all of the international media agendas show high levels of kurtosis.

In Figure 3 below, we provide a frequency distribution of all of the weekly percent changes for Twitter. For this histogram, we truncated the data at 400. Again, this agenda is obviously highly skewed with respect to the weekly percent changes. The frequency distribution here supports that it is the most skewed agenda of the three presented.

(Insert Figure 3 about here)

Overall, the combined l-kurtosis value for all Twitter accounts in our data across all topics is 0.39, just slightly higher than the overall blogosphere value. Again, most individual Twitter account l-kurtosis values are clustered around the combined value and range from 0.14 to 0.58. ABC's Twitter account with the minimum l-kurtosis value of 0.14 is just slightly leptokurtic and close to the value of a normal distribution, making it not very explosive. However, again we have observed that Twitter is a highly skewed and explosive media agenda just like the blogs and to a lesser extent, the *Times*. So, we again disconfirm the utopian vision of hypothesis 3a and confirm that of 3b.

Conclusions

We find little support for any of our three Internet Utopia hypotheses. The dystopian worldview of the Internet holds in all three of our tests. In the cases of all news coverage, policy coverage, and attention volatility/friction, we find that Internet media acts similarly to traditional media. The spread of attention on topics in Internet and traditional media in both policy and non-policy content are highly similar, with little evidence of a systematic movement to expand outside of traditional topics. Furthermore, attention in the Internet world works similarly to that in the traditional world. Despite the ability to act quickly on stories, attention shifts in the Internet world are not smooth. Instead, they are similar to those in the traditional media mold, with little change in most topics, and some topics receiving massive, sudden movements in coverage.

The result is that with the decline of traditional reporting, there has been a rise in mimicking practices. What once consisted of multiple reporters covering the same event is now one blogger or journalist covering an event, and many other sources using that one piece of coverage as a jumping-off point for their own coverage. Networks, rather than encouraging new diversity in coverage, instead focus energies on similar coverage, changing over time to deal with the “hot” issue at hand. Therefore, we have actually descended beyond where information was in traditional media. The presence of reporters at least led to original coverage of events. Now, tweaked mimicry rules the Internet dystopia.

Tables and Figures

Table 1: Percentage of Coverage by Topic

Topic	<i>NYT</i>	Blogs	Twitter
Macroeconomics	3.11%	2.90%	1.40%
Civil Rights	2.95%	6.09%	0.95%
Health	5.80%	3.05%	2.96%
Agriculture	0.54%	0.20%	0.37%
Labor and Immigration	2.41%	2.45%	1.44%
Education	2.94%	1.02%	0.97%
Environment	1.14%	0.78%	0.54%
Energy	0.96%	0.78%	0.43%
Transportation	1.91%	0.67%	1.18%
Law & Crime	6.73%	8.37%	6.69%
Social Welfare	0.88%	0.51%	0.08%
Housing	1.32%	0.11%	0.65%
Banking & Domestic Commerce	4.02%	2.69%	3.33%
Defense	14.43%	5.16%	1.61%
Science & Technology	2.32%	1.51%	8.62%
Foreign Trade	0.82%	0.23%	0.31%
International Affairs	20.47%	19.49%	10.81%
Government Operations	12.75%	20.75%	5.48%
Public Lands	0.87%	0.11%	0.07%
State and Local Government	2.30%	5.92%	0.85%
Weather	1.85%	0.11%	1.09%
Fires	0.42%	0.18%	0.81%
Arts and Entertainment	2.48%	2.36%	18.34%
Sports	4.10%	1.89%	10.14%
Death Notices	0.86%	0.42%	0.73%
Religion	1.06%	2.09%	0.37%
Miscellaneous	0.55%	10.16%	19.77%
Entropy	0.83	0.77	0.77

Table 2: Entropy Values and Policy Relevant Coverage

Source	N	Entropy	Percent Policy Coverage	Entropy Policy Relevant Coverage
Danish News Radio	191,564	0.74	94.33%	0.73
Combined Blogs	15,856	0.77	82.78%	0.75
Combined Twitter	18,903	0.77	48.75%	0.80
<i>Neue Zürcher Zeitung</i>	8,558	0.81	99.05%	0.85
<i>El Mundo</i>	44,858	0.82	82.30%	0.79
<i>NYT Front Page</i>	31,034	0.83	88.68%	0.83
<i>El Pais</i>	50,770	0.86	81.96%	0.85
<i>NYT Index</i>	49,126	0.88	81.27%	0.88
<i>Times of London</i>	21,844	0.89	81.74%	0.88

Note: Table is sorted by entropy of total observations, from least to most diverse going top to bottom.

Table 3: Weekly Kurtosis Values

Source	N	K	L-K
<i>New York Times</i> Front Page	6,430	30.36	0.25
<i>New York Times</i> Index	15,081	10.12	0.26
<i>El Mundo</i>	2,302	26.49	0.26
<i>El Pais</i>	2,378	24.24	0.27
<i>Neue Zürcher Zeitung</i>	2,417	11.86	0.24
<i>Times of London</i>	6,365	13.2	0.31
Danish News Radio*	-	-	-
Combined Blogs	303	44.14	0.36
Combined Twitter	228	78.34	0.39

*Due to issues with dates in the original dataset, kurtosis calculations were not possible.

Figure 1: Frequency Distribution of Weekly Percent Change in Coverage on Blogs

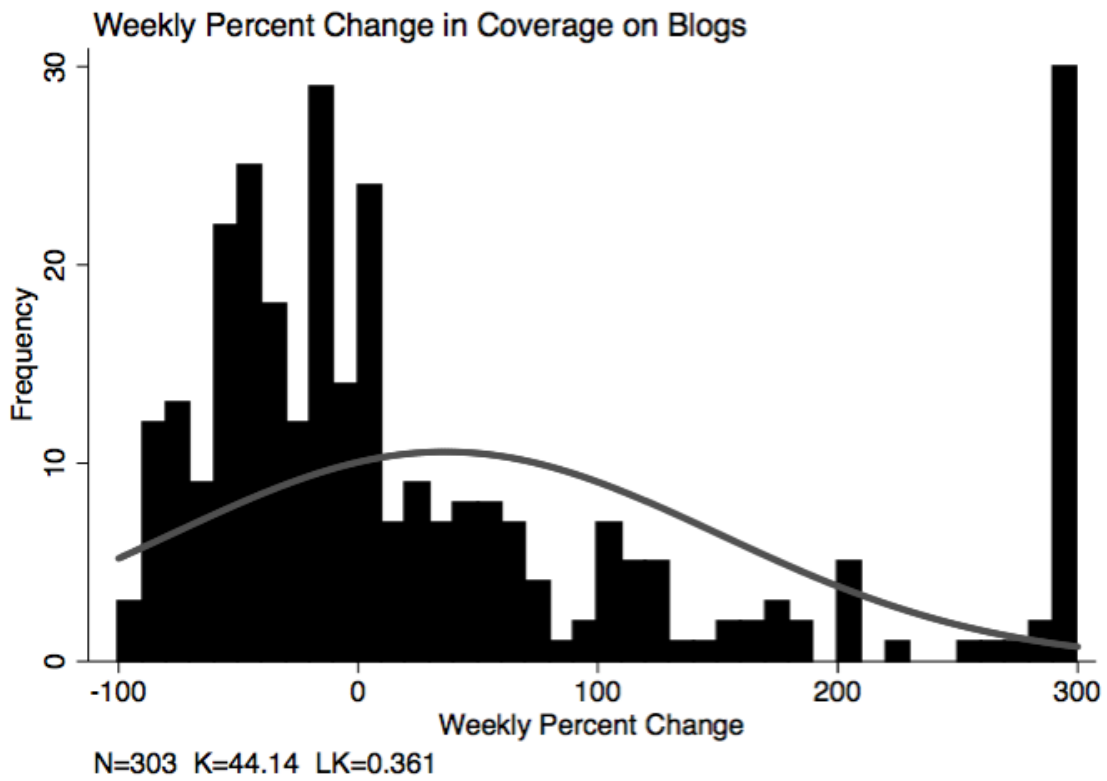


Figure 2: Frequency Distribution of Weekly Percent Change in Coverage on *NYT* Front Page

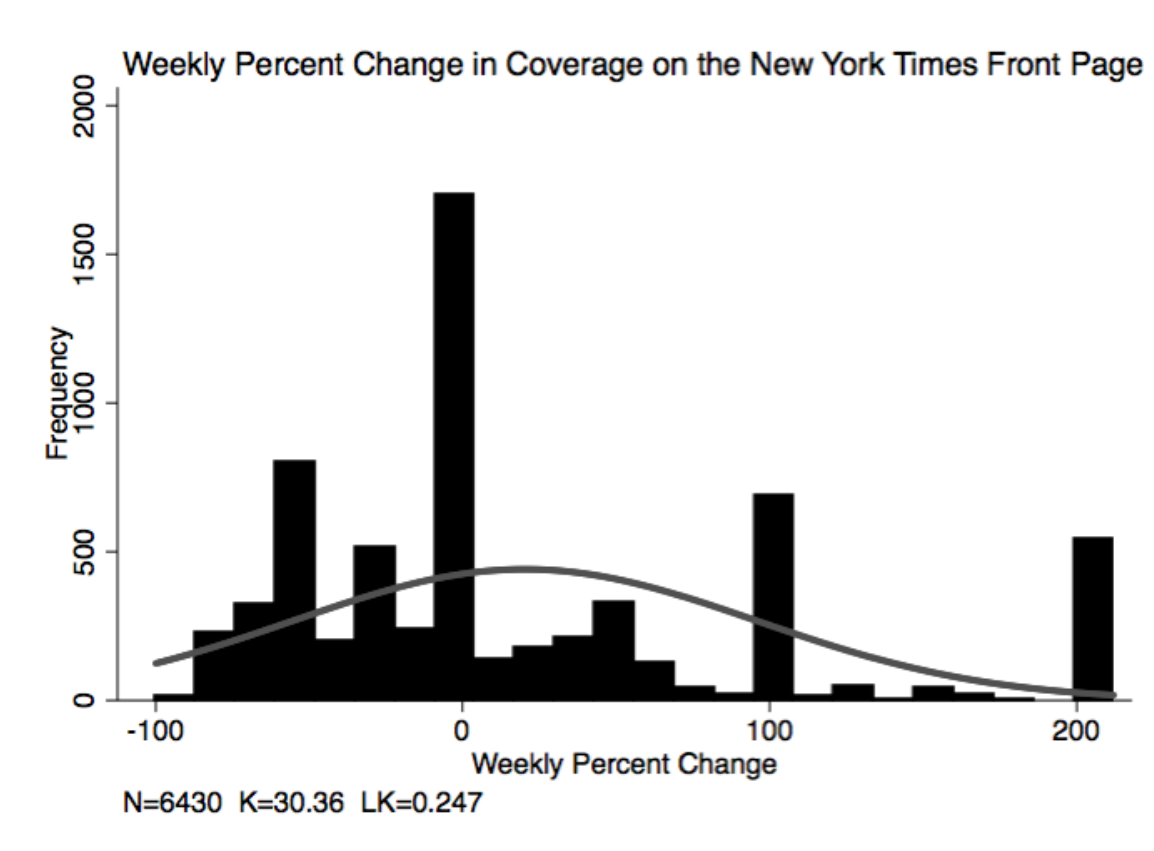
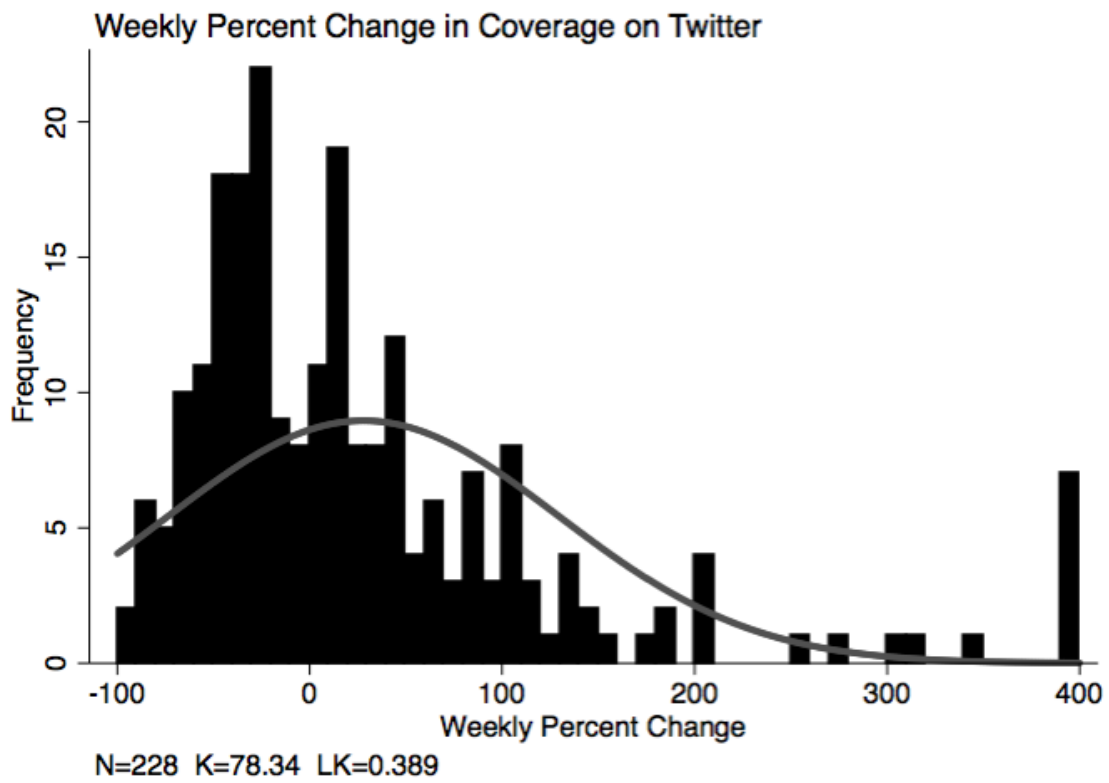


Figure 3: Frequency Distribution of Weekly Percent Change in Coverage on Twitter



References

- An, Jisun, Meeyoung Cha, Krishna Gummadi, and Jon Crowcroft. 2011. "Media Landscape in Twitter: A World of New Conventions and Political Diversity." Presented at the Fifth International AAAI Conference on Weblogs and Social Media.
- Atkinson, Mary Layton, John Lovett, and Frank R. Baumgartner. 2014. "Measuring the Media Agenda." *Political Communication*. Forthcoming.
- Baumgartner, Frank R., and Bryan D. Jones. [1993] 2009. *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.
- Baumgartner, Frank R., and Laura Chaqués Bonafont. 2014. "All News is Bad News: Newspaper Coverage in Spain." *Political Communication*. Forthcoming.
- Baumgartner, Frank R., Christian Breunig, Christoffer Green-Pedersen, Bryan D. Jones, Peter B. Mortensen, Michiel Nuytemans, and Stefaan Walgrave. 2009. "Punctuated Equilibrium in Comparative Perspective." *American Journal of Political Science* 53(3): 603-620.
- Boczkowski, Pablo J. 2009. "Technology, Monitoring, and Imitation in Contemporary News Work." *Communication, Culture & Critique* 2: 39-59.
- Boydston, Amber E. 2013. *Making the News: Politics, the Media, and Agenda-Setting*. Chicago: University of Chicago Press.
- Boydston, Amber E., Anne Hardy and Stefaan Walgrave. N.d. "Two Faces of Media Attention: Media Storms vs. General Coverage." *Political Communication*. Forthcoming.
- Boydston, Amber E., Shaun Bevan, and H.F. Thomas III. N.d. "The Importance of Attention Diversity and How to Measure It." Working Paper.
- Boydston, Amber E., Jon Moody, and Herschel F. Thomas III. 2010. "Same Day, Different Agenda? A Comparison of News Coverage Across print, Television, and Online Media Outlets." Presented at the Comparative Policy Agendas Conference, June 17-19, 2010, Bainbridge Island, WA.
- Coleman, Stephen. 1975. *Measurement and Analysis of Political Systems: A Science of Social Behavior*. New York: John Wiley & Sons, Inc.
- Farrell, Henry, and Daniel W. Drezner. 2008. "The Power and Politics of Blogs." *Public Choice* 134: 15-30.
- Grabowicz, Paul. 2014. "The Transition to Digital Journalism." UC-Berkeley School of Journalism. <http://multimedia.journalism.berkeley.edu/tutorials/digital-transform/print-editions-decline/>

- Jones, Bryan D. and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago: University of Chicago Press.
- Jones, Bryan D., Frank R. Baumgartner, Christian Breunig, Christopher Wlezien, Stuart Soroka, Martial Foucault, Abel François, Christoffer Green-Pedersen, Chris Koski, Peter John, Peter B. Mortensen, Frédéric Varone, and Stefaan Walgrave. 2009. "A General Empirical Law of Public Budgets: A Comparative Analysis." *American Journal of Political Science* 53(4): 855-873.
- Karpf, David. 2012a. "Blogosphere Authority Index." Blogosphere Authority Index. February 26. <http://www.blogosphereauthorityindex.com/> (October 15, 2012).
- Karpf, David. 2012b. *The MoveOn Effect: The Unexpected Transformation of American Political Advocacy*. New York: Oxford University Press.
- Lasora, Dominic L., Seth C. Lewis, and Avery E. Holton. 2012. "Normalizing Twitter: Journalism Practice in an Emerging Communication Space." *Journalism Studies* 13: 19-36.
- McCombs, Maxwell and Jian-Hua Zhu. 1995. "Capacity, Diversity, and Volatility of the Public Agenda: Trends from 1954 to 1994." *Public Opinion Quarterly* 59: 495-525.
- Meraz, Sharon. 2011. "Using Time Series Analysis to Measure Intermedia Agenda-setting Influence in Traditional Media and Political Blog Networks." *J&MC Quarterly* 88: 176-194.
- Nahon, Karine, Jeff Hemsley, Shawn Walker, and Muzammil Hussain. 2011. "Fifteen Minutes of Fame: The Power of Blogs in the Lifecycle of Viral Political Information." *Policy & Internet* 3(1): 2.
- Phillips, Angela. 2010. "Old Sources: New Bottles." In *New Media, Old News: Journalism & Democracy in the Digital Age*, ed. Natalie Fenton. Los Angeles: Sage.
- Pole, Antoinette. 2010. *Blogging the Political: Politics and Participation in a Networked Society*. New York: Routledge.
- Romero, Daniel M., Wojciech Galuba, Sitaram Asur, and Bernardo A. Huberman. 2011. "Influence and Passivity in Social Media." *Proc. European Conference on Machine Learning and Knowledge Discovery in Databases*.
- Shannon, C.E. 1948. "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27(3): 379-423.
- Wallsten, Kevin. 2007. "Agenda Setting and the Blogosphere: An Analysis of the Relationship Between [US] Mainstream Media and Political Blogs." *Review of Policy Research* 24(6): 567-587.

Woodly, Deva. 2008. "New Competencies in Democratic Communication? Blogs, Agenda Setting and Political Participation." *Public Choice* 134: 109-123.

Appendices

Appendix 1: Traditional Media Data

Name	Country	Date Range	Number of Stories
<i>The New York Times</i> (Front page)	United States	Jan. 1996-Dec. 2006	31,034
<i>The New York Times</i> (Index)	United States	Jan. 1946-Dec. 2008	49,126
<i>El Pais</i>	Spain	Jan. 1996-Dec. 2009	50,770
<i>El Mundo</i>	Spain	Jan. 1996-Dec. 2009	44,858
Danish radio news	Denmark	Jan. 1984-Dec. 2003	191,564
<i>The Times of London</i>	United Kingdom	Jan. 1960-Dec. 2008	21,844
<i>Neue Zürcher Zeitung</i>	Switzerland	Jan. 1996-Dec. 2003	8,558

Appendix 2: Full List of Influential Political Blogs

Ace Of Spades HQ	DownWithTyrrany!
Althouse	EconLog
AMERICAblog Gay	Economist's View
AMERICAblog News	Elder of Ziyon
American Power	Empire Burlesque
American Spectator	Eschaton
American Thinker	Fire Dog Lake
Andrew Sullivan	FiveThirtyEight
Ann Althouse	Foolocracy
Atlas Shrugs	FP Passport
Atrios	Gateway Pundit
Balkinization	GetReligion
Balloon Juice	Glenn Greenwald
Ballot Access News	Global Voices Online
Betsy's Page	GrEaT sAtAn"S gIrLfRiEnD
Big Government	Hit & Run
Bleeding Heart Libertarians	Hot Air
BLT Blog of Legal Times	Hotline On Call
Bookworm Room	Huffington Post
Business Insider	Hullabaloo
BuzzFeed	Informed Comment
Campaign for America's Future	Instapundit
Cato @ Liberty	J. Bradford DeLong's Grasping
Challah Hu Akbar	Jawa Report
CiF Watch	Jezebel
CNA Daily News	Jihad Watch
CNA Daily News-US	Joe. My. God.
CNN Political Ticker	John Redwood
ConservativeHome's Platform	Jonathan Turley
County Fair	JOSHUAPUNDIT
Creeping Sharia	Juan Cole
Crooked Timber	LA NOW
Crooks and Liars	LabourList
Daily Kos	Lawyers, Guns, and Money
Daily Pundit	Legal Insurrection
Danger Room	Little Green Footballs
Daniel Drezner	Lynn Sweet
Dean Esmay	Marathon Pundit
Democratic Underground	Matt Yglesias
Doug Ross @ Journal	Media Matters for America
Doug Wead The Blog	Mediaite

Memeorandum
Metro Weekly
Michelle Malkin
Moneybox
Mother Talkers
My DD
naked capitalism
NationalJournal Hotline On Call
New Civil Rights Movement
Newsbusters
NewsOne
Nice Deb
No More Mister Nice Blog
NYT The Caucus
OpenMarket.org
OpenSecrets Blog
Our Future
Outside The Beltway
Pandagon
Pat Dollard
Patterico's Pontifications
PinkNews.co.uk
Pirate's Cove
Policy Beta Blog
PoliPundit
Political Commentator
PoliticalWire
PoliticMo
Politics, Power, and Preventative Action
Powerline Blog
Pressure Points
Red State
Rhymes with Right
Riehl World View
Right Wing News
Right Wing Watch
Rising Hegemon
Say Anything
Scared Monkeys
SCOTUSblog
Shadow Government
Shark Tank
Simply Jews
Stephen M. Walt
Street Prophets
Sultan Knish
Taegan Goddard's Political Wire
Talk Left
Talking Points Memo
Taylor Marsh
Tbogg
Techdirt
The Agonist
The American Prospect Articles
The Baseline Scenario
The Blaze
The Cable
The Classic Liberal
The Colossus of Rhodey
The Diplomat
The Diplomat-China
The Foundry
The Incidental Economist
The Inquisitr
The Lid
The Lonely Conservative
The Long War Journal
The Mental Recession
The Moderate Voice
The New Civil Rights Movement
The Political and Financial Markets
Commentator
the sad red earth
The Shark Tank
The Volokh Conspiracy
The YES! Weekly Blog
thetorydiary
Think Progress
Threat Level
Tom Tomorrow
Towleroad News
Townhall
TPMMuckraker
Truth on the Market
TruthDig
TruthHugger

Uppity Wisconsin
Via Meadia
Virginia Right
Volokh Conspiracy
Washington Monthly
Western Journalism
ZeroHedge

White House Dossier
White House.gov Blog
Winds of Change
Wizbang Blog
Wonkette
YID With LID

Appendix 3: Two-digit Policy Agendas Topic Codes

Code	Topic
1	Macroeconomics
2	Civil Rights, Minority Issues, and Civil Liberties
3	Health
4	Agriculture
5	Labor, Employment, and Immigration
6	Education
7	Environment
8	Energy
10	Transportation
12	Law, Crime, and Family Issues
13	Social Welfare
14	Community Development and Housing Issues
15	Banking, Finance, and Domestic Commerce
16	Defense
17	Space, Science, Technology and Communications
18	Foreign Trade
19	International Affairs and Foreign Aid
20	Government Operations
21	Public Lands and Water Management
24	State and Local Government Administration
26	Weather and Natural Disasters
27	Fires
28	Arts and Entertainment
29	Sports and Recreation
30	Death Notices
31	Churches and Religion
99	Other, Miscellaneous and Human Interest

Appendix 4: Full List of Influential News Accounts on Twitter from Romero et al. (2011)

@mashable	@mckquarterly	@msnbc
@cnnbrk	@enews	@wnycradiolab
@big_picture	@nprnews	@cnnlive
@theonion	@usatoday	@davos
@time	@mtv	@planetmoney
@breakingnews	@freakonomics	@cnetnews
@bbcbreaking	@boingboing	@politico
@espn	@billboarddotcom	@tvnewser
@harvardbiz	@empiremagazine	@guardiannews
@gizmodo	@todayshow	@yahoo news
@techcrunch	@good	@seedmag
@wired	@gawker	@tvguide
@wsj	@msnbc_breaking	@travandleisure
@smashingmag	@cbsnews	@newyorkpost
@pitchforkmedia	@guardiantech	@discovermag
@rollingstone	@usweekly	@sciencenewsorg
@whitehouse	@life	
@cnn	@sciam	
@tweetmeme	@pastemagazine	
@peoplemag	@drudge_report	
@natgeosociety	@parisreview	
@nytimes	@latimes	
@lifehacker	@telegraphnews	
@foxnews	@abc7	
@waitwait	@arstechnica	
@newsweek	@cnnmoney	
@huffingtonpost	@nprpolitics	
@newscientist	@nytimesphoto	
@mental_floss	@nybooks	
@theeconomist	@nielsenwire	
@emarketer	@io9	
@engadget	@sciencechannel	
@cracked	@usabreakingnews	
@slate	@vanityfairmag	
@bbclicke	@cw_network	
@fastcompany	@bbcworld	
@reuters	@abc	
@incmagazine	@themoment	
@eonline	@socialmedia2day	
@rww	@slashdot	
@gdgt	@washingtonpost	
@instyle	@tpmmedia	

Appendix 5: Detailed Blog Information

Blog	Total Observations	HHI	N_HHI	Entropy	Total Policy Observations	Entropy-Policy Only	Percent Policy Coverage	Weekly Observations	Kurtosis	L-Kurtosis
American Thinker	240	0.16	0.13	0.73	221	0.73	92.08%	28	12.38	0.26
Andrew Sullivan	614	0.14	0.10	0.74	310	0.76	50.49%	100	8.66	0.31
Ann Althouse	795	0.12	0.09	0.76	515	0.76	64.78%	111	10.07	0.24
Atlas Shrugs	798	0.25	0.22	0.59	583	0.47	73.06%	86	27.41	0.35
Blog of Legal Times	765	0.18	0.15	0.64	687	0.63	89.80%	71	7.69	0.25
Daily Kos	564	0.31	0.28	0.56	493	0.53	87.41%	80	21.09	0.43
Digby/Hullabaloo	649	0.17	0.13	0.68	532	0.64	81.97%	21	4.92	0.24
Drezner	416	0.23	0.20	0.56	352	0.50	84.62%	33	11.20	0.40
Elder of Ziyon	790	0.67	0.66	0.23	654	0.05	82.78%	15	6.46	0.34
FiveThirtyEight	250	0.43	0.41	0.39	226	0.36	90.40%	77	12.83	0.26
Hot Air	572	0.21	0.18	0.66	477	0.66	83.39%	76	9.20	0.24
Informed Comment	357	0.27	0.24	0.59	312	0.55	87.39%	41	4.37	0.13
Lawyers, Guns, and Money	625	0.11	0.07	0.80	391	0.82	62.56%	114	55.44	0.51
Long War Journal	765	0.49	0.47	0.24	758	0.24	99.08%	28	3.16	0.15
Mental Recession	524	0.12	0.09	0.71	456	0.70	87.02%	35	9.14	0.31
Michelle Malkin	752	0.18	0.14	0.71	666	0.71	88.56%	65	4.35	0.17
New Civil Rights Movement	715	0.21	0.18	0.64	579	0.55	80.98%	66	5.09	0.18
Newsbusters	795	0.14	0.11	0.74	646	0.70	81.26%	117	6.12	0.25
Patterico's Pontifications	330	0.20	0.17	0.63	285	0.59	86.36%	40	4.57	0.23
Redstate	624	0.17	0.14	0.70	494	0.72	79.17%	92	16.95	0.41
Sultan Knish	343	0.15	0.12	0.68	265	0.62	77.26%	28	5.37	0.26
Talking Points Memo	520	0.19	0.16	0.61	485	0.60	93.27%	54	8.23	0.20
The Cable	416	0.48	0.46	0.35	405	0.34	97.36%	25	4.16	0.21
The Lonely Conservative	495	0.15	0.12	0.73	459	0.76	92.73%	77	5.41	0.21
The Tory Diary	420	0.94	0.94	0.05	408	0.00	97.14%	14	6.56	0.30
TruthDig	781	0.11	0.07	0.77	622	0.78	79.64%	121	28.68	0.33
Uppity Wisconsin	402	0.65	0.63	0.28	396	0.28	98.51%	20	16.82	0.70
Virginia Right	324	0.11	0.07	0.75	268	0.72	82.72%	22	3.77	0.19
Yid With the Lid	215	0.14	0.11	0.68	181	0.65	84.19%	34	4.57	0.16
Combined	15,856	0.11	0.08	0.77	13,126	0.75	82.78%	303	44.14	0.36
Mean	547	0.26	0.24	0.59	453	0.57	84.00%	58.31	11.19	0.28
Median	564	0.18	0.15	0.66	459	0.63	84.62%	54	7.69	0.25
Min	215	0.11	0.07	0.05	181	0.00	50.49%	14	3.16	0.13
Max	798	0.94	0.94	0.80	758	0.82	99.08%	121	55.44	0.70

Appendix 6: Detailed Twitter Information

Account (@)	Total Observations	HHI	N_HHI	Entropy	Total Policy Observations	Entropy-Policy Only	Percent Policy Coverage	Weekly Observations	Kurtosis	L-Kurtosis
ABC	740	0.11	0.08	0.77	495	0.73	66.89%	104	4.44	0.14
big_picture	427	0.23	0.20	0.52	138	0.26	32.32%	20	12.52	0.45
breakingnews	760	0.22	0.19	0.65	559	0.51	73.55%	93	6.23	0.21
CNN	715	0.11	0.08	0.77	502	0.67	70.21%	103	12.72	0.27
cnmney	708	0.11	0.08	0.76	635	0.77	89.69%	99	6.14	0.23
davos	193	0.13	0.09	0.68	149	0.64	77.20%	10	3.36	0.37
ESPN	674	0.86	0.85	0.10	2	0.23	0.30%	10	3.25	0.27
gizmodo	705	0.26	0.23	0.51	415	0.43	58.87%	58	9.35	0.29
instyle	723	0.64	0.63	0.20	8	0.25	1.11%	22	10.00	0.33
io9	685	0.30	0.27	0.46	167	0.49	24.38%	48	4.91	0.23
latimes	738	0.08	0.04	0.83	464	0.81	62.87%	126	9.04	0.21
life	622	0.23	0.20	0.56	167	0.66	26.85%	39	8.46	0.25
mental_floss	693	0.25	0.22	0.59	197	0.77	28.43%	75	42.28	0.34
newscientist	679	0.22	0.19	0.58	455	0.53	67.01%	59	13.67	0.39
newyorkpost	744	0.14	0.11	0.71	394	0.66	52.96%	92	6.96	0.16
nprnews	719	0.09	0.05	0.83	513	0.79	71.35%	124	24.55	0.29
nytimesphoto	620	0.23	0.20	0.58	325	0.51	52.42%	24	8.68	0.43
politico	713	0.34	0.32	0.55	325	0.51	45.58%	78	13.26	0.24
rollingstone	692	0.84	0.83	0.14	29	0.61	4.19%	19	4.65	0.28
slashdot	740	0.23	0.20	0.62	654	0.59	88.38%	90	11.07	0.22
socialmedia2day	726	0.53	0.51	0.34	184	0.57	25.34%	33	12.82	0.41
telegraphnews	740	0.43	0.41	0.47	623	0.34	84.19%	62	8.93	0.26
todayshow	694	0.19	0.16	0.61	138	0.67	19.88%	51	38.67	0.58
tvguide	711	0.90	0.89	0.09	9	0.42	1.27%	15	3.32	0.27
usweekly	684	0.79	0.78	0.16	19	0.30	2.78%	25	6.29	0.29
whitehouse	639	0.08	0.05	0.82	468	0.82	73.24%	51	4.35	0.14
wired	724	0.17	0.13	0.65	393	0.62	54.28%	69	9.07	0.17
yahoonews	695	0.11	0.07	0.78	465	0.70	66.91%	102	9.96	0.29
Combined	18,903	0.11	0.08	0.77	9216	0.80	48.75%	228	78.34	0.39
Mean	675	0.32	0.29	0.55	318	0.57	47.23%	60.75	11.03	0.29
Median	706.5	0.23	0.20	0.59	359	0.60	53.62%	58.5	8.98	0.27
Min	193	0.08	0.05	0.09	2	0.23	0.30%	10	3.25	0.14
Max	760	0.90	0.89	0.83	654	0.82	89.69%	126	42.28	0.58