

Agenda-setting in the one-step flow: Evidence from Facebook in the 2012 election

Deen Freelon

School of Communication, American University

Abstract: Agenda-setting is one of the best-substantiated theories in communication research. But as the digital age has reoriented the media landscape, it has called into question some of the theory's most basic underlying assumptions. The mass media, once by far the most powerful and prominent agenda-setting force at work, must now compete with other interests wishing to set their own agendas. Politicians, once relegated to communicating with citizens through mass media, can now connect with them directly via social media. Such a direct, targeted connection is known as the *one-step flow* of communication. This study examines how agenda-setting works under one-step flow conditions by applying a lexicon analysis to Barack Obama's and Mitt Romney's Facebook posts and the comments appended to them during the 2012 US presidential campaign (N = 858,307). Results indicate that Romney's agenda is more similar to his audience than Obama's is to his, both audiences are more interested in religion than either candidate, and the audiences emphasize the same attributes as the candidates for some issues but different attributes for others.

Social media and the prospects for expanded democratic participation in national policy-setting

workshop, Boston University

April 9, 2015

Agenda-setting in the one-step flow: Evidence from Facebook in the 2012 election

Politicians can communicate with citizens in two ways. They can take the indirect route by funneling their thoughts to citizens through the news media, which may alter them in unanticipated ways. Alternatively, politicians are increasingly turning to digital media to communicate directly with citizens, cutting the news media out of the equation entirely. The rise in popularity of such disintermediated or “one-step flow” communication pathways (Bennett & Iyengar, 2008; Blumler & Kavanagh, 1999; Katz, 1988) introduces intriguing new possibilities for well-known political communication theories.

Agenda-setting is one of the best-substantiated theories in communication research (McCombs, 2005). Its original formulation assumed the existence of a more or less monolithic news media whose decisions about story salience would be reflected in perceptions of issue importance among audience members. Recent work on agenda-setting in the digital age has continued to treat news media as central (Neuman, Guggenheim, Jang, & Bae, 2014; Vargo, Guo, McCombs, & Shaw, 2014). Less explored has been the potential for politicians to set their constituents’ agendas without the participation of the media. Audience members may to varying degrees follow along the agenda path set by their leaders, or attempt to inject their own pet concerns onto the agenda. Digital media have given them unprecedented opportunities for such agenda injection, but few studies have investigated how often they actually do so.

The current study provides evidence in answer to these questions by examining over 850,000 messages posted to Barack Obama’s and Mitt Romney’s official Facebook pages during the 2012 US presidential campaign. It uses a lexicon-based approach to compare the issue agendas of both candidates to those of their followers. Results indicate substantial agenda overlap between all parties with several consequential exceptions.

Agenda-setting, past and present

Agenda-setting is truly a theory that needs no introduction, having been tested in hundreds of academic studies since its classic formulation in the early 1970s (McCombs & Shaw, 1972). Whether performed by the mass media, politicians, or ordinary citizens, agenda-setting is ultimately concerned with power: the power to influence which issues are discussed and acted upon and which are ignored. In the 20th century agenda-setting power was primarily vested in the mass media and various elites. Research has focused on how the media set audience agendas (McCombs & Shaw, 1972), how political elites set media agendas (Entman, 2009), how media sources set each other's agendas (Tedesco, 2001, 2005), how nonelites set one another's agendas (Ragas & Roberts, 2009; Shaw, McCombs, Weaver, & Hamm, 1999), and how all these parties influence perceptions of issue attributes (Ghanem, 1997; Weaver, 2007).

Since the beginning, the mass media have been the premier agents in setting nonelite agendas. While many agenda items originate with politicians and others considered to be politically newsworthy, they have traditionally been transmitted to citizens through journalistic intermediaries (Bennett, Lawrence, & Livingston, 2007; Entman, 2009; Hayes & Guardino, 2010). Most agenda-setting research of the pre-digital age incorporates one or more mass media outlets as objects of study (Weaver, McCombs, & Shaw, 2004). As the digital age has matured, the media sphere has grown more crowded, but recent rumors of the end of mass media relevance (Castells, 2007; Katz & Scannell, 2009) have been somewhat exaggerated (Aday et al., 2013; Shehata & Strömbäck, 2013; Vargo et al., 2014). Money and name recognition still account for a substantial share of the agenda-setting capacity at work in the American political sphere, as evidenced by the record-breaking campaign expenditures of the last few US presidential elections (Center for Responsive Politics, n.d.).

Still, recent research has recognized that agenda-setting power is dispersed among a much greater diversity of parties than at any time in history. Journalists now jockey for audience attention with bloggers, celebrities, activists, corporations, and academics, among others. The distinction between "horizontal" (general-interest) and "vertical" (niche) media categorizes outlets according to their

intended audiences, but in the process elides critical distinctions within the categories (Ragas & Roberts, 2009; Vargo et al., 2014). For example, those amateurs referred to as “gatewatchers” (Bruns, 2005), “grassroots intermediaries” (Jenkins, Ford, & Green, 2013), and “networked gatekeepers” (Meraz & Papacharissi, 2013) occupy a particular stratum of horizontal media positioned between those who mainly create content and those who mainly consume it. By curating content from a variety of primary sources, these nonelite individuals and groups help to set their audience’s agendas.

The One-step flow

Politicians and parties have obvious reasons for wanting to set their constituents’ agendas—if nothing else, it is an essential task in the campaigning process. The media obviously play a major role in this as explained above, but politicians have always had direct communication pathways to citizens, including campaign advertisements, media spectacles, televised speeches, and election debates. Digital media have given them even more ways to reach out to the electorate, with social media being one of the most popular of these. Nearly every elected member of Congress and the president have accounts on both Facebook and Twitter (Shpayher, n.d.). These accounts offer politicians opt-in, persistent, unfiltered, and interactive communication channels between themselves and their followers that have no close analogues in the pre-internet era. From the politician’s perspective, the purpose of such channels is to cultivate a “controlled interactivity” that mobilizes supporters to spread preapproved messages to one another (Stromer-Galley, 2014). Yet politicians enjoy no monopoly on power in these spaces, and agenda contestation between elites and nonelites is both possible and measurable, as we shall see.

The general phenomenon of unfiltered, targeted communication between political elites and citizens has long been known to political communication scholars, who label it the “one-step flow” of communication (Bennett & Manheim, 2006) or “disintermediation” (Aday et al., 2013; Katz, 1988). It does not mean that news media outlets have been rendered completely irrelevant, but rather that new,

digitally-mediated communication pathways now exist between politicians and their supporters. And unlike prior direct pathways, social media platforms afford both two-way communication and convenient archiving of message content. Thus, they allow for communications between politicians and citizens to be analyzed in greater detail than ever before. Few existing studies have examined how politicians set citizens' agendas directly, in part because of the overwhelming preeminence of the mass media in the 20th-century agenda-setting process.

It should be noted that the concept of agenda-setting undergoes a slight transformation as it transitions to a one-step flow/social media context. Critically, the "agenda" of which we speak must be reconstructed based on participant conversations rather than survey self-reports. Instead of asking individuals directly about the issues they believe to be most important, the current analysis assumes that the issues they discuss most often are most important to them. This method has the advantage of avoiding observer- or social desirability biases that might accompany survey questions about issue importance. Its main disadvantage is a lack of external validity stemming from the unrepresentativeness of social media users, a point that will be explored further in the discussion.

The prospect of agenda-setting in a one-step flow context raises a new set of empirical questions. One might first ask about the extent to which different politicians actually set the agenda(s) of their audiences. Much has been made of ordinary citizens' technologically-enhanced abilities to set their own agendas with less intervention from mass media outlets (Castells, 2007; Shaw et al., 1999). Politically-interested citizens may be even more empowered to do so than the median citizen given their subject-matter expertise and strong motivations to participate in politics. It is therefore not certain that all politicians will be able to set agendas equally effectively in online one-step flow contexts. Individual politicians may differ in their agenda-setting abilities just as distinct media outlets do (Golan, 2006; Walgrave & Van Aelst, 2006).

- RQ1: To what extent do Barack Obama and Mitt Romney set the agendas for their respective social media followers?

An important related question addresses what might be called the “agenda gaps” between politicians and audiences. These are the differential foci of the two respective interests on each issue: for example, Romney might be more concerned with the economy than his audience, who might care more about religion (which happens to be the case). On issues for which the gap is small, candidate and audience might be said to concur on issue importance; but major differences show us exactly where their respective agendas diverge. This phenomenon has been noted in previous research (e.g. in Neuman et al., 2014) but its theoretical relevance has not been adequately considered. For example, issues that are popular among audiences but not candidates may evince “agenda melding” or agenda-setting between nonelites (Ragas & Roberts, 2009; Shaw et al., 1999).

- RQ2: Which issues do the candidates and their followers emphasize proportionately and disproportionately?

Third, there is the question of how second-level agenda attributes differ between candidates and audiences within a single first-level issue. It is quite possible that candidates may discuss issues in ways that differ fundamentally from how audiences discuss them. This may be one reason certain candidates at times find themselves accused of being “out of touch” with the electorate. The second level of agenda-setting is another domain within which audience members may exercise the power to resist elite agenda-setting or even to set the terms of the debate among themselves. Neuman et al. (2014) briefly broach this possibility but do not specifically address the elite/audience distinction, nor is their lexicon sufficiently comprehensive to capture a broad range of audience concerns. Hence the current extension of their work.

- RQ3: For the most prominent issues, to what extent do the candidates discuss the same issue attributes as their followers?

Methods

Lexicon-based approaches have a long history in the agenda-setting literature (Neuman et al., 2014; Tedesco, 2001, 2005). It is a fitting approach to examine one-step agenda-setting in social media because of the ready availability of an extensive textual record of communication between citizens and politicians. While previous studies have inferred agenda-setting processes on the basis of time-lagged correlations between source and target texts (e.g. Neuman et al., 2014; Tedesco, 2001, 2005), social media offer key affordances that render such inferences unnecessary. When sources (politicians in this case) post to their social media accounts, followers can append comments directly to the original messages. Therefore, any correspondence in subject matter between original messages and replies is assumed to be the result of a top-down agenda-setting process. Of course, citizens who follow politicians on social media may also attempt to steer the agenda toward their own concerns, especially because they know that attaching their comments to messages posted by well-known individuals is an effective means of attracting attention.

Agenda salience is measured by the prevalence of keywords related to political issues. As such, we can measure not only the extent to which candidates set the agendas of their followers, but also which issues are relatively more and less relevant to each candidate compared to his audience. Moreover, we can compare the specific terms used by each candidate and his audience to examine qualitatively which issue attributes each discusses most commonly. This expands the analysis onto the second level of agenda-setting (Ghanem, 1997; Weaver, 2007), which is concerned with the specific attributes of broader issues (e.g., if the issue is the economy, specific aspects could include taxes, GDP, debt, and jobs). This study's methods thus afford an examination of not only which issues candidates and their followers discuss, but also of how they discuss them.

The data to be analyzed were collected from Facebook. They include all messages posted to Barack Obama's and Mitt Romney's official Facebook pages between April 25, 2012 (the day the

Republican National Committee officially endorsed Romney) and November 5, 2012 (the day before the election).¹ During this period the Obama campaign posted 268 messages while Romney's posted 584. While both campaigns allowed only official messages to be posted to each account's Facebook wall, visitors could append comments to these messages. All publicly-available comments appended to the official candidate messages posted during the study period were collected on November 10, 2012 from the Facebook Graph API using a custom PHP script developed by the author.² In total, 233,129 Obama comments and 624,326 Romney comments were collected. The combined N of posts (official messages and comments) analyzed in this study is 858,307.

Previous lexicon-based agenda-setting research has used prepackaged lexicons (Tedesco, 2001; Vargo et al., 2014) or short, ad-hoc keyword lists (Neuman et al., 2014), but such techniques are less than ideal. Because social media users may use a wide array of slang, neologisms, and other distinctive language, fully-automated or "canned" methods may miss much of theoretical interest (Petchler & Gonzalez-Bailon, forthcoming). Therefore, this study's lexicon analysis used as a starting point the Lexicoder Policy Agendas lexicon (LPA), an open-access list of 1,402 public policy-related terms grouped into 28 issue categories (Albaugh, Sevenans, & Soroka, 2013; Albaugh, Sevenans, Soroka, & Loewen, 2013). Although the LPA contained many terms relevant to the present case, a cursory comparison between it and the data revealed that many essential terms were lacking. To remedy this, the author and a research assistant each read through 1,004 Facebook posts—306 candidate messages and 698 user comments—in search of additional terms to add to the LPA. These posts had been coded as containing issue-relevant content by other coders as part of an earlier, unpublished project. Those coders analyzed all official candidate messages and 3,080 randomly-sampled comments from the same larger dataset used in this study. Thus, the subset of messages re-examined for issue terms in this study represents all those from the previous project's sample that were judged to mention one or more political issues.

Each of the new lexicon terms culled from the Facebook posts was added to the most appropriate issue category. Terms that did not unambiguously denote a single category were discarded (including a few from the original LPA), and terms sharing a single word stem were merged (e.g. “taxpayer” and “taxes” were combined as “tax”). In addition, several of the original LPA categories were merged for the sake of parsimony, resulting in a final total of 1,476 terms across 18 categories (see Appendix A for a complete list of categories and terms). Counting all new term additions, mergers, truncations, and spelling corrections, a net total of 305 terms were added to the original LPA.

Once the categories were finalized, the data were preprocessed for analysis. Rigorous data preprocessing is extremely important in lexicon analysis, as its absence may result in unacceptable incidences of Type I and/or II error (Petchler & Gonzalez-Bailon, forthcoming). First, all the multi-word lexicon terms were concatenated with underscores in the data so that they would not be broken apart by the tokenization process. Next, each post was tokenized, or separated into its component terms and stripped of punctuation (excluding underscores), leaving all relevant multi-word terms remaining intact. Most of the tokenized terms were then lemmatized (reduced to their most canonical forms) using the NLTK package for the Python programming language (e.g. instances of “taxed” were changed to “tax”). A few terms were not lemmatized because their non-canonical forms or tenses had distinct political meanings (see Appendix B for a complete list). A representative example is the term “illegals”: the plural is often used by conservatives to refer to undocumented immigrants, but the lemmatized singular “illegal” does not share this meaning. In addition to the lemmatization process, certain lexicon terms were designated to be stemmed, or converted to a specific reduced form that may or may not be an actual word (such as “racis” which matched “racist,” “racists,” and “racism”). This list of words to stem was created manually because existing general-purpose stemmers did not consistently produce the desired stems. The complete custom stem list can be found in Appendix C.³

After the data were preprocessed in the above manner, the lexicon analysis could begin. Upon completion, it yielded a set of per-post counts for each issue category and term. That is, the lexicon output contains the number of posts in which each candidate and his commenters 1) mentioned each individual term and 2) mentioned at least one term from each issue category. These counts form the basis of the empirical analysis that follows.

Results

Before we address the research questions, let us first consider some of the more relevant descriptive statistics. In total, 81 of Obama's 268 posts (30.2%) contained at least one issue term while 292 (50.0%) of Romney's 584 did. Similarly, Obama's commenters discussed policy issues proportionally less often (80,795 posts of 233,129, 34.6%) than Romney's did (266,335 posts of 624,326, 42.7%). Thus both Romney and his followers were more focused on policy than their Democratic counterparts, at least on Facebook. Of course, many posts mentioned more than one term or issue.

Figures 1a-d reveal each subset's relative issue emphasis. Each chart's y-axis indicates the proportion of all issue category mentions. Figure 1a, for example, shows that Obama's two most-mentioned issues were civil rights and finance/economics/labor (FEL), followed by healthcare, with no other issue exceeding 10% of all issue mentions. In Figure 1b we see that Romney's messages were dominated by economic concerns and that the next most popular issue, healthcare, accounted for only 7% of issue mentions. Unsurprisingly given their much larger post volumes, the commenters' issue mentions were more evenly distributed than the candidates', although this was more the case for Obama's commenters than Romney's (Figures 1c-d). Both were most interested economics with religion and civil rights following close behind, as the remaining issues gently tapered off.

RQ1

The first research question asks about how effectively each candidate set his commenters' respective agenda. To test this, each group of commenters' issue proportions were entered as a dependent

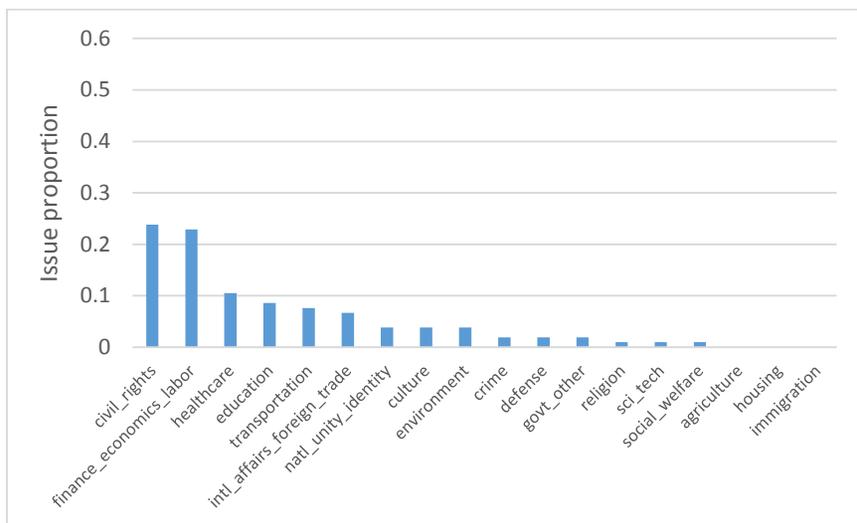


Figure 1a: Obama's issue proportions

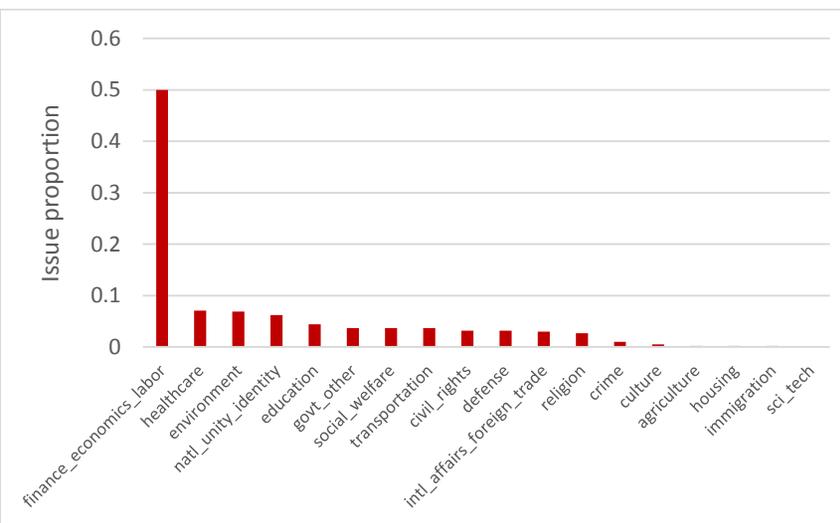


Figure 1b: Romney's issue proportions

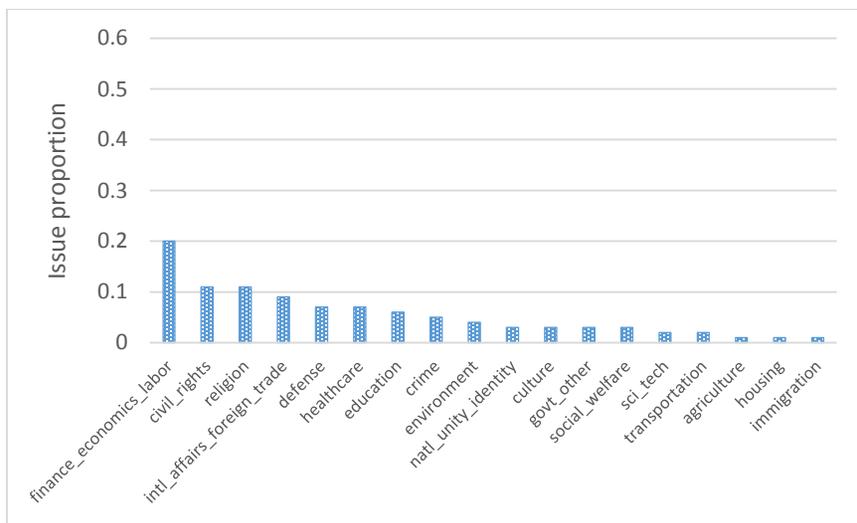


Figure 1c: Obama commenters' issue proportions

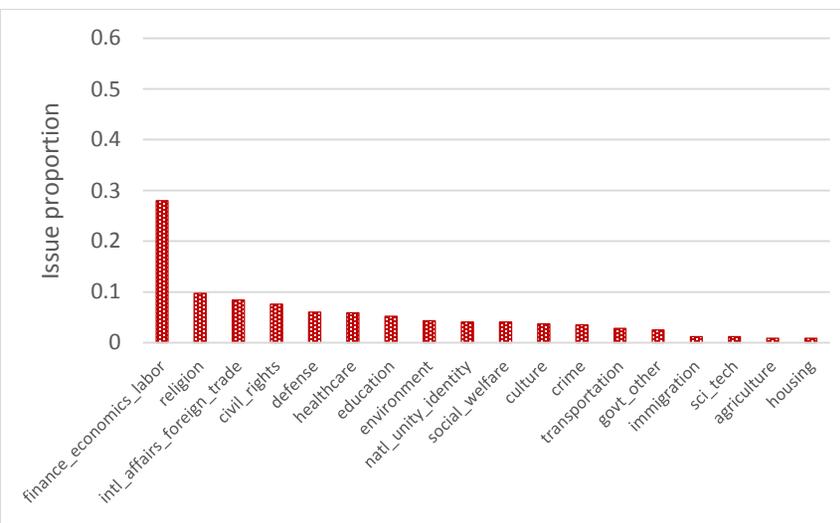


Figure 1d: Romney commenters' issue proportions

variable into two OLS regressions: one using only the corresponding candidates' issue proportions as the sole predictor, and the other adding the remaining two issue proportion subsets as predictors. This hierarchical approach includes the focal test of each candidate's agenda-setting power over his audience in two separate regressions to lower the likelihood that the results are spurious. The unit of analysis in all four regressions is the issue category ($n = 18$).

Table 1a contains the output of the regressions of Obama's commenters, while Table 1b contains those for Romney's commenters. Each table's Model 1 includes the candidate's issue proportions as the sole predictor, while Model 2 adds the remaining two subsets as controls. For the current purposes the standardized regression coefficients are the appropriate quantities to compare between models and candidates; unstandardized coefficients are included for the sake of completeness.

<i>Predictor</i>	Model 1			Model 2		
	B	Beta	SE	B	Beta	SE
Obama issues	0.515***	0.762	0.109	0.12**	0.178	0.039
Romney issues	.	.	.	-0.303***	-0.708	0.047
Romney commenters' issues	.	.	.	1.161***	1.474	0.094
Constant	0.026	.	0.01	0.001	.	0.003
R ²	0.581***			0.976***		
ΔR ²	.			0.395***		

Table 1a: OLS regressions of Obama commenters' issue proportions.

* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

<i>Predictor</i>	Model 1			Model 2		
	B	Beta	SE	B	Beta	SE
Romney issues	0.505***	0.928	0.051	0.276***	0.508	0.023
Obama issues	.	.	.	-0.082*	-0.095	0.036
Obama commenters' issues	.	.	.	0.790***	0.622	0.064
Constant	0.0276	.	0.006	0.001	.	0.003
R ²	0.861***			0.99***		
ΔR ²	.			0.129***		

Table 1b: OLS regressions of Romney commenters' issue proportions.

* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$

In answer to RQ1, comparisons between Obama's Model 1 and 2 betas and the corresponding coefficients in Romney's models reveal that the latter was a much more effective agenda-setter. Interestingly, the addition of the other subsets as predictors substantially increases the overall predictive power of Model 2 in both cases. The Model 2s also show that the two audiences' agendas were more similar to each other than either was to its corresponding candidate, although this phenomenon is more pronounced in Obama's case than in Romney's.

RQ2

To address the question of which issues the two candidates and their respective audiences emphasized differentially, the corresponding issue proportions between each pair were subtracted and graphed (Figures 2a and 2b). Both figures clearly show that the disparities in issue interest between each pair of subsets is heavily concentrated among a few issues: for the vast majority, the disparity falls below five percentage points. But the exceptions are worth exploring further.

Obama talks about his most-emphasized issue, civil rights, proportionally much more than his audience does. As the next section shows, this is primarily due to his strong interest in women's issues. His commenters, on the other hand, are disproportionately more interested in matters of religion, which the president mentions only once in all of his 268 posts. While Romney discusses religion somewhat more than Obama, his commenters care about it even more than he does. And although both the governor and his audience mention economic terms more than any other category, the former's interest exceeds the latter's by 22 percentage points.

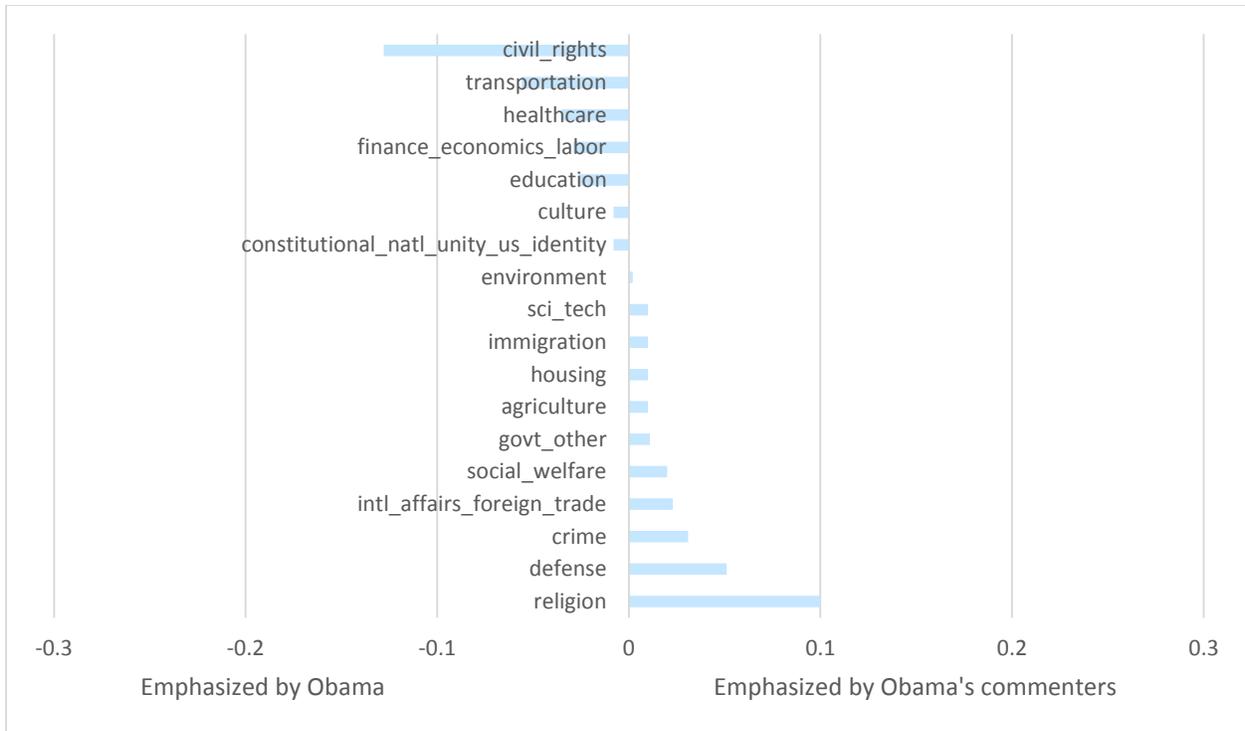


Figure 2a: Agenda gaps between Obama and his commenters

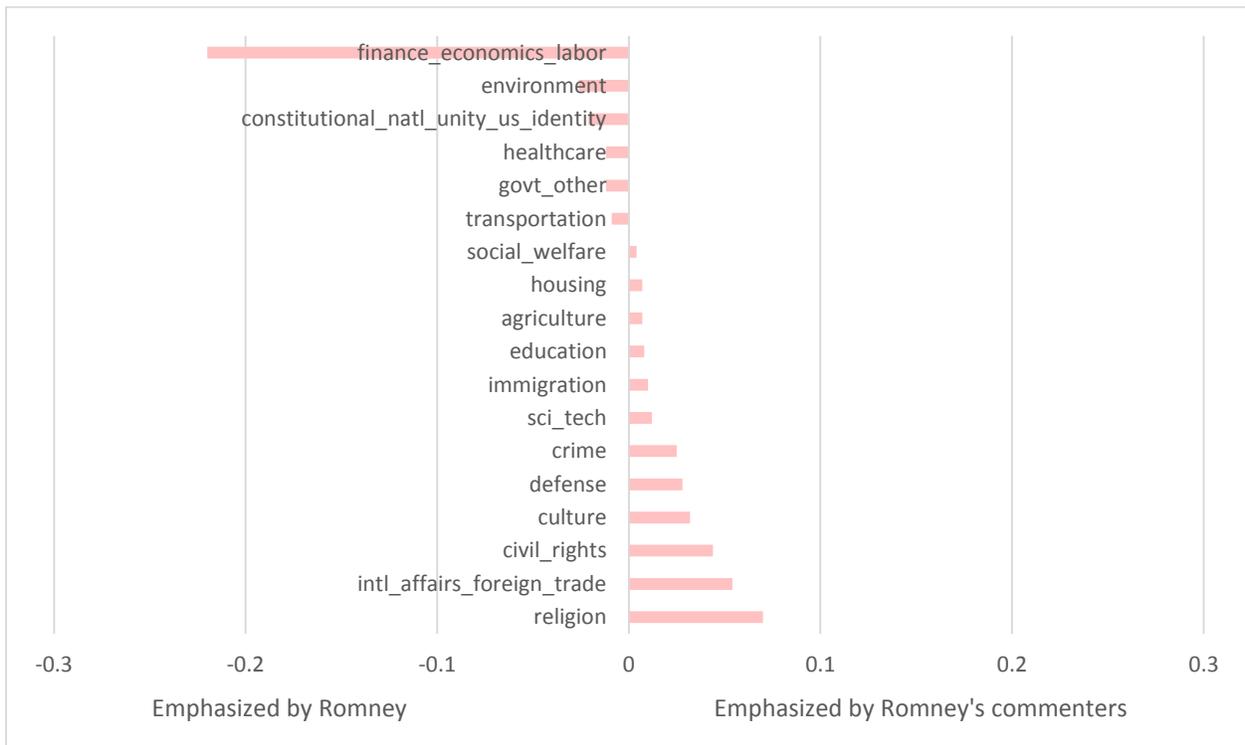


Figure 2b: Agenda gaps between Romney and his commenters

RQ3

The final research question concerns term usage differences within issue categories between the subsets. Because space limitations prohibit an exhaustive exploration of all issues, this section will examine the four most-frequently-mentioned issues across all four subsets: FEL, civil rights, healthcare, and religion. For each of these issues, Table 2 presents each subset’s top five most-used terms.

Perhaps the most readily apparent pattern in this table is the strong similarity of terms within most issue categories. This is probably most clear for FEL and healthcare, within which many of the same terms are repeated across subsets (“job,” “tax,” “obamacare,” “medicare”). Those terms that are not repeated do not seem to indicate a consistent focus on fundamentally different issue attributes than those that are.

Issue	Rank	Obama	O comments	Romney	R comments
Civil rights	1	woman	woman	woman	woman
	2	daughter	rights	mother	rights
	3	mother	daughter	rights	racis
	4	latino	abortion	latino	daughter
	5	voter_registration	racis	voting_right	abortion
Finance/ economics/ labor (FEL)	1	job	job	job	job
	2	econom	tax	econom	tax
	3	tax	econom	middle_class	econom
	4	middle_class	dollar	tax	middle_class
	5	private_sector	debt	unemploy	dollar
Healthcare	1	health	health	obamacare	obamacare
	2	obamacare	obamacare	medicare	health
	3	medicare	medicare	health	medicare
	4	copay	medical	nurs	medical
	5	preventive	doctor	medicin	doctor
Religion	1	faith	god	prayer	god
	2	-	muslim	faith	muslim
	3	-	islam	religio	religio
	4	-	religio	god	islam
	5	-	christian	catholic	mormon

Table 2: Ranked terms for four issue topics across candidate and audience data subsets

Civil rights represents a partial departure from this pattern. While all three subsets are substantially focused on women’s issues—almost certainly a result of candidate agenda-setting—the

two candidates are in near-total lexical lockstep with one another, with nearly every term having an obvious analogue on the opposing side. But while the commenters share this concern, they also both decisively inject abortion and racism onto the agenda item's second level, two issue attributes neither candidate mentions at all. The current methods do not permit any claims about which sides of each issue are most popular, but it is interesting to see the commenters raising issues the candidates would prefer to avoid.

Nowhere is this agenda injection phenomenon more evident than on the issue of religion. Obama makes only one passing reference to it, in an exhortation to his supporters to "keep the faith," which carries at best only a vague religious connotation. Romney mentions religion more often, but usually with nonspecific terms like "prayer" and "god," with the exception of "catholic" which he uses only once. Meanwhile, the commenters do not hesitate to discuss the specifics of religions that played controversial roles in the campaign, especially Islam. Romney's religion of Mormonism was a popular topic among his commenters, albeit less popular than Islam. Overall, these differences in how and how much audiences and candidates discuss religion speak volumes about which aspects of the topic they feel are most consequential.

Discussion

This study offers a summary impression of the issue agendas expressed by two presidential candidates and their social media audiences during the 2012 US presidential campaign. Its main findings are as follows: first, the fit between Romney's agenda and that of his audience is closer proportionally than for Obama. This outcome, combined with the pride of place and time precedence of candidate messages as compared to audience comments, offers strong evidence that Romney set his Facebook audience's agenda more effectively than Obama set his. Of course, the outcome of the 2012 election demonstrates that agenda-setting power on Facebook is neither necessary nor sufficient to win. Part of the disparity

may be explained by the fact that Romney's audience was simply more engaged on policy issues overall, as their greater proportion of issue mentions compared to Obama's audience attests.

The second major finding is that each candidate's agenda generally resembled his audiences' in terms of proportional emphasis with a few conspicuous exceptions. Obama's strong interest in civil rights, and specifically women's issues, far outstripped that of his audience. Their top concern was FEL issues, which also topped the lists of the other two subsets. While Obama spent almost as much time discussing both issues, his audience felt that FEL was almost twice as important as civil rights. For his part, Romney's concern with the economy exceeded his audience's, which also prioritized it over all other issues. Both audiences raise the issue of religion in greater proportions than their respective candidates. This is a clear example of what might be called "agenda injection"—an attempt to introduce an unsanctioned topic into a conversation ostensibly about something else. And while it is unlikely that this injection attempt noticeably affected the candidates' agendas, it may have played some role in shifting the agendas of audience members and their Facebook followers.

We see further evidence of agenda injection within some issues in the results of the RQ3 analysis. On the issues of FEL and healthcare, the difference between the four subsets is minimal—the same terms appear time and again, and unique terms do not indicate a fundamental shift in focus onto a different attribute of the issue. This is also true to some extent for civil rights; however, the presence of "racis" and "abortion" on both audience lists shows a willingness of nonelites to court controversies the candidates would prefer to ignore. Specifically on the topic of women's issues, it is telling that the candidates speak in nebulous terms of "women's health" and "women's issues," while the audience directly references the specific procedure that most concerns them: abortion. On religion the candidates again remain either vague (Romney) or silent (Obama), in contrast to the audience, which invokes specific religious identities they feel are germane to the campaign. In recent American campaign history,

strategic vagueness is a well-known technique through which candidates avoid offending key constituencies (Carey, 1997). With much less at stake, audiences are bound by no such rhetorical fetters.

The candidate/audience issue disparities reported here are of course particular to this empirical case. Of greater interest is the possibility that the methods described here may allow us to generalize about how candidate and audience agendas converge and diverge. The very preliminary finding that citizens are willing to address controversy and specificity to a much greater extent than candidates accords with prior research on presidential campaigning (e.g. Carey, 1997, p. 232; Stromer-Galley, 2000). But additional research is necessary to substantiate this point more completely, and to discover other potential differences between agenda items and attributes emphasized differentially between the two.

The methods used in this study are particularly relevant for exploring the role of digital media in setting national policy agendas. Lexicon analysis is a highly apt technique for determining what people are talking about online, and can be applied to millions of messages just as easily as thousands. If the data are representative, the issues that rise to the top of the agenda can be considered to legitimately reflect the priorities of a broader population. However, representativeness is rarely guaranteed with online observational data: although 71% of Americans are regular Facebook users (Duggan, Ellison, Lampe, Lenhart, & Madden, 2015), this study's audience members almost certainly come from a much more politically engaged subpopulation than the median user. Even so, this group is likely more inclusive than a comparable Twitter sample, since only 23% of Americans use that service regularly (Duggan et al., 2015). The main challenge in identifying national citizen agendas is procuring nationally representative data—the analytical methods themselves are comparatively robust and intuitive.

Aside from the general value of lexicon analysis, this study also offers several methodological advances over prior lexicon-based agenda-setting research. Such studies should always strive to include as many relevant terms as possible, as every relevant term omitted raises the incidence of Type II error by an unknown degree. This is especially important in studies of social media, where linguistic

innovation advances at an unrelenting pace (Eisenstein, O'Connor, Smith, & Xing, 2014), rendering one-size-fits-all lexicons unacceptably incomplete. While the current method of qualitatively identifying relevant terms from a representative sample of messages does not guarantee completeness, it offers a replicable balance between comprehensiveness and achievability. It is likely that many of the most-used (and thus most important) issue terms will appear at least once in such a sample. Qualitative data inspection is also essential to identify terms that should and should not be stemmed and/or lemmatized based on the context in which they are used. Canned lexical analysis solutions are insensitive to context and therefore risk inappropriately including or excluding terms whose different forms may have different meanings.

This study's main conclusion that agenda-setting occurs in one-step flow contexts is perhaps not especially surprising. Yet due to the historical role of the mass media as chief agenda-setter for the public, the capacity of candidates to set public agendas directly has been under-studied. The increasing importance of digital media in political communication will hopefully prompt researchers in the area to address this empirical oversight. Additional questions remain, such as: how do politicians' agenda-setting powers compare with that of the mass media? To what extent do the differences discovered here between incumbent and challenger (or Democrat and Republican) generalize to other contexts? What kinds of individuals are most and least susceptible to politician agenda-setting, and who is most likely to attempt and succeed at agenda injection? The current study underscores the importance of these questions and presents a solid empirical foundation upon which future studies can build to answer them.

References

- Aday, S., Farrell, H., Freelon, D., Lynch, M., Sides, J., & Dewar, M. (2013). Watching From Afar: Media Consumption Patterns Around the Arab Spring. *American Behavioral Scientist*, *57*(7), 899–919.
doi:10.1177/0002764213479373
- Albaugh, Q., Sevenans, J., & Soroka, S. (2013). Lexicoder Topic Dictionaries. McGill University. Retrieved from <http://www.lexicoder.com/download.html>
- Albaugh, Q., Sevenans, J., Soroka, S., & Loewen, P. J. (2013). The Automated Coding of Policy Agendas: A Dictionary-Based Approach. Presented at the 6th annual Comparative Agendas Project (CAP) conference, Antwerp.
- Bennett, W. L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication*, *58*(4), 707–731.
- Bennett, W. L., Lawrence, R. G., & Livingston, S. (2007). *When the press fails: Political power and the news media from Iraq to Katrina*. Chicago: University of Chicago Press.
- Bennett, W. L., & Manheim, J. B. (2006). The One-Step Flow of Communication. *The ANNALS of the American Academy of Political and Social Science*, *608*(1), 213–232.
doi:10.1177/0002716206292266
- Blumler, J. G., & Kavanagh, D. (1999). The third age of political communication: Influences and features. *Political Communication*, *16*(3), 209–230.
- Bruns, A. (2005). *Gatewatching: Collaborative online news production*. New York: Peter Lang.
- Carey, J. W. (1997). The press, public opinion, and public discourse: On the edge of the postmodern. In E. S. Munson & C. A. Warren (Eds.), *James Carey: A critical reader* (pp. 228–260). Minneapolis: University of Minnesota Press.
- Castells, M. (2007). Communication, power and counter-power in the network society. *International Journal of Communication*, *1*(1), 238–266.

- Center for Responsive Politics. (n.d.). The Money Behind the Elections. Retrieved February 21, 2015, from <https://www.opensecrets.org/bigpicture/>
- Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015, January 9). Social Media Update 2014. Retrieved from <http://www.pewinternet.org/2015/01/09/social-media-update-2014/>
- Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2014). Diffusion of Lexical Change in Social Media. *PLoS ONE*, *9*(11), e113114. doi:10.1371/journal.pone.0113114
- Entman, R. M. (2009). *Projections of Power: Framing News, Public Opinion, and U.S. Foreign Policy: Framing News, Public Opinion, and U.S. Foreign Policy*. University of Chicago Press.
- Ghanem, S. (1997). Filling in the tapestry: The second level of agenda setting. *Communication and Democracy: Exploring the Intellectual Frontiers in Agenda-Setting Theory*, 3–14.
- Golan, G. (2006). Inter-media agenda setting and global news coverage. *Journalism Studies*, *7*(2), 323–333. doi:10.1080/14616700500533643
- Hayes, D., & Guardino, M. (2010). Whose Views Made the News? Media Coverage and the March to War in Iraq. *Political Communication*, *27*(1), 59–87. doi:10.1080/10584600903502615
- Jenkins, H., Ford, S., & Green, J. (2013). *Spreadable Media: Creating Value and Meaning in a Networked Culture*. New York: NYU Press.
- Katz, E. (1988). Disintermediation: Cutting out the middle man. *Intermedia*, *16*(2), 30.
- Katz, E., & Scannell, P. (Eds.). (2009). *The end of television? Its impact on the world (so far)*. Thousand Oaks, CA: Sage.
- McCombs, M. (2005). A Look at Agenda-setting: past, present and future. *Journalism Studies*, *6*(4), 543–557. doi:10.1080/14616700500250438
- McCombs, M., & Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly*, *36*(2), 176–187. doi:10.1086/267990

- Meraz, S., & Papacharissi, Z. (2013). Networked Gatekeeping and Networked Framing on #Egypt. *The International Journal of Press/Politics*. doi:10.1177/1940161212474472
- Neuman, W. R., Guggenheim, L., Jang, S. M., & Bae, S. Y. (2014). The Dynamics of Public Attention: Agenda-Setting Theory Meets Big Data. *Journal of Communication*, 64(2), 193–214. doi:10.1111/jcom.12088
- Petchler, R., & Gonzalez-Bailon, S. (forthcoming). Automated content analysis of online political communication. In S. Coleman & D. Freelon (Eds.), *Handbook of Digital Politics*. Northampton, MA: Edward Elgar.
- Ragas, M. W., & Roberts, M. S. (2009). Agenda Setting and Agenda Melding in an Age of Horizontal and Vertical Media: A New Theoretical Lens for Virtual Brand Communities. *Journalism & Mass Communication Quarterly*, 86(1), 45–64. doi:10.1177/107769900908600104
- Shaw, D. L., McCombs, M., Weaver, D., & Hamm, B. J. (1999). Individuals, Groups, and Agenda Melding: A Theory of Social Dissonance. *International Journal of Public Opinion Research*, 11(1), 2–24. doi:10.1093/ijpor/11.1.2
- Shehata, A., & Strömbäck, J. (2013). Not (Yet) a New Era of Minimal Effects A Study of Agenda Setting at the Aggregate and Individual Levels. *The International Journal of Press/Politics*, 18(2), 234–255. doi:10.1177/1940161212473831
- Shpayher, J. (n.d.). United States House of Representatives - Congress Social Media Wiki. Retrieved from <http://govsm.com/w/House>
- Stromer-Galley, J. (2000). On-line interaction and why candidates avoid it. *Journal of Communication*, 50(4), 111–132. doi:10.1111/j.1460-2466.2000.tb02865.x
- Stromer-Galley, J. (2014). *Presidential campaigning in the Internet age*. New York: Oxford University Press.

- Tedesco, J. C. (2001). Issue and Strategy Agenda-Setting in the 2000 Presidential Primaries. *American Behavioral Scientist*, 44(12), 2048–2067. doi:10.1177/00027640121958483
- Tedesco, J. C. (2005). Issue and Strategy Agenda Setting in the 2004 Presidential Election: exploring the candidate–journalist relationship. *Journalism Studies*, 6(2), 187–201.
doi:10.1080/14616700500057270
- Vargo, C. J., Guo, L., McCombs, M., & Shaw, D. L. (2014). Network Issue Agendas on Twitter During the 2012 U.S. Presidential Election. *Journal of Communication*, 64(2), 296–316.
doi:10.1111/jcom.12089
- Walgrave, S., & Van Aelst, P. (2006). The Contingency of the Mass Media’s Political Agenda Setting Power: Toward a Preliminary Theory. *Journal of Communication*, 56(1), 88–109.
doi:10.1111/j.1460-2466.2006.00005.x
- Weaver, D. (2007). Thoughts on agenda setting, framing, and priming. *Journal of Communication*, 57(1), 142–147.
- Weaver, D., McCombs, M., & Shaw, D. L. (2004). Agenda-setting research: Issues, attributes, and influences. *Handbook of Political Communication Research*, 257–282.

Appendixes TK!

Notes

¹ <http://facebook.com/barackobama> and <http://facebook.com/mittromney>, respectively.

² An unknown number of comments were omitted from the dataset because the commenters' privacy settings prevented them from being collected.

³ The discerning reader may wonder why some tokens were lemmatized but not stemmed. In short, the lemmatization process was intended to collapse different forms of the same word together (e.g. "income" and "incomes"). The only terms that were stemmed were those that shared a word stem in common, referenced the same political issue, and would not match irrelevant terms (e.g. the stem "econom" was used to match terms like "economy" and "economics," but "incom" was not used as a stem because it would also match the irrelevant term "incoming"). Lemmatization only produces true words, not non-word stems.