



2018

# Exploring Thematic Diversity In News Coverage And Social Media Activity Of Political Candidates Using Unsupervised Machine Learning

Dror Walter

*University of Pennsylvania*, [drorvalt@gmail.com](mailto:drorvalt@gmail.com)

Follow this and additional works at: <https://repository.upenn.edu/edissertations>

 Part of the [Communication Commons](#)

---

## Recommended Citation

Walter, Dror, "Exploring Thematic Diversity In News Coverage And Social Media Activity Of Political Candidates Using Unsupervised Machine Learning" (2018). *Publicly Accessible Penn Dissertations*. 2763.

<https://repository.upenn.edu/edissertations/2763>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/2763>

For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

# Exploring Thematic Diversity In News Coverage And Social Media Activity Of Political Candidates Using Unsupervised Machine Learning

## **Abstract**

The relationship between media and politics has been at the core of communication research for over a century. Previous research has examined the impact of both volume and tone of news coverage of political candidates on their electoral success, and the relationship between the volume of candidates' social media activity (though not its tone) and electoral success. While past research found a positive relationship between these features and electoral success, recent criticisms have called into question the independent nature of these media factors. Moreover, while past research has paid some attention to volume and tone, researchers have yet to examine other key features of discourse represented in candidates' coverage as a whole. One such feature is the extent to which a political discourse is unidimensional or multidimensional in nature, referred to in this study as thematic diversity. This is due, in part at least, to the complex nature of thematic diversity making its estimation challenging.

Analyzing over 120,000 Tweets written by 142 U.S. Senate candidates during the 2012-2016 election cycles, as well as over 420,000 news articles covering 330 U.S. Senate candidates during the 2008-2016 election cycles, this study systematically explores the relationship between electoral success of political candidates and the volume and tone of their news coverage and social media activity. Using a wide array of controls, this study explores the independent (or dependent) nature of these media features. More importantly, this study goes beyond these previously studied media features, to systematically and empirically explore the relationship between thematic diversity in both candidates' news coverage and social media activity, and their electoral success.

Drawing on the conceptualization of diversity in various fields from biology, to physics and information sciences, and using two unsupervised machine learning methods, semantic network analysis and topic modeling, this study offers a novel approach to the conceptualization and estimation of thematic diversity, accounting for the variety, balance and disparity of various themes in a given corpus. Using these methods, this study offers evidence for a significant, negative, and semi-independent relationship between thematic diversity and electoral success, in both news media and social media.

## **Degree Type**

Dissertation

## **Degree Name**

Doctor of Philosophy (PhD)

## **Graduate Group**

Communication

## **First Advisor**

Michael X. Delli Carpini

---

**Keywords**

News, Political Campaigns, Semantic Network Analysis, Social Media, Thematic Diversity, Topic Modeling

**Subject Categories**

Communication

EXPLORING THEMATIC DIVERSITY IN NEWS COVERAGE AND SOCIAL  
MEDIA ACTIVITY OF POLITICAL CANDIDATES USING UNSUPERVISED  
MACHINE LEARNING

Dror Walter

A DISSERTATION

in

Communication

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

---

Michael X. Delli Carpini, Ph.D.

Professor of Communication and Walter H. Annenberg Dean

Graduate Group Chairperson

---

Joseph Turow, Ph.D., Robert Lewis Shayon Professor of Communication

Dissertation Committee

Joseph N. Cappella, Ph.D., Gerald R. Miller Professor of Communication

Sandra González-Bailón, Ph.D., Assistant Professor of Communication

EXPLORING THEMATIC DIVERSITY IN NEWS COVERAGE AND SOCIAL  
MEDIA ACTIVITY OF POLITICAL CANDIDATES USING UNSUPERVISED  
MACHINE LEARNING

COPYRIGHT

2018

Dror Walter

## ACKNOWLEDGMENTS

This dissertation is dedicated to my grandfather. Though he never had the opportunity of attending a higher education institution he always strived to be a life-long learner. My intellectual development has been the result of his dedication, support and influence.

I was extremely fortunate to have Dr. Michael Delli Carpini as an advisor. This project was developed in our weekly meetings over three years. His dedication and deep involvement with this project in its various iteration is evidenced in every part of this dissertation. This project would not be possible without the help of my committee members, Dr. Joseph Cappella and Dr. Sandra González-Bailón. It's through working with them that this research project was initially formed, and their suggestions and attention to the smallest details improved this dissertation greatly.

I would also like to thank other Annenberg faculty members whose courses and mentoring shaped my scholarly work. I am especially thankful to Dr. Krippendorff, Dr. Moehler, Dr. Mutz, and Dr. Allison. In addition, I owe a debt of gratitude to the Annenberg School staff. Without them I would not be able to tackle the challenges of this research program. Among many, I am thankful to the administrative staff for their help (especially Kelly Fernandez and Joanne Murray), the IT staff (especially Etienne Jacquot and Tejash Patel) and the Library staff (especially Sharon Black).

I would not be able to attend Annenberg in the first place without the help of my previous mentors at the Hebrew University of Jerusalem and Ben-Gurion University. I am grateful for Dr. Tamir Sheafer who guided me in my first major research project

introducing me to the nature of academic work in general, and in communication specifically. I would also like to thank Dr. Zvi Reich from whom I learned how to theorize and strategize massive research projects. Lastly, I would like to thank Dr. Lilach Nir. It was through her courses that I was first introduced to the fields of communication and political psychology. Her support enabled me to apply for PhD programs in the United States, and it is her guidance that provided me with the tools to do so.

Coming to a new country, it was my colleagues in Annenberg who made me feel immediately welcomed. I am especially thankful to my cohort, Alex Williams and Allie Hallmark, the first Americans we met in this country, to Tim and Dana Libert for the long evenings and road-trips, to Aaron Shapiro for his help with this dissertation and the hourly breaks, to Yotam Ophir for numerous lunches, and to Natacha Yazbeck, Nour Halabi, and Eleanor Marchant for their friendship and endless hours of political arguments. I also thank Sijia Yang, Josh Becher and Devon Brackbill for their help and critical advice at the various stages of this project. I was extremely lucky to have all of you, as well as the other students in Annenberg as friends and colleagues.

I am forever grateful to my mother and father for their unwavering support. Their encouragement and investment during almost four decades is what pushed me towards higher education and pursuing a PhD degree, which will be the first in my family. Lastly, to Tania my wife, who followed me here and supported me throughout this challenging process, as well as many other past challenges. Without you this adventure would not have been even remotely possible. After 14 (or 18) years together, I am excited about our new adventures. I hope the next five years would be even half as wonderful and fun as the last five.

## ABSTRACT

### EXPLORING THEMATIC DIVERSITY IN NEWS COVERAGE AND SOCIAL MEDIA ACTIVITY OF POLITICAL CANDIDATES USING UNSUPERVISED MACHINE LEARNING

Dror Walter

Michael X. Delli Carpini

The relationship between media and politics has been at the core of communication research for over a century. Previous research has examined the impact of both volume and tone of news coverage of political candidates on their electoral success, and the relationship between the volume of candidates' social media activity (though not its tone) and electoral success. While past research found a positive relationship between these features and electoral success, recent criticisms have called into question the independent nature of these media factors. Moreover, while past research has paid some attention to volume and tone, researchers have yet to examine other key features of discourse represented in candidates' coverage as a whole. One such feature is the extent to which a political discourse is unidimensional or multidimensional in nature, referred to in this study as thematic diversity. This is due, in part at least, to the complex nature of thematic diversity making its estimation challenging.

Analyzing over 120,000 Tweets written by 142 U.S. Senate candidates during the 2012-2016 election cycles, as well as over 420,000 news articles covering 330 U.S. Senate candidates during the 2008-2016 election cycles, this study systematically explores the relationship between electoral success of political candidates and the volume and tone of their news coverage and social media activity. Using a wide array of controls, this study explores the independent (or dependent) nature of these media features. More importantly, this study goes beyond these previously studied media features, to systematically and empirically explore the relationship between thematic diversity in both candidates' news coverage and social media activity, and their electoral success.

Drawing on the conceptualization of diversity in various fields from biology, to physics and information sciences, and using two unsupervised machine learning methods, semantic network analysis and topic modeling, this study offers a novel approach to the conceptualization and estimation of thematic diversity, accounting for the variety, balance and disparity of various themes in a given corpus. Using these methods, this study offers evidence for a significant, negative, and semi-independent relationship between thematic diversity and electoral success, in both news media and social media.



## TABLE OF CONTENTS

ACKNOWLEDGMENTS .....	iii
ABSTRACT.....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	x
1. INTRODUCTION .....	1
1.1 Thematic Diversity in Political Campaigns .....	2
1.2 Structure of the Dissertation .....	7
2. THEORETICAL FRAMEWORK .....	14
2.1 News Media and Elections.....	16
2.1.1 News Media and Elections: Volume.....	19
2.1.2 News Media and Elections: Tone .....	24
2.2 Social Media and Elections.....	28
2.2.1 Social Media and Elections: Volume.....	35
2.2.2 Social Media and Elections: Tone .....	41
2.3 Tone and Volume: Limitations .....	46
2.4 Thematic Diversity.....	50
2.4.1 Thematic diversity in elections: Monothematic message strategy .....	63
2.4.2 Thematic diversity in elections: Multi-thematic message strategy.....	68
3. METHODOLOGICAL FRAMEWORK .....	75
3.1 Challenges in Thematic Diversity Measurement.....	76
3.2 Thematic Diversity - Conceptual Considerations .....	84
3.3 Unsupervised Machine Learning Approaches to Thematic Diversity .....	91
3.3.1 Topic modeling .....	92
3.3.2 Diversity in topic models .....	97
3.3.3 Semantic network analysis.....	102
3.3.4 Diversity in semantic networks.....	107
4. METHOD .....	116
4.1 Non-Media Data.....	117
4.2 Media Data Gathering.....	120
4.3 Media Measures: Volume and Sentiment.....	123

4.4 Measuring Diversity.....	128
4.4.1 Topic Modeling.....	128
4.4.2 Semantic Networks Analysis .....	140
4.4.2.1 Constructing the networks for the news data.....	141
4.4.2.2 Constructing the networks for the Twitter data. ....	149
4.4.2.3 Community structure and diversity estimation.....	152
4.4.2.3 Using random networks as benchmark for diversity estimation.....	156
4.5 Statistical Modeling Strategy .....	161
5. RESULTS .....	163
5.1 News Coverage and Electoral Success .....	164
5.2 Social Media Activity and Electoral Success .....	179
5.3 The Antecedents of Thematic Diversity .....	189
5.4 Correlation and Multicollinearity Analysis .....	192
6. SUMMARY AND DISCUSSION.....	203
6.1 Volume and Tone: News .....	203
6.2 Volume and Tone: Social Media .....	206
6.3 Thematic Diversity: News .....	207
6.4 Thematic Diversity: Social Media .....	214
6.5 Thematic Diversity and Electoral Success – Possible Mechanisms .....	217
6.6 Summary of Contributions.....	220
6.7 Limitations and Future Directions .....	223
6.8 It’s the Structure.....	230
REFERENCES .....	233

## LIST OF TABLES

Table 1: Simplified topic-document matrix .....	93
Table 2: Simplified topic-word matrix .....	94
Table 3: Hypothetical topic-document matrix for two candidates and two topics.....	97
Table 4: Hypothetical topic-word matrix for two candidates and two topics .....	99
Table 5: A hypothetical adjacency matrix (13x13).....	110
Table 6: The top 5 negative and positive tweets identified in the corpus (non-text data removed).....	128
Table 7: Summary of multiple regression analysis of news media factors on Republicans vote share.....	169
Table 8: Summary of multiple regression analysis of news media factors and non-media factors on Republicans vote share .....	172
Table 9: Summary of multiple regression analysis of news media factors and non-media factors on Republicans vote share with relative news coverage as volume indicator.....	175
Table 10: Summary of multiple regression analysis of news factors and non-media factors on Republicans vote share .....	177
Table 11: Summary of multiple regression analysis of social media factors on Republicans vote share .....	184
Table 12: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share .....	186
Table 13: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share (using random network benchmark for diversity estimation) .....	188
Table 14: Summary of multiple regression analysis of non-media factors on thematic diversity in candidates' news coverage. ....	190
Table 15: Summary of multiple regression analysis of non-media factors on thematic diversity in candidates' social media activity.....	191
Table 16: A correlation matrix for the non-media and news media factors used in the regression models predicting electoral success. ....	193
Table 17: A correlation matrix for the news media thematic diversity factors used to predict electoral success. ....	194
Table 18: A correlation matrix for the non-media and social-media factors used in the regression models predicting electoral success. ....	195
Table 19: A correlation matrix for the social media thematic diversity factors used to predict electoral success. ....	<b>Error! Bookmark not defined.</b>
Table 20: Summary of multiple regression analysis of social media factors and non-media factors on candidate vote share, separated by party .....	197

Table 21: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share (using natural log transformation for the topic modeling diversity estimation).....	199
Table 22: A correlation matrix for the social media thematic diversity factors used to predict electoral success, with outlier observations removed .....	200
Table 23: Summary of multiple regression analysis of social media factors and non-media factors on candidate vote share, separated by party .....	201
Table 24: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share (using topic modeling for diversity estimation).....	202

## LIST OF FIGURES

Figure 1: A Hypothetical corpus thematic structure. ....	81
Figure 2: A hypothetical topic distribution for two candidates (3 and 5 topic structure). 86	86
Figure 3: A hypothetical topic distribution for two candidates (5 topic structure).....	87
Figure 4: A hypothetical topic distribution for two candidates (3 topic structure).....	88
Figure 5: A network representation of the hypothetical simplified semantic network... 111	111
Figure 6: A network representation of hypothetical and simplified semantic networks for two candidates (Community 1 in yellow, community 2 in red, community 3 in gray). ....	113
Figure 7: The distribution of volume of news coverage (as number of articles) received by each candidate (2008-2016). ....	124
Figure 8: The distribution of volume of social media activity (measured as the number of tweets shared) for all candidates 2012-2016. ....	125
Figure 9: The distribution of mean sentiment in news coverage per candidate. ....	127
Figure 10: The distribution of mean sentiment in Twitter activity per candidate. ....	127
Figure 11: Model fit indicators for models ranging from k=10 to k=100 (by 10).....	130
Figure 12: Semantic coherence and semantic exclusivity per k level for k=10 to k=100 (by 10). ....	131
Figure 13: Model fit indicators for models ranging from k=50 to k=70 (by 1).....	133
Figure 14: Semantic coherence and semantic exclusivity per k level for k=50 to k=70 (by 1).....	133
Figure 15: Model fit indicators for models ranging from k=20 to k=200 (by 20).....	137
Figure 16: Semantic coherence and semantic exclusivity per k level for k=20 to k=200 (by 20). ....	137
Figure 17: Model fit indicators for models ranging from k=100 to k=140 (by 5).....	138
Figure 18: Semantic coherence and semantic exclusivity per k level for k=100 to k=140 (by 5). ....	139
Figure 19: Co-occurrence matrices for the sentence “Ant Bat Cat Dog Eel Fox Goat Hog Ibis Ant,” using a 2-word window for co-occurrence. ....	144
Figure 20: Change In top (x) co-occurring word dyads as a function of window size and length of top dyads list (x).....	146
Figure 21: The term-document matrix for the three short sentences (“Ant Bat Cat Dog,” “Dog Eel Fox Goat,” “Goat Hog Ibis Ant”)......	150

Figure 22: The term-document matrix for the three short sentences (“Ant Bat Cat Dog,” “Dog Eel Fox Goat,” “Goat Hog Ibis Ant”).	151
Figure 23: The normalized co-occurrence matrix for the three short sentences (“Ant Bat Cat Dog,” “Dog Eel Fox Goat,” “Goat Hog Ibis Ant”) and the corresponding semantic network.	152
Figure 24: Relationship between the modularity score of the networks drawn from the Twitter dataset and the diversity estimation using Rao’s coefficient (Pearson’s $r=0.78$ ).	155
Figure 25: Relationship between the number of communities in the networks drawn from the Twitter dataset and the diversity estimation using Rao’s coefficient (Pearson’s $r=0.8$ ).	156
Figure 26: A scatterplot for the diversity estimations for the observed and random-generated networks.	159
Figure 27: A histogram for the values of (Observed Diversity – Random Diversity)....	160
Figure 28: A scatterplot for the diversity gap between observed and random-generated networks, and the number of articles from which data was drawn.	160
Figure 29: Summary of linear regression analysis for social media variables predicting Republican candidates’ electoral success, with different (k) topic model per facet (N=165).	164
Figure 30: A sample of 9 topics from the final topic model (k=69).	166
Figure 31: The semantic network for candidate Al Franken drawn from the candidate’s news coverage during the 2008 Senate Elections (size of nodes is determined by weighted degree; color of node represents community membership; Layout created using the Force Atlas algorithm).	167
Figure 32: The correlation between the vote share won by a republican candidate and the candidate’s absolute (left) and relative (right) volume of news coverage.	174
Figure 33: Summary of linear regression analysis for social media variables predicting Republican candidates’ electoral success, with different (k) topic model per facet (N=77).	179
Figure 34: A sample of 9 topics from the final topic model. Size and color indicate word frequency (k=69).	180
Figure 35: The semantic network for candidate Bernie Sanders drawn from the candidate’s social media activity during the 2012 Senate Elections (Size of nodes is determined by weighted degree; color of node represents community membership; Layout created using the Atlas Force algorithm).	182

Figure 36: The semantic network for candidate Ted Cruz drawn from the candidate’s social media activity during the 2012 Senate Elections (Size of nodes is determined by weighted degree; color of node represents community membership; Layout created using the Fruchman-Reingold Algorithm). ..... 183

Figure 37: Scatter plot of the thematic diversity for the Democratic and Republican candidates ..... 198

## 1. INTRODUCTION

A few seconds before the credits roll at the end of the documentary “The War Room,” which follows the 1992 Clinton presidential campaign, the camera fades from the election-night victory party to the now-deserted campaign headquarters in Little Rock, Arkansas. The camera pans slightly and focuses on a whiteboard at the center of the room. On the whiteboard, under “Rules,” three sentences are listed: “Change vs. more of the same,” “Don't forget healthcare,” and “The economy, stupid.”

In the 1992 U.S. Presidential Elections, candidate Bill Clinton (then governor of Arkansas) was competing against George H. W. Bush in a campaign built on criticism of the struggling economy under the incumbent Bush, while dodging allegations raised weekly regarding Clinton’s foreign affairs expertise, draft dodging, and inappropriate relationships with various women (Jamieson, 1996). The three “rules” presented prominently on the whiteboard at the campaign headquarters were aimed at reminding those working for the campaign to stay focused on these specific issues at all times. In a phone call during the last week of the race, documented in “The War Room,” campaign advisor James Carville can be seen imploring Clinton: “Stay focused! Talk about things that matter to people. You know? It’s the economy stupid, OK?” In essence, Carville’s strategy was to connect every possible message opportunity to this theme, so as to take advantage of the struggling economy.

Although the campaign had three main foci, the phrase “it’s the economy stupid”—and the idea of focusing on a single issue—is the one that entered the American political lexicon. The phrase has been cited countless times, spoofed in popular culture, repeated



over and over by media pundits, and used in the titles of numerous scholarly works in various fields. While the economy might not always be the one key issue, this emphasis on one specific theme for a campaign is similar to the common belief held by political strategists, consultants and researchers that it is important to keep the campaign message coherent, succinct, and as unidimensional as possible (Benoit et al., 2011; Bradshaw, 2004; Conway III et al., 2012a). But is it true? The systematic and empirical examination of this common wisdom is at the core of this study.

### **1.1 Thematic Diversity in Political Campaigns**

Drawing on various fields of research, from biological diversity (Solow, Polasky, & Broadus, 1993), to agenda diversity (Chaffee & Wilson, 1976; Peter & De Vreese, 2003; Tan & Weaver, 2013) and frame complexity (Kleinnijenhuis, Schultz, & Oegema, 2015), this study examines the extent to which a political discourse, such as campaign strategy or news coverage, is unidimensional or multidimensional, which I refer to as “thematic diversity.” Thematic diversity describes the variety and interconnectivity of themes in a given corpus, whether they are issues, actors, or viewpoints on a subject (Kleinnijenhuis et al., 2015). Most simply, thematic diversity can be seen as nothing more than the number of topics present in a specified discourse. However, borrowing from conceptualizations of diversity from other scientific fields, from biology to physics and information sciences (Stirling, 2007), one can argue that the concept of thematic diversity requires more than just asking *how many* themes are presented in a corpus (i.e., variety). Rather, it includes the distribution of these themes over the discourse (balance), and the extent to which themes contained in the same discourse differ from one another

(disparity). I elaborate more on these components in the methodological framework chapter.

While some research on the role of thematic diversity in election campaigns does exist, it suffers from several limitations that will be addressed in this dissertation. Much of the writing on “It’s the economy, stupid” (or simply “staying on message,” as it is often referred to) relies on case studies and anecdotal evidence. This is due, in part at least, to the complex definition of thematic diversity, which makes its estimation challenging, especially when trying to capture thematic diversity in large and varied corpora and when no reliable a priori lists of topics for the specific context exist for a given corpus. Additionally, the systematic research that does exist on the role of thematic diversity (Benoit et al., 2011; Bradshaw, 2004; Sellers, 1998) is largely limited to national, as opposed to state-level, election campaigns. Existing research has also focused on more traditional forms of political advertising, with much less attention paid to the role of thematic diversity in either campaign news coverage or in newer types of direct communication, such as candidates’ social media activity.

This dissertation addresses these shortcomings in several ways. First, it extends the current literature on the relationship between electoral success and monothematic versus multi-thematic campaign message strategies in traditional political advertising to candidates’ coverage by the news media. While the electoral impact of campaign news coverage has been studied a great deal, these studies focus largely on the *volume* and *tone* of this coverage (Bélanger & Soroka, 2012; De Vreese, 2010; Hopmann, Vliegenthart, De Vreese, & Albæk, 2010; Norris, Curtice, Sanders, Scammell, & Semetko, 1999) rather than on the diversity of coverage. This research has shown that volume of

coverage and the tone of coverage (positive or negative) can impact candidate electoral success (Balmas & Sheaffer, 2010; Boomgaarden, Vliegenthart, & de Vreese, 2012; Coleman & Wu, 2010; Eberl, Boomgaarden, & Wagner, 2017; Geers & Bos, 2017; Geiß & Schäfer, 2017; Hopmann et al., 2010; Johann, Kleinen-von Königslöw, Kritzinger, & Thomas, 2017; Kioussis, Mitrook, Wu, & Seltzer, 2006; Lengauer & Johann, 2013; M. McCombs, Llamas, Lopez-Escobar, & Rey, 1997; Norris et al., 1999; Oegema & Kleinnijenhuis, 2000).

Recent criticism has called into question the extent to which these relationships are truly independent, rather than the result of a spurious or mediated process related to other non-media factors (Bélanger & Soroka, 2012). This dissertation assesses the electoral impact of thematic diversity in news coverage relative to both the volume and tone of coverage. It does so while controlling for a large host of non-media factors, thus providing additional evidence for the nature of these relationships and exploring the factors that shape thematic diversity in candidates' news coverage.

Second, this study fills a gap in the literature regarding the relationship between candidates' activity on social media and their electoral success. With the rising importance of social media as a campaign tool in the past decade (Stromer-Galley, 2014), researchers have started paying closer attention to how candidates use online tools for campaigning (Bright et al., 2018; Jungherr, 2016). Extant research has also examined the within-campaign processes that shape candidates' online behavior, as well as differences in the use of online tools between different candidates based on various candidate-level and race-level factors (Evans, Cordova, & Sipole, 2014; Gilmore, 2012; Jungherr, 2016; Peterson, 2012; Vergeer, Hermans, & Sams, 2013). In addition, a barrage of studies in

recent years has attempted to predict elections from the general Twitter “chatter” regarding the political arena (Beauchamp, 2017; Gayo-Avello, 2013; Jungherr, 2016; Tumasjan, Sprenger, Sandner, & Welppe, 2010)—that is, to see whether the number of mentions a candidate receives on the platform and the nature of these mentions can impact electoral success.

However, only a handful of studies have explored the relationship between candidate social media activity and candidate electoral success (Bright et al., 2018; Jungherr, 2016), and most of these focus on mere usage as a predictor. Very few studies have examined the impact of volume of usage (Bright et al., 2018; LaMarre & Suzuki-Lambrecht, 2013; Vergeer, Hermans, & Sams, 2011), and even these were limited in terms of context, time-frame, and lack of appropriate controls. Moreover, I was unable to locate *any* study examining the relationship between sentiment (or tone) in candidate social media activity and candidate electoral success, or the impact of thematic diversity and message strategy on candidate electoral success. Thus, the impact of candidates’ social media activity remains poorly understood (Jungherr, 2016). This dissertation addresses this shortcoming by assessing the electoral impact of thematic diversity in candidates’ social media use, relative to that of both volume and tone, and doing so while controlling for a large host of non-media factors (Beauchamp, 2017) and exploring the factors that shape thematic diversity in candidates’ social media activity.

As mentioned above, the lack of systematic empirical research on thematic diversity might be the result of challenges in the operationalization, conceptualization, and measurement of thematic diversity. The traditional measurement of thematic diversity requires considerable resources and some hard decisions about the level of

resolution at which thematic separation takes place. Traditional measures struggle with accounting for variety, balance, and disparity at once, and often require an a priori understanding of each of the corpora being compared. To address this, the third contribution of this dissertation is methodological: it conceptualizes and estimates thematic diversity using two different unsupervised machine learning methods. These methods do not require an a priori assumption of possible themes, issues, or topics in a given corpus prior to the analysis. They allow for comparisons of different discourses by using the same identical procedure over all corpora. They enable the researcher to account not only for the number of categories but also the interconnectivity of these categories. This eliminates the need to make binary decisions about whether topics are similar or different, as it brings a more fine-grained perspective to estimate the extent to which two themes are not entirely independent but also not entirely identical (as will be elaborated when addressing the measurement of disparity). The models account not only for the number or extent of differences between themes, but also for their distribution. As such, these methods offer a relatively cost-effective way to estimate thematic diversity that does not rely on human coders, who are costly to use for large corpora, and even impossible in extremely large datasets, such as the ones examined in this study.

Finally, from a wider methodological perspective, this study also contributes to existing research on semantic network analysis. While research on semantic networks has grown considerably in the last decade, and while researchers often use semantic network analysis as a tool for analyzing various discourses and corpora, relatively few researchers have extended their analysis from a single network perspective to multiple networks, or a between-network perspective (Baden, 2010; Carley & Palmquist, 1992; Danowski,

2012a; Doerfel & Connaughton, 2009a; Eberl, Jacobi, & Schlögl, 2014a; Qin, 2015a; Shim, Park, & Wilding, 2015). Even these studies tend to be limited to a small number of graphs and to a more basic set of methods for comparison, i.e., focusing on more qualitative comparisons. While applicable to small-scale analyses, such methods are inadequate for the comparison of a larger set of semantic networks. Therefore, this study follows in the footsteps of studies such as Eberl et al. (2014) or Doerfel and Connaughton (2009a), extending semantic network analysis scholarship by utilizing prominent network graph-level indicators as a method for a large-scale comparison of multiple semantic networks and their impact. By focusing on a prominent set of measures related to network cohesion and partitioning, taking advantage of the role of sub-graphs in research on semantic networks in communication, and providing a novel method of estimating diversity in networks, this dissertation addresses the need to advance the study of thematic diversity in semantic network analysis (Eberl et al., 2014a) and enhances the comparative capacity of semantic network analysis as a research tool.

## **1.2 Structure of the Dissertation**

I begin with a review of the theoretical framework developed for this study in chapter 2. I first address the relationship between the volume of coverage that candidates receive in the news media and their success in the polls, as well as the relationship between the tone of that coverage and candidates' electoral success. I then review the theoretical explanations for these relationships, existing findings, and criticisms and limitations of the existing research, focusing on the direct/indirect nature of these relationships. Following this, I turn to discuss these two features in a different context: candidates' direct communication with voters via social media.

Following the discussion of volume and tone, I turn to the concept at the core of this study, thematic diversity in candidates' news coverage and social media activity. I begin by discussing how diversity has been conceptualized in existing research on content and in communication research on rhetorical diversity (Chaffee & Wilson, 1976; Kleinnijenhuis et al., 2015). Drawing on this varied body of research, I define thematic diversity as the range of topics discussed in the media, the balance (or distribution) of these topics, and the extent to which these topics are interconnected with each other. I then turn to a discussion of the advantages of monothematic and multi-thematic message strategies. Based on these arguments, I offer two competing hypotheses, one in support of a monothematic message strategy (based on theories related to issue ownership, repetition, and media attention; Allport & Lepkin, 1945; Hänggli & Kriesi, 2012; Petrocik, 1996), and one in support of a multi-thematic message strategy (focusing on the advantages of message flexibility and the phenomenon of issue convergence; Sigelman & Buell, 2004).

Chapter 3 lays out the methodological framework of the study. It differs from the methods chapter by focusing not on “how” to estimate thematic diversity estimation, but on the “why”. In contrast with the methods chapter, which discusses the specific procedures, parameters, and tools used in the study, as well as descriptive data regarding the corpora, this chapter focuses on the conceptualization of thematic diversity from a larger theoretical perspective.

I begin by addressing current approaches used to estimate diversity in communication research and their limitations. I then discuss emerging research on diversity in other scientific fields outside of communication, such as biology, physics,

and public policy, to identify meaningful structural features that an adequate diversity estimation must address. These include, variety (number of categories), balance (the distribution of categories), and disparity (the relationship of these categories to each other). From a thematic diversity standpoint, I conceptualize thematic diversity as asking three related questions of a given corpus: How many themes are discussed? How equal are they in their prominence? And how similar or different are these themes from each other? I then apply these features to a hypothetical and simplified example of candidates' topic structure to explain how each feature adapted from different areas of research, such as biological diversity, can be applied to the issue of thematic diversity in political rhetoric.

Following this more high-level discussion, I explain how the concepts of variety, balance, and disparity can be applied to the unsupervised machine learning methods used in this study. I first describe the method of topic modeling in general and its application to political discourse. This is followed by a conceptualization of variety, balance, and disparity, using the overall topic structure data drawn from past studies. I also discuss the research on semantic network analysis in general, and particularly in political communication, including the limitations of the current state of the field. I then discuss applications of the logic of variety, balance, and disparity to network analysis, including dilemmas in the operationalization of the concept and its limitations, using concepts and methods drawn from the research on topic modeling.

Following the methodological framework chapter, I turn in Chapter 4 (Methods) to discuss in more detail the procedures carried out in this study. I begin by describing the sample of U.S. Senate candidates used for the analysis, and the non-media data that were



used either as a dependent variable or as control variables, and present summary statistics for these variables. I then describe the methods used for gathering the media data, including the different processes used for scraping candidates' coverage in the news and social media activity.

Following the discussion of scraping procedures, I describe the analysis procedures for the media data. First, I address the measurement of volume and tone in candidates' social media activity and news coverage. I then turn to discuss the details of the two unsupervised machine learning methods used for diversity estimation. I detail the processes used for the topic modelling estimation of the Twitter dataset and the news dataset, including model fit statistics and the process through which decisions were made about the number of topics to be included in the models. I then turn to discuss details of the semantic network analysis, for both the Twitter dataset and the news dataset. These processes were somewhat different for the two datasets as a consequence of their size and the assumptions made regarding the thematic structure of the documents (i.e., a document-level approach vs. a moving-window based approach). Lastly, I address the statistical approach used to explore the relationships between candidates' electoral success and the volume, tone, and thematic diversity of their social media activity and news coverage.

Chapter 5 presents the results of the analyses carried out for this study. First, I present results on the relationship between candidates' news coverage and their electoral success. Following a descriptive presentation of the corpus, I present results of analyses using topic modeling as the basis for estimating thematic diversity and examine the impact that the number of topics in the possible models ( $k$ ) has on model performance,

the results for models containing only the media variables, and the results of a more elaborate model, controlling for non-media factors. Following this, I present the results for models in which semantic network analysis was used to estimate diversity, including examples of more diverse and less diverse semantic networks drawn from candidates' coverage. I present the results of the raw scores for semantic network diversity, as well as the results of more elaborate techniques using random configuration models as benchmarks for observed network thematic diversity. Finally, I present the results of a model that combines thematic diversity estimations measured using topic modeling with thematic diversity estimations using semantic network analysis.

I then turn to a presentation of the results regarding the relationship between candidates' social media activity and their electoral success. Here I repeat the order in which results were presented for the news coverage analyses, starting with models using topic modeling-based diversity estimation, and ending with full models that incorporate both semantic network analysis and topic modeling-based diversity estimations. I end the chapter by presenting results from various regression models, in which thematic diversity serves as a dependent variable, exploring the factors that shape thematic diversity in both news and social media.

Finally, in Chapter 6, I discuss the implications of these results more broadly from a theoretical and a methodological perspective. First, I summarize and review the results on the relationship between the volume and tone of news coverage of political candidates and their electoral success. I reflect on the role played by these features, as well as the independent, or non-independent, nature of these relationships. I also discuss the results for the relationship between political candidates' social media activity and their electoral

success, situating the discussion in the larger context of political advertising's impact on electoral success. I then turn to discuss thematic diversity in both news and social media. Aside from more general conclusions drawn from the findings of this study, I also address the importance of these features in comparison to more simplistic measures of volume and tone, and in terms of their independence from other non-media factors. The discussion is also informed by the results for the relationship between non-media factors and thematic diversity in candidates' news coverage, and the factors that shape thematic diversity in candidates' social media activity.

Following this, I address the study's contributions in light of the existing literature on these subjects and the findings presented here. These include both theoretical contributions to the research on the relationship between electoral success and the media (news and social), as well as methodological contributions related to the unsupervised methods used to estimate thematic diversity. I also address the limitations of this study and the questions that it raises or leaves open. Therefore, and in a similar vein, I also offer possible directions for future research, first, in the context of campaign communication beyond thematic diversity, and second, in the context of thematic diversity beyond elections and political communication, with suggested applications of the theoretical and methodological lessons from this research to additional contexts in which thematic diversity is theorized to play a key role. These include questions of a more normative nature, such as the contribution of media thematic diversity to the media's ability to both represent and inform the public, as well as more effects-oriented research into the antecedents and consequences of thematic diversity in the representation of foreign countries in U.S. media, and the role of diversity in conflict escalation.

I end my discussion by revisiting the debate on the role of thematic diversity in political rhetoric, and especially the common wisdom that monothematic message strategies are advantageous. Reviewing the various theories and arguments connected to this debate, from agenda setting to issue ownership and issue convergence, and building on the results of the present study, I argue that when discussing political discourse during elections campaigns, Carville's common adage is correct in essence, but it might be "the structure" rather than the economy, immigration or other specific issues by themselves.

## 2. THEORETICAL FRAMEWORK

Does news media coverage of political candidates, and candidates' direct communication with potential voters, influence candidates' electoral success? These questions have been at the core of political communication research for almost a century (Bernays, 1928; Katz & Lazarsfeld, 1955; Lippmann, 1922), focusing on various media, content features, effect processes, and a wide range of outcomes. In this chapter, I review the growing literature on this subject and present the theoretical framework for the present study. I begin by discussing past research on media and elections. I focus on two key features of media content: volume and tone (Norris et al., 1999).

I first address the relationship between the volume of coverage that candidates receive in the news media and their success in the polls. I rely mainly on theories of mere exposure and agenda setting (Geiß & Schäfer, 2017). I then turn to discuss the relationship between the tone of the coverage that candidates receive in the news media and candidates' electoral success (Hopmann et al., 2010), relying on theories such as second-level agenda setting and affective priming (Sheafer, 2007). Following this discussion, I introduce these two features in a different context: candidates' direct communication with voters via social media. Existing research examines how candidates use social media (Bode & Dalrymple, 2016; Borah, 2016; Evans et al., 2014; Kruikemeier, 2014a; Stromer-Galley, 2014), and the relationship between candidate success and the volume of general Twitter chatter (Beauchamp, 2017; Caldarelli et al., 2014; Gayo-Avello, 2013; Jungherr, Jürgens, & Schoen, 2012; Metaxas, Mustafaraj, & Gayo-Avello, 2011; Murthy, 2015; Tumasjan et al., 2010). However, only a handful of studies have examined the relationship between the volume of a candidate's own social

media activity and the candidate's success (Bright et al., 2018; Kruikemeier, 2014a; LaMarre & Suzuki-Lambrecht, 2013; Vergeer, Hermans, & Sams, 2011). Moreover, I was unable to locate any studies that addressed this relationship from the perspective of sentiment or tone. Regarding the volume of candidate activity, I review existing studies as well as their limitations, focusing especially on the lack of non-media factors as controls in prediction models (Beauchamp, 2017; Bright et al., 2018). I then turn to the issue of negativity and positivity in candidates' social media activity. As research is extremely lacking in this context (Jungherr, 2016), I rely on the distinct yet related area of research on negativity in televised advertising. This provides some initial insights into the relationship between negativity in candidates' direct communication with voters and their electoral success (Fridkin & Kenney, 2012; Krupnikov, 2011). I then address criticisms recently raised against these arguments and discuss the limitations of the existing research and methods, focusing on the direct/indirect nature of these relationships (Bélanger & Soroka, 2012; Soroka, Bodet, Young, & Andrew, 2009).

Following the discussion of volume and tone, I turn to the concept at the core of this study, *thematic diversity* in candidates' news coverage and social media activity. I begin by discussing the conceptualization of diversity in various scientific fields outside of communication and general social sciences (Stirling, 2007), as well as existing research on content or rhetorical diversity in communication research (Chaffee & Wilson, 1976; Kleinnijenhuis et al., 2015). Drawing on this varied body of research, I define thematic diversity as the range of topics discussed in the media, the balance (or distribution) of these topics, and the extent to which these topics are interconnected with each other. I then discuss the advantages of monothematic and multi-thematic message

strategies. In this discussion, I rely on research on social media strategy and news coverage, as well as related literature from the fields of strategic political communication and political advertising, which, while not based on social media and not addressing thematic diversity directly, can still offer insights from more traditional modes of campaign communication related to general message selection strategy in political campaigns (Benoit et al., 2011; Sellers, 1998). Finally, based on these bodies of literature, I offer two competing arguments, one in support of a monothematic message strategy (based on theories related to issue ownership, repetition, and media attention; Allport & Lepkin, 1945; Hänggli & Kriesi, 2012; Petrocik, 1996), and one in support of a multi-thematic message strategy (focusing on the advantages of message flexibility; Sigelman & Buell, 2004).

While this chapter establishes the hypotheses and research questions for the study, the next chapter presents the methodological framework and addresses the larger issue of thematic diversity, including challenges in diversity measurements, the application of diversity estimations adapted from various scientific fields (from biology to physics), and the measurement of thematic diversity using unsupervised machine learning methods.

## **2.1 News Media and Elections**

The study of the relationship between news coverage of political candidates and their electoral success has been at the core of political communication research for almost a century (Bernays, 1928; Katz & Lazarsfeld, 1955; Lippmann, 1922). Studies on this relationship were conducted in multiple countries, in the context of general elections as well as more limited campaigns, such as referendums or primary elections, and even the impact of candidates' news coverage in foreign outlets and their impact on foreign public

opinion (Balmas & Sheaffer, 2010; Boomgaarden et al., 2012; Geers & Bos, 2017; Johann et al., 2017).

The common argument holding together numerous hypotheses about the impact of news media on voting behavior (and hence election results) is that most citizens do not interact with candidates or parties directly, but rather almost solely through the media (De Vreese, 2010; Lippmann, 1922). Because the media serve as the primary conduit for information on the different options that voters have to choose from, the ways in which candidates and parties are presented in the media can potentially impact voters' perception of political actors and thus their voting decisions (De Vreese, 2010; Hopmann et al., 2010).

This, of course, does not mean that voters and candidates do not come in direct communication with each other. Candidates reach voters by means of direct communication through televised advertising (Bradshaw, 2004), to direct mailing (Gosnell, 1926; Green & Zelizer, 2017) and rally attendance (Althaus, Nardulli, & Shaw, 2002). Moreover, new technologies, and especially social media, gradually enable more efficient and cost-effective direct communication between candidates and voters, thereby increasing voters' exposure to candidates' messages unfiltered by the news media (Bode et al., 2016a; Borah, 2016; Stromer-Galley, 2014).

These various modes of communication will be discussed in greater detail in the second part of this chapter (focusing on direct communication via social media). However, even with the current abundance of direct communication channels between politicians and their constituencies, many voters are still exposed to political information through the news media (Gottfried, Barthel, Shearer, & Mitchell, 2016). Moreover, the



impact of news media can be direct or indirect (Hopmann et al., 2010), utilizing various media features (Norris et al., 1999), and linked to different political outcomes (De Vreese, 2010). Most importantly, the news media also interact with the increasing number of direct communication channels and interpersonal relationships to impact voting behavior (Johann et al., 2017; Katz, 1957).

While this study focuses on voting behavior, voting is just one of various outcomes of news coverage of elections, as evidenced by the extensive literature on the subject. News media have been shown to impact general perceptions of the political system, especially focusing on cynicism and efficacy (Cappella & Jamieson, 1997; De Vreese & Semetko, 2002; Schuck, Boomgaarden, & de Vreese, 2013), cognitive and affective outcomes, such as candidate evaluation and political knowledge (often using agenda setting, priming and framing as theoretical frameworks; Hansen & Pedersen, 2014; Iyengar, Peters, & Kinder, 1982), and behavioral and participatory outcomes, including engagement, information seeking, turnout/mobilization, and lastly, vote choice (De Vreese & Semetko, 2002; Hayes & Lawless, 2015).

While, the general area of research on media and politics has been prominent, research on the direct impact of media on voting decisions, especially in terms of volume and tone, has been somewhat rare until as recently as the last two decades, perhaps as a rejection of the strong effects paradigm and a turn toward more nuanced and limited media effect processes (Hopmann et al., 2010). In addition, while more research has been conducted on this issue in the last two decades, results are still mixed at best, and recent research raises some important questions regarding the independent impact of these features that have yet to be fully answered (Bélanger & Soroka, 2012).

The first step to understanding the direct impact that news media coverage may have on voting behavior is to define what news media features need to be examined. Norris et al. (1999) have offered three such features to define media coverage of political actors that might relate to voting decisions. These include stop-watch coverage (the volume of coverage), directional coverage (the tone of coverage), and agenda coverage (the topics chosen for coverage). This division can also be applied to the wide-range of studies conducted on the issue in the last two decades, as well as the general framework for this study. I will refer to these three features as volume, tone, and thematic structure.

Studies of agenda coverage examine the topics that are addressed and highlighted in media coverage, utilizing some of the field's most prominent theories, including framing, agenda setting, and priming (Scheufele & Tewksbury, 2007). For example, the literature has considered the impact of issue priming on assessments of presidential efficacy (Iyengar, 1987), and how the prevalence of issues positively or negatively related to candidates might affect their general assessment by the public (Druckman, 2004; Iyengar, 1987; Pan & Kosicki, 1997; Petrocik, 1996; Scheufele & Tewksbury, 2007; Sheafer & Weimann, 2005). The theories related to issue ownership, issue convergence, and priming will be revisited later in the dissertation, when the impact of thematic structure is addressed through the lens of monothematic and multi-thematic message strategies. The literature on the impact of coverage volume and tone will be reviewed in the following sections.

### **2.1.1 News Media and Elections: Volume**

As mentioned above, multiple studies have addressed the impact of volume of coverage on candidates' electoral success in the last two decades. These studies,

however, often do not offer an elaborate theoretical framework. For example, Geers & Bos (2017) argue that candidates' visibility in the media is a necessary condition for electoral success, as a major pre-requisite for any voter's choice of a candidate is an awareness of that candidate's existence. This is especially true in multiparty political systems, where the large number of options that voters can choose from means that without the knowledge that a party is competing in the elections, a voter cannot decide to vote for that party (although this effect might be limited in the case of a two-party system). Other studies rely on a "mere exposure" argument (Geiß & Schäfer, 2017). This argument suggests that the mere frequency of contact with an object positively influences the perception of that object. Thus, the more coverage that a candidate or party receives, the more likely that the candidate or party will be evaluated more favorably.

Some studies offer a more elaborate theoretical explanation, connecting agenda setting theory with the impact of candidate visibility (even if not doing so expressly; Geiß & Schäfer, 2017). The original argument of the agenda setting theory is that an issue's salience in the media can affect its accessibility and, as a result, its perceived importance by individuals who are exposed to this media (Funkhouser, 1973; Iyengar et al., 1982; McCombs & Shaw, 1972). For example, intense coverage of foreign affairs or economic issues will make these more accessible in viewers' minds, thus rendering them more important to those viewers.

The relationship between volume of coverage and accessibility can also be the result of several competing cognitive process, both conscious and unconscious. Exploring these alternative explanations can aid in the application of this theory to the relationship between the volume of coverage that a candidate receives in the news media and his or

her electoral success. The most prominent explanation for this phenomenon is accessibility bias (Iyengar, 1990). In line with agenda setting theory, this argues that an issue's prominence in the media affects the accessibility of this information in viewers mind. Thus, when asked, for example, to identify the biggest challenge facing the country, viewers will tend to name the issue that is most salient in the news coverage at the time (Funkhouser, 1973; McCombs & Shaw, 1972). This is the explanation that is most widely used by researchers when applying agenda setting theory as a standalone theory to the perception of political actors (in contrast with studies that apply agenda setting in tandem with priming theories, as will be discussed in the thematic structure section; Scheufele & Tewksbury, 2007). In accordance with this explanation, candidates who are more prominent in the media become more favorable, as voters are more likely to recall the candidate when asked to name someone that they consider voting for.

In contrast with this more commonly cited unconscious process, evidence has also been presented for a more conscious means by which individuals infer the importance of issues from media salience. From this perspective, the impact of media depends not (only) on a heightened accessibility of information, but rather on agenda cueing or agenda reasoning processes (Pingree & Stoycheff, 2013). According to agenda cueing, audiences actively and consciously estimate the importance of issues from media coverage, as they assume that if an issue has been mentioned more, it must be more important. As a result, this effect can be mediate by variables that are distinct from those related to accessibility bias. For example, the trust that viewers place in the media channel through which information is received can impact the extent to which they infer issue importance from media salience (Pingree & Stoycheff, 2013).

From an agenda reasoning perspective, when an issue is covered more often in the media, it is usually followed by some justification for its importance and the reason for it being featured in the specific news article or program. Thus, rather than an unconscious influence of media salience on accessibility, or the conscious inference made from mere media salience, the agenda setting effect can be a direct result of the reasons that the media itself provides for the importance of an issue (Pingree & Stoycheff, 2013). While these two conscious explanations for the impact of agenda setting may help explain the relationship between volume of media coverage and electoral outcomes, it should be noted that they have only been examined in the context of the perceived importance of political issues and have not been applied to the importance of political actors.

In order to further elaborate on the relationship between the effect of candidate visibility on electoral success and agenda setting, the three explanations for agenda setting outlined above (accessibility bias, agenda cueing, and agenda reasoning) can be used in tandem to argue that candidates who receive more media coverage will be considered by the viewers as more important or more successful, thus increasing their favorability with the voters (Hopmann et al., 2010; Oegema & Kleinnijenhuis, 2000). From an accessibility bias perspective, salience in the media can increase the ease by which candidates are accessible in voters' memory, thus increasing the rate at which voters invoke a candidate when primed (similar to the "visibility as necessary condition" argument mentioned earlier; Geers & Bos, 2017). However, agenda cueing can also explain the nature of this relationship, as candidates who are more salient in the media will be considered more important, leading voters to assume that media coverage implies, at least to some extent, that candidates are more viable or important. Finally, from an

agenda reasoning perspective, candidates who are more frequently featured in the media will be mentioned in tandem with reasons for why they are featured and justifications for why they are important enough in the political arena to be mentioned; candidates can leverage such rationales to enhance their viability as political actors, thus increasing their favorability with voters.

As mentioned earlier, only a limited number of studies have directly addressed the impact of volume of coverage on candidate performance, although these have seen a steady rise in recent decades (Boomgaarden et al., 2012; Eberl et al., 2017; Geers & Bos, 2017; Geiß & Schäfer, 2017; Hopmann et al., 2010; Johann et al., 2017; Lengauer & Johann, 2013; Norris et al., 1999; Oegema & Kleinnijenhuis, 2000). This small but growing literature has generally found a modest positive relationship between the volume of coverage that candidates receive in the media and their electoral success. However, some studies have found a more limited effect, moderated by the type of media (stronger for elite media) and type of voters (stronger for less ambivalent voters; Johann et al., 2017). With the exception of the framework offered by Norris et al. (1999), in which volume and tone are considered as separate features, there is evidence to suggest that volume and tone might not be completely independent, as volume might moderate the impact of tonality (Eberl et al., 2017). Finally, there is some evidence to suggest that the impact of volume might be even more limited when taking into account the influence of non-media factors on media coverage and electoral success concurrently. These limitations will be further elaborated after discussing the impact of tone in candidates' coverage, with a discussion of the casual mechanism between these relationships and the impact of direct and indirect effects of news coverage.

Considering the existing evidence for the impact of volume of coverage on voters' perceptions and evaluations, and building on the theoretical framework and explanations for the agenda setting effect, I offer the following hypothesis:

*H1: The volume of news coverage that a candidate receives is positively correlated with that candidate's electoral success, with higher coverage being related to higher success.*

### **2.1.2 News Media and Elections: Tone**

Another group of studies has used the tone of news media coverage of a candidate as a predictor of his or her electoral success and public opinion of the candidate. Using both hand coding (Hopmann et al., 2010; Soroka et al., 2009) and, in recent years, automated sentiment analysis (Bélanger & Soroka, 2012), researchers have argued that the valence of coverage can influence voters, with more positive coverage contributing to electoral successes (Hopmann et al., 2010; Soroka et al., 2009). For example, Kleinnijenhuis et al. (2007) examined four news types: issue positions, real-world developments, support and criticism of political actors, and news on the success and failure of political actors. The last two of these offer direct positive or negative assessment of candidates from different perspectives, attributing affective judgement to the candidates themselves, their positions, or their actions.

Evidence has shown that the effect of tone might be weaker for voters who were already decided on their favored candidate before the elections (Fournier, Nadeau, Blais, Gidengil, & Nevitte, 2004) and stronger for initially undecided voters who selected their candidate during the campaign. This is similar to the differentiation between crystallizers

and early deciders that Lazarsfeld et al. (1968) draw out, and offers an interesting contrast with evidence showing that advertising tone can have a stronger impact on already-decided voters for the candidate that they support (this will be discussed later in the context of tone's impact on electoral success in direct communication channels; Krupnikov, 2011).

While most studies in this area have used cross-sectional research designs and time-series models, experimental evidence for the effect has also been provided (Norris et al., 1999). Experimental findings support the notion that the impact of tone of coverage is not uniform. First, it can differ for different types of voters, as mentioned above in the context of early-deciders and those who decide during the campaign. Second, experimental findings point toward an effect for positive coverage but not for negative coverage. As recent evidence shows, the tone of coverage can have a “backfiring” effect in cases when negativity in coverage collides with voters' pre-dispositions (Geiß & Schäfer, 2017). However, it should also be noted that the extent to which effects found in experimental studies are long-term and enduring, and therefore meaningful in the context of real elections, remains an open question.

While many studies provide empirical support for the link between tone of coverage and voting behavior (often in ways that mirror research on the impact of volume of coverage for electoral success), theoretical elaborations on the nature of this effect has been rather minimal. Most researchers simply point to the apparent and logical connection between positive coverage and voters' positive dispositions toward a candidate, as it seems self-evident that more positive coverage will make a candidate more favorable while negative coverage will make a candidate less favorable. However,



as mentioned above, there is also evidence to contrary, showing that the impact of news coverage tone on electoral success is a more complicated phenomenon moderated by various features of the voters, the political actors, and the media outlet. As a result, researchers have recently begun to offer a more elaborate framework for understanding the impact of tone on electoral success.

Researchers have taken up the contested theory of second-level agenda setting as a theoretical framework for understanding the relationship between tone of coverage and electoral success (Balmas & Sheafer, 2010; Coleman & Wu, 2010; McCombs, Llamas, Lopez-Escobar, & Rey, 1997). While first-level agenda setting explores the impact of issue salience on perceived issue importance, second-level agenda setting focuses on the salience of issue or object attributes and the extent to which these are presented in a positive, negative, or a neutral manner. For example, first-level agenda setting might focus on how coverage of a candidate's character might affect the perception of that candidate, second-level agenda setting would explore what features of this candidate's character are highlighted and whether these are presented as favorable or unfavorable to that candidate. For instance, highlighting the trustworthiness or untrustworthiness of a candidate will have a different impact, even though both highlight the same feature in a candidate's character.

Second-level agenda setting thus suggests that it is necessary to examine not only the cognitive and substantive attributes of an object, but also the affective attributes connected to it—in other words, to look not only at the volume of coverage, or the topics covered, but also to the manner in which these topics are covered (McCombs et al., 1997). Moreover, there may be interactions between these two sets of attributes, or the

two levels of agenda setting, with negative affect attached to an issue, for example, increasing its prominence with viewers due to a negativity bias (Sheafer, 2007).

Another theoretical framework that has been offered to explain the impact of tone of coverage on candidate electoral success is affective priming (Sheafer, 2007). This examines the direct and indirect impact of the affective attributes attached to various issues. Priming (or at least some interpretation of priming used in political communication) can be seen as an extension of agenda setting (Scheufele & Tewksbury, 2007). If agenda setting examines the impact of object salience in the media on the object's perceived importance, then priming is the result of this heightened perceived importance on object evaluation. By changing the importance of a political issue, the media can influence the standard by which a candidate is evaluated. For example, changing the standard of evaluation from foreign affairs issues to economic issues might negatively or positively affect candidate favorability. However, this leaves an important question open—how does issue prominence impact evaluation? Why do some issues influence an evaluation in a positive manner and others in a negative manner?

Some explanations, which will be elaborated on later, relate to issue ownership and candidate performance as possible answers to these questions. Highlighting an issue such as social welfare might automatically aid one side of the political map more if that party or candidate is considered a priori more credible or capable in treating this issue (see later discussion on issue ownership; Petrocik, 1996; Sigelman & Buell, 2004). In other cases, some issues might be affectively charged in advance, thus lending candidates a positive or negative effect by dint of having highlighted particular issues. Finally, some issues might not naturally lend themselves to a positive or negative evaluation in this more

automatic manner. In these instances, the affective attribute attached to an issue can change its priming effect. For example, while economic issues can be highlighted to increase their centrality to candidate evaluation, presenting the issues in a positive (economic growth) or negative (economic decline) light can change the impact that the priming of economic issues will have on voters' evaluation of an incumbent (or a challenger). Thus, the tone not only applies to the candidates themselves, but to the contexts in which they are mentioned and the issues to which they attach themselves (Sheafer, 2007).

Based on these previous findings on the impact of tone of coverage on candidates' electoral success and the theoretical explanations offered above, I formalize the following hypothesis:

*H2: The tone of news coverage that a candidate receives is positively correlated with that candidate's electoral success, with more positive coverage being related to higher success.*

## **2.2 Social Media and Elections**

While news media have often been hypothesized to influence voter behavior, they are far from the only channel through which candidates can communicate with voters. Technological advancements have reduced the costs of direct communication with voters, especially for modern campaign technologies that rely on online and electronic communication. Stromer-Galley (2014) details changes in the usage of digital technologies in political campaigns over two decades. Early attempts at digital campaigning can be found in the Dole vs. Clinton elections in 1996 and have since

become increasingly sophisticated, from the basic html websites used in the 1996 elections to more interactive and audience engagement-based campaigns. Notable examples are Howard Dean's campaign for the Democratic presidential nomination in 2004, Ron Paul's campaign for the Republican presidential nomination in 2008, and Barack Obama's campaign in 2008. Following the increased popularity of social media websites, such as Twitter and Facebook, campaigns since (at least) 2012 have paid increased attention to these channels as a means of directly reaching and engaging potential voters, as well as means to understanding public opinion and identifying issues that voters care about (Stromer-Galley, 2014).

In current campaigns, politicians use social media as a tool for crafting their character, as well as for promoting their activities, campaign efforts, and to mobilize voters (Borah, 2016; Bright et al., 2018; Jackson & Lilleker, 2011). Social media allows politicians to connect not only with the public directly, but also with journalists, making platforms like Facebook or Twitter vital communication tools to influence intermedia processes (Verweij, 2012). Studies have shown that social media communication can give audiences the impression of a direct conversation, thus increasing politicians' favorability (Lee & Shin, 2012). Twitter activity, especially when personalized, can induce a stronger sense of intimacy with audiences and lead to better message recall. The most common social media platforms among politicians currently are Facebook (Williams & Gulati, 2012) and Twitter (Lee & Oh, 2012). As these platforms are also used by many media consumers as a central news gathering platform (Petrovic, Osborne, Mccreadie, Macdonald, & Ounis, 2013), they can serve as an efficient vehicle to inform voters about candidates and the campaigns. Evidence shows that politicians attempt to

use these channels not only to recruit supporters and energize opinion leaders, but to actively raise funds needed for their current campaigns. However, social media activity has also been found to be somewhat distinct from other types of political advertising, at least in terms of the topics that candidates focus on when communicating through social media channels (Borah, 2016). This is especially true for more extreme candidates who eschew mainstream channels or for local or state candidates to solicit out-of-state donations (Hong, 2013).

Unlike news media coverage, political candidates are able to more directly control the messages they convey through social media. In their direct communication with voters via Facebook or Twitter, for instance, candidates or their social media managers can work to create carefully crafted messages, with the medium often serving as a platform only. Candidates' direct communication via social media may thus have more in common with press releases than media coverage. Therefore, while this study builds on the growing research on the role of social media in political campaigns, I also rely on more traditional PR literature, especially as it relates to the impact of negative tone in social media activity on electoral success, a context in which social media research has been lacking.

In tandem with the rise of social media channels and strategies for political campaigns, research in political communication on social media use during campaigns has also grown sharply (Bode & Dalrymple, 2016; Bright et al., 2018; Jungherr, 2016). While this research has been largely interdisciplinary in nature, drawing on various fields of research, from communication and political science, to computer science, researchers

from these different fields offer different sets of questions, employ different methods, and utilize different concepts and theories (Jungherr, 2016).

Jungherr (2016) identifies three prominent questions in the current research. First, what are the factors that influence candidate behavior in social media? Studies responding to this question examine the features of candidates who are quick to adopt these tools in their campaigns and the factors that predict candidates' ignoring these tools. Second, what do candidates do with social media? Studies that ask this question often adopt content analysis methods, as well as interview-based approach, in an attempt to understand what roles social media play within political campaigns, what messages are delivering on social, and the extent to which these strategies differ from previous and more traditional campaigning strategies. The third question is, what impact does candidates' social media activity and digital campaign strategies have on their electoral success? Unlike research responding to the first two questions, this last group is more likely to focus on the relationship between general Twitter chatter and candidates' success in the polls, rather than a candidate's own activity on social media (Bright et al., 2018).

Several factors have been shown to influence the rate at which campaigns adapt to the changing digital landscape. It is often suggested that because underdog candidates are more willing to take risks, they will be more likely to adopt newer technologies, in part due to economic necessity (social media channels are far less expensive than traditional channels) or to leverage the medium itself to signal the candidates' originality and innovativeness (Stromer-Galley, 2014). As such, research has found that candidates in the opposition are more likely to be using Twitter (Vergeer et al., 2013). However, there is

also contrasting evidence to suggest that the opposite might be true. Running an efficient and impactful social media campaign requires adequate funding and skilled manpower, which is more readily available to major candidates (Evans et al., 2014). With evidence drawn from various countries over the last decade, from the US to Norway, India to Brazil, research has shown that there are a number of additional factors that might influence the likelihood of a candidate using Twitter. Among these are party affiliation (though which party is more likely to use Twitter can change over time and between countries; Peterson, 2012; Vergeer & Hermans, 2013), age (Larsson & Kalsnes, 2014), gender (Evans et al., 2014; Gilmore, 2012), and race competitiveness\_ (Evans et al., 2014). Building on this research, the present study will use several of these factors to explore not only what factors shape the adoption of social media as a campaigning tool, but what factors shape thematic diversity for candidates who use social media as a tool for campaign communication.

Studies of what candidates do with social media explore different aspects of the activity, from the topics that candidates choose to discuss to the rate of direct interactions between candidates and voters (Bode et al., 2016b; Borah, 2016; Kruikemeier, 2014b). Generally, candidates seem to use Twitter most often for information dissemination and calls for action, rather than for policy-related discussions. In addition, candidates use these channels to help construct their image and brand, often as a way of humanizing the candidates. For example, Rick Santorum's campaign in the 2012 Republican primaries opened a Twitter account for Santorum's sweater vest (#fearricksvest). This was designed to show a more playful side of the candidate. Interestingly, while social media is interactive in nature, most studies suggest that campaigns adopt a more traditional

broadcast model for crafting their social media messages, rather than interacting directly with audiences (Jungherr, 2016). Though notable exceptions exist (for example, Obama's Facebook and Reddit town hall interviews during the 2012 campaign, Clinton's online interview during the 2008 elections, or Santorum's sweater vest account), the dominance of the broadcast model in candidates' social media use suggests the relevance of research on more traditional political advertising strategies. I revisit this issue when discussing the negative or positive tone of social media activity, a topic for which no direct evidence currently exists, but which has received much attention in the traditional campaign strategy literature.

The third strand of research that Jungherr (2016) identifies, on the relationship between social media activity and a candidate's electoral success, is the most closely related to the present study. The research that Jungherr (2016) identifies, however, focuses most intently on the relationship between general chatter on Twitter and candidate electoral success, treating social media in general, and Twitter in particular, as a measure of public opinion (this is no different from campaign professionals themselves, who might view the platform as a "naturally occurring focus group"; Stromer-Galley, 2014). This conversation is important, as it provides several useful guidelines and benchmarks for exploring the relationship between candidate's own social media activity (rather than the general chatter about the candidate) and his or her electoral success, especially when considering the scarce research existing on this latter question (candidate activity), and the abundant research on the former (general Twitter chatter).

One of the first studies to explore the relationship between political success and mentions on social media was conducted in Germany and was quite successful in



predicting the distribution of votes in the 2009 German national parliament elections based on a dataset of over 100,000 political tweets (Tumasjan, Sprenger, Sandner, & Welpe, 2010). However, this study was heavily criticized for the seemingly arbitrary choice of timeframe and the failure to include specific political actors (Jungherr et al., 2012). Subsequent studies have been carried out in different electoral systems and countries, from the UK (Murthy, 2015) to Italy (Caldarelli et al., 2014) and the US (Metaxas et al., 2011). However, these studies have provided mixed results, at least regarding the utility of analyzing volume and tone in social media mentions of candidates and parties as predictors of electoral success. Two relatively recent reviews provide insights into the problems and pitfalls with existing research. Both Gayo-Avello (2013) and Beauchamp (2017) find the claims regarding the predictive power of social media to be highly exaggerated.

According to Beauchamp (2017), while there is reason to be cautious, there is also cause to be optimistic, as he provides evidence for the predictive capacity of social media data when more advanced machine-learning methods are used. He thus provides several guidelines for how to adequately support a claim regarding such predictions. First, these claims should be supported by statistical testing rather than descriptive data, and for that goal, studies may need to limit their scope to races that provide larger sample sizes (such as Senate elections). Second, these claims should not be evaluated independently, but rather in comparison with other benchmarks of election predictions. This will help assess whether social media data can provide additional insights beyond existing measures. Third, considering the complexity of online communication, researchers may need to go beyond mere sentiment (tone) and volume as predictors of electoral success.

In regard to the statistical testing requirement, this dissertation offers an analysis of multiple Senate races, thus allowing for an estimation of the models' predictive power. In regard to the requirement for adequate benchmarks, this study offers a wide set of controls when examining the relationship between candidates' social media activity (and news coverage) and electoral success, with an adjusted R-square estimation for the controls-only model being ~0.7. Finally, this study goes beyond mere volume and sentiment to explore the relationship between thematic diversity and electoral success.

### **2.2.1 Social Media and Elections: Volume**

While Beauchamp (2017) provides useful guidelines for research design, and while this body of research provides some initial insights regarding the relationship between candidates' presence on social media and their electoral success, this study focuses not on the relationship between electoral success and what is said *about* politicians in Twitter, but rather on the relationship between electoral success and *what candidates themselves* say on Twitter. Current evidence on this topic is limited in terms of breadth, methodology, and, most of all, in terms of the features of social media activity researchers focus on. Generally, and similar to the mixed results regarding the relationship between volume of candidates' mentions in general Twitter chatter and electoral success, the picture drawn from these studies is that Twitter use has a small effect, if at all, on electoral success, that the effect is likely to be indirect, and that it might aid some candidates in specific contexts more than others.

Some early studies have identified a positive relationship between the volume of social media activity by candidates and their electoral success. Examining the 2009 European Parliament elections in the Netherlands, Vergeer, Hermans, and Sams (2011)

explored the impact that the mere adoption of Twitter as a campaigning tool, as well as the volume of tweets, has on electoral success. According to these authors, opposition candidates (though from parties in the parliament) tend to be more active on social media, while more established party candidates and candidates from fringe parties tend to be less active. They also found more progressive candidates to be early adopters of Twitter relative to conservative candidates (though both sides of the political spectrum were similar in their eventual adoption rate), and that more progressive candidates tweeted more often. Unexpectedly, they found that higher prioritized candidates within the various parties made more use of Twitter, rather than those who were more disadvantaged resource-wise. They also found that candidates who tweeted more frequently received a higher share of the votes, with a higher advantage for an increased volume of activity toward election day. However, given that there may be strong relationships between the factors that shape social media activity and those that impact electoral success, the relationships that Vergeer, Hermans, and Sams detected between social media activity and electoral success might have been spurious or at least mediated. This is similar to criticism raised earlier in the context of the relationship between the volume of news coverage that candidates receive and their electoral success (Bélanger & Soroka, 2012).

Attempting to replicate the results of Tumasjan et al. (2010) mentioned earlier, but in the context of candidate activity rather than general Twitter chatter, LaMarre and Suzuki-Lambrecht (2013) examined the relationship between Twitter adoption by candidates for the 2010 U.S. House of Representatives, as well as their tweeting frequency, and electoral success. Unlike Vergeer, Hermans, and Sams (2011), LaMarre

and Suzuki-Lambrech (2013) did not identify a difference in activity between incumbents and challengers. While they were able to find a relationship between political candidates' adoption of Twitter and their electoral success, they did not find a significant relationship for the frequency of their tweeting. Similarly, and in the context of the 2012 Dutch elections, Kruikemeier (2014a) also found a positive relationship between candidates' Twitter adoption and their electoral success. Furthermore, they also identified a positive relationship between the amount of interactivity in candidates' social media activity and their electoral success, though in a more limited manner.

The most recent study available for the relationship between candidates' social media activity and electoral success (Bright et al., 2018) sought to correct some of the methodological problems that had been raised by using panel data for social media activity and standing in the polls. Examining candidates' Twitter use in the 2015 and 2017 U.K. general elections, Bright et al. (2018) found that the volume of social media activity during elections can have a significant impact on candidates' success in the polls. However, this impact is likely to be very small in absolute terms (with a modest contribution to the models' R-square of roughly 3%). Put in real-world terms, Bright et al. (2018) argue that an increase of 175% in social media activity is needed in order to generate a change of 1% in vote share. However, as Bright et al. (2018) argue and as I maintain when addressing the present study's findings, such small changes can be impactful in close races.

Given these findings, the studies reviewed here also highlight the problematic nature of exploring the impact of new communication technologies in political campaigns. The take-away lessons from many of these studies may have already been

obsolete in 2016, or even in 2012. In the current election environment, in which virtually every candidate establishes a presence on Twitter (even if a modest one), the predictive capacity of mere activity as a campaign tool might be questionable (due to the lack of variability in this factor). However, the theoretical and methodological lessons of these studies might be applicable to thinking about more complex forms of social media activity as predicting factors. Thus, the question that needs to be asked is why the adoption of Twitter was found to be related to electoral success and how it reflects the benefits of social media as a means of direct communication between candidates and voters.

Several arguments have been offered in support of a relationship between social media activity and candidates' success (Bright et al., 2018; Jungherr et al., 2012). Early adoption of Twitter helped candidates seem more modern and up-to-date. Achieving this perception, for example, was explicitly stated by Barack Obama's campaign team (Stromer-Galley, 2014). However, it is not only a candidate's mere usage of social media but also the frequency of his or her activity can contribute to their image as tech-savvy. In campaigns where voters are hard to reach, these technologies were previously celebrated as tools that might level the playing field, allowing candidates with fewer resources to reach a greater number of voters. From this perspective, using social media to inform the public about candidates' opinions might increase their electoral success, especially for non-incumbent candidates. In light of evidence that Twitter is now used regularly by both incumbents and more prominent candidates, this argument can be called into question (Vergeer et al., 2011), as it is now clear that a well-organized social media campaign requires an investment of resources as well (Stromer-Galley, 2014). On the other hand,

there is also evidence to show that more fringe or extreme candidates can indeed benefit from social media, especially in terms of fundraising (Hong, 2013).

Social media also allows candidates to engender a perception of transparency for their campaigns, as well as a feeling of personal connection with the candidates (Bright et al., 2018). It has been observed that candidates tend to share more personal information in Twitter, likely as part of a strategy to humanize the candidate and create an emotional connection (Evans et al., 2014; Kruikemeier, 2014a; Stromer-Galley, 2014), a strategy that might be more effective online than on television or newspaper advertising (Lee, 2013; Lee & Jang, 2013; Lee & Shin, 2014). Such connections can, in turn, affect recognition, recall, favorability, and imagined intimacy with the candidate (Lee & Oh, 2012). However, the translation of these feelings of intimacy and favorability into actual voting behavior is still in need of further research (Kobayashi & Ichifuji, 2015).

Most of these arguments do not offer an elaborate theoretical schema, but rather rely on arguments similar to the “mere exposure” premise mentioned previously in the context of news coverage’s impact on electoral success. The logic is that candidates who appear in users’ feeds more often will be seen more favorably as a result of increased exposure, or as a result of the information and arguments disseminated by these candidates becoming more visible. However, additional, more elaborate and more indirect explanations have been offered as well (Bright et al., 2018).

First, some theories use the two-step flow theory (Katz, 1957) as a way of arguing for the benefits of social media for candidates. From this perspective, the strength of social media is in its networked structure. While more politically interested social media users are more likely to follow candidates on social media, they, in turn, might forward

candidates' opinions and messages to other individuals in their social circle who might be less interested in politics or not connected to with the candidate on social media (Park, 2013).

As was made clear during the 2016 U.S. presidential elections when Donald Trump's tweets seemed to dictate the media's agenda, social media activity itself has the potential to increase news coverage of a candidate. This suggests that the arguments discussed in the previous sections on the impact of news coverage on electoral success is also relevant to the social media context as well. News media content creators, such as journalists and editors, can themselves be influenced by the social media activity of candidates, thus allowing candidates to impact the media's agenda through a process referred to as intermedia agenda setting (Golan, 2006; Parmelee, 2014). While this was easily observed during the 2016 election cycle, the relationship between journalists and candidates via social media has been observed in previous research exploring the relationships between different media systems in different countries and the behavior of journalists assigned to the "Twitter beat" (Broersma & Graham, 2012; Parmelee, 2014; Verweij, 2012).

Finally, there are also arguments for a reduced relationship between social media activity and electoral success. Aside from major scandals, most of these arguments suggest not a negative impact of social media activity on electoral outcomes, but rather a null effect. First, especially in the context of Twitter, these platforms are often limited to specific demographics (Blank, 2016). Those who follow politicians on social media are more likely to be politically involved and more politically informed (Bode & Dalrymple, 2016), suggesting that exposure is limited to those who already hold strong political

opinions and reducing the likelihood that social media activity might have an effect on their voting choices (Prior, 2013; Zaller, 1992). Moreover, social networks have been argued to function as echo-chambers, in which those voters who already support a candidate comprise the majority of his or her online followers, leading the candidate to broadcast messages to the already-convinced (Bright et al., 2018).

Despite these somewhat mixed findings, the limitations in prior research, and the available counterarguments, the bulk of the previous research still suggests that a positive relationship between candidates' social media activity and electoral success can be expected. More formally:

*H3: The volume of a candidate's social media activity is positively correlated with that candidate's electoral success, with higher activity being related to higher success.*

### **2.2.2 Social Media and Elections: Tone**

While several studies explore the relationship between the mere usage of Twitter, or the frequency of tweeting and electoral success, no current studies have explored the relationship between the tone of a candidate's social media activity and his or her electoral success. However, arguments regarding the relationship between the two can be drawn from a related body of research on more traditional forms of direct communication between candidates and voters, such as televised advertising.

While the usage of social media in election campaigns is a relatively new phenomenon (Stromer-Galley, 2014), televised ads have been a core component of election campaigns since the presidential elections of 1952 between Eisenhower and



Stevenson and have since accounted for some of the largest expenditures of major campaigns (Benoit, Leshner, & Chattopadhyay, 2007). Given that ads are a crucial component of modern campaigns, televised ads have garnered much scholarly attention exploring the impact of both tone and volume of messages on voting outcomes.

For example, a recent meta-analysis has explored existing knowledge regarding the impact of televised advertising on various election-related outcomes (Benoit et al., 2007). They found that existing research supports a significant relationship between political advertising and political knowledge, perception of candidate character, attitudes toward candidates, agenda positions, and interest in the campaign (Benoit et al., 2007). Most importantly, the meta-analysis has estimated the effect size of televised advertising on vote choice to be  $r=.19$ . In a similar manner to the results that Bright et al. (2018) present for the social media context, televised advertising can have a significant but modest effect on candidate political success by influencing the vote choice of potential voters. Thus, while not identical in nature, one can argue that as the closest proxy, research on televised advertising can serve as a useful baseline for thinking about the relationship between candidates' activity on social media and their electoral success.

This is an important point as research on televised advertising has also paid attention to the impact of tone in political advertising on electoral outcomes, a relationship that has thus far been ignored in the context of candidate social media activity. For instance, for televised ads, some researchers have argued that negative messages can have a stronger impact than positive ones (Benoit et al., 2007; Krupnikov, 2011; Lau, Sigelman, & Rovner, 2007), owing to the tendency of negative information to

be more memorable and more likely to be incorporated into decision-making processes due to its perceived relevancy (Krupnikov, 2011).

The impacts of negative and positive advertising, and the difference in effect sizes between the two, has been explored in connection to various politically relevant outcomes, including knowledge acquisition, mobilization, general view of the political system, and, most important for this study, attitudes toward political candidates (Benoit et al., 2007). While results for the last outcome are generally mixed, a trend toward a moderate positive impact for negative advertising can be identified. An earlier meta-analysis by Lau et al. (1999) has argued that no evidence can be identified for the relative effectiveness of negative advertising. Furthermore, some scholars have argued that negative advertising might not only be ineffective, but that it can also have a negative impact on the candidate using negative messages through a backfire effect. In other words, candidates might not only fail to reduce their opponents' favorability through negative tactics, but they can also be penalized for using what voters might judge to be unethical campaign methods. However, a more recent meta-analysis by the same authors concluded that recent evidence suggests that a change in attitude towards the attacked candidates due to negative advertising is detectable, although the effect size may be very small, with a mean of 0.14 (Lau et al., 2007).

Complicating this research further is the fact that negative advertising is not a monolithic strategy: different types of negative advertising might have different levels of potency (Niven, 2006). The extent to which negative information is relevant to the decision-making process, for example, might affect the potency of negative advertising (Fridkin & Kenney, 2008). Different candidates have been found to benefit differently, or

be penalized to a greater or lesser extent, from negative advertising. For example, incumbents might be judged more harshly than their challengers for using negative tactics, therefore reducing their “net gain” from using a negative messaging strategy due to this backlash (Lau et al., 2007). The timing of negative advertising has also been shown to be critical. Specifically, negative advertising might deter voters from actively supporting a candidate under two temporal conditions: first, that the message is received after an individual has already decided on a candidate they favor, and second, that the message they are exposed to targets that exact candidate (Krupnikov, 2011). Thus, negative advertising during the last stretches of the campaign might be more impactful than negative advertising at the beginning of the campaign.

Given this prior research on television advertising, it is reasonable to expect a positive relationship between negative rhetoric on Twitter and electoral success—or, to put it differently, a negative relationship between tone of social media activity and electoral success. This is especially true given the timeframe from which the data collected for this study were drawn: the last 6 months of each election cycle. However, there are also counter-arguments that could be raised, especially for the difference between television and social media, and the direct and personal nature of social media as a mode of communication. As mentioned earlier, candidates use social media as a way of crafting their image. As such, and given that candidates’ official accounts are often named after them (even if it is a campaign aid or social media strategist who crafts the message), avoiding a backfire effect by distancing a candidate from a negative message might be much more challenging. Thus, candidates might choose to use positive messages more frequently. However, social media (and Twitter especially) might impose

different rules on this type of communication. With Twitter feuds becoming an increasingly common occurrence, with the critical nature of much of the rhetoric in Twitter, and with the tendency of negative messages to be more shareable online, it might be that negative messages will flourish more in this message ecosystem and thus have a stronger impact (although perhaps more so for within-party sharing rather than between party lines; Brady, Wills, Jost, Tucker, & Van Bavel, 2017). The existing evidence on the relationship between negative sentiments toward a candidate on general Twitter chatter and his or her success provides a compelling argument that if a negative message about a candidate goes viral, it can have an impact on observers' voting preferences.

In summary, current research on social media provides little evidence for the effect of message tone on electoral outcomes. However, the relationship between the sentiment of Twitter chatter *about* candidates (rather than *by* them) and their success might indicate that a widespread negative message strategy can have an impact. Moreover, existing research in the different yet related context of televised advertising tends to lend support for a negative relationship between the tone of candidate rhetoric and electoral success, especially when timeframes close to an election are considered. Keeping in mind that the difference in medium requires some caution, the following hypotheses is put forward:

*H4: The tone of a candidate's social media activity is negatively correlated with that candidate's electoral success, with more negative sentiment being related to higher success.*

### 2.3 Tone and Volume: Limitations

While evidence have been presented for the impact of volume and tone of news coverage on candidates' electoral success, as well as volume and tone in general Twitter chatter about candidates, and the volume of their own social media activity, there are some caveats to these evidence and arguments that must be considered. These stem from to two major issues, the direct vs. indirect nature of these effects, and, relatedly, the causal nature of the relationship and the independence of media predictors. Both issues have been raised by different scholars in the context of both social media and news coverage (Beauchamp, 2017; Bélanger & Soroka, 2012).

**Direct/indirect effect.** Volume and tone in the news coverage of candidates have been argued to offer both a direct and an indirect effect on electoral success (Hopmann et al., 2010). Both relate to the relationship between voters' perception of candidates and the volume or tone of coverage that the candidates receive in the media. However, the direct effect relates only to the tone and volume of the media that voters are exposed to. This effect follows the various explanations presented above, for example, regarding accessibility bias. In contrast, and based on theories related to the importance of interpersonal communication in mediating the impact of the media (Lazarsfeld et al., 1968; Mutz, 2006), Hopmann et al. (2010) present an argument for indirect exposure, in which individuals can be affected by the general media environment, not only the media they consume personally.

The direct exposure perspective focuses on the ways in which features of news coverage can change the opinion of those who are exposed to it. For example, more exposure to a problem for which a candidate has high credibility will increase voters'

support for that candidate. More favorable coverage of a candidate in the media content that a voter consumes will improve his or her opinion of that candidate. By contrast, being exposed to the inner workings and instrumental maneuvers of the campaign will influence voters' cynicism toward the political system, and so on.

The indirect exposure perspective, by contrast, posits a mediated process between media content and some voters. This mediation can be social, with more news-attentive opinion leaders propagating messages to which they are exposed to other individuals who may be less attentive to the news. This premise has been explored in a rich body of research branching from the two-step flow theory (Brosius & Weimann, 1996; Katz, 1957). More recent studies offer similar hypotheses regarding the capacity of new media—and especially social media—to re-transmit news content to voters who were not initially exposed to it directly (Choi, 2015; Park, 2013). Other indirect explanations discuss the impact of the information environment on candidates' credibility, funding, and other electoral assets, all of which in turn contribute to electoral success (Bélanger & Soroka, 2012). Thus, even if a candidate is not prominent in the information that a specific voter consumes, or even if that voter refrains from consuming any political information at all, the prominence of the candidate in the general information environment can still have an effect on one's vote choice. Of course these two modes of exposure are not mutually exclusive; rather, they are best understood as complementary explanations, with direct exposure influencing certain voters at the same time that the general information environment affects others more subtly or indirectly.

The relationship between candidates' coverage in the media environment and their electoral success can be accounted for with different explanations. To better understand

these requires consideration of the causal ordering of this relationship, the independence of media predictors, and whether the media (both news and social) actually affects voter intentions or if it merely “captures and arranges in a readily quantifiable form the evolving mood of the campaign” (Soroka et al., 2009).

**Independent vs non-independent effect.** Much of the evidence supports a positive correlation between volume and tone of news coverage and electoral success. Similarly, the volume of a candidate’s coverage in social media activity and general Twitter chatter has been shown to correlate with electoral success. Nonetheless, there is no consensus over the direction, strength, or even the existence of the media’s impact on the political arena. Some scholars, utilizing what Bélanger & Soroka (2012) call the “historical model,” attribute much less power to the media, especially compared with other “real-world” indicators, such as the state of the economy or past election performance. Thus, the causal direction of the effect might actually be the opposite of these “real-world” indicators shaping coverage, e.g., when more successful and popular candidates receive more positive and a greater proportion of news coverage.

Considering the often cross-sectional nature of these studies (Bright et al., 2018), the relationship between electoral results and media coverage or social media activity, may indeed be spurious, with the media simply tapping general public opinion. Thus, instead of media leading public opinion, they would instead be seen as capturing trends in the general public (Soroka et al., 2009). In other words, non-media factors, such as incumbency status, candidate experience in office, political leaning, or election timing, may be influencing the volume of media coverage that a candidate receives. Incumbents, for example, might receive more media coverage as their actions in office provide them

with more credibility and provide the media more material to address in this coverage. Incumbents might have more resources, which allows them to engage with social media more rapidly and effectively. Incumbency also offers an advantage to candidates in terms of vote share as well (Ban, Llaudet, & Snyder, 2016; Greene, 2016; Hummel & Rothschild, 2014). One could reasonably argue, then, that the impact of volume of coverage or of social media activity on electoral success is spurious, with non-media factors influencing the volume of coverage or activity and electoral success concurrently, or with media coverage mediating the effect of non-media factors.

To make this question even more complex, different non-media factors might have different relationships with electoral success and media coverage. For example, consider the case of states' political leaning and candidate political experience. The political leaning of a state can have a large impact on a candidate's predicted success, with more conservative states offering an advantage for Republican candidates. In addition, in those states, Republican candidates are more likely to have a higher volume of coverage and more positive coverage. Similarly, candidates with more political experience are more likely to receive a higher volume of coverage and more likely to be successful at the polls concurrently.

These two relationships are not equally independent from the media. Voters might not need the media to know what party a candidate belongs to (a detail often appearing on ballots) or the political leaning of the state they live in. However, in regard to political experience, one could ask, if a candidate does not mention their experience, or if his or her experience is never addressed by the media, can that experience have an independent effect on their electoral success? Thus, some non-media factors might have a spurious



relationship with media predictors and electoral success, while others might have a mediated relationship.

In more practical terms, this debate highlights the importance of including non-media controls in models estimating the impact that volume of media coverage has on electoral success, a practice sometimes ignored by researchers (Beauchamp, 2017; Bélanger & Soroka, 2012). The lack of adequate controls, and the lack of attention paid to the possibility of the non-independent nature of media variables, is even more excessive in recent work on the relationship between candidates' social media activity and their electoral success. Thus, in the models presented in this study, I control for the timing of elections (midterm or not), the conservative leaning of the state, candidate funding, and candidate experience in office. These controls help the model measure not only the relationship between media and electoral success, but also the independent impact of media coverage, that is, independent of critical non-media factors' effects. I elaborate more on these controls and the rationale for their selection in Chapter 4.

## **2.4 Thematic Diversity**

The concept of diversity is prevalent across a number of scientific fields, from biology to physics, information sciences to economics, public policy, and more (Stirling, 2007). This rich literature is conceptually linked to research on diversity in communication, as it enumerates diversity in terms of specific objects in a given environment. It is also methodologically linked to research on diversity in communication via the concept of information entropy (Shannon & Weaver, 1949). This is often used by researchers in political communication to estimate the diversity of issues, themes, and topics prevalent in media content, although as a framework it may also be

considered limited due to reasons that will be discussed in Chapter 3 (Chaffee & Wilson, 1976; van Hoof, Jacobi, Ruigrok, & van Atteveldt, 2014; Wanta, King, & McCombs, 1995).

I begin by addressing the issue of linguistic complexity and distinguishing it from thematic diversity, the concept at the center of this study. I then turn to the definition of thematic diversity itself and its relationship to theories such as agenda setting and framing. Additionally, I elaborate on the importance of these concepts, both normatively and in the context of communication research development. Finally, I discuss the research identifying the antecedents and consequences of thematic diversity. After discussing thematic diversity in more general terms, the following section discusses thematic diversity in the specific context of media and elections.

Research on the thematic diversity of issues and frames in communication can be traced as far back as the 1970s. Following the then-novel idea of agenda setting, researchers examined whether the media might not only be successful in telling people *what* issues to think about but *how many* issues to think about (Chaffee & Wilson, 1976; Wanta et al., 1995). Researchers examined issue diversity in both the media and public opinion. Shannon and Weaver's (1949) concept of H entropy was often used as a measure to examine the rate of diversity, typically with the assistance of content analysis (almost solely based on hand-coding), which would make the media content quantifiable.

Thematic diversity (Kleinnijenhuis et al., 2015) should be differentiated from linguistic complexity, which examines the complexity of a language system itself or its usage in a given text (Tolochko, 2016). For example, examining the complexity of news and its relationship to the knowledge gap, Kleinnijenhuis (1991) used the Felsch Reading

Ease Test (Flesch, 1948), a measure of readability that takes into account the average sentence length and average word length, with lower scores indicating lower readability and hence higher complexity. Shorter words and sentences, by contrast, result in higher scores, indicating higher readability. Readability, in turn, according to Kleinnijenhuis, interacts with education to explain why different audiences gain different levels of knowledge from media outlets with varying levels of linguistic complexity.

Alternative metrics incorporate various linguistic concepts, such as the ratio between unique words in a text and the total number of words (the less words tend to repeat the more complex the text is) or syntactic measures (for example, sentence length or the number of clauses in a sentence). In other words, the focus of linguistic complexity is not on the variety of issues and topics but on the complexity of the language used to convey the information regarding those topics and issues.

Thematic diversity, by contrast, does not address the complexity of the language itself, but rather the variety and interconnectivity of the themes in a given corpus, whether they are issues, actors, or viewpoints on a subject (Kleinnijenhuis et al., 2015). This refers not only to the number of issues, topics, or themes, but to their distribution and interconnectivity as well. Simply put, the estimation of thematic diversity seeks to answer how many issues, topics, or themes are present in a group of texts, how closely related those themes are to each other, and how equal their distribution is within a given corpus.

This section describes several related concepts that will later be used to conceptualize a method for measuring thematic diversity, including: agenda diversity, frame complexity, and integrative complexity. These concepts are distinct, as they

address a variety of different components (as can be understood by their name). Agenda diversity focuses on the variety of issues, frame complexity focuses on the variety of frames (broadly defined), and integrative complexity, being a theory originating not in communication but in psychology, focuses on the relationship between different arguments without addressing the agenda/frame distinction.

However, while these concepts are different in nature, they can nonetheless contribute to a conceptualization of the variety of themes in a given corpus (be they frames or issues or arguments). In addition, the logic and research on the antecedents and consequences of these concepts can be applied to understand the impact of thematic diversity in candidates' news coverage and social media activity on electoral outcomes. Finally, their methodological shortcomings serve as a worthwhile guide and benchmark for the methodological innovations offered in this study.

Agenda diversity is an elaboration on the theory of agenda setting. Agenda setting theory was influenced by Walter Lipmann's view of the media as a "beam of a searchlight" (Lippmann, 1922) and Bernard Cohen's argument that while the media might not be very successful in telling us how to think, they are "stunningly successful in telling their audience what to think about" (Cohen, 1965). Agenda setting originates with the observational studies of McCombs and Shaw (1972) and Funkhouser (1973), which were aimed at measuring the correlation of issue salience in the news media and issue salience in public opinion, findings later supported on the individual level and experimentally by Iyengar et al. (1982). The media agenda is often measured using hand-coding of the mentions of different issues in a given corpus (Chaffee & Wilson, 1976; van Hoof et al., 2014), and public opinion is often measured using the "most important

problem” survey question (though it was suggested that this measure might be too narrow or inaccurate; J. K. Lee, Choi, & Kim, 2014). As discussed earlier, accessibility bias as well as more conscious alternative mechanisms, such as agenda cueing and agenda reasoning, were offered as the mechanisms at the core of the agenda setting effect (Iyengar, 1990; Pingree & Stoycheff, 2013; Scheufele, 2000).

Paraphrasing Cohen’s (1965) argument, research on agenda diversity explored whether the media can be successful in telling people not only on *what* to think but also *how many things* to think about. Thus, agenda diversity as a measure can be applied to public opinion, or how many issues are deemed important by individuals, as well as media content, or how many issues are presented in the media. Shortly after the initial uses of agenda setting theory, Chaffee and Wilson (1976) examined the relationship between media ownership diversity and agenda diversity in public opinion. Moving beyond the sheer number of issues and using Shannon and Weaver’s H entropy measure (Shannon & Weaver, 1949), Chaffee and Wilson (1976) defined agenda diversity (in public opinion) as the equality (or inequality) in the distribution of opinions over the “most important problem” question. When the distribution of issues in the agenda was more skewed—that is, dominated by a small number of issues receiving ample attention—diversity was considered to be lower. When the agenda was comprised of various issues all receiving equal attention, diversity was considered to be high. Subsequent studies have explored the agenda diversity of public opinion over time (McCombs & Zhu, 1995) and between countries (Peter & De Vreese, 2003). Aside from the public’s agenda diversity, a smaller set of studies (likely limited by the need for increased resources to code textual materials) addressed agenda diversity in the media,

exploring the causes and effects of media agenda diversity (Jennings et al., 2011; Peter & De Vreese, 2003). As I discuss later, the estimation of diversity in these studies can be argued to be somewhat crude, given that their emphasis on the number of issues or their distribution does not account for how related the issues are to each other—or, in other words, their interconnectivity (also referred to as disparity in studies outside of communication and the social sciences; Stirling, 2007).

The theories regarding frame complexity are based on framing theory (as well as the less popular theory of second-level agenda setting theory). Unlike agenda setting, which has a relatively widely accepted definition, the framing literature comprises various and often conflicting definitions. In psychology, Tversky and Kahneman (1981) define frames as “the decision-maker's conception of the acts, outcomes, and contingencies associated with a particular choice.” However, while frequently mentioned in communication research, this notion of frame is rarely used as an operationalization guidance in framing studies (Matthes, 2009). Surprisingly, this is also true for the competing definition, originating in sociology, by Goffman (1974), who saw frames as “definitions of a situation [that] are built up in accordance with principles of organization which govern events, at least social ones, and our subjective involvement in them.” Similarly, Gamson and Modigliani’s definition of frame as a “central organizing idea or storyline that provides meaning to an unfolding strip of events” is also widely cited, with framing operating in media analyses in a parallel to the role of schemas in cognitive psychology (Gamson, Croteau, Hoynes, & Sasson, 1992). Lastly, in contrast with these abstract definitions, a more practical and concrete definition (Matthes, 2009; Scheufele & Iyengar, 2012) is Entman’s, which sees framing as the selection of aspects of a perceived

reality to “promote a particular problem definition, causal interpretation, moral evaluation, and treatment recommendation” (Entman, 1993). With its highly modular and practical structure, this definition has served as methodological guidance to various studies performing content analysis of media (Matthes, 2009). These definitions are still constantly debated in the field, with some researchers adopting a broader definition and others advocating for limiting the concept to sets of competing messages where the information is completely equivalent (Cacciatore, Scheufele, & Iyengar, 2016).

In accordance with these various definitions frame complexity has been conceptualized in several different ways. For example, Kleinnijenhuis et al. (2015) examined changes to the complexity of news frames in the coverage of the 2008 economic crisis in different stages of the crisis, from 2007-2012. Relying on the more modular definitions of framing elements (Benford & Snow, 2000; Entman, 1993), this was done by identifying nine groups of issues and stakeholders and measuring the co-occurrence of the different issue and actor groups in the text. Similarly, Huang (2010) examined “the central organizing ideas,” identifying facts, contexts, attributions, consequences, and framing devices, such as metaphors, exemplars, or catchphrases in the news coverage relating to the “On Taiwan” controversy and the “Fourth Nuclear Power Plant” dispute. Similar to agenda diversity studies, researchers then quantify diversity as the number of central organizing ideas as well as the level of inequality in their distributions using Shannon and Weaver’s H entropy measure (Huang, 2010; Shannon & Weaver, 1949).

Although I discuss this issue at length in the next chapter, it should be noted that the lack of attention to the interconnected nature of topics and themes is a major

methodological gap in most of the studies considered here. However, this issue has received some theoretical attention in the form of nominal and thematic diversity. Measuring thematic diversity, both from the agenda and framing perspectives, researchers have often differentiated between nominal and thematic diversity (Kleinnijenhuis et al., 2015; Peter & De Vreese, 2003), with nominal diversity defined as the number of issues found in the media or in public opinion and thematic diversity being the semantic or categorical diversity of issues in the media or in public opinion. To illustrate this, consider two candidates exhibiting equal rates of nominal diversity, each addressing three issues. If candidate A's issues are all related to the economy (employment, the deficit, and infrastructure, for example) and candidate B discusses issues relating to concerns about the economy, the environment, and national security, then candidate B's thematic diversity would be considered higher, even though from a nominal perspective both might be identical. This is an important distinction, as it raises the challenge of differentiating between these issue categories, which the methods suggested in this study address directly.

Studies on thematic diversity have explored both the causes and effects of diversity in different contexts. Researchers have suggested variables relating to demographics (age or education), civic engagement, and political interest as factors affecting public agenda diversity (Huang, 2010; Lee et al., 2014; McCombs & Zhu, 1995; Peter & De Vreese, 2003). In accordance with framing and agenda setting theory, researchers have also suggested that the amount and types of media used can explain public agenda diversity. Diversity in the media, in turn, has been hypothesized to be affected through a number of channels. First, the media system in a country was suggested as possibly shaping media



thematic diversity. For example, media richness can affect the diversity of information received by audiences, with more media outlets resulting in higher agenda diversity (Chaffee & Wilson, 1976). Second, the nature of the outlet itself can affect the thematic diversity of the outlet's publication—for example, smaller outlets vs. bigger outlets (Jacobi, 2016; Voakes, Kapfer, Kurpius, & Chern, 1996), or online vs. print media outlets (Carpenter, 2010).

A line of research originating in psychology has identified “real world” factors, such as crises, as possible antecedents shaping media integrative complexity. Integrative complexity stems from research on conceptual complexity and the measures developed in that context for assessing complexity in thought processes. Conceptual complexity is often measured using paragraph completion tests on various issues (for example, “when a friend acts differently toward me...”; Suedfeld, 2010). The responses to these tests are scored in terms of differentiation and integration—that is, the extent to which the text acknowledges other possible points of view and to which it integrates them into a coherent viewpoint. Different individuals, it is argued, exhibit different levels of complexity. However, unlike conceptual complexity, integrative complexity views complexity as a product not only of individual- or personality-related features, but also as a product of contextual features. For example, complexity can be affected by goals, time limits and stress or crisis, or as a result of resource depletion processes (Suedfeld, 2010). In an example of personality traits effect and contextual effect, the level of complexity of U.S. presidents' State of the Union addresses was found to vary between presidents, but also to change within a president's term, decreasing over time as presidents' term nears an end (Thoemmes & Conway, 2007).

Additionally, while conceptual complexity often relies on paragraph completion tests, integrative complexity often relies on the analysis of archival content. One of its more popular uses is to assess the complexity of powerful individuals' thought processes that researchers might be unable to reach directly, such as political actors at the center of large-scale events. By analyzing the level of complexity in texts (broadly defined) produced, for example, by leaders, researchers then attempt to assess the level of thought-process complexity characteristic of different leaders (Thoemmes & Conway, 2007) or the effects of the stage of presidential term and crisis on the level of complexity (Suedfeld, Cross, & Brcic, 2011; Thoemmes & Conway, 2007).

In over 100 studies published under the theoretical framework of integrative complexity, this method has been used by researchers in political psychology to explore the role of complexity in the context of wars (Stewart & Suedfeld, 2012; Suedfeld & Tetlock, 1977), revolutions (Suedfeld & Rank, 1976), and elections (Conway III et al., 2012), as well as the role of integrative complexity in more day-to-day settings, such as the creative and professional success of bi-cultural individuals living abroad (Tadmor, Galinsky, & Maddux, 2012), scientists' thinking on research and teaching (Feist, 1994), and the impact of positive and negative life-events (Suedfeld & Bluck, 1993).

The most relevant theoretical branch of integrative complexity for the purposes of this study comes from an unlikely sub-specialty: the analysis of political actors' integrative complexity levels as it relates to revolutions, international crises, and violent conflicts. The first study to use integrative complexity as a predictor of conflict escalation and de-escalation was Suedfeld & Tetlock's (1977) research on the diplomatic correspondences between countries in six international conflicts, from 1911-1962. Their

study found that higher complexity in diplomatic correspondences correlated with peaceful conflict resolution. They argued that this is due to prolonged stress leading to reduced complexity of thought processes and, hence, message complexity. However, while international crises resulting in war are characterized by lower levels of communicative complexity than those that were resolved peacefully, the relationship was correlational rather than causal due to the cross-sectional nature of their study. This theory and research design has been replicated across a number of contexts, from the Cold War (Suedfeld, 1992; Tetlock, 1988) to Korea (Koo & Han, 2007), the Persian Gulf (Wallace, Suedfeld, & Thachuk, 1993), Central America (Liht, Suedfeld, & Krawczyk, 2005) and the Middle East (Maoz & Astorino, 1992). The effects of conflict were measured on prime ministers, U.N. representatives, leaders of countries that share a border with countries that are involved in conflicts (Walker & Watson, 1994), and on discussants in online forums regarding conflicts (Abe, 2012). Further, a variety of conflict types have been examined, including wars and civil wars (Stewart & Suedfeld, 2012; Suedfeld & Jhangiani, 2009), revolutions (Suedfeld & Rank, 1976), global terrorism (Smith, Suedfeld, Conway, & Winter, 2008), and international negotiations (Liht et al., 2005).

Scoring of integrative complexity has been a major obstacle to many scholars and practitioners, as training often requires a three-day workshop, with further practice and instruction, to reach an adequate level of reliability (Conway, Conway, Gornick, & Houck, 2014). Even after such training, hand-coding can still be unreliable or systematically biased (Tetlock, Metz, Scott, & Suedfeld, 2014). Hand-coding can also limit the breadth of materials and size of corpora on which research is possible. While

this rudimentary coding may indeed be applicable when studying a handful of leaders and their correspondence during a conflict, it might be too lengthy and complicated to apply to large-scale analysis of elite discourse, media discourse, and public opinion. Recent approaches used dictionary methods to solve this “bottleneck” problem, but results were found to be sub-optimal (Tetlock et al., 2014); this might be expected when using generic off-the-shelf dictionaries and when applying dictionary methods in contexts to which the dictionaries were not originally adapted (González-Bailón & Paltoglou, 2015). Such studies measured complexity as it is manifest by the inclusion of words that indicate nuance, such as “possibly” and “alternative,” as opposed to words that indicate low complexity, such as “always” and “unquestionably” (Stewart & Suedfeld, 2012). These limitations in methods, as well as in context, are addressed in Chapter 3 of this study.

The importance of thematic diversity in media discourse has been underscored by research on the relationship between the public agenda and the media agenda, as well as by more normative arguments about the role of the media in modern democratic society. As with the effects themselves, this debate can be divided into arguments, one related to the diversity of issues and corresponding with agenda diversity, the other with to a diversity of viewpoints and corresponding with frame complexity.

The arguments for agenda diversity relate to both information dissemination and representation. The first is a group of arguments regarding the media’s role in informing citizens about the political process (Carpenter, 2010; Jennings et al., 2011). As Graber (2003) argues, “Democracies need citizen monitors, but not everyone needs to monitor the same thing.” With many issues of importance on the political agenda, and with different constituencies requiring information about different issues, media agenda

diversity is necessary to ensure that all important issues are being monitored by the public.

The second type of arguments relates to the media's function in representing the interests of various publics (Huang, 2010; Jennings et al., 2011). Diverse media agendas help ensure that different groups' interests are considered by the public, and thus "contribute to the fairness of public discourse, the products of which will also be more justified" (Huang, 2010). This argument also applies to the issue of frame complexity, as a more diverse set of views in the media allows for the representation of more diverse publics, increases political competition, and encourages a richer public debate regarding various issues (Huang, 2010; Jennings et al., 2011). In addition, given the complexity of political issues themselves and the ideal that citizens should be well informed on these issues, oversimplification of issues in the media can be problematic for the political process (Jacobi, 2016). However, a counter-argument can also be made that the complexity of political issues could negatively impact comprehension by various individuals and publics with differing education levels, ages, or other demographic traits. Thus, outlets that offer more simplified coverage might be beneficial to informing, for example, audiences with lower levels of education (Kleinnijenhuis, 1991).

Thus, the importance of thematic diversity is underscored both by normative arguments about the role of the media and by empirical arguments about the effects of thematic diversity on public opinion. However, the research on thematic diversity still suffers from several shortcomings, some of which can be addressed by applying unsupervised machine learning methods to measure thematic diversity, as detailed in the next chapter. More importantly, while the research has addressed many and various

contexts, it has yet to examine directly the relationship between the thematic diversity of political candidates' news coverage and direct communication with voters, and candidates' electoral success.

#### **2.4.1 Thematic diversity in elections: Monothematic message strategy**

This study goes beyond the impact of political candidates' specific agenda issues covered in the news media, or the volume of news coverage and its valence, to explore the impact of agenda diversity on candidates' performance in the polls. Drawing on various fields of research, from biological diversity (Solow et al., 1993) to agenda diversity (Chaffee & Wilson, 1976; Peter & De Vreese, 2003; Tan & Weaver, 2013) and frame complexity (Kleinnijenhuis et al., 2015), the extent to which a political discourse, such as a campaign strategy or news coverage, is unidimensional or multidimensional is referred to in this study as "thematic diversity."

From a thematic diversity point of view, this study aims to address a lacuna in existing research regarding the relationship between thematic diversity of political candidates' news coverage and direct communication with voters, and candidates' electoral success. More formally, this study addresses two research questions:

*RQ1: Does the thematic diversity of political candidates' news coverage correlate with candidates' electoral success?*

*RQ2: Does the thematic diversity of political candidates' direct communication with voters via social media correlate with candidates' electoral success?*

Current research on thematic diversity, however, does not offer much insight into these questions. While thematic diversity per se has not been studied often in the context

of campaigns, the impact of message choice in campaign strategy has been addressed extensively by previous research. This somewhat distinct line of scholarship has examined the role of message choice in candidates' strategic communication often advocating for "staying on message" (Benoit et al., 2011).

This stance is perhaps best exemplified by Bill Clinton's strategist James Carville's famous adage: "It's the economy, stupid." This line was prominently presented on a whiteboard at Clinton's 1992 presidential campaign headquarters in Little Rock, Arkansas,<sup>1</sup> reminding those who worked in the campaign headquarters to stay focused on this one single issue. Carville's idea was to connect every possible message opportunity to the theme of economy and thereby take advantage of economic sluggishness during the presidency of incumbent G. H. W. Bush (though the campaign had three main focuses: "Change vs. more of the same," "The economy, stupid" and "Don't forget health care"; Kelly, 1992). Simply put, this phrase, often repeated by media pundits and political strategists and lending its name to thousands of scholarly works, resonates with the argument that it is important to keep the campaign message coherent, succinct, and as unidimensional as possible.

However, much of the writing on "staying on message" often lacked in empirical evidence, relying mostly on case studies and anecdotal evidence. Additionally, past empirical examinations of the role of thematic diversity in campaigns (Benoit et al., 2011; Bradshaw, 2004; Sellers, 1998) devoted less empirical attention to state-level elections than to federal campaigns. Most importantly, they devoted less empirical

---

<sup>1</sup> This sign can be seen in Hegedus and Pennebaker documentary "The War Room" (1993) which followed Clinton's 1992 presidential campaign.

attention to media channels that are not controlled directly by candidates, such as candidates' news coverage. However, findings and theories gathered from this body of research can offer useful insights in thinking about thematic diversity in the context of U.S. Senate candidates' news coverage and direct communication with voters via social media.

One of the theories most often invoked in support of a monothematic message strategy is the theory of issue ownership (Benoit et al., 2011; Doerfel & Connaughton, 2009b; Sellers, 1998). As a concept, issue ownership is directly connected to the previously discussed theory of priming. Researchers have used priming theory as an approach to study the effects of media on electoral success. In this perspective, priming can be seen as an extension of agenda setting. If agenda setting examines the impact of object salience in the media on the object's perceived importance, priming is the result of this heightened perceived importance on observers' evaluations. For example, by heightening or decreasing the importance of social welfare issues or foreign affairs issues by giving them less air time, the media can influence the standard by which a candidate is evaluated—and hence, this candidate's favorability with the voters. Similar to the earlier mentioned theory of affective priming, studies of issue ownership theory are aimed at explaining which issues are most beneficial to candidates' image and, more importantly, why (Petrocik, 1996; Sheafer, 2007).

Issue ownership theory argues that parties “own” different political issues for which they are considered to be more capable and on which their message is considered to be more credible. Issue ownership often stems from historical conditions, cleavages, and disputes that gave rise to a party in the first place, as well as the social base of



support for a party (Walgrave, Lefevere, & Nuytemans, 2009). Parties often tend to emphasize only these issues when crafting their campaign messages. Given that constituencies believe their party to be capable of promoting their best interest and that the party often reciprocates by focusing on these issues even more (hence strengthening its record on them), issue ownership tends to be a long-term and self-reinforcing phenomenon (Benoit et al., 2011; Petrocik, 1996). The Democratic party in the U.S., for example, tends to “own” the issue of healthcare and welfare (as is the case with European social-democratic parties as well; Walgrave et al., 2009), while the Republican Party “owns” issues such as crime reduction and national security.

When voters assess a candidate, they factor in their own calculations of the candidates’ performance on various issues as well as the importance or weight that they give each issue (Druckman, 2004; Scheufele, 2000). When crafting their campaign strategy, candidates can choose between changing voters’ opinions on issues or altering the weight that they give to these different issues. Taking into account the relatively stable nature of issue ownership (at least in the national level and for salient issues; Walgrave et al., 2009), when a candidate or a party chooses a campaign theme, they will often focus on the issues that their party already owns, rather than addressing issues they have not been historically associated with, whether that means trying to change voters’ opinions on these issues or changing voters’ minds on the candidates’ relevant legislative voting or policy record. In other words, party or candidate attempts to “trespass” into another party’s owned issue may often prove futile, and is therefore avoided (Norpoth & Buchanan, 1992).

From a diversity and complexity point of view, the logic of issue ownership seems to suggest that focusing on a single message can benefit candidates. As each party owns only a limited set of issues, candidates find it beneficial to limit themselves to those issues in their message strategy (Benoit et al., 2011). It should be noted, however, that this choice is most often made by the candidate or campaign managers and as such, touches mostly on campaign-generated messages and communication. However, mass media often plays a part in this process as well, as various outlets emphasize the issues that different political actors own. When an issue that a candidate owns receives higher salience in the media, assessment of that candidate can often change accordingly (Druckman, 2004; Iyengar, 1987; Petrocik, 1996).

Several additional arguments can be offered in support of this stance. Bradshaw (2004) argues that a campaign message need only be either a rationale for the candidate's election or a rejection of his or her opponent. As voters often do not engage in politics too deeply, it is hard to get more than one argument or talking point through to them. Therefore, campaigns find it beneficial to focus on the single point at all times. As repetition and reinforcement are critical for message effect (Allport & Lepkin, 1945; Henkel & Mattson, 2011), the campaign should focus only on a small set of messages with as little variation as possible. Lastly, as audiences are not always attentive, during the moments when they are paying attention, the message delivered must be the strongest the campaign has to offer (Benoit et al., 2011).

In a similar manner and from the media's perspective, the media itself is limited in the number of issues to which it can devote attention at any given time. Therefore, at the same time, political actors and campaign managers should not overload the processing

capacity of the media (Shoemaker & Reese, 2014) by promoting too many frames and arguments that draw on a set of issues that is too diverse (Hänggli & Kriesi, 2012).

Based on these arguments, the following two hypotheses can be drawn regarding research questions 1 and 2:

*H5a: The thematic diversity of political candidates' news coverage negatively correlates with candidates' electoral success.*

*H6a: The thematic diversity of political candidates' direct communication via social media negatively correlates with candidates' electoral success.*

#### **2.4.2 Thematic diversity in elections: Multi-thematic message strategy**

The research on issue ownership seems to assume that candidates have complete control over their message strategy. As a consequence of this assumption, competing arguments can offer support for the opposite position—that is, that a more thematically diverse strategy could be beneficial at least when it comes to media coverage and local elections. These arguments are, in part, grounded in criticism of issue ownership theory, the pressures of the political and media landscape, and the need for message flexibility.

First, it may be the case that the assumption that issue ownership is an inflexible phenomenon is inaccurate. In an experiment conducted by Walgrave et al. (2009) in the context of the 2007 Belgian general elections, participants watched artificial news clips in which various candidates commented on their party's owned issues, other parties' owned issues, and issues not owned by any party. The results of the experiment show that addressing issues not owned by any party—or even issues owned by other parties—can change issue ownership attributions by the participants. In other words, media coverage

of a party in connection with an issue, and especially an unowned issue, can help a party strengthen its issue ownership. Therefore, while some scholars argue that a campaign should focus only on a single issue already owned by the party or candidate, an expansion of these themes might prove fruitful, especially when messaging expands into issues that are not owned by the opposing party, issues that the opposing party performs poorly on, or issues that can offer substantial electoral payoff.

Indeed, research shows that in reality most candidates do not stay on message, but rather tend to offer a set of messages on various topics and issues (Benoit et al., 2011; Norpoth & Buchanan, 1992; Walgrave et al., 2009). Moreover, while issue ownership offers compelling arguments for monothematic message strategies, evidence for a competing process of issue convergence in U.S. election campaigns has also been offered (Sigelman & Buell, 2004). Rather than discussing completely different sets of issues during a campaign, as would be predicted by issue ownership, U.S. presidential candidates were found over several decades to address a similar set of issues (Sigelman & Buell, 2004).

The explanations for this similarity (Sigelman & Buell, 2004), and for candidates' tendency to go off message (Benoit et al., 2011; Norpoth & Buchanan, 1992; Walgrave et al., 2009), emphasize the pressures that the media and political systems place on candidates' message strategy. Candidates are often forced to follow the media's agenda (rather than setting it) and are often asked for their reactions to opponents' stances on various issues rather than address issues of their own choosing. For example, if a terrorist attack takes place in Europe during a campaign, no candidate can allow themselves to continue to address only issues related to domestic social equality; every candidate would

likely be asked to comment on issues of national security. Candidates who fail to provide adequate responses, it is argued, can be penalized in the polls or considered as detached (Sigelman & Buell, 2004). However, it is important to note that this requirement that candidates be able to offer arguments and messages on a wide range of topics may only be applicable to national presidential races, rather than local elections.

From a media perspective, addressing one's unowned issues can be imperative in specific media arenas in which the choice of topics is not controlled completely by the candidates. For example, in an interview or during a debate, candidates are often required to react to whatever issue they are presented with. Though "spinning" the question is a possibility, candidates might not always be able to reliably turn to this strategy. However, candidates' ability to address a variety of issues depends on their credibility in the topics and their record on the issues. Candidates prefer to address issues on which they have established record (Sellers, 1998). This record can be established by issue ownership (as discussed earlier), by their political past—but also, at least partially, by the media and by candidates' own message strategy. The thematic diversity of candidates' news coverage and message strategy can therefore provide them with a larger pool of issues to choose from, with a larger record to draw on, and a larger variety of issues to credibly "own," when in need.

Thus, the first explanation for the relationship between a more thematically diverse media coverage of candidates and electoral success is indirect—diversity in candidates' news coverage provides them with necessary flexibility in message strategy. There are plenty of situations in which candidates need to stray from their central message (Benoit et al., 2011; Hänggli & Kriesi, 2012), including the failure of the candidate's original

strategy, “narrowcasting” (offering different messages to different publics), debates and interview contexts, and sudden changes in the political environment. A more thematically diverse set of news coverage can may thus candidates to draw on a wider range of issues for which they can be awarded credibility (Norpoth & Buchanan, 1992). A candidate who never associated him- or herself to a specific issue domain, or who has never spoken about an issue, might find it hard to “trespass” into such a domain owned by his or her opponent if they are all of a sudden required to do so by changing campaign circumstances (Norpoth & Buchanan, 1992).

Further consideration of candidates’ lack of control over the media’s agenda from the perspective of agenda setting and priming provides a more direct explanation for the effect of media thematic diversity on candidates’ performance as well. In line with issue ownership theory (Petrocik, 1996) and priming (Iyengar et al., 1982), publics’ assessment of candidates is shaped by the standards or contexts in which they are judged. If voters are more concerned about environmental or economic issues, this preference can impact their general assessment of candidate’ performance and character. However, different constituencies might value such issues differently. If one group of voters deems an issue to be of great importance but a candidate’s image lacks credibility or experience on this specific issue, then voters’ assessment of this candidate can be negatively impacted (Sigelman & Buell, 2004). Thus, if environmental issues are critical to one group of voters and job creation is the standard by which another group judges candidates, a candidate whose news coverage connects him to both issues will be able to cast a wider “priming net” and not be seen as inexperienced. A more diverse record for candidates, as established by the media, may not only affect their flexibility in terms of direct

messaging, but could also impact their public perception directly, putting them in a more advantageous position relative to multiple potential media agendas.

Relatedly, candidates might want to vary their message when addressing different audiences (Benoit et al., 2011; Jacobs & Shapiro, 2005). While keeping the message monothematic when addressing each specific group directly, candidates need to be able to narrowcast—that is, to tailor their message specifically to the various constituencies that they hope to attract. As with the earlier example, candidates hoping to be credibly associated with multiple issues should have a more varied and complex image to draw on—an image established by lengthy, repeated, and diverse exposure to news media (Sellers, 1998).

The evidence and arguments for monothematic message strategy's advantages often rely on studies of campaigns at the national level or which focus on the strategy for crafting candidate-produced messages (Benoit et al., 2011; Bradshaw, 2004; Hänggli & Kriesi, 2012; Sellers, 1998). However, the case for more local elections, such as U.S. Senate elections, and the case of news media coverage, rather than candidate-produced messages, might be different and for several reasons. Candidates may not be able to keep their message in the news as constantly as they can in their own produced messages. Unlike ads that candidates produce or the speeches that they give in carefully orchestrated rallies, in mediated channels, candidates must cater to the needs and values of journalists (Hänggli & Kriesi, 2012) and therefore may need to stray from their main message if the discussion requires it. Moreover, while issue ownership is a long-term phenomenon (Petrocik, 1996), evidence suggests that it may be possible for candidates to change or claim new issues, especially when they are less salient or not previously owned

(Walgrave et al., 2009). It is also possible that concerns at the local level are less salient than national-level issues and that issues are owned by a specific party only to a small extent. Local campaigns might thus offer more opportunities for candidates to trespass into others' messages.

The relationship between thematic diversity and electoral success can also be explained from a media logic perspective. Diversity in a candidate's news coverage, for example, can be the product of a favorable standing in the polls. Given on-air time constraints, when newsrooms devote more resources to coverage of successful candidates, an increase in overall resources might translate not only to a larger volume of coverage (with the upper bound limited), but also to a larger number of issues and contexts in which a candidate is discussed. Similarly, and from a political experience perspective, the media may reflect the political reality. An incumbent candidate with more political experience might receive more diverse coverage due to his or her political record being more expansive and touching on a wider range of issues and themes. Incumbent candidates thus benefit from their activity across varied political contexts, as this means that they are *covered* in a wider variety of contexts (Sellers, 1998). This explanation can also be extended to other political factors, such as a state or district's political leaning, or midterm election cycles that are often perceived as a referendum on the president's party (Grofman, Brunell, & Koetzle, 1998). Diversity can be affected both by candidates' message strategy and political factors concurrently. For example, given its inherent benefits in terms of diversity (Greene, 2016), incumbency might contribute to diversity of news coverage, not only because it offers more issues to which a candidate



can be connected, but also because incumbents' message strategy offers a more diverse array of themes.

Lastly, the impact of diversity in news coverage on support for candidates can also be conceptualized from an argument quantity perspective. Evidence show that the quantity of arguments, rather than merely their quality, can affect arguments' persuasiveness and opinion formation (Petty & Cacioppo, 1986). Evidence shows that this effect is even stronger for specific demographics, such as older voters, and for low salience or low involvement issues (Wang & Chen, 2006). Thus, by touching on a larger set of issues, thematic diversity in news coverage can provide a greater number of reasons for supporting a candidate.

Therefore, and more formally, the following two competing hypotheses can be drawn regarding the role of thematic diversity in electoral campaigns:

*H5b: The thematic diversity of political candidates' news coverage negatively correlates with candidates' electoral success.*

*H6b: The thematic diversity of political candidates' direct communication via social media negatively correlates with candidates' electoral*

The next chapter elaborates on the methodological considerations required to test these hypotheses against each other, focusing on the measurement of thematic diversity in the current literature, the limitations of current measurements, the logic of diversity in disciplines outside of communication and the social sciences, and the application of this logic to unsupervised machine learning method.

### 3. METHODOLOGICAL FRAMEWORK

This chapter discusses issues relating to the conceptualization and measurement of thematic diversity using unsupervised machine learning methods. Whereas the next chapter covers methods and discusses specific procedures, parameters, and tools used in this study, as well as descriptive data for the corpora, this chapter focuses on the conceptualization of thematic diversity from a broader theoretical perspective.

I begin by addressing current approaches used to estimate diversity in communication research and their limitations. I then discuss emerging research on diversity in other scientific fields, such as biology, physics, public policy, and information sciences, to identify meaningful structural features that an adequate diversity estimation must address. These include variety, balance, and disparity. I then apply these features to a hypothetical and simplified example of candidates' theme structure to explain how each feature can be adapted from diverse research areas to thematic diversity in political rhetoric.

Following this more high-level discussion, I explain how the concepts of variety, balance, and disparity can be applied through the unsupervised machine learning methods used in this study. I first describe the method of topic modeling in general and its application to political discourse, followed by a conceptualization of variety, balance, and disparity using the overall topic structure data. Similarly, I discuss research on semantic network analysis in general, and particularly in political communication, including the limitations of the current state of the field. I then discuss the application of the logic of variety, balance, and disparity to semantic network analysis, including dilemmas and

limitations. These discussions preface the methods section in the following chapter, in which I discuss the specific procedures carried out in this study in greater detail.

### **3.1 Challenges in Thematic Diversity Measurement**

As reviewed in the previous chapter, various researchers have developed estimations for diversity to examine the antecedents and consequences of thematic diversity. These researchers commonly use manual content analysis to estimate the number of topics. The most complex of such measures is a method that Suedfeld and Tetlock (1977) developed in their studies of integrative complexity, which has been adapted by more than 100 subsequent studies. This theoretical framework has been popular as a measurement of cognitive complexity and is aimed especially at measuring complexity in archival material produced by political leaders to examine their decision-making processes. Such studies have examined integrative complexity in times of peace, looking at supreme court judges (Gruenfeld, 1995) and presidents (Suedfeld et al., 2011; Thoemmes & Conway, 2007), but also times of conflict—for example, integrative complexity in the writing of decision-makers involved in the Cuban Missile Crisis or the Gulf War (Suedfeld & Rank, 1976; Suedfeld, Wallace, & Thachuk, 1993).

Despite the popularity of the integrative complexity framework, a number of critics have argued that the training requirements to perform this type of analysis are too cumbersome, leading to its under-utilization (Conway et al., 2014). The manual for coding integrative complexity lays out the process of measurement and estimation (Baker-Brown et al., 1992), which focus both on differentiation and integration on a 1-7 scale. For each paragraph, the coder estimates whether the author of the paragraph:

perceives only one variable or process in decision-making and argumentation (1), recognizes two different variables to pay attention to (3), perceives the interaction between these two variables (5), and perceives not only the interaction but also the interdependence of these two variables (7). It then goes on to detail examples of each of the levels. Coder training requires a relatively intensive process and evaluation that multiple coders are required to pass in order to perform the analysis (for reasons of reliability establishing, time constraints, and research staff turnover). Preprocessing also takes considerable resources, as materials need to be prepared for analysis, separated into paragraphs, and anonymized manually to prevent coder bias.

Similarly, a series of studies in communication focusing on agenda diversity has used manual content analysis to estimate topic diversity in the media (McCombs & Zhu, 1995; Peter & De Vreese, 2003). In these studies, the documents under examination were coded for the appearance of 12 specific topics: “jobs/unemployment,” “welfare,” “money,” “public spending,” “law and order,” “government/political decision-making,” “social relations,” “environment/food,” “technology/research,” “EU-related problems,” “foreign policy/affairs,” and “miscellaneous.”

Both methods highlight some of the challenges in the manual estimation of diversity.

### ***Problem 1: A priori knowledge***

The first problem is the need for an a priori knowledge of the issues, themes, or topics in a corpus. Most of these studies rely on hand-coding to identify themes or frames in media coverage and public opinion; as a consequence, researchers must have a pre-

existing coding schema at the outset. This may be problematic when analyzing corpora in which the potential topics are unknown. For example, looking at the categories offered by McCombs & Zhu (1995), the researchers assume that these 12 topics are a complete set of all possible topics in the corpus. If other issues arise, diversity will be underestimated. Further, if unacknowledged topics exist only for some cases or for some media but not others, then the diversity estimation is likely to be skewed, reducing the validity of the results. Therefore, the methods suggested in this study rely on unsupervised machine learning, which is well suited to drawing topics and themes organically and inductively from the corpus rather than using a pre-existing coding schema.

### ***Problem 2: Comparing different corpora***

The potential for skewed data persists even when researchers have valid a priori knowledge of the possible themes in the texts under examination. From a comparative point of view, existing methods might be limited if researchers aim to analyze corpora in which the possible set of topics changes drastically from one group of texts to the other as this renders the application of a unified coding schema to all corpora impossible.

In order to reproduce or compare results across studies, all would need to apply the same unified coding schema. However, topics and issues are likely to change over time, with new topics introduced into the agenda at different historic points in time, or even across geographies as different counties, states, or countries might have different possible agenda topics; issues could also be medium-specific, with some media types concentrating on different topics, or unique to a specific political context, for example, as the set of issues pertinent to presidential elections diverge from the agenda of mayoral elections.

Such discrepancies could also persist for individual studies comparing corpora from different contexts. So, for example, the range of issues in the context of Senate elections in California is likely to be different, at least to some extent, from coverage of Senate elections in Idaho. These issues are further exacerbated by the unequal distances between topics (i.e., some issues are more closely related than others) and the interconnectedness of various possible agenda topics.

***Problem 3: Interconnectivity of themes***

The method commonly employed to study agenda diversity (McCombs & Zhu, 1995; Peter & De Vreese, 2003) implicitly assumes that all topics are independent, or at least different to a similar extent from each other. Thus, if two corpora have an equal distribution of two topics, these corpora will be similar in diversity, even if the two topics in one corpus are more or less similar to each other than in the other corpora.

Despite the overall lack of methodological attention to the interconnectedness of certain topics and themes, this issue has received some theoretical treatment in the form of nominal and thematic diversity. Measuring thematic diversity both from the agenda and framing perspectives, researchers have differentiated between nominal and thematic diversity (Kleinnijenhuis et al., 2015; Peter & De Vreese, 2003), with nominal diversity defined as the number of issues found in the media or in public opinion and thematic diversity as the semantic or categorical diversity of issues in the media or in public opinion. To illustrate this, imagine two candidates. Candidate A discusses unemployment, welfare, and public spending. Candidate B discusses law and order, environmental issues, and foreign policy. Both candidates exhibit equal nominal diversity (three issues). However, from a thematic diversity perspective, candidate B is more

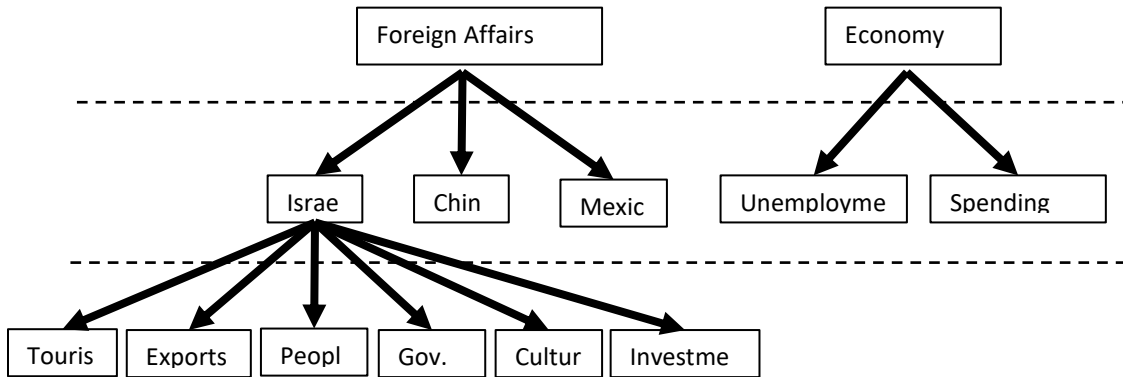
diverse in his issue choices. This is an important distinction, as it raises the challenge of differentiating between issue categories, an obstacle that the methods used in this dissertation directly address.

As a solution, or to estimate thematic rather than nominal diversity, Peter and De Vreese (2003) grouped several categories into larger thematic contexts. The problem is that this potential solution reduces the resolution that the analysis can provide and assumes the relationships between these topics a priori, which in turn leads to the same problems mentioned above. For example, according to Peter and De Vreese's (2003) method, a candidate who discusses unemployment will be considered as diverse as a candidate who discusses unemployment and welfare. These issues can be identical in some contexts, states, or media, or they can be very different in scope and meaning.

Despite this pitfall, the distinction between thematic and nominal diversity is valuable for conceptualizing an estimation for thematic diversity. The main contribution of the distinction is that it emphasizes accounting not only for the number of topics in a given corpus or set of corpora, but also their interconnectivity. While a given corpus might exhibit many topics, these can be closely related, thus reducing the corpus' thematic diversity. However, a corpus could also have fewer topics that are drastically different from each other, thus enhancing thematic diversity. Therefore, the methods suggested in this study account not only for the distribution or quantity of the themes, but also estimates their interconnectivity.

**Problem 4: Defining the resolution of theme structure**

A final challenge is how to determine the level at which a thematic structure should be estimated. Figure 1 shows a hypothetical topic structure for a corpus. This corpus can be divided into issues of foreign affairs and issues relating to the domestic economy. However, both issues can be further divided into sub-topics.



*Figure 1: A Hypothetical corpus thematic structure.*

The issue of economy can be divided into unemployment and spending, as in the coding framework suggested by McCombs & Zhu (1995). The issue of foreign affairs can be subdivided into issues relating to Israel, China, or Mexico, all of which are foreign affairs issues, but nonetheless relate to different foreign actors. Each of these foreign affairs topics can be further subdivided, for example, to theoretically driven components of nation brands (Anholt, 2006), such as a country’s tourism potential, culture, or government. And, of course, these sub-topics themselves can be further divided to smaller and more specialized issues.

The issue of resolution determination poses several problems that relate to the challenges already discussed. First, it is clear that the distances between topics at different levels is not identical. The difference between issues of foreign affairs and the



economy is much larger than the difference between Israel's government and the potential for investment in Israel. Second, even at the same resolution level, the distance between each dyad of topics might not be similar. For example, is the difference between culture and investment identical to the thematic distance between tourism and exports? Third, diversity estimations for the relationships between topics at each of these levels changes drastically. If we were to zoom-out too much, then we might miss important distinctions between thematically similar topics; conversely, if we were to zoom-in too closely, then we might risk overestimating diversity. Finally, as the different levels cannot be applied uniformly over all topics and all corpora, it is impossible to quantify whether the same level of resolution was achieved, for example, in economy-related issues and foreign affairs issues.

***Addressing these challenges using unsupervised machine learning approaches:***

In this study, two different unsupervised machine learning approaches are used to explore thematic diversity in large corpora. Machine learning is a broad class of methods that leverage observed data to make predictions on unknown or future data (Grimmer & Stewart, 2013). Supervised machine learning methods identify a set of known categories in existing data, such as political affiliation, sentiment, or objects in an image. For example, logistic regression models are a basic form of supervised machine learning method, with the dependent variable serving as a known category to classify.

Unsupervised machine learning methods, by contrast, are aimed at clustering existing data into a set of *unknown* categories. These methods are well-suited to address the challenges to thematic diversity measurement surveyed in this section. First, unsupervised learning methods do not require a priori assumptions of possible themes,

issues, or topics. Second, they allow for comparison of different corpora by using the same identical procedure over all corpora. Third, they enable the researcher to account not only for the number of categories, but also the interconnectivity of these categories; this elides the potential bias of researchers' binary decisions on the similarity or difference between topics, as it provides a more fine-grained estimation of the extent to which two themes are not entirely independent but also not entirely identical (as will be elaborated when addressing measurements of disparity).

The first method, topic modeling, estimates the topic structure over all corpora at once, thus enabling an examination of diversity in all cases using a single standard. In addition, by exploring not only the distribution of all topics but also the extent to which these topics share a vocabulary, topic modeling also accounts for both the nominal and thematic perspectives. The second method, semantic network analysis, estimates the number of themes and their interconnectivity. By applying the same benchmarks and procedures over the various corpora used in this study, semantic network analysis allows for comparison of dramatically different corpora. By exploring community structure diversity, this technique estimates not only the number of topics in and across corpora, but also the extent to which these topics relate to one another from a shared vocabulary point of view.

In the following sections I review both methods and their usage for the conceptualization of thematic diversity in large corpora.

### 3.2 Thematic Diversity - Conceptual Considerations

As mentioned earlier, the concept of diversity is especially interesting due to its prevalence in various scientific fields, from biology to physics, information sciences, economics, public policy, and more (Stirling, 2007). This rich body of research helps us to think more broadly about the nature of diversity measurement and allows us to incorporate lessons learned from various scientific fields—for example, the measurement of biological diversity in an eco-system (May, 1990)—to rethink diversity in the context of thematic structure.

Stirling (2007) points toward three features of diversity that are prominent across almost all of these fields. These are *variety*, *balance*, and *disparity*. Variety measures the number of categories to which system elements can be divided. For example, how many animal species exist in a specific ecological system. All other features being equal, the more species that are in a system, the more diverse that system is. A system with tigers, elephants and snakes is more diverse than a system with only tigers and elephants. However, this categorization does not take into account the importance of these categories.

The second feature, balance, asks how equal the distribution of categories is. This type of estimation has of course been central to communication research and is the focus of Shannon and Weaver's (1949) H-entropy measurement of diversity. Continuing with the ecological example, a system with 1,000 elephants and one tiger is less diverse than a system with 501 elephants and 500 tigers. Thus, with variety being equal, higher balance (or in Shannon and Weaver's terms, higher entropy) indicates higher diversity. However, Shannon and Weaver's measurement (which is the most commonly used in

communication studies) does not take into account a third critical feature of diversity: disparity.

Disparity is the extent to which categories differ from each other. “An ecological community comprising 20 varieties of beetle is less diverse than the one comprising less than 20 species drawn from different insect, reptile, and mammalian taxa” (Stirling, 2007, p. 710). Thus, to understand and measure diversity, we need to ask not only how many categories are in a system, or even how balanced their distributions are, but also whether and to what extent these categories differ from each other.

The following formula, accounting for variety, balance, and disparity, was offered by Stirling (2007) based on Rao’s diversity coefficient (Rao, 1982). ( $\delta_{ij}$ ) indicates the distance between each two categories, with ( $p_i$ ) indicating the prevalence of category (i) in the corpus:

$$D = \sum_{ij(i \neq j)} (\delta_{ij}) \cdot (p_i \cdot p_j)$$

These three features should be kept in mind when discussing any estimation of thematic diversity. In addition, each of these concepts, as well as variables ( $\delta$ ) and ( $p$ ), needs to be translated into an operational procedure for estimation. However, before discussing the methods and calculations that were used to estimate diversity (using both topic modeling and semantic network analysis), it is important to discuss the applications of these diversity dimensions to the context of theme or topic analysis, and the logic of this formula.

The first and simplest consideration is variety, or in this case, the number of themes in a given corpus. In accordance with Stirling's (2007) formula, more themes (*ceteris-paribus*) should result in higher diversity. This can be seen by the example in Figure 2.

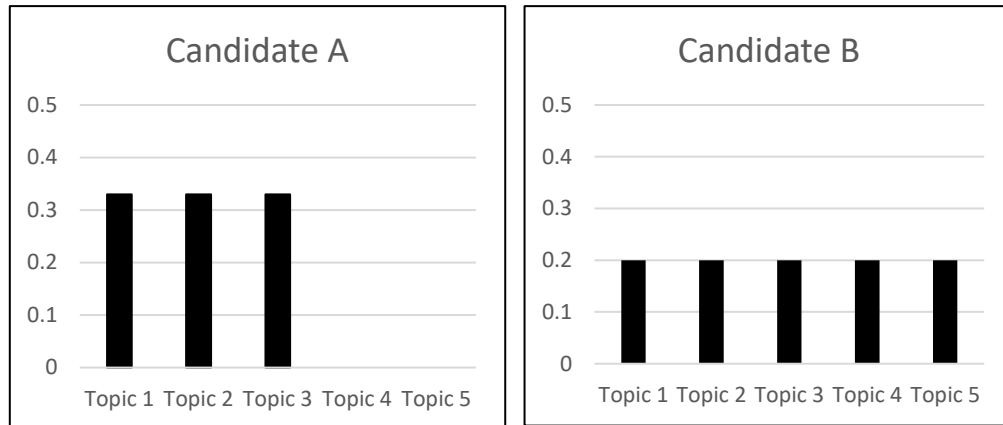


Figure 2: A hypothetical topic distribution for two candidates (3 and 5 topic structure).

Figure 2 visualizes the topical structure for the coverage of two hypothetical candidates. Here, candidate A discusses three topics in equal proportion (33%), while candidate B discusses five topics in equal proportion (20% each), thus keeping the balance between topics constant for both of the two candidates. With all else being equal (we also assume these topics are similarly independent from each other), we can argue that the diversity estimation for candidate B should be higher than the diversity estimation for candidate A.

However, this is not the only relevant consideration. From a balance perspective, while two candidates might discuss the same number of topics, a candidate can mention a topic prominently, thus giving it more weight, or sparingly and in passing, thus making the topic less relevant. Thus, aside from the number of themes, one should also consider the distribution of topics. In a second example given in Figure 3, both candidates are

equal in the number of topics they discuss. However, the distribution of the topics is different.

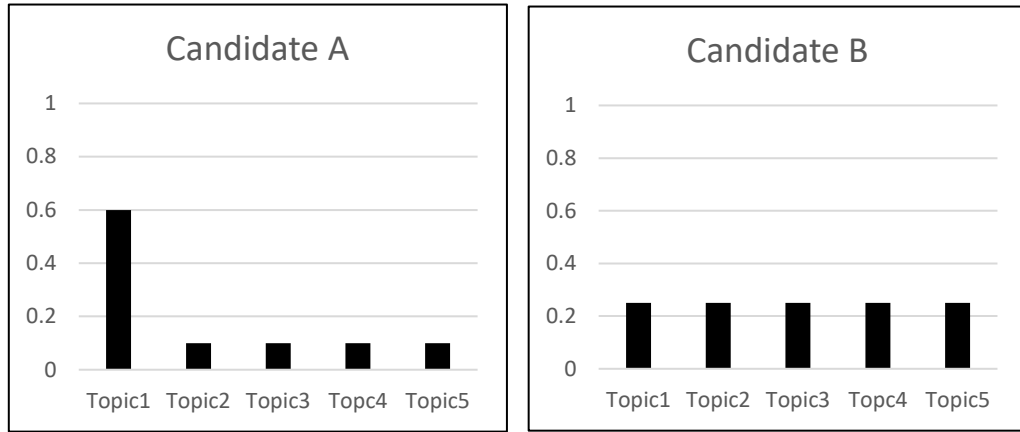


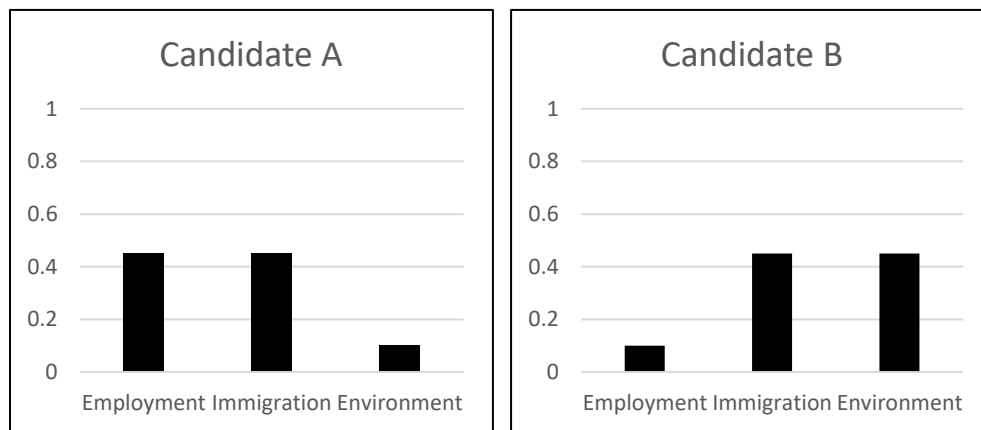
Figure 3: A hypothetical topic distribution for two candidates (5 topic structure).

Here, candidate A discusses topic 1 six times more often than topics 2-5. Candidate B, on the other hand, pays relatively equal attention to all five topics. Thus, from a diversity point of view, candidate B should be considered to be more thematically diverse than candidate A, even though they are identical in terms of the number of topics they address. Thus, an adequate diversity estimation should address the distribution of themes, with a more equal distribution resulting in higher thematic diversity estimate.

In order to address these two considerations (variety and balance), the second part of the formula presented above can be used:  $D = \sum_{ij(i \neq j)} (p_i \cdot p_j)$ . The more equal the distribution of topic proportion is, the higher the diversity estimation will be. For candidate A, for example, the summation of all topic proportion multiplications will be equal roughly to 0.6. For candidate B, the summation of all topic proportions multiplications will be roughly equal to 0.8. In addition, the maximum value of diversity

rises with the number of topics. The maximum value for a two-topic equal distribution is  $0.5 (= 2 \cdot 1 \cdot 0.5^2)$ , the maximum value for a 5-topic equal distribution (as in this case) is  $0.8 (= 5 \cdot 4 \cdot 0.2^2)$ , the maximum value for a ten-topic model is  $0.9 (= 10 \cdot 9 \cdot 0.1^2)$ , and the maximum value for a 100-topic model is  $0.99 (= 100 \cdot 99 \cdot 0.01^2)$ . Thus, the first part of the equation accounts for both variety and balance. More equal distributions with higher number of categories will receive a higher score in terms of diversity.

Lastly, and building on the nominal and thematic diversity discussion as well as the concept of disparity, attention should be paid to the interconnectivity of different topics. That is, we need to account not only for the number of topics a candidate discusses, or how equal their distribution is, but also how connected these topics are to each other.



*Figure 4: A hypothetical topic distribution for two candidates (3 topic structure).*

In Figure 4, both candidates discuss the same number of topics (three) and the distribution of these topics is identical between the two candidates (two topics at 45% and one topic at 10%). Thus, they are equal in terms of variety and balance. However, the topics that each candidate emphasizes are different. Candidate A pays more attention to employment and immigration, which, for the sake of this argument, are closely related, as

immigration-related rhetoric often addresses issues of employment. Candidate B emphasizes two different issues, immigration and environmental issues. Again, for the sake of this argument, I will assume that these two topics are completely unrelated to each other, or at least much less similar to each other than employment and immigration (though of course, in reality these two topics might be more closely related, and this relationship should be measured empirically from the data as explained in the topic modeling and semantic network analysis sections).

In this example, the extent to which each topic is similar or different from each other plays a key role in diversity estimation. Candidate A pays much more attention to two very similar topics. Thus, the contribution of these two topics to their thematic diversity estimation should be rather minimal, as they serve almost as an identical topic. Candidate B, however, pays attention to two very different topics, thus maximizing the extent to which they diversify their rhetoric. Thus, an adequate diversity estimation should account not only for the distribution of topics or their number, but also how similar the topics are to each other. A candidate might discuss a large number of topics, but if all topics are near-identical, then a candidate that discusses a smaller set of unrelated issues might be considered more diverse (similar to the biological example mentioned earlier regarding spiders and mammals).

In order to address this, the diversity formula offered by Stirling (2007) adds a value ( $\delta$ ) as a weight for each pairwise topic proportion multiplication. This value estimates the distance between each pair of categories. Therefore, if categories (i) and (j) are extremely similar, then  $(p_i \cdot p_j)$  will be multiplied by roughly 0 (no distance), such that their contribution to the diversity estimation will be virtually nullified (as they almost



do not constitute two categories, but rather one due to their similarity). Two completely independent categories will have a ( $\delta$ ) score equal to 1, and hence the value of ( $p_i \cdot p_j$ ) will be maximized.

In the example above, the diversity for candidate A can be formally presented as:

$$(\delta_{12}) \cdot (0.4 \cdot 0.4) + (\delta_{13}) \cdot (0.4 \cdot 0.2) + (\delta_{23}) \cdot (0.4 \cdot 0.2) = .16\delta_{12} + .08\delta_{13} + .08\delta_{23} = .08(2\delta_{12} + \delta_{13} + \delta_{23})$$

The diversity for candidate B can be formally presented as:

$$(\delta_{12}) \cdot (0.2 \cdot 0.4) + (\delta_{13}) \cdot (0.2 \cdot 0.2) + (\delta_{23}) \cdot (0.4 \cdot 0.4) = .08\delta_{12} + .08\delta_{13} + .16\delta_{23} = .08(\delta_{12} + \delta_{13} + 2\delta_{23})$$

Assuming that employment and immigration are proximal topics, and that immigration and environmental issues are distal, we can then state that:  $\delta_{12} < \delta_{23}$ . Looking at the estimation above, we can conclude that:  $.08(2\delta_{12} + \delta_{13} + \delta_{23}) < .08(\delta_{12} + \delta_{13} + 2\delta_{23})$ , or in other words, that the diversity estimation for candidate B is larger than the estimation for candidate A. Hence, the formula offered here can account not only for the number of topics and their distribution, but also for the specific interrelationship between each pair of topics.

In the next section I discuss the two unsupervised machine learning methods used to extract the themes and topics from the corpora under consideration. I will then turn to address how variety, balance, and disparity can be calculated using these theme structures. Although the method of calculation differs between the two methods due to their distinct outputs and data structure, they are both guided by the same three considerations and the same basic formula.

More specifically, I will focus on the two critical estimations needed to address the three features of variety, balance, and disparity: the calculation of topic prominence ( $p$ ) and the calculation of similarity or distance between each two topics ( $\delta$ ).

### **3.3 Unsupervised Machine Learning Approaches to Thematic Diversity**

Computational content analysis describes an extremely diverse group of methods. While some researchers exploit the efficiency of computational tools to reduce the costs of content analysis, others use such tools to discover latent features in texts that are not naturally interpretable by humans. What unites this set of methods is that it relies on quantitative models of language, which, despite being flawed and reductionist, are useful—accurate enough to give researchers useful insights into media content (Blei, 2011; Grimmer & Stewart, 2013). Most methods in computational content analysis can be divided to two major groups: supervised and unsupervised (Petchler & Gonzalez-Bailon, 2013). This categorization is largely based on the distinction between methods for classifying texts into known categories (supervised), and those that classify texts into unknown categories (unsupervised).

Supervised methods are often used in a more deductive manner, assigning documents or terms to pre-determined and known categories. For example, in Bélanger and Soroka's (2012) study, dictionary-based methods are used to assess whether texts are negative or positive in sentiment—that is, a limited set or range of known categories. The dictionary-based approach utilizes lists of tokens that are annotated for their value on different scales, for example, their positive and negative valence. Combining the frequency of words in a set of documents with information on those words' attributes from a sentiment dictionary, each document is assigned a specific value to estimate the

prevalence of negative and positive words and, hence, its total sentiment. Limitations of such methods will be discussed in the conclusion chapter (González-Bailón & Paltoglou, 2015).

Unsupervised methods, by contrast, do not assume a pre-determined set of categories for a given corpus. Unlike supervised methods, unsupervised methods are more inductive in nature, extracting the set of possible categories from the corpora themselves. These approaches are especially useful in contexts where the set of possible topics for coding is unknown a priori, or when addressing (as this study does) the structure of these topics rather than their nature. In addition, these methods are extremely cost effective (Blei, 2011; Petchler & Gonzalez-Bailon, 2013; Roberts, Stewart, & Tingley, 2014), thus enabling analysis of large corpora that would otherwise be impossible to address using manual content analysis. This study utilizes two unsupervised methods for data analysis: topic modeling and semantic network analysis.

### **3.3.1 Topic modeling**

Topic models are a broad class of unsupervised text analysis methods aimed at providing cost-efficient and automated procedures for classifying texts into a set of latent categories, which are referred to as topics. Despite the label “unsupervised”, this procedure is not independent from researcher choices and decisions. For example, specific hyper-parameters can be chosen by the researcher, as well as the number of topics in the model. At the basis of this method (as well as semantic network analysis, which will be reviewed next) is an understanding of a text’s meaning as relational—words that appear together are assumed to be thematically related.

Using co-occurrence as an assessment of meaningful relationships, these methods do not take into account other language features, such as syntax, narrative, or document structure. Instead, they rely almost exclusively on the “bag of words” approach. Despite the seeming simplicity, flaws, and reductionism of topic modeling approaches, they have nonetheless proven to be a powerful tool in various social science-related fields (Blei, 2011; Grimmer & Stewart, 2013). These tools aid researchers by reducing large amounts of data to a more easily interpretable sets of matrices related to topic structure and document structure, with an initial study by Blei, Ng, and Jordan (2003) cited in thousands of academic publications in the first decade following its publication.

The basic intuition behind topic modeling is the view of documents as mixtures of topics, and topics as a cluster of words that tend to co-occur in these documents. Thus, the various algorithms used in this set of methods are designed to estimate the latent unobserved structure of topics based on the observed words, documents, and word-document distributions. This intuition can be further exemplified by a generative process.

We assume each document in the corpus has a specific mixture of topics. Table 1 offers an extremely minimal and simplified version of such a topic distribution.

*Table 1: Simplified topic-document matrix*

	Topic 1	Topic 2
<i>Document 1</i>	0.7	0.3
<i>Document 2</i>	0.8	0.2
<i>Document 3</i>	0.9	0.1
<i>Document 4</i>	0.5	0.5
<i>Document 5</i>	0.01	0.99

In the matrix presented in Table 1, the rows represent the documents in a given corpus and the columns represent two possible topics. In the hypothetical example, the

first document presents a larger probability for topic 1 (0.7) and a smaller probability for topic 2 (0.3). Document 4, by contrast, contains an equal probability for topics 1 and 2 (0.5). Given the distribution for document 1, and (n) word length for this document, a hypothetical author draws a random topic for each of the (n) words in this given text. As can be seen from table 1, and as can be understood from the generative process example, the sum of probabilities for all topics in a given document must equal 1, as these probabilities encompass all possible alternatives for a topic choice.

Drawing a topic for the first word in the first document, there is a higher probability the author will draw topic 1 rather than topic 2. Following the drawing of a random topic for each word space, the author then randomly chooses a word from a topic-word distribution as exemplified in the matrix in Table 2:

*Table 2: Simplified topic-word matrix*

	Topic 1	Topic 2
<i>Word 1</i>	0.4	0.1
<i>Word 2</i>	0.3	0.1
<i>Word 3</i>	0.1	0.1
<i>Word 4</i>	0.1	0.3
<i>Word 5</i>	0.1	0.4

Table 2 presents the probability of choosing a word based on the choice of topic. For example, choosing topic 1, word 1 is more likely to be drawn than word 4. Thus, if the topic drawn from the first word-space was topic 1, then a random word will be drawn from the normal distribution of words for topic 1. Some words will be drawn with a higher probability and some with a much lower probability. In addition, similar to the choice of topics, and given that a word must be selected, the sum of probabilities for all words over a topic must equal 1. Moreover, each word has a positive probability of being drawn, even if this probability is infinitesimally small. Through this iterative process,

every word in the text is chosen, first by a drawing a random topic from the document-topic distribution, and then by drawing a random word from the word-topic distribution. The summation of all topic choices and all word choices to 1 is an important property of these distributions for the calculation of diversity, as will be shown later.

In this regard, and from a more conceptual perspective, it should be noted that topics are, in essence, considered a mere distribution of words. They do not contain any inherent meaning and their interpretation is often subjective, relying on the most prominent words in each word cluster. Thus, while the distributions are referred to as topics, issues, or even frames, one should be careful in interpretation.

The objective of topic modeling algorithms is to find the parameters that are most likely to generate the observed corpus. In other words, to estimate the topic-word and document-topic distributions that best approximate the set of documents. The result is an estimation of the latent topic-structure that characterizes a group of documents. The conditional distribution of hidden variables, given the observed variables, is computed through an iterative process of random topic-assignment and word-assignment to maximize various model evaluation criteria, as will be discussed in the methods section when addressing the choice of topic number ( $k$ ) used for each corpora in this study.

The choice of the appropriate number of topics ( $k$ ) is still a contested issue in the topic modeling literature. Some researchers offer a face-validity examination of the topic structure, comparing the results of the process to the researcher's knowledge of the corpus and context. However, this solution is problematic in studies like this one, when the size of corpus limits such an examination.

Moreover, the notion of “true” topic structure is problematic, as discussed in Section 3.1 and exemplified in Figure 1. While a topic structure can be divided to  $(k)$  topics, resulting in a stable and usable model, it is often the case that each of these can be further divided into subtopics. The choice of resolution level, therefore, does not have a “right” answer, and should be guided by the needs of the study and the theoretical framework. Such decisions require that the researcher decide on a trade-off between features such as an exclusivity of topics and their coherence. On the one hand, choosing a  $(k)$  value that is too small might result in what is referred to as “chimera topics” (Mickel, 2016). Similar to the mythic creature as an amalgamation of different animals, “chimera topics” are topics that are constructed from a number of radically different themes and clustered together erroneously. On the other hand, choosing too high of a  $(k)$  value might result in small topics that lack theoretical interest and that are closely related to each other. I elaborate more on this tension between coherence and exclusivity in the methods section. However, this tension is important to keep in mind, as it highlights the benefits of controlling for distance between each topic-dyad when estimating thematic diversity via the balance of the topic distribution—and guides the decisions made in this study to rely on higher  $(k)$  topic structure models.

This study uses two central matrices estimated by this process of model estimation. The first is a topic-document matrix  $(\theta)$ , which details the various topic distributions for each document in the corpus. The second is the topic-word matrix  $(b)$ , which estimates the multinomial word distribution for each latent topic. The first matrix addresses what Stirling (2007) refers to as balance, while the second matrix calculates pairwise theme disparity, or the dyadic relationships between each theme in the corpus.

### 3.3.2 Diversity in topic models

While the specifics of pre-processing, model selection, and model assessment will be discussed in the methods section, here I address the logic behind the topic modeling diversity estimation used in this study, using the Beta ( $\beta$ ) and Theta ( $\theta$ ) matrices. As detailed in earlier, the general diversity estimation in this study builds on Bache, Newman, and Smyth's (2013) method, applying it to a set of documents rather than a single document, as well as on Rao's diversity coefficient (1982) and Sterling's (2007) conceptualization of diversity measurement:

$$D = \sum_{ij(i \neq j)} (\delta_{ij}) \cdot (p_i \cdot p_j)$$

According to this formula, estimating the diversity of a group of texts requires several specific inputs. These inputs are the distance between each two categories (topics or themes) in a given corpus ( $\delta_{ij}$ ), as well as the proportion of these categories in a given corpus.

To estimate the proportion of topics in the corpus for each candidate, I use the Theta ( $\theta$ ) matrix, or the document-topic distribution. As detailed earlier, rows in this matrix represent the documents in the corpus and columns represent the topics.

Table 3: Hypothetical topic-document matrix for two candidates and two topics

	Topic 1	Topic 2
<i>Document A1</i>	0.7	0.3
<i>Document A2</i>	0.8	0.2
<i>Document A3</i>	0.9	0.1
<i>Document B1</i>	0.5	0.5
<i>Document B2</i>	0.5	0.5



In the hypothetical example in Table 3, the rows represent a corpus of five documents. In this example, the documents relate to two candidates. Documents A1, A2, and A3, relate to candidate A, while documents B1 and B2 relate to candidate B. For each of the documents, the topic distribution is detailed in the document's row, with topic indicated by the column. For example, document A1 is an uneven mixture of topic 1 (0.7) and topic 2 (0.3). In contrast, document B1 is an equal mixture of topic 1 and topic 2 (0.5).

From this table, we are able to assess the share of topics per candidate—or the average topic prevalence ( $p_i$ ) for each candidate. For example, for candidate A, the average share of topic-1 ( $p_1$ ) is the average of all topic-1 loadings for that candidate. In this case, it is equal to:  $\frac{0.7+0.8+0.9}{3} = 0.8$ . Similarly, the average share of topic 2 ( $p_2$ ) for candidate A is equal to  $\frac{0.3+0.2+0.1}{3} = 0.2$ . For candidate B, both topics have an average proportion of 0.5.

As shown earlier, the more equal a distribution is, the higher the sum of  $(\sum_{i \neq j} p_i \cdot p_j)$  will be. For example, for candidate B:

$$D_B = (\sum_{i \neq j} p_i \cdot p_j) = 0.5 * 0.5 + 0.5 * 0.5 = 0.5 .$$

For candidate A, by contrast:

$$D_A = (\sum_{i \neq j} p_i \cdot p_j) = 0.8 * 0.2 + 0.2 * 0.8 = 0.32.$$

In order to accurately estimate the thematic diversity of each candidate corpus, the distance between each pair of topics ( $\delta$ ) also needs to be estimated and incorporated into the diversity formula. The Beta ( $\beta$ ) matrix, which details the multinomial distributions of

words over topics (or the probability a unique word occurs for each topic), is used for this purpose. In the table below, each row represents a unique word (out of five possible words), while each column represents one of three different topics. Based on an understanding of meaning in a text as relational, I measure topic distance as the inverse of the similarity between these topics in terms of vocabulary. In the example provided in Table 4, topic-1 and topic-2 are more similar to each other than to topic-3, as both have a high probability of word 1 and word 2 appearing, while topic 3 has a high probability of word 4 and word 5 appearing.

Table 4: Hypothetical topic-word matrix for two candidates and two topics

	Topic 1	Topic 2	Topic 3
<i>Word 1</i>	0.4	0.3	0.1
<i>Word 2</i>	0.3	0.4	0.1
<i>Word 3</i>	0.1	0.1	0.1
<i>Word 4</i>	0.1	0.1	0.4
<i>Word 5</i>	0.1	0.1	0.3

Several estimations are available to measure these distances, the most common being cosine-similarity (Ramage, Hall, Nallapati, & Manning, 2009). With an identical number of words for each topic (as each word has a non-zero probability of appearing in each topic), the cosine similarity estimation is near-identical to the Pearson correlation of the two columns. This measure calculates the cosine of the angle between two vectors—for example, the word vector for topic-1 and the word vector for topic-2. More precisely, cosine similarity can be calculated as:  $\frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$ , with A referring to items in vector 1, and B referring to items in vector 2.

Thus, with vector 1 being the word distribution of topic 1, and vector 2 the word distribution for topic 2, the similarity between topic-1 and topic-2 can be calculated as:

$$\frac{0.4*0.3+0.3*0.4+0.1*0.1+0.1*0.1+0.1*0.1}{\sqrt{0.4^2+0.3^2+0.1^2+0.1^2+0.1^2}\sqrt{0.3^2+0.4^2+0.1^2+0.1^2+0.1^2}} = \frac{0.27}{0.28} = 0.96$$

The similarity between topic-1 and topic-3 can be calculated as:

$$\frac{0.4*0.1+0.3*0.1+0.1*0.1+0.1*0.4+0.1*0.3}{\sqrt{0.4^2+0.3^2+0.1^2+0.1^2+0.1^2}\sqrt{0.1^2+0.1^2+0.1^2+0.4^2+0.3^2}} = \frac{0.15}{0.28} = 0.53$$

Lastly, the similarity between topic-2 and topic-3 is identical to the similarity between topic-1 and topic-3 and is equal to 0.53 as well.

As this is a measure of *similarity* rather than distance, ( $\delta$ ) is estimated as 1- (cosine similarity). Or, in these examples, 0.04 and 0.47, respectively. If a candidate's corpus exhibits a large proportion of topic-1 and topic-2, then the contribution of these proportions to the estimation of this candidate's thematic diversity will be multiplied by 0.04 and, hence, will be limited in its contribution. From a conceptual perspective, this owes to the fact that these two topics share a large proportion of their vocabulary and hence contribute little to the thematic diversity of the corpus. However, emphasis on topics 1 and 3, or 2 and 3, will have a much larger effect on diversity estimates, as the quantities for these topics, ( $p_i$ ) and ( $p_j$ ), will be multiplied by 0.47, to incorporate the difference in their vocabulary into the model.

From this explanation, one can observe an apparent limitation of this method. The topic model is estimated over the whole corpus at once, or in other words, on all texts related to all candidates. As each topic has a positive probability of appearing in each document (no matter how small), the number of topics ( $k$ ) will be identical for all candidates in the corpus. Hence, while this method accounts well for disparity and

balance, it does not account well for variety. Moreover, because it makes estimates for all candidates at once, a new model needs to be constructed each time we want to add new cases, unless we assume that the topic structure is identical for both new and old cases. This assumption is problematic, for example, when adding candidates from new election cycles to a model calibrated to previous election cycles.

In addition, while this method can estimate the extent to which two topics are similar to each other, it is unable to account for the inner structure of the topics. Thus, it is vulnerable, for example, to “chimera” topics, or topics that erroneously cluster together otherwise distinct themes. This consideration will guide my decisions regarding model estimations, model choice, and the preferable number of topics ( $k$ ). Specifically, because the inclusion of disparity ( $\delta$ ) allows the diversity estimation to take into account the similarity between different topics, this method of estimation is insensitive to topic-resolutions that are too high and, as a consequence, can result in redundant topics, but it is also hyper-sensitive to topic-resolutions that are too low and thus result in “chimera topics”. Therefore, it is preferable to use a larger ( $k$ ) topic models when estimating diversity in a given corpus. I present evidence for this limitation in the results section, by exploring the impact of ( $k$ ) selection on the various regression models’ predictive power.

In the next section, I turn to discuss another unsupervised method—semantic network analysis. This method offers a complementary measure of diversity, based on the structure of the network representation of word co-occurrence within a corpus.

### **3.3.3 Semantic network analysis**

The second method used to explore the issue of thematic diversity in this study is semantic network analysis—a sub-field of general network analysis and computational textual analysis. The study of networks as patterned social interactions dates as early as the 19th century (Freeman, 2004). As a topic in mathematics, it deals with the graph representation of complex relationships between sets of objects (Baronchelli, Ferrer-i-Cancho, Pastor-Satorras, Chater, & Christiansen, 2013). Dramatic growth in the availability of large data sets has aided the study of the topological properties of networks in general and in semantic networks in particular (Baronchelli et al., 2013; Steyvers & Tenenbaum, 2005). As a special case of generalized mathematical networks, semantic networks use semantic units (such as words) as nodes and the relationship between them (such as co-occurrence) as edges (Baden, 2010; Carley & Palmquist, 1992; Diesner, 2012). This is distinct from social network analysis, in which social agents are assigned as nodes and social relationships between them are represented as edges. However, it is important to stress that both semantic networks and social networks are represented and modeled under the same framework of graph theory, though distinguishable for the kind of entities and relationships they address. This similarity can provide researchers with the opportunity to use novel methods developed in rapidly growing body of research on social network analysis to advance comparative research on semantic networks.

Semantic network analysis is an established method and has been used in studies dating back as far as 1969 (Collins & Quillian, 1969; Steyvers & Tenenbaum, 2005). However, the method has grown rapidly in popularity as a tool for communication research in the last decade (Baden, 2010; Diesner, 2012; Doerfel & Barnett, 1999;

Zywica & Danowski, 2008), as it provides communication researchers with an efficient method to model and quantify discourses and corpora related to various media and messages. Studies have used this method in different contexts, such as the structure of the International Communication Association as reflected by paper titles (Doerfel & Barnett, 1999), changes in discourse in Islamic countries following the Arab Spring (Danowski & Park, 2014), online perceptions of privacy (Yuan, Feng, & Danowski, 2013), Dutch political parties framing of the EU constitutional referendum (Baden, 2010), and more.

From a methodological perspective, representing the interdependence between semantic entities can be constructed in a number of ways, with different methods requiring different sets of decisions to be made. One such decision, for example, is the definition and extraction of semantic units from the unstructured textual data to serve as nodes in the semantic network. Semantic units can be defined as words, topics, groups of words (n-grams), subset of the words in a specific context, or several words grouped together by methods of stemming and lemmatization, reducing groups of words to their basic common form (Baden, 2010; Carley & Palmquist, 1992; Yang & González-Bailón, 2015). Some data pre-processing procedures can eliminate tokens that contain little relevant knowledge. For example, a list of “stop words”, deemed a priori to contain little semantic information, can be used with measures such as PMI or TF-IDF (Aizawa, 2003) to determine a posteriori which words are “unique” to a specific context and thus more important; similarly, simple frequency counts can be used to choose the most prominent words for analysis. These and other methods aim to identify the set of concepts to be used as nodes in the semantic network.

Other decisions relate not to the operationalization of semantic units but to the operationalization of edges, or links, between these units—for example, co-occurrence. Different studies have operationalized co-occurrence in various ways, for instance by using different levels for co-occurrence (document, sentence, moving window, etc.) and with different methods of calculation, thresholding, normalization, and significance testing. Such decisions affect the structure of the graph objects. For example, defining the strength of a word’s connection with another word as a function of their collocation in the same text (based on the assumption that words that appear together more frequently are also more strongly connected), often results in a non-directed network. This implies that the relationship between the dyads of collocated words is symmetrical (the extent to which word A appears with word B is identical to the extent to which word B appears with word A). Such semantic representation is often used for natural language processing tasks, such as the construction of search engines (Turney, Pantel, & others, 2010), text summarization tools (Nenkova & McKeown, 2012), and more. However, it is important to note that such methods all use the “bag of words” approach and may thus be somewhat limited, or at least reductionist, relative to other methods that construct edges between semantic units in a non-symmetrical fashion. These result in directed networks (Carley & Palmquist, 1992; van Atteveldt, Kleinnijenhuis, & Ruigrok, 2008) and offer a more sophisticated measure of co-occurrence windows (Baden, 2010). The specifics of the method used in this study, such as stemming and TF-IDF measurement for token filtering, a moving window for co-occurrence measurement, the Ochiai coefficient for normalization, and utilizing processes of backbone extraction, will be discussed and explained in the next chapter.

In terms of analysis, most studies in communication use the network structure as a means to describe specific discourses. As such, these “maps” can aid researchers by reducing large corpora into a more manageable graphic representation or as material for the statistical analysis of different graph properties, from the centrality of different nodes to the density of a complete graph.

Additionally, many studies in communication take advantage of clustering techniques in network analysis to group words together and create what are often referred to as frames, issues, themes, or topics. Such methods enable researchers to observe the frame or topic structure of the analyzed content without the need for pre-existing coding schema or an a priori assignment of frames and topics (Baden, 2010; Qin, 2015b; Quinn & Powers, 2016). Such analysis can aid researchers in understanding not only the possible set of themes in a discourse, but also the relationships between these themes and their evolution over time and across media, as will be shown in this study.

This method of exploring the frames, topics, or themes in discourse over issues has not only been used to analyzed news media coverage, but also other types of corpora. For example, a study by Baden (2010) used cluster identification in semantic networks as a method for analyzing elite discourse on the Dutch EU referendum campaign. Analyzing Dutch parties’ and politicians’ direct communication, Baden (2010) identifies several frames prominent in elite discourse on the issue. Such frames relate, for example, to the referendum’s economic consequences, its cultural aspects, and its implications for environmental and human rights issues. Similarly, Kim and Kim’s study (2015) re-analyzed open-ended survey responses on embryonic stem cell research collected in 2006 by the UK Department of Health (DH). These researchers used the Girvan-Newman



grouping method to evaluate respondents framing of embryonic stem cell research. Their method was designed to identify salient frames, including the therapeutic purpose of embryonic stem cell research and concerns regarding the destruction of human embryos. A study by Quinn & Powers (2016) analyzed comments on *New York Times* articles related to online sharing, using cluster analysis to extract four separate networks, or themes, related to online sharing (including communality, surveillance, the public sphere, and information distribution).

As can be seen even from this very limited review, various network construction and cluster identification techniques have been used by researchers in diverse contexts. However, while these researchers and others conceptualize discourse by dividing it into multiple graphs or sub-graphs, only a limited amount of research has extended the analysis from a single network perspective to a multiple- or between-network perspective (Baden, 2010; Carley & Palmquist, 1992; Chewing, 2015; Danowski, 2012b; Qin, 2015b).

Even these rare comparative studies are limited to a small number of graphs and to a more basic set of methods for comparison, typically in a more qualitative form of analysis. For example, Baden's research on elite discourse on the Dutch EU referendum campaign compared four semantic networks of four Dutch parties and their statements related to the referendum. While such an approach is applicable to small-scale comparisons, it is inadequate for comparing a larger set of semantic networks. This study thus offers an approach that draws on general network analysis to develop a method better suited to comparative analysis.

Despite semantic networks being a unique case of general network theories, indicators developed for analyzing other types of networks—such as social, physical, or purely mathematical networks—can be useful in analyzing and comparing semantic networks as well. Such indicators might uncover latent features of the textual corpus that could be related to corpus features of interest to communication researchers and real-world outcomes. Most importantly, methods and indicators from outside communication research can help facilitate a more comparative approach, focusing not on a single map, or discourse, but by juxtaposing and assessing the structural features of multiple corpora. In this dissertation, I focus on network community structure as a measure of thematic diversity.

### **3.3.4 Diversity in semantic networks**

While past research has tended to focus on more general features, such as graph centrality and density, in the present study, semantic network analysis becomes applicable to measuring thematic diversity through the identification of clusters and community structure. In network analysis, techniques for cluster identification and community detection seek to identify areas of heightened density in a network—or in other words, grouping together nodes that share a stronger connection with each other distinct from what they share with the rest of the nodes in the network (Rubinov & Sporns, 2010). Studies in communication using semantic network analysis often apply clustering techniques drawn from general network analysis to group words. The primary assumption of this method is that words that co-appear frequently are thematically related. Building on this logic, we can develop analysis to focus on the community structure indicators.

From a broader perspective, the translation of texts into a graph structure enables researchers to draw on methodological developments in the dynamic field of network analysis, which offers several novel approaches to measure structural semantic features. For example, developments in the analysis of multi-layered networks might enable researchers to analyze multiple semantic networks, whether drawn from different sources or different time-frames, concurrently, thus offering insights into changes in the thematic landscape of a given discourse over time or between sources. However, taking advantage of such methodological developments requires the conceptual “translation” of graph structural features into discourse features. An example of such conceptual “translation” from community structure to thematic diversity of discourse is detailed in this section.

Communities (or modules) are sub-graphs or sets of nodes that exhibit a high density of internal links when compared with links to nodes in other communities. A variety of algorithms are used to find the community structure that maximizes the within-group links, while minimizing between-group links (Lancichinetti & Fortunato, 2009). Group distinctiveness is a ratio of internal and external links, and is measured by a modularity score (Rubinov & Sporns, 2010). Thus, modularity measures the extent to which communities in the network are interconnected within, yet separated from, other groups, with higher values indicating a better division. It is reasonable to use the modularity score as an independent variable, alongside the number of communities, to represent both the volume and distinctiveness of sub-graphs in a given semantic network—thereby addressing both the nominal number of word groups and the thematic distinctiveness of each group. Based on the notion of word communities as topics or

frames (Baden, 2010), multi-thematic coverage is expected to exhibit a more communities and higher modularity.

However, this estimation might be lacking in terms of simplicity, as well as in terms of validity. To understand these issues, we must return to the prism of variety, balance, and disparity, as offered by Stirling (2007): while the modularity estimation accounts for variety, its estimation of disparity (the distance between categories) is done on a global level, for the whole graph, rather than at the community level (that is, between each dyad of communities). In addition, while this measure accounts for variety, and to some extent for disparity, it does not account for the balance of these communities in terms of community size or importance.

Thus, I offer an estimation of network thematic diversity that combines information gathered using community detection, but incorporating this framework into the logic of diversity estimation offered by Stirling (2007) as well as with the clustering techniques in semantic networks as a means of identifying themes, as detailed in previous studies.

As explained earlier, Rao's diversity coefficient is calculated as (Rao, 1982):

$$D = \sum_{ij(i \neq j)} (\delta_{ij}) \cdot (p_i \cdot p_j)$$

In accordance with this formula, to estimate the diversity of a group of texts, several specific inputs are needed. These are the distance between each two categories or themes ( $d_{ij}$ ), as well as the proportion of these themes in our corpus.

For each candidate, an independent semantic network is created from the co-occurrence matrix, drawn from the corpus relevant to the candidate (either their news

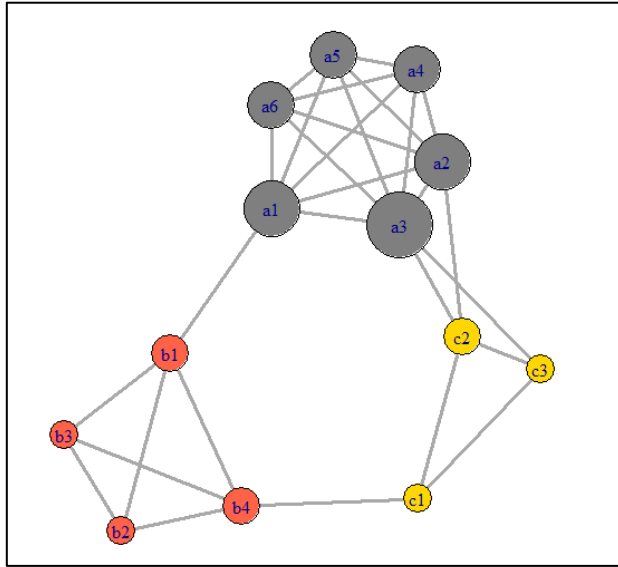
coverage or their social media activity). This contrasts with the topic modeling method, where a model was drawn for all candidates at once. The procedures are applied uniformly, over all corpora, to facilitate comparisons between the different networks. I elaborate more on the specifics of the method and the comparative procedures in the following chapter.

For the sake of simplicity, I demonstrate the method using a minimal example of a network drawn from the adjacency matrix presented in Table 5. This adjacency matrix describes the relationship between 13 words (A1-C3). The matrix is symmetrical over the diagonal (resulting in an undirected network). Relationships between words are represented by 1 or 0, with 1 indicating the existence of a relationship between two words, and 0 representing the lack of such relationship, resulting in a non-weighted network (unlike the actual data, which uses weighted edges as well as up to 1000 nodes in each network).

Table 5: A hypothetical adjacency matrix (13x13)

	<i>a1</i>	<i>a2</i>	<i>a3</i>	<i>a4</i>	<i>a5</i>	<i>a6</i>	<i>b1</i>	<i>b2</i>	<i>b3</i>	<i>b4</i>	<i>c1</i>	<i>c2</i>	<i>c3</i>
<i>a1</i>	1	1	1	1	1	1	1	0	0	0	0	0	0
<i>a2</i>	1	1	1	1	1	1	0	0	0	0	0	1	0
<i>a3</i>	1	1	1	1	1	1	0	0	0	0	0	1	1
<i>a4</i>	1	1	1	1	1	1	0	0	0	0	0	0	0
<i>a5</i>	1	1	1	1	1	1	0	0	0	0	0	0	0
<i>a6</i>	1	1	1	1	1	1	0	0	0	0	0	0	0
<i>b1</i>	1	0	0	0	0	0	1	1	1	1	0	0	0
<i>b2</i>	0	0	0	0	0	0	1	1	1	1	0	0	0
<i>b3</i>	0	0	0	0	0	0	1	1	1	1	0	0	0
<i>b4</i>	0	0	0	0	0	0	1	1	1	1	1	0	0
<i>c1</i>	0	0	0	0	0	0	0	0	0	1	1	1	1
<i>c2</i>	0	1	1	0	0	0	0	0	0	0	1	1	1
<i>c3</i>	0	0	1	0	0	0	0	0	0	0	1	1	1

Figure 5 presents a visualization for this network after undergoing a process of community detection. In this figure, labels indicate node name, size indicates the degree of the node (the sum of all edges of that node), and color indicates community membership for each node, with three different communities detected.



*Figure 5: A network representation of the hypothetical simplified semantic network.*

In order to estimate graph-level diversity, the distance between each two categories or themes ( $\delta$ ) needs to be estimated first. The relationship between the groups of nodes can be estimated using the edges connecting these communities to each other. For example, community B (including nodes b1, b2, b3 and b4) is loosely connected to the other communities with only one edge to each. However, community C and community A are more strongly related, as these have three edges between them. The importance of these between-community edges also depends on the size of the communities themselves. For example, larger communities have a higher probability of having between-community edges. From a different perspective, having three edges between two communities of size ( $n=3$ ) is much more meaningful in terms of community relationship

than having three edges connecting two communities of size ( $n=300$ ). Therefore, the number of edges between each two communities is normalized by the observed number of edges within these communities (in a similar logic to that of the modularity estimation; Newman, 2006). Lastly, as ( $\delta$ ) is a measure of distance, the normalized share of between edges is inversed.

More formally, the measure of distance is defined as:  $\delta_{ij} = 1 - \frac{\sum E_{ij}}{\sum E_i + \sum E_j + \sum E_{ij}}$ ,

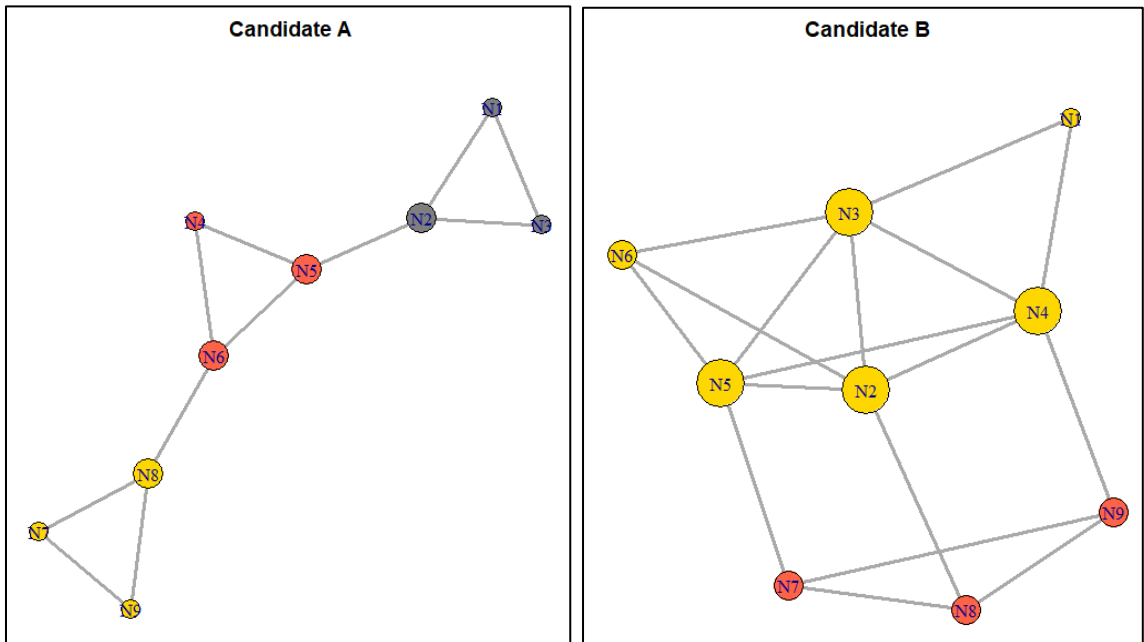
with  $\sum E_{ij}$  defined as the sum of edges between each two communities (i) and (j),  $\sum E_i$  defined as the sum of edges within community (i), and  $\sum E_j$  defined as the sum of edges within community (j).

The second indicator needed for the estimation of diversity is the size of each community in the general network structure ( $p_i$ ). As edges in the network represent the co-occurrence of words in the network, the sum of edges in a single community represents the combined co-occurrence of nodes belonging to that community in the corpus. Thus, the size of each community can be defined as:  $p_i = \sum E_i$ , or the sum of all edges for all nodes in a given community (this measure is also roughly equal to the cumulative weighted degree of all nodes in given network). Finally, similar to the measurement of the between-edges, and in order for the sum of topic proportions to be 1, the total within edges for each community is normalized by dividing it by the sum of all total edges for all communities in the network:  $p_i = \frac{\sum E_i}{\sum_1^j \sum E_j}$

In order to calculate the diversity for each network, the formula offered by Stirling (2007) is iterated over all possible community dyads. As an example, consider these two highly simplified and hypothetical examples of semantic networks. Although these

networks differ from the semantic networks used in this study by size and weight (being much smaller and unweighted networks), they can be useful in exemplifying the logic behind the proposed diversity estimation.

The two networks in Figure 6 offer the same number of nodes ( $n=9$ ). However, the network for candidate A contains three distinct and equally sized communities, with each connected by a single edge to other communities. The network for candidate B, however, contains fewer communities, which are skewed in size and highly connected to each other.



*Figure 6: A network representation of hypothetical and simplified semantic networks for two candidates (Community 1 in yellow, community 2 in red, community 3 in gray).*

Thus, from the perspectives of variety, balance, and disparity, the estimation for the thematic diversity for candidate B should be lower than that of candidate A.

Looking first at candidate A, the values for ( $\delta$ ) will be equal to:



$$\delta_{12} = 1 - \frac{\sum E_{ij}}{\sum E_i + \sum E_j + \sum E_{ij}} = 1 - \frac{1}{3+3+1} = 0.86$$

$$\delta_{13} = 1 - \frac{\sum E_{ij}}{\sum E_i + \sum E_j + \sum E_{ij}} = 1 - \frac{0}{3+3+1} = 1$$

$$\delta_{23} = 1 - \frac{\sum E_{ij}}{\sum E_i + \sum E_j + \sum E_{ij}} = 1 - \frac{1}{3+3+1} = 0.86$$

Incorporating both ( $p_i$ ) and ( $\delta$ ) for all communities, the diversity measurement will sum to:

$$D_A = \sum_{ij(i \neq j)} \delta_{ij} \cdot p_i \cdot p_j = 2 \cdot \left( 0.86 \cdot \frac{4}{13} \cdot \frac{5}{13} + 1 \cdot \frac{4}{13} \cdot \frac{4}{13} + 0.86 \cdot \frac{5}{13} \cdot \frac{4}{13} \right) = 0.59$$

Looking at candidate B, the values for ( $\delta$ ) will be equal to:

$$\delta_{12} = 1 - \frac{\sum E_{ij}}{\sum E_i + \sum E_j + \sum E_{ij}} = 1 - \frac{3}{3+11+3} = 0.82$$

Incorporating both ( $p_i$ ) and ( $\delta$ ) for all communities, the diversity measurement will sum to:

$$D_B = \sum_{ij(i \neq j)} \delta_{ij} \cdot p_i \cdot p_j = 2 \cdot \left( 0.82 \cdot \frac{6}{20} \cdot \frac{14}{20} \right) = 0.34$$

As can be seen, due to candidate A having higher variety (number of themes), more balance (themes equal in size), and higher disparity (less between-theme connections), the diversity estimation for Candidate A is higher than the diversity estimation for candidate B.

Unlike topic modeling, diversity estimation using semantic network analysis accounts for disparity, balance, and diversity, as the number of communities is not set a priori. In addition, because the topic model needs to be estimated over the full corpus,

adding new documents after estimating topic models is problematic, as it assumes that the latent topic structure for the new documents is identical to earlier topics. With the semantic network analysis approach, however, new documents can be added as long as the procedure and parameters are kept identical.

Semantic network analysis also suffers from some limitations. For example, semantic network analysis is much less developed and validated than topic modeling, which remains the most prominent unsupervised machine learning method in the social sciences. In addition, the single community membership assumption of the method used in this study is problematic: as it assumes each word can only belong to one community. In reality, words can share different themes with differing levels of relatedness to each theme. I elaborate on these subjects in the conclusion chapter and offer future steps that might be helpful in improving both methods, specifically by applying features of one to the other or by using both in tandem. The following chapter discusses the specifics of the data gathering, topic modeling, semantic network construction, and statistical modeling used in this study.

## 4. METHOD

In this chapter, I detail the various methods and measures used in the present study. I begin by describing the sample of U.S. Senate candidates included for analysis and the non-media data used in this study, either as a dependent variable or as control variables, and present summary statistics for these variables. I then turn to describe the methods that were used for gathering the media data, including the different processes used for scraping the candidates' social media activity and their news coverage.

I then turn to discuss the analysis of the media data. First, I address the measurement of volume and tone in candidates' social media activity and news coverage. I then turn to discuss the details of the two unsupervised machine learning methods used to estimate diversity. Specifically, I outline the processes for topic model estimation for both the Twitter and the news datasets, including model fit statistics and the process for making decisions on the number of topics to include in the models. I then discuss details of the semantic network analysis of both the Twitter and the news datasets. These processes were somewhat different as a result of the size of the data and the assumptions that I made regarding the thematic structure of the documents. I detail the processes used to estimate diversity for both methods, including the use of diversity in randomly generated networks as benchmarks for observed estimations.

Finally, I address the statistical approach used to explore the relationship between candidates' electoral success and the volume, tone, and thematic diversity of their social media activity and news coverage.

## 4.1 Non-Media Data

Data for analyzing the relationship between news coverage and electoral success was gathered for U.S. Senate candidates between the years 2008 and 2016 (n=330). Data for social media activity was gathered for U.S. Senate candidates between the years 2012 and 2016 (n=142). This smaller window is due to the larger availability of Twitter activity during and after the 2012 elections.

While I began with a larger number of candidates, some candidates needed to be removed from the sample for several reasons. The main issue was that the statistical models treated each race as a single observation, measuring the success of the Republican candidate as an outcome of the various features of the race, the state, the candidate, and their opponent (I elaborate more on this in Section 4.5). Elections in which only one candidate competed were removed from the sample. For example, these included the 2014 Alabama Senate elections, in which incumbent Republican candidate Jeff Sessions ran unopposed. Another issue was races that featured more than two major candidates; these are defined as races in which the conservative and liberal candidate together gained less than 80% of the total votes. Such instances were removed as well. For example, this included the 2010 U.S. Senate elections in Florida, in which the two frontrunners were a Republican candidate (Marco Rubio) and a former Republican-turned-independent (Charlie Christ), followed closely by a Democratic candidate (Kendrick Meek). In addition, there were some rare cases in which both candidates were from the same party, as was the case for the 2016 California Senate elections, where both main candidates were from the Democratic party. A handful of races were removed due to a lack of substantial news coverage of the candidate or their opponent. For the purposes of the

analysis, there needed to have been at least ten articles for both candidates from which to draw the networks. Similarly, candidates that lacked substantial Twitter activity, which was defined as having at least 50 tweets on their feed during the timeframe for the study, were also removed from the sample (hence the larger  $N$  for the news data compared with the Twitter data). The remaining 330 candidates across 165 races for the news study, and 142 candidates across 71 races for the social media study, were used in the following analysis.

For the dependent variable, the percentage of votes gained by each candidate was gathered using data files available from the FEC website.<sup>2</sup> Thus, the models were designed to help explain not only whether a candidate won or lost the election, but the actual share of votes gained by that candidate. As expected, with most candidates and their opponents gathering together a little less 100% of the votes, the average percentage of votes gathered by each candidate between 2008 and 2016 was 48.2 ( $SD=12.9$ ). The mean and standard deviation for the races between 2012 and 2016 was nearly identical ( $M=48.1$ ,  $SD=12.4$ ).

Based on prior research on the factors that influence electoral success, several additional variables were included as controls in the regression models. These variables were used to examine the extent to which the media content predictors, such as volume, tone, and thematic diversity, were in fact independent predictors of electoral success, as well as predictors in the models in which thematic diversity was used as a dependent variable. First, the election cycle was marked as either midterm or quadrennial, in accordance with evidence for a disadvantage in midterm elections for candidates from the

---

<sup>2</sup> <http://www.fec.gov/disclosure.shtml>

sitting president's party (Grofman et al., 1998). In the sample, 134 candidates competed in elections defined as midterm, and 196 competed in races that took place during non-midterm elections (this is a higher number, as three out of the five election cycles used for this study were non-midterm elections).

In addition, the ideological leaning of a state has a strong effect on candidates' success and the competitiveness of the race. A control variable for state ideology based on Hummel and Rothschild's (2014) research was thus also included. State conservativeness was measured as the mean American Conservative Union Foundation's (ACU) ratings<sup>3</sup> given to both senators of a state in the year prior to the elections. The mean ACU rating for candidates running in years 2008-2016 was 43.2, with a standard deviation of 34.1 (or  $M=40.7$  and  $SD=32.8$ , for candidates running in 2012-2016).

Next, due to the impact of campaign contributions on electoral success (Jamieson, 1996; Magee, 2012), data regarding the contributions gathered by each candidate were collected from the FEC website for all candidates.<sup>4</sup> The mean value for campaign contributions was \$6.45 million, with a standard deviation of \$5.8 million. The values for candidates running in 2012-2016 was somewhat higher ( $M=7.56m$ ,  $SD=6.4m$ ), which is likely explained by inflation and the rising cost of political campaigns.

Lastly, with existing evidence for the impact of incumbency as well as general experience in public office on candidates' success (Hummel & Rothschild, 2014), candidate experience was also entered as a factor in the models. Details regarding the experience of candidates were gathered from the candidates' Wikipedia page, Ballotpedia

---

<sup>3</sup> <http://acuratings.conservative.org/acu-federal-legislative-ratings/>

<sup>4</sup> <http://www.fec.gov/disclosure.shtml>

page, as well as their campaign website if one was available. Experience in elected office included: experience as a senator (usually the incumbent, though in some races former senators who were not incumbents competed in the race), experience as a U.S. congressman, experience as a governor, and other experience (at the local or state level). These were entered separately given that some forms of experience, such as having already served as a senator, are expected to be more valuable than other types of experience, for example, holding an office in local government.

## **4.2 Media Data Gathering**

The various features of candidates' media content, including both news coverage and social media activity, are at the core of this study. Due to the size of modern day media data needed for such an analysis, the process for gathering the data was complex, requiring a combination of manual and computational tools. Moreover, the strategy used to gather news media and social media data required different methods and posed unique challenges. The following sections detail the processes used to gather data for both media channels.

In order to analyze the news coverage for the candidates in the 2008-2016 U.S. Senate elections, all coverage of these candidates was downloaded using LexisNexis. For each candidate, a search was performed in the LexisNexis database using the candidate's full name, with a time-frame of six months prior to the elections. A python script was then used to parse the results into different articles, collecting article-level meta-data as well.

Search was performed in the LexisNexis database for all U.S. Newspapers and Wire Services (Lexis-Nexis code: 140954). This strategy was chosen specifically, as opposed to the more common strategy of searching only major newspapers, which is unlikely to identify local coverage, and for Senate elections local coverage might be of critical importance. However, it is also important to note that the database may have some limitations that need to be considered. LexisNexis is by no means identical to print editions of newspapers and as such may provide a somewhat biased sample of news coverage (Ridout et al., 2012). The database does not contain all outlets publishing in the U.S. and might thus be missing both large and small news outlets. However, this problem is more common for wire services data and for international news. Therefore, in the context of this study, for which I rely on local news, such problems should be less acute. Further, while some local sources might be missing from the sample, given that candidates are compared to their counterparts, any biases in the LexisNexis database are likely to influence both candidates in similar ways, thus limiting the overall bias in estimating media features.

Additional limitations are the inclusion of duplicate items (for example, a wire service article that was printed verbatim by another news outlet), and server test items in the database. Test items were identified by excluding extremely short items from the database. Duplicate articles were identified and removed using a random 200-character string taken from the middle of the article. If that exact 200-character string was found in another already archived article, then the article was deemed to be duplicate and was not archived again. I elaborate more on these issues when addressing the pre-processing procedures.



Data for social media activity was gathered for a more limited time-frame due to the limited availability of the data (2012-2016). Several off-the-shelf tools and packages are available for mining social media data in general and Twitter data in particular. These tools allow researchers and developers to search the Twitter database (Rest API) as well as observe the ongoing stream of tweets that are constantly uploaded to the website (Streaming API). However, both services come with some rate and size limitations, and research has shown that they might deliver non-representative samples for the requested content (Tromble, Storz, & Stockmann, 2017). Thus, a non-API approach was chosen for the data retrieval in this study.

In order to gather content for all candidates, the Twitter username (handle) of all candidates needed to be obtained. This was done using a combination of methods offered in previous studies (Bode et al., 2016; Bright et al., 2018; Jungherr, 2016). First, official Twitter pages were gathered from Wikipedia, Ballotpedia, and the candidates' websites. A Google search was also performed using the candidate's name, state, and the keywords "Twitter," "campaign" and the year of the race. The first two pages of Google search results were manually examined to identify additional viable Twitter pages to assess whether they related to the candidate, or whether they related to another individual with the same name, a parody account, a hijacked account, or other non-genuine campaign pages. For this task, timing was found to be critical. Several candidates' pages were removed from Twitter or hijacked by a third party by the time the search was conducted, as can be viewed from the content of the page feed. For example, the Twitter handle "@SadlerTX" was previously attached to candidate Paul Sadler but has since been

claimed by a Russian speaking individual. Therefore, I decided to focus on more recent elections (2012-2016) for which more pages still existed online.

Following the identification of candidate-related Twitter usernames, all activity in these pages (for all tweets written by the user) was downloaded and parsed into separate tweets using a custom-built python script. These included the textual data of the tweets along with any tweet-level metadata supplied by the page. I chose to use this more direct approach as opposed to other search methods as these can limit the amount of data gathered from Twitter pages or even skew data search results due to unknown criteria for inclusion (Tromble, Storz, & Stockmann, 2017). This is especially the case for candidates with a large volume of Twitter activity during the elections. Finally, as was done for the news data, duplicate Tweets were identified. These tweets were not removed from the data at all stages and for all methods, as will be elaborated in the topic modeling section in this chapter.

### **4.3 Media Measures: Volume and Sentiment**

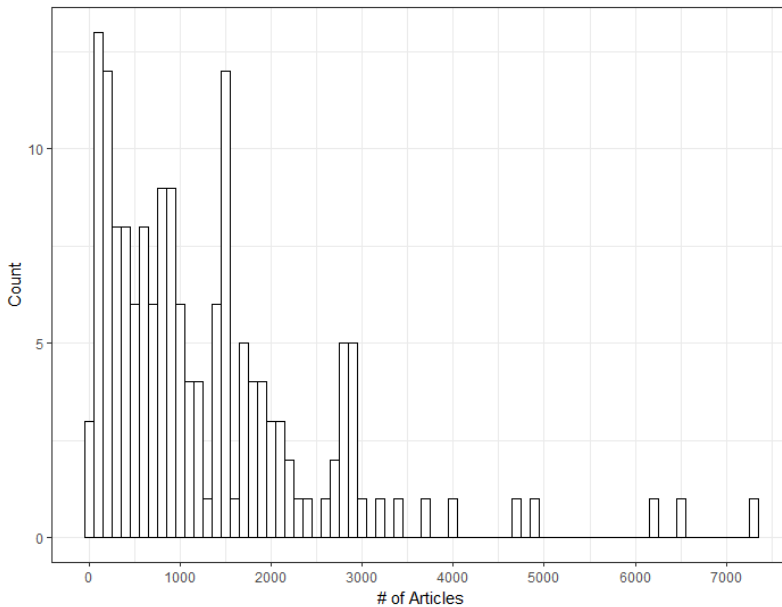
The volume of news coverage was measured as the *number of articles* mentioning the candidate ( $M=1326$ ,  $SD=1410$ ) in the six months prior to the elections. The smallest number of articles per candidate in the sample was 11 (for candidate Rob Tingle, who ran in the 2008 Senate elections in Rhode Island). Joe Biden captured the largest number of articles, 17,659, in his 2008 Senate election campaign in Delaware. This is likely due to Biden competing for both a Senate seat and the vice-presidency at the same time (as is allowed by Delaware's constitution).<sup>5</sup> The second highest number of articles, 9,684, was

---

<sup>5</sup> In addition to the analysis presented in the results section, additional models were constructed were this observation was dropped to make sure the results stand even when discarding this more unique case.

gathered by John Kerry in his 2008 Massachusetts Senate campaign (likely due to his status as a former presidential nominee). The mean number of articles for Republican and Democratic candidates was found to be surprisingly similar (1323 vs. 1285 respectively).

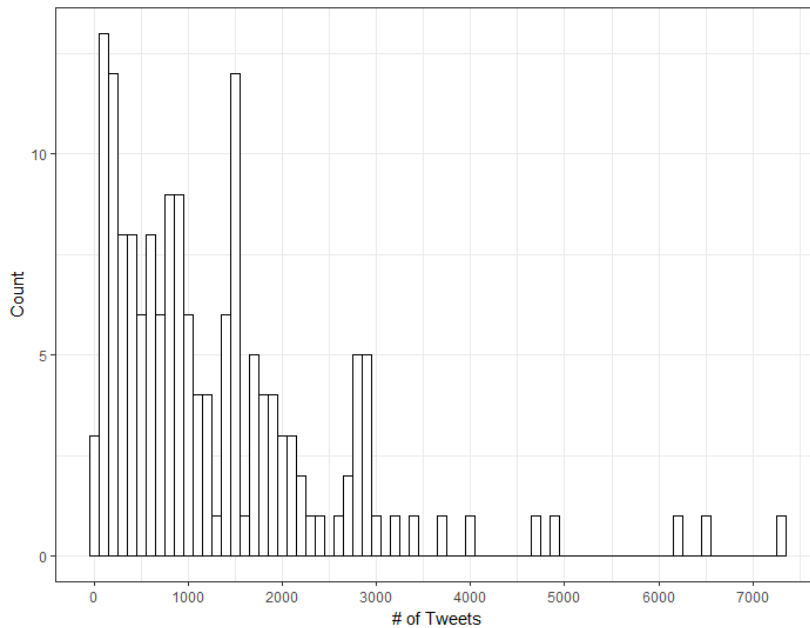
Figure 7 presents the distribution of the number of articles covering the candidates:



*Figure 7: The distribution of volume of news coverage (as number of articles) received by each candidate (2008-2016).*

Volume of social media activity was measured as the number of tweets written by candidates on their Twitter pages during the six months prior to an election. The mean number of tweets per candidate was 1,294, with a standard deviation of 1,228 tweets. The most prolific tweeter was Marco Rubio (FL, 2016) who posted 7,333 tweets during his Senate campaign in 2016, though the candidate was also running in the Republican presidential primaries at the same time, which might explain his intensive activity. He is followed by Scott Brown (MA, 2012), who tweeted 6,470 times, and then Mitch McConnell (KY, 2014), who tweeted 6,243 times. In accordance with Republican candidates populating the top of the prominent tweeters list, the average number of tweets

per Republican candidate was 1,422, while the average number of tweets per Democratic candidate was slightly lower at 1,167, although this difference was not found to be significant. Figure 8 presents the distribution of the number of tweets broadcasted by the various candidates:



*Figure 8: The distribution of volume of social media activity (measured as the number of tweets shared) for all candidates 2012-2016.*

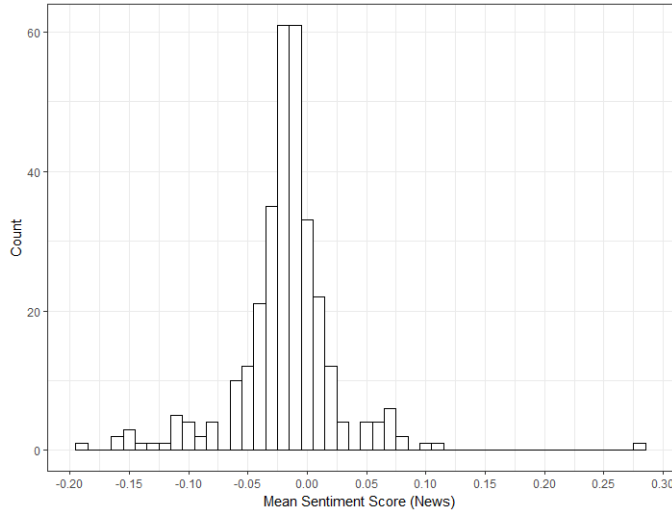
Automated sentiment analysis for each candidate’s corpus was performed using a custom python script that incorporated the Sentistrength sentiment dictionary (Guo & Vargo, 2015; Thelwall et al., 2010). The Sentistrength sentiment dictionary provides a score for a set of words that assess negativity or positivity on a scale of -4 (highly negative) to +4 (highly positive). To calculate the sentiment score for each candidate, each document mentioning his or her was divided into independent sentences (splitting the text on characters such as .,” “!” and “?”). A search was then performed to determine whether each individual sentence included the name of a candidate. As suggested by

previous studies, only sentences containing a candidate name were subsequently used to measure the sentiment in a news article, (Bélanger & Soroka, 2012). For each token in the sentence, an attempt was made to match it to the Sentistrength dictionary, identifying both positive and negative sentiment words. The number of occurrences for each of these words was logged and multiplied by the sentiment score of that word (-4 to +4). These counts were tallied at the document level and divided by the total number of tokens to prevent any bias stemming from document length.

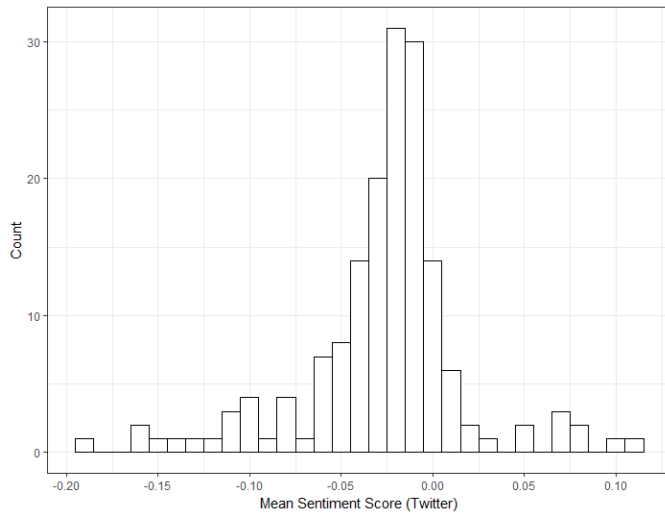
Finally, a sentiment score was calculated at the candidate-level by averaging the sentiment score of all documents mentioning this candidate (again to prevent the interaction between volume of coverage and sentiment). This process resulted in a single sentiment score for each candidate that could be used for subsequent regression models. The mean sentiment per candidate for news coverage was slightly negative ( $M=-0.02$ ,  $SD=0.04$ ). A small but significant difference ( $p=0.014$ ) was found between the sentiment of news coverage of Republican candidates ( $M=-0.024$ ) relative to coverage of Democratic candidates ( $M=-0.012$ ). The distribution of sentiment per candidate can be seen in Figure 9.

The process used to score the sentiment in candidate tweets was nearly identical, although this did not require a candidate's name to appear in each sentence due to the small size of the documents and the apparent relevancy of the text to the candidates. As with news coverage, sentiment was calculated on a per-word basis and averaged over all tweets written by a candidate. Figure 10 presents the distribution of Twitter sentiment over all candidates. Similar to news coverage, mean sentiment for candidates' tweets was

slightly negative ( $M=-0.027$ ,  $SD=0.044$ ). No significant difference in sentiment was found between Republican and Democratic candidates.



*Figure 9: The distribution of mean sentiment in news coverage per candidate.*



*Figure 10: The distribution of mean sentiment in Twitter activity per candidate.*

Table 6 presents the most negative and positive Tweets identified in the corpus:

Table 6: The top 5 negative and positive tweets identified in the corpus (non-text data removed).

Top 5 Negative Tweets	Top 5 Positive Tweets
[1] "awful to hear of terror attacks at the airport in istanbul prayers to all who were hurt and killed by this appalling violence"	[1] "thanks to all brave ct soldiers serving in the nationalguard special thanks to the friends & families as well for their love & support"
[2] "another devastating terror attack today we need to confront this global war on terror head on with a comprehensive strategy to defeat it"	[2] "thank u loved the cupcakes mtbakedinacupcg thanks for coming in & supporting a local family owned business hope u liked the cupcakes"
[3] "from the threat of domestic violence sexual assault child abuse and violent crimes (3/9)"	[3] " jebdakhaptyn great meeting you as well thanks for the kind greeting hope you had a great birthday"
[4] "josh mandel's opposition to the auto rescue is wrong for ohio wrong for the middle class and wrong for veterans"	[4] "aquarius0211 i wish you grace peace and love i hope your father continues to recover well that u continue to excel toward your dreams"
[5] "(2/2) this is yet another senseless & violent attack in wisconsin that is tragic and heartbreaking we all mourn this horrible loss of life"	[5] "we hope you get to enjoy friends family good food and ca's great outdoors today -- have a safe and happy independence day everyone"

## 4.4 Measuring Diversity

### 4.4.1 Topic Modeling

The first method used to estimate the level of thematic diversity across the different corpora was topic modeling. The topic modeling processes for the Twitter and news media datasets were slightly different. First, in order to analyze the Twitter data, all documents for all candidates were loaded together (N=124,557 tweets). Following this, the tweets were pre-processed with an initial cleaning of hyperlinks, referral data (VIA/RT), and visual data such as photos and videos.

At the second step, a small number of extremely short tweets (having less than 10 characters, which is the equivalent of less than two words) were removed from the data (N=650). In terms of duplicates, some candidates were found to use the same tweet several times, while in other cases, different candidates used the same tweet. Though the number of duplicate tweets was not large relative to the size of the data (N=3,912), such duplication might nonetheless skew results by artificially inflating the co-occurrence of

specific terms in the corpora. Therefore, duplicated tweets were removed from the data at the modeling stage (to be restored later at the diversity measurement stage). At the end of this process, 119,995 tweets were prepared for analysis using the Structural Topic Modeling package for R (STM; Roberts, Stewart, Tingley, et al., 2014). All numbers and other non-alphabetical characters were also removed from the texts. The texts were then separated into tokens, and all stop-words were removed from these token-lists. Finally, all terms were stemmed using the text mining package for R (TM; Feinerer, 2017).

While the size of the data was not extremely large relative to the news data, size still posed some pragmatic limitations for an efficient analysis given the hardware available for this study. Thus, as is common in this type of analysis, sparse tokens were removed from the matrix (325,096 out of a total of 1,264,456 in the full data). In extreme cases, this process may result in the removal of all tokens in a given document (especially in short documents, such as tweets); as such, the number of tokens used for the topic modeling procedure was reduced slightly from  $n=119,995$  to  $n=119,567$ .

The next step required estimating the adequate number of topics to use in the model. First, several models ranging from  $k=10$  to  $k=100$  were estimated. Figure 11 shows the various model-fit indices for the resulting topic-models:



### Diagnostic Values by Number of Topics

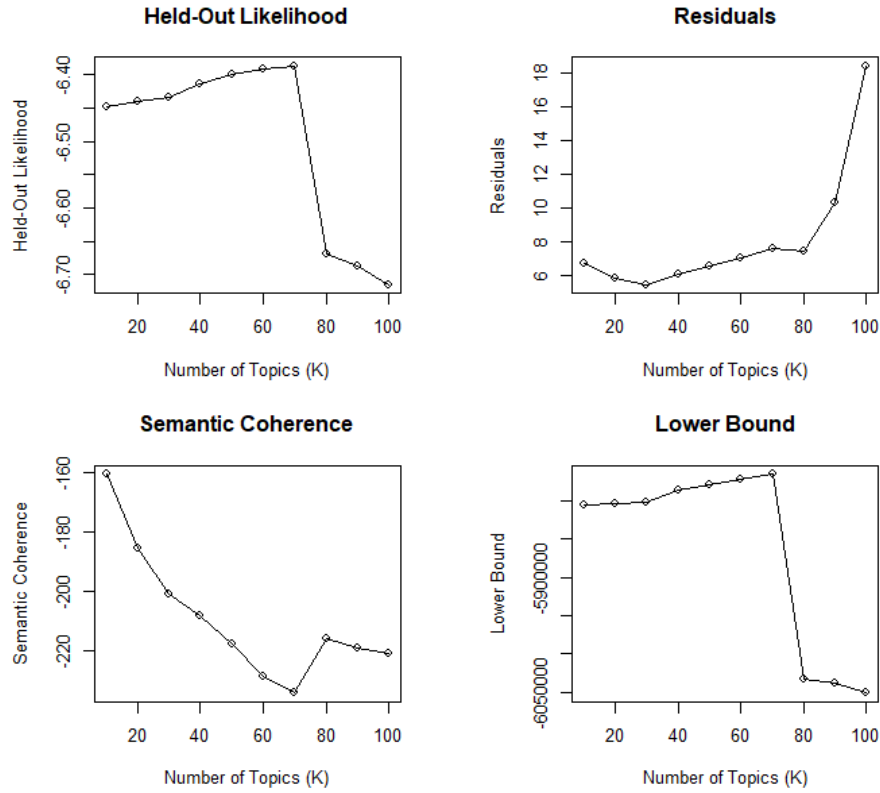


Figure 11: Model fit indicators for models ranging from  $k=10$  to  $k=100$  (by 10).

As can be seen from the held-out likelihood and lower bound plots, there was a sharp reduction in model effectivity beyond the  $k=70$  level. Semantic coherence also reduced gradually, as might be expected for larger  $k$ -levels as this can result in different sets of topics sharing similar prominent words (as discussed in section 3.3.1). To further understand the differences in topic structure, Figure 12 presents the contrasting considerations between semantic exclusivity and semantic coherence—per each level of  $k$ .

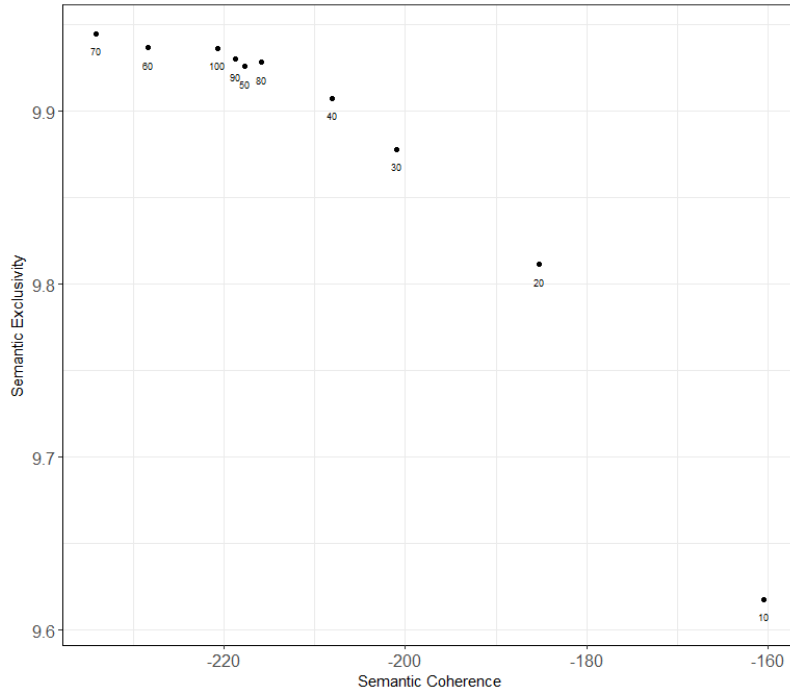


Figure 12: Semantic coherence and semantic exclusivity per  $k$  level for  $k=10$  to  $k=100$  (by 10).

As mentioned earlier, the choice of  $k$  is complicated and can be influenced by several conflicting considerations. While it is clear from Figure 11 that  $k > 70$  leads to lower quality results, the range of  $k$  between 10 and 70 offers different valid options. The smaller the chosen  $k$  is, the higher the chance to get “high-level” or general topics, which might resemble the earlier mentioned “chimera topics.” This can be seen by the lower exclusivity for high  $k$  values, meaning that there is similarity between the probability of various tokens appearing in different topics. However, these topics are also more coherent, meaning that words in a topic do tend to appear together more frequently. In other words, the higher the level of  $k$ , the more independent each topic becomes relative to all others (up to  $k=70$ ), but they also become less stable in term of coherence.

While decisions about which topic structure to choose can be difficult, in this case the choice was made somewhat easier by the fact that the diversity estimation is the end-purpose of these models. The diversity estimation incorporates topic similarity ( $\delta$ ) into the measurement and can therefore account for the problems stemming from including too many similar topics. I thus opted to err on the side of caution by using a higher number of  $k$ , rather than having too few topics. The problems that can occur due to too large a number of topics— several topics that are too similar to one another—can be corrected by the estimation. In contrast, the problems that might result from too few topics— the amalgamation of two topics into one artificial topic—cannot be corrected by the diversity estimation and therefore should be avoided.

I therefore chose to further focus on a range of  $k$  around 70 topics, which allowed me to more closely examine the model fit indices of more fine-grained choices. Similar to earlier discussions, Figures 13 and 14 present the model fit indices for topics ranging from  $k=50$  to  $k=70$  (by 1). Based on the considerations outlined above, I chose to focus on  $k=69$  as the optimal model, in terms of both held-out likelihood as well as the trade-off between coherence and exclusivity. Examples for the topics identified by this model can be seen in Section 5.2 of the following chapter. The final model of 69 topics was used to calculate diversity in each candidate's corpus, using the method elaborated in Section 3.3.2 on diversity in topic models.

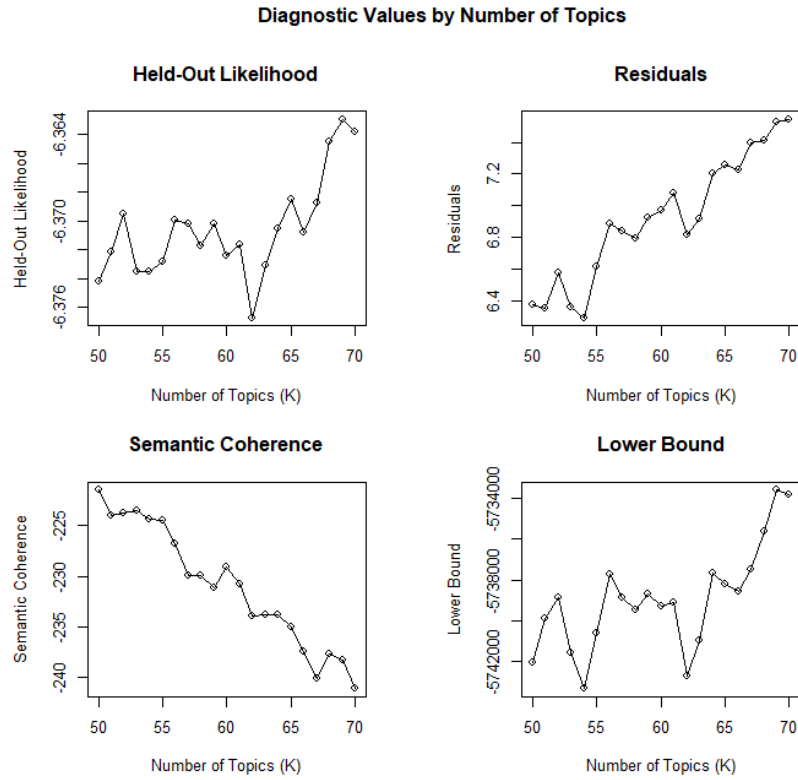


Figure 13: Model fit indicators for models ranging from  $k=50$  to  $k=70$  (by 1).

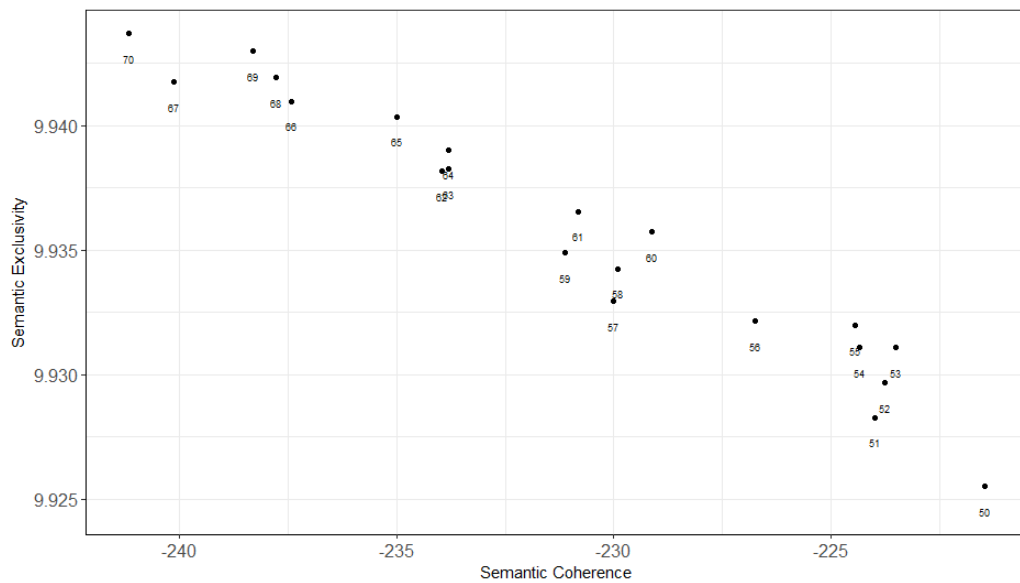


Figure 14: Semantic coherence and semantic exclusivity per  $k$  level for  $k=50$  to  $k=70$  (by 1).

First, duplicated tweets were re-added to the sample (to avoid skewing any of the individual candidates' results), using the topic loading estimation from the existing pool of tweets.  $(\delta_{ij})$  was then measured over the whole topic structure for all candidates. In other words, topic distance was shared between all candidates and defined as the dissimilarity in vocabulary between every two-topic dyad (for all of the 2,346 possible dyads). More precisely, the exponentiated beta matrix (word-topic) was extracted from the topic model, and the inverse of cosine similarity was calculated for each topic dyad and stored in a matrix size  $(k*k)$ .

The topic proportion  $(p_i)$ , by contrast, was calculated for each candidate separately using the Theta matrix (topic-document), limited only to the specific candidate's tweets. For example, for candidate 1 and topic 1,  $(p_1)$  was equal to the average loading of the first topic over all tweets written by this candidate. This was calculated for each of the topics and for each of the candidates. Lastly, the values of  $(p_i)$  and  $(\delta_{ij})$  for each topic were entered to the earlier mentioned formula  $(D = \sum_{ij(i \neq j)} (\delta_{ij}) \cdot (p_i \cdot p_j))$  to calculate the thematic diversity for the candidate.

Performing topic modeling over the news corpus proved to be more of a challenge. This owed to the massive amount of data included in this corpus. Therefore, the process was carried in a similar manner to that for the Twitter corpus, but with several alterations. First, and similar to the Twitter corpus analysis, all documents for all candidates were loaded together ( $n=425,201$  articles). Whereas hyperlinks, referral data (VIA/RT), and visual data posed challenges for analyzing Twitter corpus, the news corpus had a different problem stemming from the size of the data and the use of non-standard

characters (for example, names that were spelled using non-English characters). Therefore, the data was first converted to an encoding containing only alpha-numeric characters in the English language. At the second step, all extremely short articles (having less than 100 characters, which is the equivalent of less than 20 words) were removed from the data for analysis (N=76,749). These documents included server test-broadcasts, extremely short updates, and database errors.

Second, it was observed that some articles appeared multiple times in the corpus. This is understandable, as articles mentioning a candidate might also mention their opponent and therefore be included in the corpora for both. Additionally, reliance on several media outlets on the same wire service can increase the similarity of coverage. Such duplication in the data can skew the results by artificially inflating the co-occurrence of specific terms in these articles. Therefore, multiple appearances of the same item in the Lexis-Nexis database had to be discarded. Duplicate articles were identified and removed using a random 200-character string taken from the middle of the article. If that exact 200-character string was found in another already-archived article, then the article was deemed to be a duplicate and was not entered into the model (N=111,282). These documents, however, were later added back into the model following the post-modeling stage, using the theta matrix (document-topic) for identical items. Ultimately, 237,170 articles were prepared for analysis using the Structural Topic Modeling approach (STM; Roberts, Stewart, Tingley, et al., 2014).

Similar to the Twitter corpus, all numbers and other non-alphabetical characters were removed from the strings, all stop-words were removed, and all strings were tokenized and stemmed. As the size of data was still massive, and as is common in this

type of analysis, sparse tokens were removed from the matrix. However, this was done in a more aggressive manner than for the Twitter dataset with a sparsity level of 0.99 (as opposed to the 0.999 level, which was used for the Twitter data), removing 17 million tokens out of 97 million words in the corpus. However, due to the size of the documents, data loss was minimal with only 5 out of 237,170 documents deleted due to this process, a number which is extremely unlikely to skew the results.

The next step required estimating the adequate number of topics to be used in the model. However, due to the size of the corpus and the extreme times involved in its modeling (with average sized models requiring an average of a day or more to converge), this estimation had to be performed in a somewhat less fine-grained manner than was the case for the Twitter data. At the first step, ten models, ranging from  $k=20$  to  $k=200$ , were estimated. Figure 15 show the various model-fit indices for the resulting topic-models. Unlike in the case with the Twitter data, there was no sharp reduction in model effectivity beyond a specific  $k$  level. However, there were also diminished returns for held-out likelihood and semantic coherence as  $k$  was increased. It is important to keep in mind that minimizing  $k$  was desirable, as this offers much lower run-time for the process. To further understand the differences in topic structure, Figure 16 presents contrasting considerations between semantic exclusivity and semantic coherence—per each level of  $k$ .

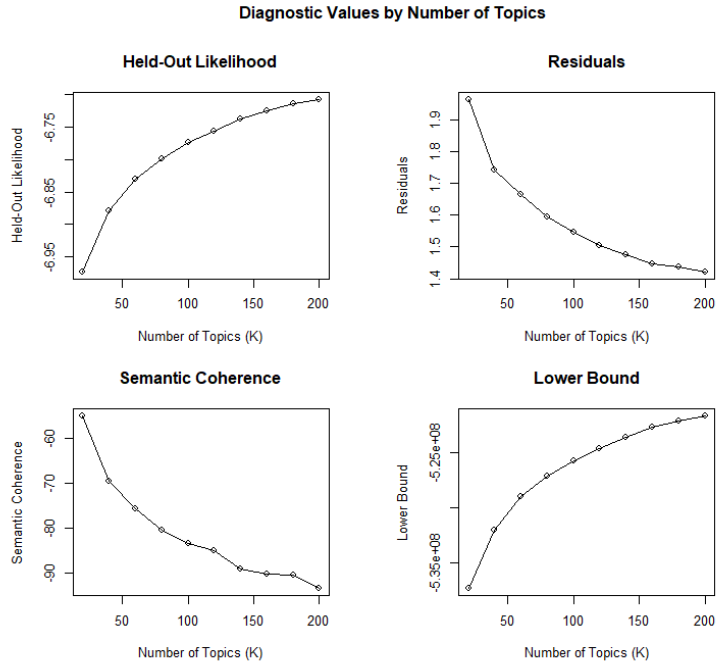


Figure 15: Model fit indicators for models ranging from  $k=20$  to  $k=200$  (by 20).

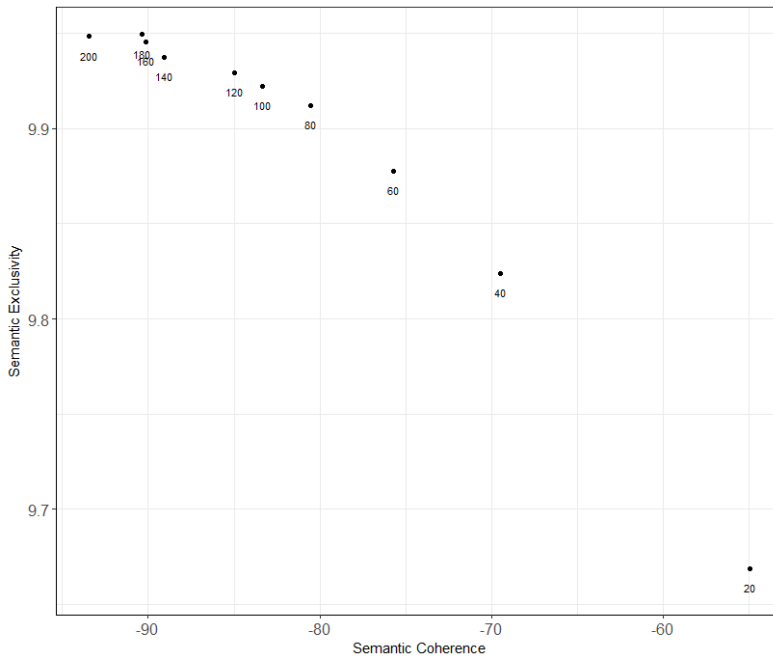


Figure 16: Semantic coherence and semantic exclusivity per  $k$  level for  $k=20$  to  $k=200$  (by 20).



As can be seen in Figure 16, enlarging the number of topics from 20 to 40, or from 40 to 60, improved model performance substantially. However, these improvements became smaller and smaller as  $k$  continued to increase. However, as mentioned in Section 3.3.2, there is a strong argument to be made in the case of diversity measurement for erring on the side of too many topics, i.e., for sacrificing coherence for exclusivity, as the diversity estimation incorporates ( $\delta$ ) into the calculation, thereby negating the impact of topics breaking into similar sub-topics. Considering these arguments, I decided to focus on a range of around 120 topics to more closely examine the model fit indices of more fine-grained choices. Similar to earlier discussions, Figures 17 and 18 present the model fit indices for topics ranging from  $k=100$  to  $k=140$  (by 5).

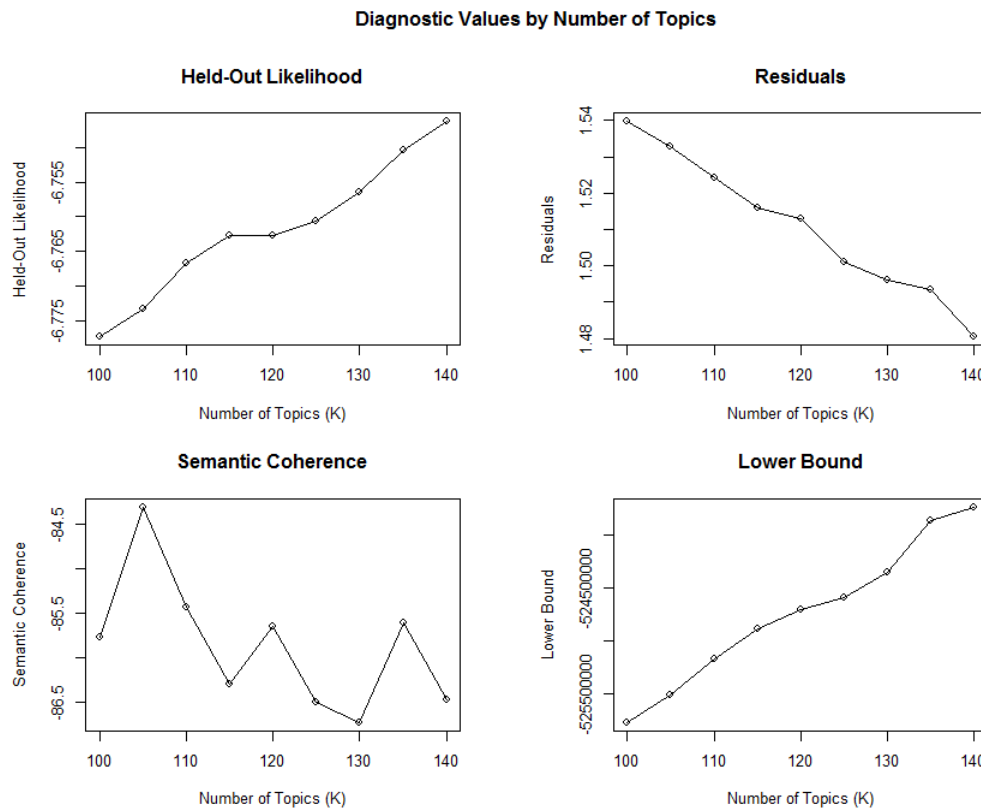


Figure 17: Model fit indicators for models ranging from  $k=100$  to  $k=140$  (by 5).

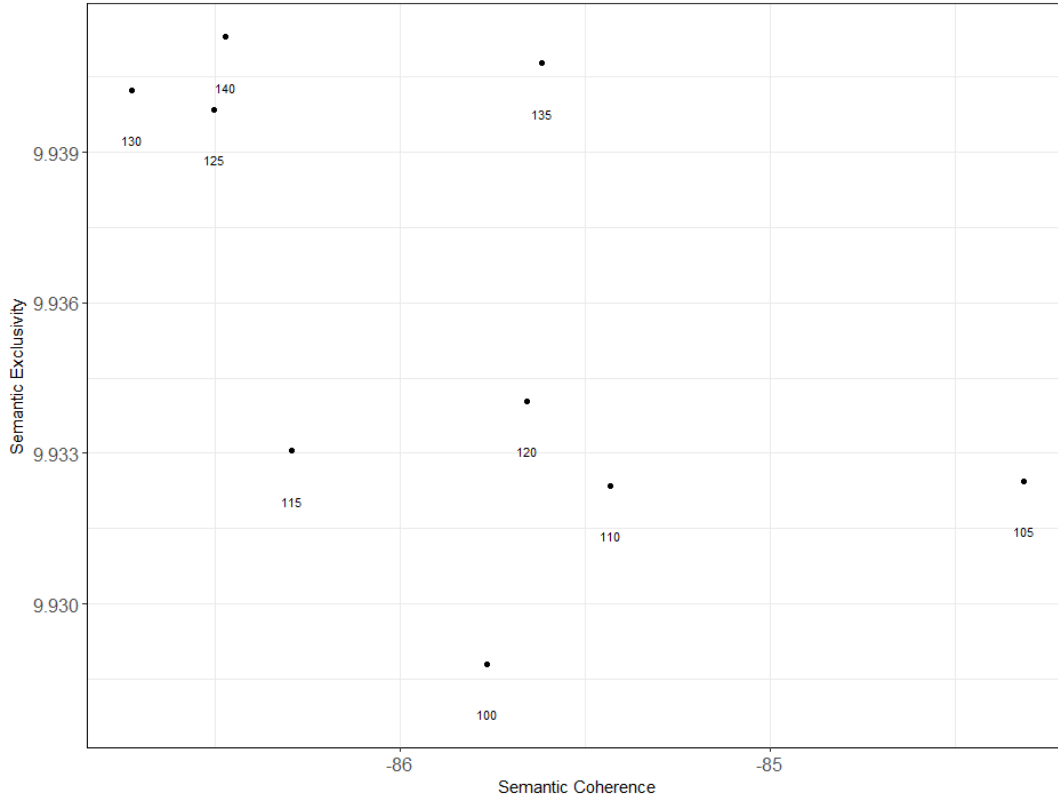


Figure 18: Semantic coherence and semantic exclusivity per  $k$  level for  $k=100$  to  $k=140$  (by 5).

I chose to focus on 135 topic models as the main model for analysis. As can be seen in Figures 17 and 18, a  $k$  of 135 offers near-optimal held-out likelihood and an efficient compromise between coherence and exclusivity (increasing  $k$  to 140 offers marginal improvement to exclusivity but at the cost of coherence). A sample of these topics can be viewed in Section 5.1 of the following chapter.

Of course, compared with their usage for the diversity estimation, the nature of these topics is of less interest for the present study. The final model with a  $k$  set to 135 topics was used to calculate diversity in each candidate's corpus using the method elaborated in Section 3.3.2 on diversity in topic models. These diversity estimations for

both Twitter activity and news coverage were used in the models presented in the results in the following chapter.

#### **4.4.2 Semantic Networks Analysis**

The second method used to explore thematic diversity is semantic network analysis. As a special case of generalized mathematical networks, semantic networks use semantic units (such as unique words) as nodes and the relationship between them (such as co-occurrence) as edges (Baden, 2010; Carley & Palmquist, 1992; Diesner, 2012).

Unlike the process for topic modeling, semantic networks were created individually for each candidate in the corpus. In order to construct the semantic networks, first, all coverage of a candidate in U.S. newspapers and social media was retrieved using LexisNexis (see Section 4.2). A python script was then used to parse the results into different articles and tokenize the texts—that is, converting each text into an ordered vector of words. The TFIDF measure was used to identify the most important words in each candidate’s corpus, with the top 100 words for social media and the top 1,000 words for news coverage kept for analysis. For social media activity, co-occurrence was measured at the document-level, using a cosine similarity transformation over the term-document matrix to create a word co-occurrence matrix. For the analysis of news media, the pure bag of words approach used in topic modeling and the social media analysis was jettisoned. Word order was preserved, and co-occurrence was not measured at the document-level, but in a moving window sized at 50 words. These matrices were then used to construct the semantic network, followed by community detection and diversity estimation. The details of these processes are described in the following sections.

#### 4.4.2.1 Constructing the networks for the news data.

*Choosing the tokens.* First, the data was pre-processed, including a lemmatization of the tokens (using NLTK WordNetStemmer; Bird et al., 2009) and the removal of duplicate entries (per each candidate), stop-words, and seed nodes (the search terms used to retrieve the articles). These sets of pre-processed word vectors were then used for the construction of the semantic networks, with a separate network drawn for each candidate.

One of the main challenges for this analysis was the large volume of unique words in each corpus. This number increases exponentially over the run time for the various tools and scripts. Moreover, while stop-word lists can remove some of the low information words in the corpus (such as “and,” “if,” “or,” etc.), some low-information words were context-specific and hence not listed in pre-existing stop-word lists. Finally, due to the large number of unique words in the corpora, the network representations were often much too massive for analysis and visualization using the available hardware. Thus, I focused analysis on a set of 1,000 unique words in each corpus, chosen using the TFIDF indicator.

The TFIDF measure is one of the most prominent word-level measures in natural language processing, used in various commercial applications such as search-engines, and is designed to identify the most “important” words in a corpus (Sparck-Jones, 1972). “Important” here refers to words that are frequent but still unique to specific contexts, thus providing more information on the document. TF refers to term frequency, or how many times a word appears in each document. The more prominent a word is, the higher the TF\*IDF total value will be. IDF refers to inverse document frequency. That is, 1 divided by the number of documents that a word appears in. Thus, a word appearing in

more documents will receive a lower total TF\*IDF score, as it is less “unique” to specific contexts.

This measure assigns more value to words that are frequent in some documents but not in all and are therefore considered to contain more relevant information. For example, a word that appears once in the corpus will be considered low information, as it is not frequent enough to invest resources into measuring its co-occurrence. A word like “and” might be extremely frequent; however, it also appears in every document and is thus less useful for the purpose of extracting semantic information from the corpus. After calculating the TFIDF score for each unique token in the corpus, the top words were compiled as a set of tokens from which the co-occurrence matrix was built.

***Defining co-occurrence.*** Co-occurrence was analyzed using the following process. First, a matrix containing all possible dyads of the top 1,000 tokens in the whole corpus was created (1,000x1,000). Both the rows and columns in the matrix were comprised of the list of the top 1,000 TFIDF unique words appearing in a candidate’s coverage (that is, each word appears once and only once as a row and as a column). Therefore, each cell in the matrix represents a theoretically possible co-occurrence of two words in the corpus. A python script was then used to enumerate the actual co-occurrences of these word dyads in the texts, with co-occurrence defined as words that appear together in a moving window the size of (n) words. However, the selection of the correct value for (n) was far from trivial.

Some studies view words as related if they appear in the same document in a simple bag of words approach. This assumption was used to construct the semantic networks for the Twitter data. Because tweets often contain a single thematic idea, it is

possible to argue that two words appearing together within the same tweet are also semantically related and should be considered as “co-occurring.” However, news articles are much more complex forms of text than tweets. The articles included in the sample often contained several themes. Assuming that two words are related even if one word appears at the beginning of an article and the second word appears some 1,000 words later is unrealistic (Baden, 2010; Lee, 2014). Another possibility is to consider only words that appear in the same sentence as thematically related. However, such a definition can be too strict, as words that appear in one sentence are likely still to be thematically related to words appearing in the previous or following sentences. Therefore, an alternative approach would be to view words as co-occurring not if they appear within the same document, or the same sentence, but if they appear within a certain pre-determined distance from each other.

To illustrate, the following sentence can be used: “Ant Bat Cat Dog Eel Fox Goat Hog Ibis Ant.” This sentence contains 10 words (A,B,C,D,E,F,G,H,I,A), or a set of 9 unique words (A,B,C,D,E,F,G,H,I). Using a 2-word window, the word “Ant” will be registered as co-occurring with the words Bat and Cat (“Ant Bat Cat Dog Eel Fox Goat Hog Ibis Ant”). Similarly, the word “Bat” will be registered as co-occurring with “Cat,” “Dog” and “Eel” (“Ant Bat Cat Dog Eel Fox Goat Hog Ibis Ant”). We then continue to go over the sentence word-by-word enumerating co-occurrences for each word (the search-forward approach was used because it is more efficient than the backwards-window; retrospective co-occurrence is accounted for by previous terms in the search-forward approach).

In Figure 19, the matrix on the left can be drawn for the words that co-occur with “Ant” and “Bat.” The matrix on the right is drawn for all co-occurrences in this sentence. As can be observed, these matrices are symmetrical, meaning that when divided along the diagonal, the upper triangle and the lower triangle are identical (which will later result in a non-directed network).

	A	B	C	D	E	F	G	H	I
A	X	1	1						
B		X	1	1					
C			X						
D				X					
E					X				
F						X			
G							X		
H								X	
I									X

	A	B	C	D	E	F	G	H	I
A	X	1	1	0	0	0	0	1	1
B		X	1	1	0	0	0	0	0
C			X	1	1	0	0	0	0
D				X	1	1	0	0	0
E					X	1	1	0	0
F						X	1	1	0
G							X	1	1
H								X	1
I									X

Figure 19: Co-occurrence matrices for the sentence “Ant Bat Cat Dog Eel Fox Goat Hog Ibis Ant,” using a 2-word window for co-occurrence.

Already, this example raises the critical question of how large the moving window should be for the actual analysis of news data. Previous studies have suggested a large range of possible (n) values for the size of the word-distance window within which two words are considered to co-occur. For example, focusing on Dutch news media, Baden (2010) used a window of 30 words as a maximal distance for two words to be considered related. In contrast, examining the text of theoretical writing by scholars associated with the Frankfurt School, Lee (2014) considered words to be related if they appear within 120 words from each other, or roughly one paragraph. The selection of these (n) word distance windows stems from the different contexts of these text sources. In the first

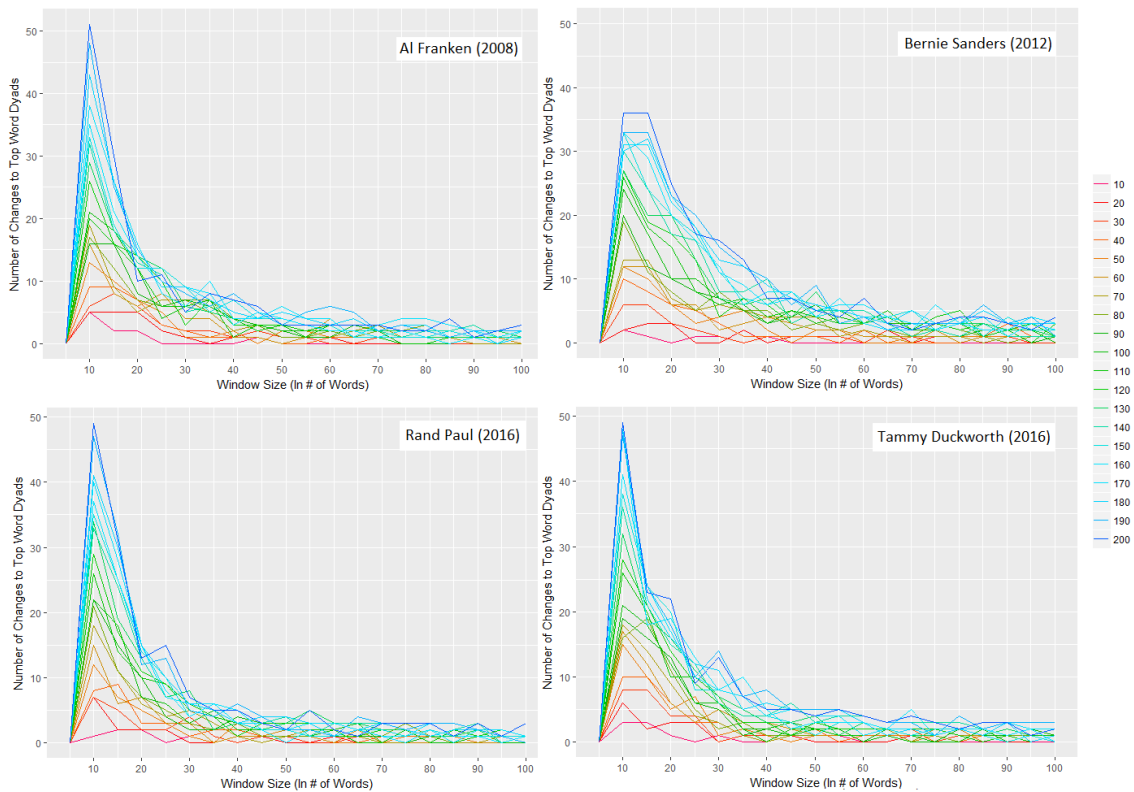
example, in which texts were written in Dutch and gathered from Dutch newspapers, one could reasonably expect the writing to be more succinct and the thematic units simpler, spanning only a handful of sentences. By contrast, when examining texts written by scholars associated with the Frankfurt School (although Lee (2014) used English translations, these were still influenced by the German in which many of these documents were written), one would expect a wordier presentation of ideas, with larger and more complex thematic units spanning several paragraphs. The lesson of this comparison is that the right size window for co-occurrence needs to be determined by observing and tailoring to the context in which the study is conducted.

In order of support the decision to create a 120-word window, Lee (2014) suggests observing changes to word dyads as a result of changing the size of window for co-occurrence. She observed that changes to word dyads declined rapidly after reaching roughly the 104-word mark. I used a similar method to observe the reliability of window-size decisions, rather than relying on the stated window size of 120 words offered by Lee (2014). This was based on the assumption that the rate of change will be different for news articles than for dense, academic philosophy translated from the German.

First, I chose four candidates whose corpus length ranged from slightly above the mean number of documents-per-candidate (1304) to slightly below the median number of documents-per-candidate (900). Then, for each of the four candidates, I created a semantic network using varying word-window sizes, ranging from five to 100 words (punctuation and paragraph breaks were considered by the algorithm as a multiple-word-space, as will be detailed later). This allowed me to observe changes to the pattern of word co-occurrence for each candidate. I started by examining changes to the top ten



most frequently co-occurring words as a result of changes to the word-window sizes. I did this by comparing the top ten most frequently co-occurring words when using a ten-word word-window size with the top ten most frequently co-occurring words when using a five-word word-window size and registering the number of dyads that were different between the two. I then repeated this process for all window sizes (comparing the 15-word window with the ten-word window, the 20-word window with the 15-word window, and so on), and for several  $n$ -top dyads (examining changes to top 20 most co-occurring dyads, top 30 dyads, and so on, up to the top 200 dyads).



*Figure 20: Change In top ( $x$ ) co-occurring word dyads as a function of window size and length of top dyads list ( $x$ ).*

Figure 20 presents the results of this analysis for all four candidates. The x-axis represents different possible window sizes (from five to 100 words). The y-axis

represents number of changes to the top (x) co-occurring dyads when moving from one window size to a larger one. Finally, the colors represent the scope of the analysis, from examining the differences in the top ten co-occurring dyads to examining the difference in the top 200 co-occurring dyads. The larger the size of the top dyads, the more stringent the test is, as dyads are more likely to change due to the increasing list size and the inclusion of less connected word dyads.

As enlarging the window for co-occurrence is costly in terms of computational resources, with larger window sizes taking a much longer time to analyze, my goal was to find the minimal window size that exhibits stable and reliable results in terms of top co-occurring dyads. Based on the figure above, I determined this window size to be 50 words. This value minimizes window size while still maintaining changes to the co-occurrence patterns due to relatively small changes to window size.

After deciding on an adequate word-window size, a python script was then used to enumerate the actual co-occurrences of word dyads in all texts for all candidates, with co-occurrence defined as words that appear together in a moving window no larger than 50 words. For every word (a) in an article, the script gathered all words (b) appearing 50 words after that word (using only forward-search to avoid counting word co-occurrences twice, as would be the case if the search was done 50 words backward as well).

Additionally, the algorithm was set to consider punctuations and paragraph breaks as a multiple-word-space, with distance “penalties” for punctuations ranging from one word (for “,”) to five words (for manual line breaks). The matrices for all articles per candidate were then summed and the resulting corpus-level co-occurrence matrix served as data for constructing the semantic network for each candidate, with the cells of the matrix

being symmetric on the main diagonal—in other words, the upper and lower triangles were identical, and each contained all of the information needed for the construction of the undirected semantic network.

*Normalization of the co-occurrence matrix.* Next, the co-occurrence matrix needed to be normalized to allow for a comparison between different matrices and networks drawn for each candidate. However, determining the correct normalization of the co-occurrence matrices is also a complex matter. Several possible alternatives were available, including Pearson correlation, cosine similarity, and other measures.<sup>6</sup> Examples for normalization are more common for a related problem, normalizing the values of an occurrence, rather than a co-occurrence, matrix. In the context of this study, an occurrence matrix is defined as the term-document matrix, or a matrix where the rows are all the documents in the corpus and the columns enumerate the appearance of all possible words in a corpus (similar to that used for the Twitter data). When normalizing this type of occurrence matrix, cosine similarity is considered to be a “best practice.”

However, when using a moving window approach, the actual theoretical occurrence matrix is essentially unobtainable. A normalization method suitable to a co-occurrence matrix was thus necessary. Using cosine similarity as a method of normalization would have been erroneous, as the word similarity would be normalized twice if this method were to be applied to a co-occurrence matrix (Zhou & Leydesdorff, 2016). An equivalent normalization, the Ochiai coefficient, was thus used instead. As Zhou and Leydesdorff (2016) show, both theoretically and empirically, the Ochiai normalization over the co-

---

<sup>6</sup> Earlier versions of this manuscript used the cosine similarity approach to normalization which has been shown empirically and theoretically to be suboptimal (Zhou & Leydesdorff, 2016).

occurrence matrix is identical to the cosine similarity normalization over the occurrence, or term-document matrix, and is thus in line with current best practices.

Following the normalization and using the iGraph package for the statistical language R (Csardi & Nepusz, 2006), the co-occurrence, or term-term matrix, was then converted into a network object. Each unique token appearing in the candidate's coverage served as a node, with the normalized co-occurrences calculated earlier (as a measure of 0-1) serving as the edges. An example of such network can be found in Section 5.1 of the following chapter.

#### **4.4.2.2 Constructing the networks for the Twitter data.**

The construction of the semantic networks from the Twitter data differed from the news media data in two important ways. First, instead of selecting the top 1,000 words, only the top 100 TFIDF words were selected to create the semantic network. This difference stems from the lower number of unique words appearing in tweets relative to most other corpora (unique documents, tweets, contain far fewer words than the average article). Second, because tweets often contain a single thematic idea, words that appear together within the same document (tweet) were determined to be semantically related or “co-occurring”—an approach that is considerably more cost-efficient than the moving window approach.

*Choosing the tokens.* The extraction of relevant tokens was done in a similar fashion to the process used for the news data. First, the data was pre-processed, including a lemmatization of the tokens (using NLTK WordNetStemmer; Bird et al., 2009) and the

removal of stop-words. Like the news media corpus, the choice of tokens was done using the TFIDF measure, by selecting the top 100 words in each candidate’s corpus.

**Defining co-occurrence.** Because tweets often contain a single thematic idea, one can reasonably argue that two words appearing together within the same tweet are also semantically related and should be considered as “co-occurring.” Thus, unlike the analysis of the news data, and due to the relatively small text size of tweets, words in the Twitter corpus were viewed as related if they appear in the same document in a simple bag of words approach. Therefore, to analyze the Twitter data, the first step was to construct a term-document matrix. To illustrate, the following three short texts can be used:

Text 1 - Ant Bat Cat Dog

Text 2 - Dog Eel Fox Goat

Text 3 - Goat Hog Ibis Ant

The correspondent term-document matrix for these three texts is:

	A	B	C	D	E	F	G	H	I
Text 1	1	1	1	1	0	0	0	0	0
Text 2	0	0	0	1	1	1	1	0	0
Text 3	1	0	0	0	0	0	1	1	1

Figure 21: The term-document matrix for the three short sentences (“Ant Bat Cat Dog,” “Dog Eel Fox Goat,” “Goat Hog Ibis Ant”).

This matrix can then be used for the construction of the word co-occurrence matrix and the network creation using the following steps.

**Normalization of the co-occurrence matrix.** A normalization of the term-document matrix was needed to allow for comparison between different matrices and

networks drawn for different candidates, and to convert the term-document matrix into a co-occurrence matrix. The normalizing using cosine similarity is considered as “best practice.” It is also equivalent to the Ochiai coefficient used for the news data and was therefore chosen for this analysis (Zhou & Leydesdorff, 2016).

For each pair of words of the 100 words used to construct the term-document matrix, cosine similarity was calculated as:  $\frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$ , with A referring to items in the column related to word 1, and B referring to items in the column relating to word 2. In more descriptive manner, for each two words, the cosine similarity measure estimated the extent to which these words “share” documents. The more documents shared between two words, the more related they are. This is also normalized by the general frequency of the words, as more frequently used words are expected to co-appear with other words more often. The resulting matrix is similar to the matrix constructed for the news media, with rows and columns representing a set of 100 unique words, and each cell representing their normalized relatedness.

	A	B	C	D	E	F	G	H	I
Text 1	1	1	1	1	0	0	0	0	0
Text 2	0	0	0	1	1	1	1	0	0
Text 3	1	0	0	0	0	0	1	1	1

Figure 22: The term-document matrix for the three short sentences (“Ant Bat Cat Dog,” “Dog Eel Fox Goat,” “Goat Hog Ibis Ant”).

To illustrate, Figure 22 presents the term-document matrix for the three short texts offered earlier. The transformation of this matrix can be done in the following manner (examples for the calculation shown only for three unique co-occurrences):

$$\text{Cosine}(A,B) = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} = \frac{1*1+0*0+1*0}{\sqrt{1+0+0} \sqrt{1+0+1}} = 0.71$$

$$\text{Cosine}(A,D) = \frac{\sum_{i=1}^n A_i D_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n D_i^2}} = \frac{1*1+0*1+1*0}{\sqrt{1+0+1} \sqrt{1+1+0}} = 0.5$$

$$\text{Cosine}(A,E) = \frac{\sum_{i=1}^n A_i E_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n E_i^2}} = \frac{1*0+0*1+0*0}{\sqrt{1+0+0} \sqrt{0+1+0}} = 0$$

This process results in the following matrix, on the left of Figure 23.

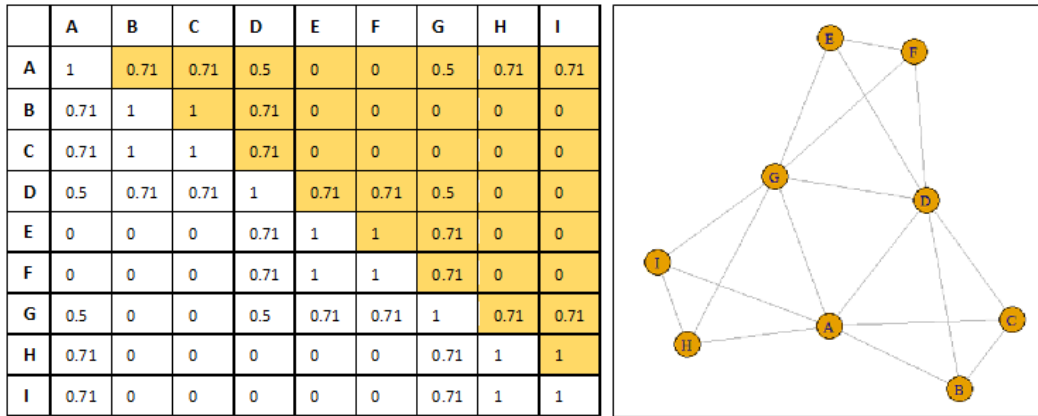


Figure 23: The normalized co-occurrence matrix for the three short sentences (“Ant Bat Cat Dog,” “Dog Eel Fox Goat,” “Goat Hog Ibis Ant”) and the corresponding semantic network.

Following the same process for all available dyads, a matrix can be drawn to represent the relationship between all unique words. This normalized co-occurrence matrix can then be converted to a graph object for further analysis, as exemplified by the network structure on the right of Figure 23. More detailed examples of actual semantic networks drawn from political candidates’ social media activity can be seen in Section 5.2 in the following chapter, focusing on networks that exhibit high and low diversity scores—the calculation of which is explained in the following section.

#### 4.4.2.3 Community structure and diversity estimation.

Following the normalizations and using the iGraph package for the statistical language R (Csardi & Nepusz, 2006), the co-occurrence, or term-term matrix, was then

converted into a network object. Each unique token appearing in the candidate's tweets or news coverage (depending on the corpus) served as a node, with the normalized co-occurrences calculated earlier (as a measure of 0-1) serving as the edges.

In order to calculate thematic diversity, first, the community structure of each given network needed to be estimated. Community detection was performed using the multi-step modularity maximization method suggested by Blondel et al., (2008), due to its high-level of accuracy while still maintaining computational efficiency (Lancichinetti & Fortunato, 2009). As an additional benefit, the community structure of all networks can be compared using this technique, as the structures all offer a community with maximized modularity.

As mentioned earlier, when dividing a network into sub-graphs or communities, the ratio of internal (within sub-graph) and external (between sub-graph) links is measured by the community structure modularity score (Rubinov & Sporns, 2010). Thus, modularity measures the extent to which communities in the network are interconnected within, yet separated from, other groups, with higher values indicating a better division. Blondel et al.'s (2008) multi-step modularity maximization method (also known as Louvain method) uses an iterative process to maximize this modularity score.

First, the algorithm assigns a separate community for each node. Then, for each node, the algorithm attempts to add the node to its neighboring communities. If the addition of a node to one of its neighbors' communities increases the modularity score of the network, then the node is added to that community permanently. This process is performed iteratively for all nodes in the network until no improvement to the modularity score can be made. The result is a new network of communities with only one to two



nodes. The algorithm then again attempts to add each community to its neighboring communities. If this change increases the modularity, then the communities are merged, and a new network of merged communities is created. This process is performed over and over until it reaches a stage in which no community can be merged with any of its neighboring communities to increase modularity. After the number of communities and the modularity were logged, community membership was added as a node-level attribute for all nodes in the network and used to estimate diversity. As will be discussed in the final chapter, this single community approach for community detection can be seen as a drawback when compared with the topic modeling approach and might be improved as future approaches and tools become available.

Thematic diversity was measured according to the conceptual processes elaborated in Section 3.3.4. First, the distance between each two communities ( $\delta$ ) was measured as the sum of edges between them, divided by the sum of all edges in a subgraph containing only these two communities. Formally, this was measured as  $(\delta_{ij} = 1 - \frac{\sum E_{ij}}{\sum E_i + \sum E_j + \sum E_{ij}})$ , with  $\sum E_{ij}$  defined as the sum of edges between each two communities (i) and (j),  $\sum E_i$  as the sum of edges within community (i), and  $\sum E_j$  as the sum of edges within community (j).

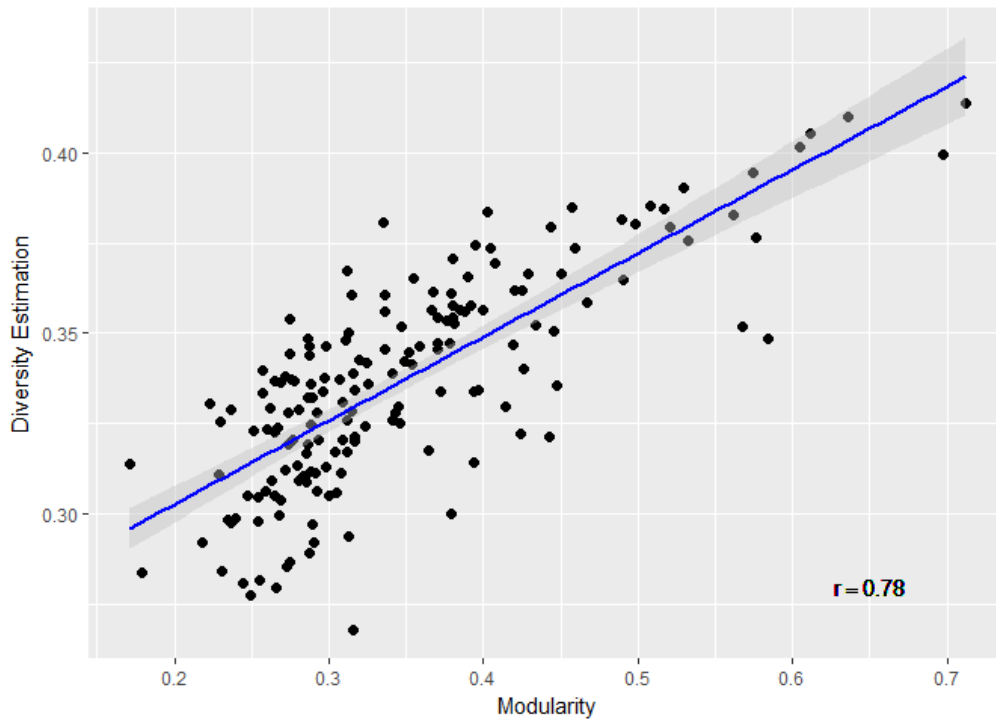
The size of each community ( $p_i$ ) was measured as the sum of edges in a single community divided by the sum of all total edges for all communities in the network:

$$p_i = \frac{\sum E_i}{\sum_1 \sum E_j}$$

Finally, these values were entered into the formula offered by Stirling (2007),

$$\text{iterated over all possible community dyads } ( D = \sum_{ij(i \neq j)} (\delta_{ij}) \cdot (p_i \cdot p_j) ).$$

Unsurprisingly, diversity captured several dimensions of the network community structure concurrently, thus performing better than modularity or community number alone (as well as more basic measures of semantic network diversity, such as mere network density, cf. Eberl, Jacobi, & Schlögl, 2014b).



*Figure 24: Relationship between the modularity score of the networks drawn from the Twitter dataset and the diversity estimation using Rao's coefficient (Pearson's  $r=0.78$ ).*

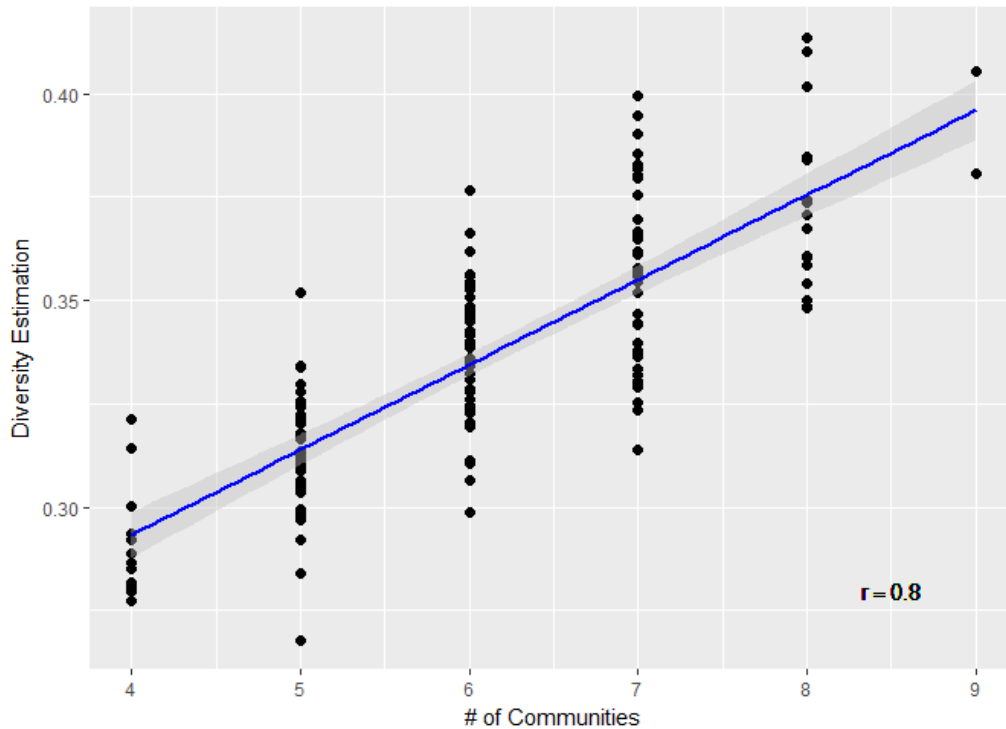


Figure 25: Relationship between the number of communities in the networks drawn from the Twitter dataset and the diversity estimation using Rao's coefficient (Pearson's  $r=0.8$ ).

#### 4.4.2.3 Using random networks as benchmark for diversity estimation.

In order to explore the relationship between thematic diversity and electoral success, different networks with different number of nodes, edges, density, and degree of distribution need to be compared—a process not without difficulties (van Wijk, Stam, & Daffertshofer, 2010). The main problem is differentiating between network properties that might be random, as these can lead to spurious relationships with the outcome variable, and network properties that result from fundamental design principles of the observed semantic networks (Maslov, Sneppen, & Zaliznyak, 2004; Squartini & Garlaschelli, 2011; Stouffer, Camacho, Jiang, & Nunes Amaral, 2007). Additionally, in both the news media and social media, the semantic networks are created from different

corpora, with different number of texts and different length-distributions for each set of texts. Such discrepancies lead one to suspect that there could be an impact on the observed relationships (despite the number of nodes being identical). Moreover, graph feature controls (such as number of texts, length of texts, or density) are expected to be correlated at least to some degree with the structural features of the network. Therefore, controlling for these differences—for example, by adding them as covariates in a regressions model—might increase multicollinearity and make the interpretation of the specific coefficients problematic (in terms of effect size and direction). Additionally, inserting these variables as covariates into the regression models assumes a linear relationship between corpus-level variables and diversity—an assumption that might not accurately characterize their relationship.

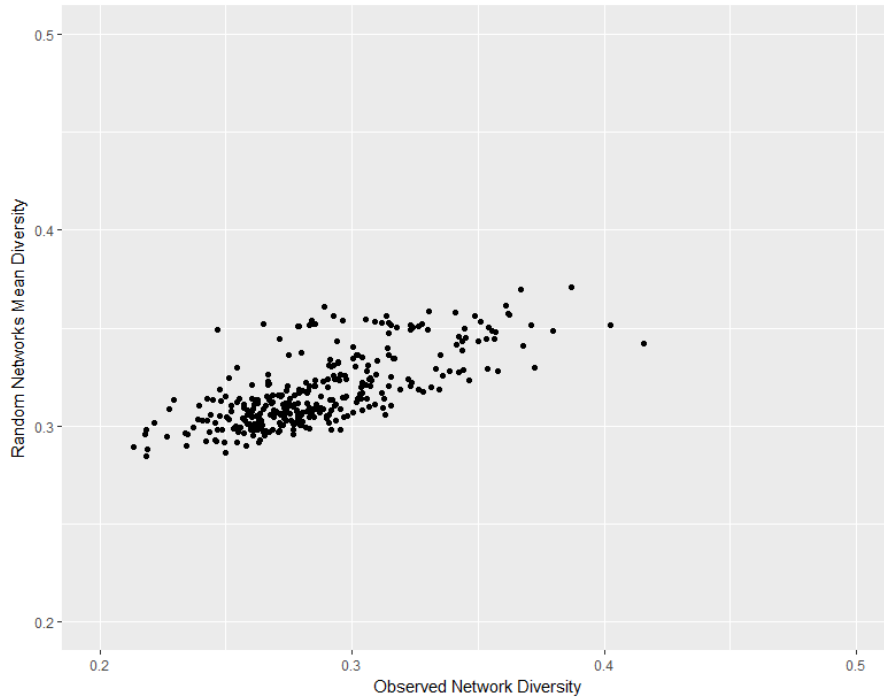
Thus, in order to address these problems and to improve the robustness of the comparative method, I use configuration models to generate random network models with an identical number of nodes and degree distributions for each semantic network to serve as a benchmark (Squartini & Garlaschelli, 2011). This strategy can help identify non-trivial and significant structural features of the semantic networks before examining their consequences and antecedents.

First, I calculated the graph-level statistics for the semantic network, as described earlier, which resulted in an observed diversity score for each network. Then, for each of these networks, I generated 100 random surrogate networks to be used as a benchmark. Each of the randomized networks was created with the configuration model method. For each network, the configuration model algorithm removes all edges but keeps the “stubs” of each edge intact. It then chooses an edge stub randomly and connects it to another

random stub. This process is iterated until all stubs are connected. Following this, all edge weights from the original network model are collected. For each re-wired stub, an edge is randomly assigned from the observed weight vector until all edges receive a weight score. The result is essentially a rewiring of the edges between all nodes and their weight scores, while keeping the number of nodes, density, and even degree and node-strength sequence constant.

This process was repeated 100 times for each network. The mean and standard deviation for each of the graph-level indicators over the ensemble of random networks was calculated. Finally, the diversity score for each graph was calculated as the Z-score for thematic diversity. This was done by subtracting the random network mean diversity from the observed network diversity and then dividing it by the standard deviation for diversity over the entire ensemble of random networks. The normalized diversity indicator in this method, therefore, represents the difference between diversity in the observed network and the expected graph-level indicators in random networks, with identical numbers of nodes and density. I present the results of this process in the next chapter following the general results of the semantic network analysis.

Interestingly, the randomized networks showed some very consistent features. First As figure 26 shows the spread of the diversity estimation was much larger for the observed semantic networks than for the random networks. This indicated that diversity of random incoherent networks has a firmer lower bound. In other words The diversity for random networks is the result only of basic network properties, which might not be dictated by the organizational features of the network



*Figure 26: A scatterplot for the diversity estimations for the observed and random-generated networks.*

This argument is strengthened by the diversity estimations for the random networks being generally higher than the diversity estimations for the observed networks in a fairly consistent manner. This can be seen in Figure 27 which shows the histogram for the gap between the diversity estimation of the observed network and the diversity estimation for the random-generated networks. The prevalence of negative values in this histogram indicates that in most cases (n=295) and aside from the smallest networks (n=35) generally the thematic diversity of randomly generated textual data was found to be higher than a more coherent text would produce. This is also supported by Figure 28 which plots the difference between the observed and random networks diversity estimations on the number of articles from which the data was driven.

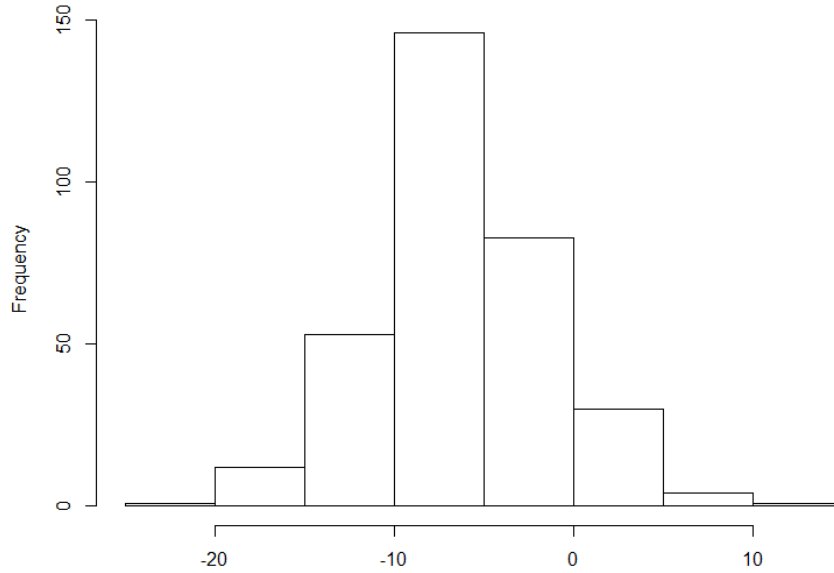


Figure 27: A histogram for the values of  $(\text{Observed Diversity} - \text{Random Diversity})$ .

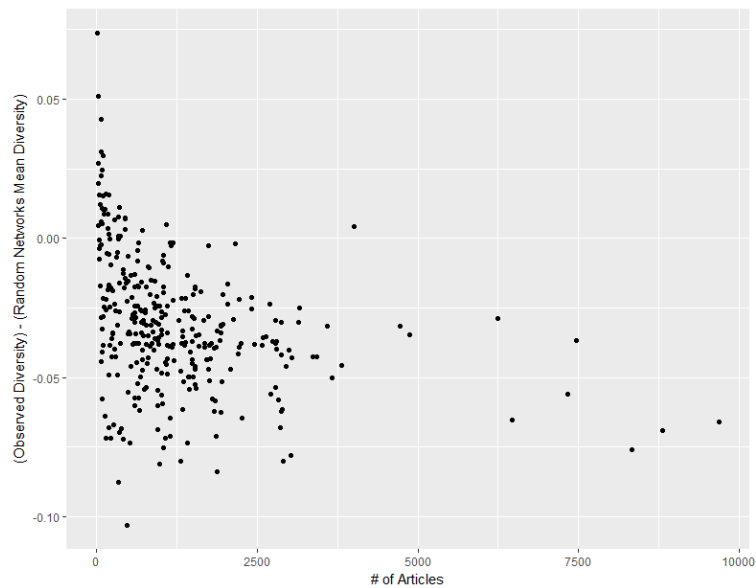


Figure 28: A scatterplot for the diversity gap between observed and random-generated networks, and the number of articles from which data was drawn.

These findings can be seen as supportive of the suggested method. If more coherent and monothematic campaigns are expected to produce less diverse semantic networks, then a randomly connected network, representative of randomly and incoherent generated

text should show high levels of diversity. I discuss these results in further details on Chapter 6 of this dissertation. In addition, the models presented in the following chapter were estimated also on a sample that includes only the network which showed lower diversity than random and the various results presented in the next chapter show similar and stronger results.

#### **4.5 Statistical Modeling Strategy**

The goal of this analysis is to examine the relationship between the volume, tone, and diversity of candidates' news coverage and social media activity and their subsequent electoral success. Given that the share of votes that each candidate received in each race is non-independent from the share of votes their opponent receives (roughly, the sum of both candidates' share of the votes is 100%), I ran separate regression models for the Republican and Democratic candidates. For each of these models, the independent variables included both media and non-media variables for the candidate and his or her opponent. Although it is statistically unjustified to treat candidates competing in the same state as independent observations, and while using two separate regression models avoids the potential for the observations to be non-independent, it is important to note that such an approach does come at the cost of weakened statistical power. Therefore, the results presented here should be interpreted as conservative estimates. While I report the results for the Republican candidates only, the results for the Democratic candidates were almost identical.

In the following chapter, I present the results for the multiple regression models, starting with the analysis of political candidates' news media coverage followed by the analysis of their social media activity. For both contexts, I present the results using topic



modeling as a diversity estimation, semantic network analysis as a diversity estimation, and the results obtained using both diversity measurements taken in tandem.

## 5. RESULTS

I begin by presenting the results on the relationship between candidates' news coverage and their electoral success. I first outline the results for the topic modeling-based diversity estimation, examining the impact of the number of topics in the possible models ( $k$ ) on model performance and then focusing on an optimal solution of  $k=135$ . Results are given for a model containing only media variables—volume, tone, and diversity—followed by the results of a more elaborate model that controls for non-media factors, such as experience and state ideology.

Following this, I present the results of the semantic network analysis diversity estimation. I first present the results of the raw scores for semantic network diversity, as well as the results of more elaborate models controlling for non-media factors. Finally, I give the results for a model of thematic diversity estimation that combines topic modeling and semantic network analysis.

I then turn to present the results for the relationship between candidates' social media activity and their electoral success. For this purpose, I repeat the same order of presentation, starting with models using topic modeling-based diversity estimation and ending with full models that incorporate both diversity estimations, semantic network analysis and topic modeling. In the next chapter, I discuss the implications of these results more broadly, from both a theoretical and methodological perspective.

## 5.1 News Coverage and Electoral Success

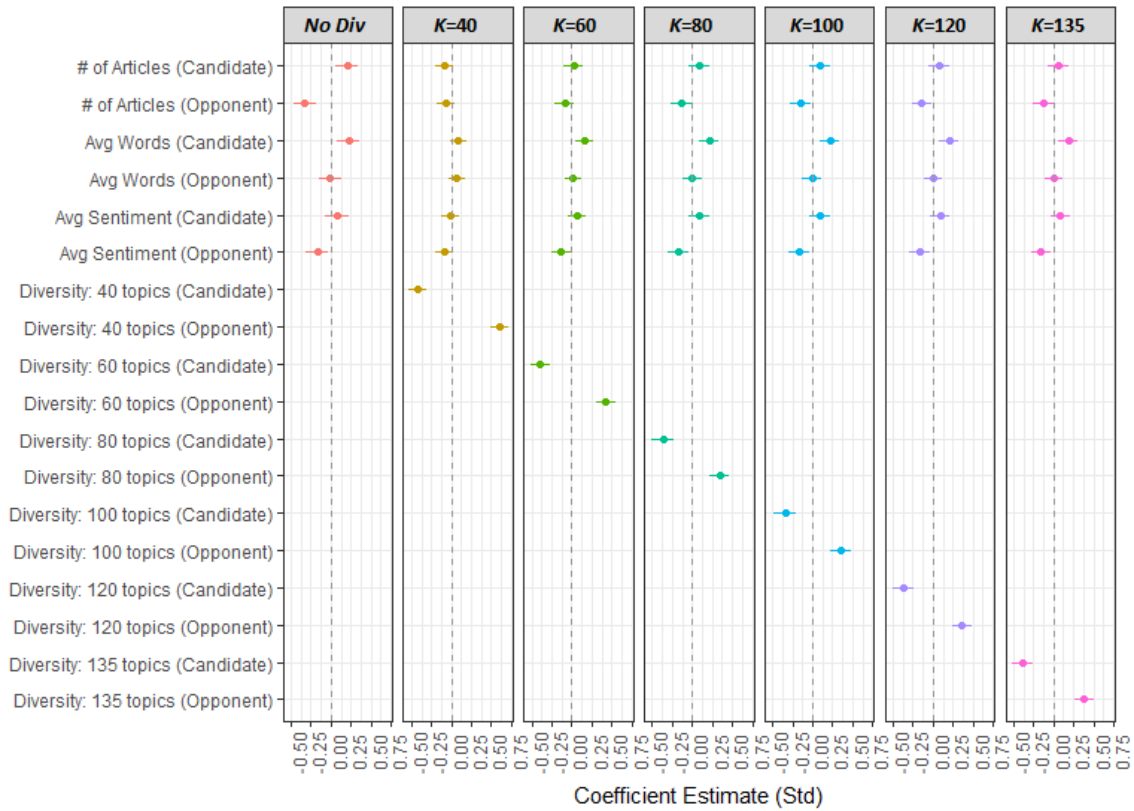


Figure 29: Summary of linear regression analysis for social media variables predicting Republican candidates' electoral success, with different ( $k$ ) topic model per facet ( $N=165$ ).

Figure 29 presents the results of the linear regression models predicting Republican candidates' electoral success. It directly compares the diversity estimation performance across the various topic models (ranging from 40 to 135 topics). First, as can be seen from the first two rows, the volume of coverage has a positive relationship with electoral success, though this relationship is not significant for all models. This is mainly the case for the regression model in which thematic diversity is not included, as well as models in which a high- $k$  topic structure was used. In other words, the number of articles covering a candidate is positively related to his or her electoral success, with the number of articles mentioning his opponent having a negative relationship on electoral success. Similarly,

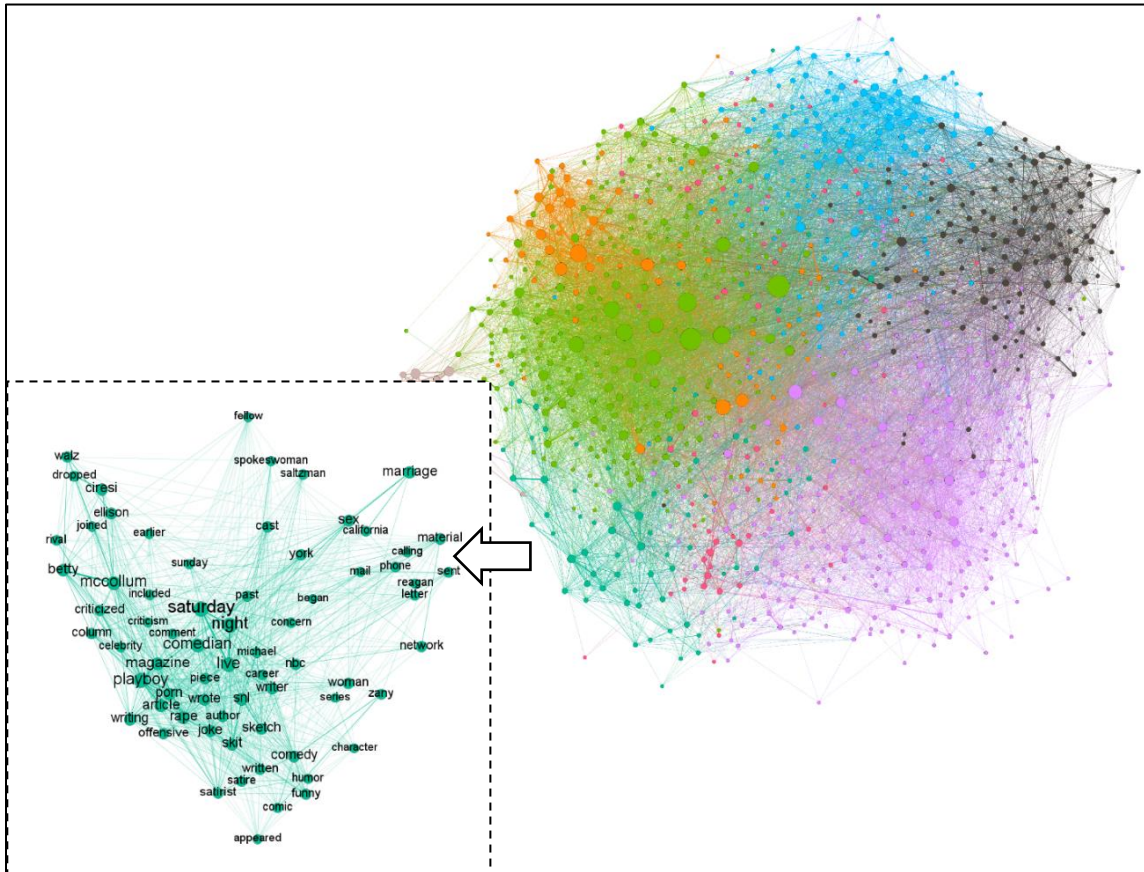
the sentiment of coverage has a positive relationship with electoral success, though this relationship weakens when the diversity variables are incorporated into the model.

Most importantly, we can see that for higher (k) topic models, diversity is negatively related to electoral success. This is true for all models ( $p < .001$ ). I therefore chose to proceed with the (k=135) maximal model in the following analyses. However, it should be noted that the results presented above are very similar regardless of the number of topics included in the model, though all topic models used in this study offer a relatively large number of topics, and that the results might vary for models with a lower number of topics (though these models are also likely to offer a much lower model fit for the topic models themselves).

Figure 30 shows some of the most interesting topics estimated by the model. For example, the three topics in the first row were the most common. This should not be surprising, given that these topics discuss the campaign itself—a topic that was relevant to all candidates in the sample and which the media tends to focus on when covering candidates and elections campaigns (Cappella & Jamieson, 1997). They also serve as an example of topics that were somewhat similar in vocabulary, especially the two rightmost topics, both of which emphasize words related to the campaign. Of more interest were topics that appeared for fewer candidates, or at least not for all candidates. These can be seen in Figure 30 in the second and third rows. For example, the rightmost theme in the second row emphasizes words such as nuclear, Iran, Russia, weapon, sanction, agreement, and negotiation. Therefore, this theme seems to be related to foreign affairs and the nuclear arms race. The next theme contains words such as transport, regional, airport, and highway, thus connecting to issues of transportation and infrastructure.



estimations and candidates' electoral success is shown in Tables 7 and 8, under "TM Diversity."



*Figure 31: The semantic network for candidate Al Franken drawn from the candidate's news coverage during the 2008 Senate Elections (size of nodes is determined by weighted degree; color of node represents community membership; Layout created using the Force Atlas algorithm).*

Figure 31 presents an example of one semantic network from the set used to estimate thematic diversity. This network is constructed from the first corpus in the data set, the news coverage of candidate Al Franken in the 2008 Senate elections in Minnesota. Presenting a dense network of 1,000-nodes in a low-resolution format is somewhat of a challenge, and therefore edges were removed using the maximal level of thresholding possible without separating the network into different components.

However, community data as well as the diversity calculations were based on the full network.

In the network shown in Figure 31, each unique token appearing in the candidate's coverage serves as a node, with the normalized co-occurrences calculated earlier (as a measure of 0-1) serving as the edge. The color represents the community to which each word belongs, and the figure highlights one particular community related to the candidate's past as a cast member for NBC's *Saturday Night Live*, and several scandals related to his past (interestingly, this coverage dates to 2008, well before the 2018 controversy around Franken's conduct with women, which led to his resignation as a senator). While the semantic networks for the news corpus were very large in terms of density and node number, for the social media data, only 100 tokens were used; the network visualization can therefore be presented in a more detailed manner. Further review of the network structure, its various communities, and a visual representation of different diversity levels is shown in Figures 35 and 36 in the next section, which discusses the relationship between thematic diversity in social media activity and electoral success. Most importantly, the diversity estimations calculated from the networks appear in Tables 7 and 8 under SNA diversity.

Table 7 presents the results for the regression models estimating the impact of various media factors on candidates' electoral success. Model 1 presents the results of the regression models containing only the basic corpus features—volume, average length, and tone. As the results of model 1 show, volume was found to be positively correlated with electoral success, with more coverage being connected to higher success ( $p < .01$ ). Tone was also found to be a significant predictor, though only for one opponent's

coverage ( $p < .05$ ): positive coverage for one's opponent was related to a smaller share of votes. However, as subsequent models show, when adding diversity as a predictor to the model, volume becomes only marginally significant at best ( $p = .10$ ). Similarly, while the impact of tone persists in model 2 ( $p < .05$ ), in the models incorporating diversity, it is found to be non-significant.

Table 7: Summary of multiple regression analysis of news media factors on Republicans vote share

Variable	model 1 (RSE)		model 2		model 3		model 4	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	.00	<.001***	.00	<.001***	.00	.007**	.00	.012*
# Articles (Rep)	.20	.002**	.07	.2	.10	.108	.07	.253
# Articles (Dem)	-.33	.007**	-.12	.251	-.04	.487	.03	.674
Doc Length (Rep)	.22	.003**	.19	.005**	.07	.197	.09	.11
Doc Length (Dem)	-.01	.926	.00	.983	.10	.093•	.08	.15
Tone (Rep)	.08	.261	.08	.227	.01	.816	.03	.578
Tone (Dem)	-.17	.047*	-.16	.047*	-.04	.539	-.06	.286
TM Diversity (Rep)			-.39	<.001***			-.23	.001***
TM Diversity (Dem)			.38	<.001***			.24	<.001***
SNA Diversity (Rep)					-.45	<.001***	-.31	<.001***
SNA Diversity (Dem)					.65	<.001***	.52	<.001***
Adj R <sup>2</sup>	.19		.43		.52		.58	

Note: • $p < .1$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$  (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=165).

As can be seen in model 2, diversity estimated with topic modeling is found to have a very strong and negative relationship with electoral success. Lower thematic diversity for a candidate is a predictor of higher electoral success; similarly, higher thematic diversity for one's opponent is a predictor of higher electoral success. In addition, the predictive capacity of the model improves greatly from an Adjusted R-Squared score of 0.19 for the corpus features only, to an Adjusted R-Squared score of 0.43 for the model that includes thematic diversity as a predictor, as estimated with topic modeling.



Model 3 is similar to model 2 but uses semantic network analysis as the basis for diversity estimation rather than topic modeling. As in model 2, diversity is also found to be a significant and negative predictor of electoral success when using semantic network analysis for estimation. This is unsurprising, as these two measures were significantly correlated ( $r = 0.37$ ,  $p < 0.001$ ). However, this is not to say that both measures are identical: each method uses different assumptions, measurements, and statistical procedures for estimation (as reviewed in the two previous chapters). First, while the results of model 3 are almost identical to model 2, the performance of the diversity estimation is better, with the model's Adjusted R-Squared score being .52 (compared with .43 for the model using topic modeling for diversity estimation). Second, both tone and volume were not found to be significant in model 3 (to rule out multicollinearity, the variation inflation factor was calculated for each model, with results being well below the accepted threshold).

As both diversity estimations focus on different aspects of the thematic diversity concept, and with each having their own drawbacks and advantages, I wanted to explore whether introducing both estimations into the model at the same time can improve the model's predictive capacity. Model 4 presents the results when diversity scores are estimated using both semantic network analysis and topic modeling. As can be seen in model 4, both thematic diversity estimations were found to be significant and negative predictors of electoral success, even when included together in the model. Moreover, including both measures in the model improved its predictive capacity relative to the models using only a single diversity estimation method, though only very moderately (with an Adjusted R-Squared of .58, compared with .52 and 0.43 scores for the semantic

network analysis and topic modeling estimations, respectively). The implications of this finding will be discussed in the next chapter.

To fully assess these results, I introduced additional control variables, as detailed in the previous chapter. As model 5 shows, the control variables were found to perform as expected, with midterm elections, conservative ratings, funding, and past political experience being significant predictors of electoral success. The model's Adjusted R-Squared score was found to be very substantial (.7), indicating that these non-media variables offer a relatively comprehensive explanation for candidates' electoral success.

When including all controls and media factors in the model, thematic diversity estimated using topic modeling was again found to be significantly and negatively related to electoral success (with a level of  $p < .05$  for candidates' own diversity and  $p = .001$  for candidates' opponents' diversity). Similar results were found when using semantic network analysis to estimate diversity, as well as when introducing both variables in tandem.

However, although significant, when compared with other predictors in the model, thematic diversity was found to have a relatively modest relationship with electoral success—smaller than state ideology or a candidate's funding, and roughly similar in magnitude to political experience (such as having served as a congressman or governor). Moreover, and in contrast with the models presented in Table 7, improvement to the predictive capacity of the models was modest, with an Adjusted R-Squared score of .74 for the models using only one diversity measure, compared with a score of .7 for the model containing the non-media variables only.

Table 8: Summary of multiple regression analysis of news media factors and non-media factors on Republicans vote share

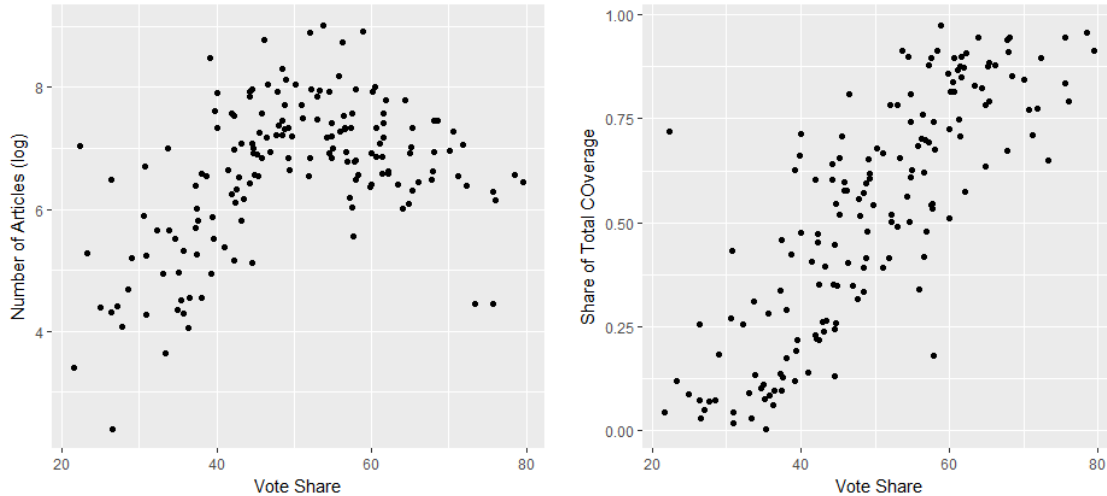
Variable	model 5 (RSE)		model 6 (RSE)		model 7 (RSE)		model 8 (RSE)	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	.00	<.001***	.00	<.001***	.00	.067•	.00	.131
Midterm (1=yes)	.19	<.001***	.21	<.001***	.22	<.001***	.23	<.001***
Conservative Rating	.22	.018*	.25	.001***	.21	.025*	.22	.005**
Funding (Rep)	.52	<.001***	.46	<.001***	.59	<.001***	.54	<.001***
Funding (Dem)	-.51	<.001***	-.46	<.001***	-.39	.001***	-.35	.001**
Funding^2 (Rep)	-.40	.005**	-.31	.002**	-.41	.001***	-.37	<.001***
Funding^2 (Dem)	.30	.005**	.28	.002**	.24	.006**	.21	.005**
Senator (Rep)	.29	<.001***	.27	.001**	.27	.001**	.24	.004**
Governor (Rep)	.12	.086•	.10	.134	.14	.058•	.11	.107
Congressman (Rep)	.16	.001***	.13	.009**	.14	.006**	.12	.011*
Other Exp. (Rep)	.06	.252	.05	.302	.08	.145	.07	.201
Senator (Dem)	-.24	.008**	-.13	.099•	-.13	.173	-.09	.283
Governor (Dem)	-.12	.021*	-.08	.102	-.07	.177	-.06	.256
Congressman (Dem)	-.16	.009**	-.09	.099•	-.08	.218	-.06	.289
Other Exp. (Dem)	-.13	.054•	-.06	.331	-.07	.346	-.05	.435
# Articles (Rep)			-.11	.052•	-.08	.11	-.07	.141
# Articles (Dem)			.03	.342	.06	.047*	.08	.004**
Doc length (Rep)			.00	.984	-.01	.782	-.01	.807
Doc length (Dem)			.05	.236	.08	.018*	.07	.051•
Tone (Rep)			-.05	.106	-.06	.083•	-.04	.205
Tone (Dem)			-.04	.392	-.01	.847	-.03	.509
TM Diversity (Rep)			-.14	.046*			-.10	.134
TM Diversity (Dem)			.18	.001***			.17	.001**
SNA Diversity (Rep)					-.10	.151	-.05	.465
SNA Diversity (Dem)					.29	<.001***	.25	<.001***
Adj R^2		.7		.74		.74		.76

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=165).

As can be seen in Model 8, when both thematic diversity estimations were used in tandem, they were not only found to be significant and negative predictors of electoral success; they also improved the predictive capacity of the models over those using only a single diversity estimation method, although very moderately, with an Adjusted R-Squared score of .76, compared with a .74 score for the models containing only one estimation method.

Finally, when all controls were incorporated into the model, volume was found to have a significant relationship with electoral success (at the  $p < .05$  level). However, surprisingly, this relationship was found to be in the opposite direction than initially expected. This unexpected finding prompted additional exploration, as it seems at odds with some of the existing literature.

Figure 32 shows the relationship between a candidate's news coverage and their electoral success, using two slightly different measures. In both scatter plots, the x-axis represents the share of votes gathered by each republican candidate in the sample. For the plot on the left, the y-axis represents the number of articles covering the candidate. As can be seen, the relationship is non-linear. Generally, candidates that won a higher share of votes received more coverage in the news media, while candidates who performed poorly received less coverage (I use the logged variable for the y-axis for a clearer visual presentation, although the relationship is identical for the non-logged variable as well). However, as can also be seen in this plot, higher coverage was also awarded to candidates that won roughly 50% of the votes. A likely explanation for this is that more competitive races—indicated here as races in which candidates won a near-equal share of votes—often receive more attention from the media. This is similar to the impact of funding, which has a positive relationship with electoral success, but tends to get higher in more competitive races.



*Figure 32: The correlation between the vote share won by a republican candidate and the candidate's absolute (left) and relative (right) volume of news coverage.*

In order to examine this argument, the figure on the left presents the relationship between vote share and relative volume of coverage. For the y-axis, volume was defined as the share of coverage the Republican candidate received out of the total coverage received by both the candidate and their opponent. As can be seen clearly in this plot, the relationship was found to be linear, meaning that when additional explanatory variables are not controlled for, the relative volume correlates positively and strongly with electoral success ( $r=.8$ ,  $p<0.001$ ).

I then re-examined the relationship between volume of coverage and electoral success in a larger context, repeating the media-only model as well as the model in which non-media factors were included as controls.

The results of these models, presented in Table 9, paint a similar picture to the previous findings. The control variables, as well as the diversity estimations, show similar coefficients and significance levels to the earlier models. However, some changes

should be noted in the relationships between the relative volume of coverage and electoral success. First, in the model containing media-only factors, relative volume was found to be much more impactful. When using relative volume instead of absolute volume of coverage, volume was found to be a significant and positive predictor of electoral success.

Table 9: Summary of multiple regression analysis of news media factors and non-media factors on Republicans vote share with relative news coverage as volume indicator

Variable	Model 9		Model 10	
	Beta	p	Beta	p
Constant	.00	.004**	13.51	.195
Midterm (1=yes)			5.92	<.001***
Conservative Rating			.06	.06.
Funding (Rep)			.00	.009**
Funding (Dem)			.00	.003**
Funding^2 (Rep)			.00	.059•
Funding^2 (Dem)			.00	.031*
Senator (Rep)			3.33	.133
Governor (Rep)			6.40	.123
Congressman (Rep)			4.10	.025*
Other Exp. (Rep)			1.41	.426
Senator (Dem)			-.78	.748
Governor (Dem)			-1.75	.554
Congressman (Dem)			-2.51	.203
Other Exp. (Dem)			-2.19	.227
<b>Relative # of Articles (Rep)</b>	.59	<.001***	13.64	.091•
Doc length (Rep)	.05	.274	.00	.747
Doc length (Dem)	.10	.037*	.00	.026*
Tone (Rep)	.01	.899	-6.49	.529
Tone (Dem)	-.03	.501	-9.40	.486
TM Diversity (Rep)	-.04	.448	.00	.641
TM Diversity (Dem)	.11	.026*	.03	.009**
SNA Diversity (Rep)	-.11	.102	-8.14	.762
SNA Diversity (Dem)	.20	.003**	67.60	.003**
Adj R <sup>2</sup>		0.69		0.77

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=165).

In addition, the predictive capacity of the model is enhanced by this change from an adjusted R-square score of 0.58 to an adjusted R-score of 0.69. However, when considering the full model, relative volume was not found to be a significant predictor of electoral success (though it was significant at the nontraditional threshold of  $p < 0.1$ ). Finally, unlike the inclusion of absolute volume of coverage, when relative volume was used, the relationship between volume and electoral success appears to be positive, as expected, and diminishes when non-media controls are included, as was previously argued by Belanger and Soroka (2012).

Finally, to further validate the results of the diversity estimation using semantic network analysis, and to offer a more accurate measurement of diversity (as discussed in Section 4.4.2.3), Table 10 presents the models in which diversity was estimated using the semantic network analysis approach, but utilizing random generated networks as benchmarks for diversity in each corpus. The predictors in these models measure not the diversity observed directly for each corpus, but the extent to which the diversity estimated for the specific network was different than the expected diversity for a random network with identical network features (such as density, number of nodes, and edge strength sequence). These indicators are referred to as RSNA (Random Semantic Network Analysis) in Table 10.

As can be seen from Table 10, even after this costly and lengthy procedure (in terms of computing power) conducted under a very stringent test for diversity estimation, the relationship between diversity measured with semantic network analysis and electoral success remains significant. However, the improvement to the model's performance is minor, with an adjusted R-squared score of .71 for the full model (including the non-

media factors), compared with .7 for the model containing only the non-media factors, and in contrast with a score of .74 for the model using topic modeling to estimate thematic diversity.

Table 10: Summary of multiple regression analysis of news factors and non-media factors on Republicans vote share

Variable	model 11 (RSE)		model 12 (RSE)	
	Beta	p	Beta	p
Constant	0.00	0***	0.00	0***
Midterm (1=yes)			0.20	0***
Conservative Rating			0.26	0.007**
Funding (Rep)			0.60	0***
Funding (Dem)			-0.50	0***
Funding^2 (Rep)			-0.42	0.002**
Funding^2 (Dem)			0.31	0.005**
Senator (Rep)			0.30	0.001***
Governor (Rep)			0.14	0.062•
Congressman (Rep)			0.16	0.002**
Other Exp. (Rep)			0.08	0.174
Senator (Dem)			-0.18	0.06.
Governor (Dem)			-0.10	0.067•
Congressman (Dem)			-0.11	0.127
Other Exp. (Dem)			-0.08	0.287
# Articles (Rep)	0.25	0.001***	-0.09	0.086•
# Articles (Dem)	-0.34	0.001***	0.02	0.764
Doc length (Rep)	0.14	0.041*	-0.01	0.839
Doc length (Dem)	0.04	0.746	0.06	0.097•
Tone (Rep)	0.07	0.292	-0.06	0.062•
Tone (Dem)	-0.11	0.156	-0.01	0.8
RSNA Diversity (Rep)	-0.19	0.019*	-0.01	0.831
RSNA Diversity (Dem)	0.28	0.001***	0.11	0.035*
Adj R^2	.3		.71	

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=165).

In summary, while tone and volume were initially shown to be significant predictors of electoral success, when additional controls and the semantic network diversity estimation were included in the model, the basic media factors were generally not found to be significant at the p<0.05 level. However, in some models where a wide



set of non-media controls were included, volume was found to have a significant relationship with electoral success, but in the opposite direction than hypothesized. As I argue in the next chapter, while this result is at odds with some existing findings, it can be explained, in part, by looking to more recent research on the relationship between volume and tone of news coverage and electoral success and the indirect nature of this effect (Bélanger & Soroka, 2012), as well as by the impact that race competitiveness has on the volume of coverage that the election receives by the news media. Indeed, including relative volume in the model instead of absolute volume shows a more positive relationship between volume and electoral success.

Thematic diversity—whether measured using topic modeling, semantic network analysis, or both, and whether used in tandem with relative or absolute volume—was found to be a strong negative predictor of electoral success ( $p < .01$ ). Even after controlling for a wide array of non-media factors, and even when used in tandem, these measures still offer a very strong predictive model for electoral success. Following this analysis, I now turn to offer a similar analysis for the relationship between candidates' social media activity and their electoral success.

## 5.2 Social Media Activity and Electoral Success

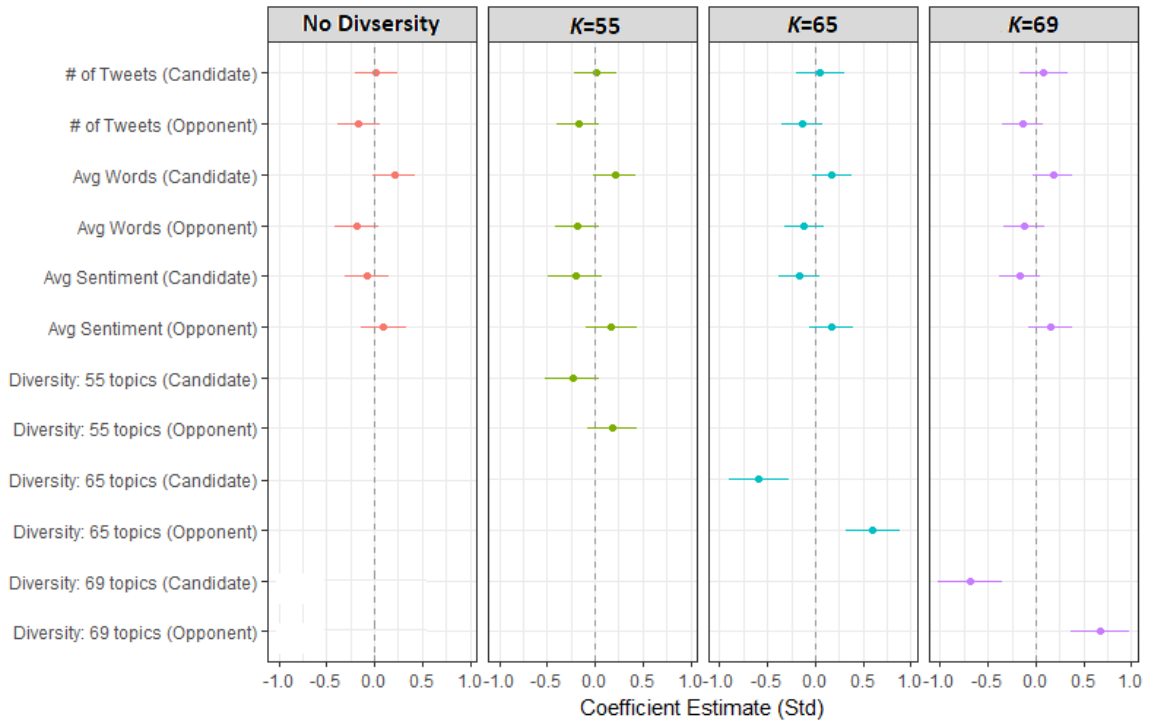


Figure 33: Summary of linear regression analysis for social media variables predicting Republican candidates' electoral success, with different ( $k$ ) topic model per facet ( $N=77$ ).

Figure 33 presents the results of the linear regression models in which the dependent variable is the share of votes gathered by the Republican candidates. It directly compares the performance of the diversity estimation across the various topic models (from 55 topics to 69 topics). First, as can be seen from the first two rows, the volume of candidates' social media activity has a somewhat positive (though non-significant by common thresholds) relationship with electoral success. Similarly, tone of social media activity had a negative relationship with electoral success, although also not significant by common thresholds.

Most importantly, and as with the relationship of news media coverage with electoral success, diversity was found to be negatively related to election outcomes, with



As with topic modeling for the news media corpus, the themes that emerge from the data touch on many topics and contexts. For example, in the example given in the top-right corner of Figure 34, the theme captures procedural campaign issues, including instructions on how to vote and get-out-to-vote appeals. The theme on the left of this topic, at the center of the top row, includes words that relate to the healthcare debate—a theme similar to one that emerged in the news corpus. Other themes that can be observed in this figure include (moving from top to bottom and from left to right): media appearances, education, veterans, energy, legislation, women’s issues, and immigration. However, more important to the purposes of this study, the diversity estimations are used in the following regression models (Tables 9 and 10) labeled as TM diversity.

Figures 35 and 36 present examples of semantic networks drawn from the set used to estimate thematic diversity. These figures present two networks, the first exhibiting a high diversity score and the second a low score. As the semantic networks used for the social media activity analysis contain only 100 nodes, a more detailed presentation of these networks is possible.

Figure 35 presents the semantic network of candidate Bernie Sanders’ social media activity during his 2012 campaign for the U.S. Senate. The nodes represent the top 100 most prominent words in the corpus (using TFIDF as a measure). The edges represent the cosine similarity between the words (calculated using cosine similarity as an edge weight). Lastly, the colors indicate community membership for each node. Among the various communities are sets of words related to agriculture (in dark green at the bottom-right corner of the network), the Supreme Court (in orange), social security (light green), energy and climate change (light green), and income inequality (in red). This network







Table 11 presents the results for the regression models estimating the relationship between social media activity factors and candidates' electoral success. Model 1 presents the results of the regression models containing only the basic corpus features—volume, average length, and tone. As the results show, model 1 performed quite poorly, with an adjusted R-squared score of .04 and an average tweet length only being close to significance (at the  $p < .1$  level), indicating a weak positive relationship between tweet length and electoral success. This poor performance for the media predictors persisted when adding thematic diversity as a predictor to the models.

Table 11: Summary of multiple regression analysis of social media factors on Republicans vote share

Variable	model 1		model 2		model 3 (RSE)		model 4	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	.00	.022*	.00	.004**	.00	.333	.00	.967
# Tweets (Rep)	-.02	.851	.05	.72	.02	.837	.12	.399
# Tweets (Dem)	-.15	.207	-.12	.251	.03	.735	-.01	.924
Tweet Length (Rep)	.20	.099•	.17	.109	.30	.02*	.22	.073•
Tweet Length (Dem)	-.23	.074•	-.17	.123	-.09	.496	-.09	.428
Tone (Rep)	.02	.884	-.06	.561	.04	.773	-.04	.696
Tone (Dem)	-.04	.743	.02	.839	.01	.928	.04	.752
TM Diversity (Rep)			-.80	<.001***			-.77	<.001***
TM Diversity (Dem)			.80	<.001***			.71	<.001***
SNA Diversity (Rep)					.09	.587	.11	.373
SNA Diversity (Dem)					.36	.049*	.18	.227
Adj R <sup>2</sup>	.04		.31		.1		.32	

Note: • $p < .1$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$  (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=71).

By contrast, as can be seen in model 2, diversity estimated using topic modeling was found to have a very strong and negative relationship with electoral success. Lower thematic diversity in candidate social media activity was related to higher electoral success; similarly, higher thematic diversity for a candidate's opponent was a predictor of higher electoral success. In addition, the model's predictions improve greatly from an

Adjusted R-Squared score of .04 for the corpus features only, to an Adjusted R-Squared score of 0.31 for the model when thematic diversity is estimated as a predictor via topic modeling.

Model 3 is similar to model 2, but uses semantic network analysis as a basis for the diversity estimation instead of topic modeling. The thematic diversity estimations calculated with semantic network analysis were found to perform significantly worse than those estimated with topic modeling—only the diversity of candidates’ opponents was found to be significant ( $p < .05$ ). Further, the two measures were not found to be significantly correlated. Finally, incorporating both estimations into the model showed almost no improvement over model 2, in which only the topic modeling estimation was used. This might indicate a lack of reliability for this measure when analyzing smaller corpora with fewer words, as these result in fewer nodes and connections. I will discuss this issue further in the next chapter.

To fully assess these results, I introduced additional control variables, as detailed in the methods section in Chapter 4. As model 5 shows, the control variables were found to perform as expected, with midterm elections, conservative rating, funding, and past political experience being significant predictors of electoral success. The model’s Adjusted R-Squared score was again found to be very substantial (.7), indicating that non-media variables offer relatively comprehensive explanations for candidates’ electoral success.

When including all controls and media factors in the model, thematic diversity estimated using topic modeling was again found to be significantly and negatively related to electoral success (with a level of  $p < .05$  for candidates’ own diversity and  $p = .01$  for



candidates' opponents' diversity). Similar to the results for models 2 and 4, however, diversity estimations calculated with semantic network analysis were found to be non-significant by traditional thresholds (the approaching significance,  $p=.066$ ).

Table 12: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share

Variable	model 5		model 6		model 7		model 8	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	.00	<.001***	.00	.001***	.00	.934	.00	.506
Midterm (1=yes)	.16	.033*	.18	.023*	.16	.061•	.16	.041*
Conservative Rating	.31	.014*	.31	.013*	.27	.039*	.30	.017*
Funding (Rep)	.72	.017*	.69	.022*	.79	.013*	.72	.019*
Funding (Dem)	-.75	.002**	-.69	.004**	-.73	.005**	-.63	.011*
Funding^2 (Rep)	-.58	.05•	-.62	.036*	-.64	.038*	-.64	.031*
Funding^2 (Dem)	.52	.019*	.53	.016*	.51	.03*	.50	.028*
Senator (Rep)	.15	.24	.08	.509	.15	.269	.07	.611
Governor (Rep)	-.03	.72	-.01	.863	-.05	.597	-.03	.757
Congressman (Rep)	.09	.307	.08	.363	.06	.528	.07	.462
Other Exp. (Rep)	.03	.757	.00	1	-.01	.927	.00	.997
Senator (Dem)	-.14	.351	-.09	.517	-.15	.31	-.13	.385
Governor (Dem)	.03	.72	-.08	.452	-.01	.891	-.07	.543
Congressman (Dem)	-.12	.312	-.10	.363	-.15	.227	-.12	.301
Other Exp. (Dem)	.01	.931	.06	.46	.02	.854	.06	.488
# Articles (Rep)			-.02	.798	-.02	.86	.01	.962
# Articles (Dem)			.08	.293	.21	.029*	.14	.135
Doc length (Rep)			.03	.638	.08	.318	.07	.365
Doc length (Dem)			-.05	.507	.00	.993	-.01	.86
Tone (Rep)			.04	.59	.10	.245	.05	.569
Tone (Dem)			-.04	.644	-.07	.434	-.03	.727
TM Diversity (Rep)			-.29	.031*			-.29	.037*
TM Diversity (Dem)			.45	.002**			.39	.008**
SNA Diversity (Rep)					.03	.689	.03	.698
SNA Diversity (Dem)					.19	.066•	.12	.237
Adj R^2	.7		.75		.71		.74	

Note: • $p < .1$ . \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$  (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=71).

Improvements to the models' predictive performance were modest for topic modeling-based diversity estimation, with an Adjusted R-Squared score of .75 for models

using only one diversity measure compared with a score of .7 for the model containing the non-media variables only. The model with thematic diversity estimated via semantic network analysis did not exhibit such improvement, however, both when entered alone and when entered along with the topic modeling-based estimation.

Finally, to further validate the results of diversity estimations using semantic network analysis, and to offer a more accurate measurement of diversity (as discussed in Section 4.4.2.3), Table 13 presents the models in which diversity was estimated using the semantic network analysis approach but utilizing randomly generated networks as benchmarks for diversity in each corpus. As can be seen in models 9 and 10, this procedure improved the performance of the diversity predictors. This likely stems from the fact that these models' predictors measured not the observed diversity in each corpus, but the extent to the diversity estimation differed from the expected diversity for a random network with identical general network features (such as density, number of nodes, and edge strength sequence). Even after this costly and lengthy procedure (in terms of computing power), improvements to the model performance seem to be only minor, with an adjusted R-squared score of .72 for the full model (including the non-media factors) compared with .7 for the model containing only the non-media factors (without diversity estimates) and .75 for the model using topic modeling only to estimate thematic diversity.

All in all, when considering political candidates' social media activity, this study finds almost no evidence for a relationship between either tone or volume of candidates' social media activity and their electoral success. However, in regard to thematic diversity, candidates who offered more monothematic message strategies were found to perform

significantly better across all models than candidates who exhibited a more multi-themed messaging strategy. The results for thematic diversity in candidates' social media activity are therefore similar to those for news media coverage.

Table 13: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share (using random network benchmark for diversity estimation)

Variable	model 9 (RSE)		model 10 (RSE)	
	Beta	p	Beta	p
Constant	.00	.289	.00	.043*
Midterm (1=yes)			.18	.033*
Conservative Rating			.27	.038*
Funding (Rep)			.77	.014*
Funding (Dem)			-.78	.002***
Funding^2 (Rep)			-.63	.039*
Funding^2 (Dem)			.55	.02*
Senator (Rep)			.17	.191
Governor (Rep)			-.04	.606
Congressman (Rep)			.06	.505
Other Exp. (Rep)			-.03	.754
Senator (Dem)			-.11	.483
Governor (Dem)			.00	1
Congressman (Dem)			-.11	.384
Other Exp. (Dem)			.03	.712
# Articles (Rep)	.05	.677	-.01	.957
# Articles (Dem)	-.01	.926	.20	.026*
Doc length (Rep)	.30	.013*	.09	.289
Doc length (Dem)	-.14	.262	-.02	.807
Tone (Rep)	.03	.798	.10	.214
Tone (Dem)	-.03	.805	-.10	.225
RSNA Diversity (Rep)	.11	.491	.05	.578
RSNA Diversity (Dem)	.33	.033*	.19	.03*
Adj R^2		.10		.72

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=71).

However, from a methodological perspective, it seems that in the case of social media activity, analyzing diversity through semantic networks was less adequate than the topic modeling approach. The lack of reliability persisted even when using a more advanced method for comparing multiple networks, with the randomly generated models

serving as a benchmark. This might result from the more limited text available on candidates' Twitter feeds for semantic network construction, which limits the method's validity and reliability. With fewer texts relative to the news media corpus, and with the text length being significantly shorter, it seems that the approach requires significant changes before it is adequate. I discuss such possible changes, as well as additional methodological and theoretical implications of these results, in the next chapter.

### **5.3 The Antecedents of Thematic Diversity**

While thematic diversity in both news coverage and social media activity was found to have a strong relationship with electoral success, it is clear from the results that this relationship is not completely independent. This is evidenced, for example, by disparities in the predictive power of the different models. While the media factors alone—and chief among them, thematic diversity—can explain a large portion of the variability in electoral success, when inserted into models controlling for non-media factors, their relationship with electoral success diminishes.

It thus seems that the media variables are related, to some extent, to the non-media factors. Table 14 presents the results of four models in which thematic diversity was used as a dependent variable and the non-media factors included in previous models here serve as predictors. These variables included the type of election (midterm vs. quadrennial), how conservative the state is, the funding that the candidate received, and whether the candidate was an incumbent or a challenger. In addition, as seen earlier for the relationship between the volume of coverage and electoral success, the competitiveness of a race can also influence media coverage greatly. Therefore, a variable was added to represent the vote gap between the candidates, calculated as the

absolute value of the difference in votes between the Republican and Democratic candidates. Higher values represent less competitive races (where the gap was larger) and lower values represent more competitive races (as indicated by a smaller gap between the candidates).

Table 14: Summary of multiple regression analysis of non-media factors on thematic diversity in candidates' news coverage.

Variable	SNA Diversity (R)		SNA Diversity (D)		TM Diversity (R)		TM Diversity (D)	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	.00	<.001***	.00	<.001***	.00	<.001***	.00	<.001***
Midterm (1=yes)	-.12	.042*	-.08	.131	-.16	.017*	-.15	.009**
Vote Share Gap	.31	<.001***	.36	<.001***	.21	.027*	.19	.206
Conservative Rating	-.29	<.001***	.20	.004**	-.21	.004**	.07	.605
Funding	-.34	<.001***	-.26	.001**	-.18	.003**	-.16	.01*
Incumbency (1=yes)	-.04	.562	-.16	.022*	-.19	.001***	-.21	.064•
Adj R <sup>2</sup>	.45		.48		.25		.20	

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). All models computed with Robust Standard Errors (N=165). SNA indicates semantic network analysis based estimation. TM indicates topic modeling based estimation. (R) and (D) indicate the Republican and Democratic candidates, respectively.

As the results of Table 14 show, several non-media factors were found to influence thematic diversity in candidate's news coverage, with a large portion of the variability in thematic diversity explained by these factors—although the models' predictions were stronger for thematic diversity estimated using semantic network analysis than using topic modeling. Midterm elections were found to have a negative impact on thematic diversity. In addition, more competitive races seemed to reduce the thematic diversity of political candidates' news coverage to a great extent. However, this impact was found to be more prominent when measuring diversity using semantic network analysis and less so when using topic modeling. The conservative rating of the state in which the race took place had a strong effect on diversity as well. For Republican

candidates, the more conservative the state was, the less diverse their coverage was in the news media. For Democratic candidates, an opposite relationship was found, with more diverse news coverage in more conservative states. Using both methods, and for both Republican and Democratic candidates, funding was found to have a negative relationship with thematic diversity, with candidates that received more contributions having less thematically diverse news coverage. Finally, and in a similar vein, incumbents were found to have a less thematically diverse news coverage. I elaborate on the possible implications of these findings in Chapter 6 (Summary and Discussion).

Table 15: Summary of multiple regression analysis of non-media factors on thematic diversity in candidates' social media activity

Variable	SNA Diversity (R)		SNA Diversity (D)		TM Diversity (R)		TM Diversity (D)	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	.00	<.001***	.00	<.001***	.00	<.001***	.00	.002**
Midterm (1=yes)	-.09	.448	.11	.377	-.12	.073•	-.17	.009**
Vote Share Gap	.40	.006**	.28	.062•	.10	.617	-.02	.866
Conservative Rating	.16	.333	.07	.622	-.32	.164	-.04	.776
Funding	.05	.735	-.23	.076•	.35	.208	.11	.371
Incumbency (1=yes)	-.06	.713	.03	.816	.10	.551	-.23	.159
Adj R <sup>2</sup>	.13		.18		.11		.03	

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). All models computed with Robust Standard Errors (N=71). SNA indicates semantic network analysis based estimation. TM indicates topic modeling based estimation. (R) and (D) indicate the Republican and Democratic candidates, respectively.

Table 15 presents the results of a similar analysis for thematic diversity in candidates' social media activity. Relative to the news media coverage, the non-media factors explained much less of the variability in thematic diversity. When using topic modeling to estimate diversity, midterm elections were found to negatively influence the thematic diversity of candidates' social media activity. As with the news media corpus, candidates in midterm elections (2014) tended to focus on a smaller set of themes than in

quadrennial elections. When using semantic network analysis to estimate diversity, the race's competitiveness was found to be positively related to thematic diversity—again, similar to the results for the news media coverage. However, it should be noted that the performance of the semantic network estimations was sub-optimal in the context of social media activity when compared to topic modeling.

In the following chapter, I summarize all of the findings of this study and offer some broader conclusions. I address the role of tone, volume, and thematic diversity in candidates' news coverage and their direct communication with voters via social media, as well as the non-media factors that shape them. I also elaborate on this study's limitations and possible future directions. I address the measurement of diversity using unsupervised machine learning methods, elaborating on their advantages and drawbacks, and refer to the “big-picture”—the extent to which media predictors serve as dependent or independent factors in the relationship between news coverage, social media activity, and candidates' political success. I end with final thoughts on how this study's unique findings for monothematic and multi-thematic message strategies fit within current theory on thematic diversity in election campaigns.

#### **5.4 Correlation and Multicollinearity Analysis**

Similarity across the coefficients for the social media thematic diversity predictors, as well as correlations between the independent variables, raise some concerns regarding the relationship between the thematic diversity of candidates' rhetoric and their opponents'. These concerns persist despite acceptable variation inflation factors found for these models. In this section, I elaborate on these concerns and offer several possible solutions to validate the results of the regression models' findings. First, to better

understand these concerns, I explored the correlation matrix for the variables predicting electoral success in the regression models using both the non-media and news media factors. Table 16 below show these correlations.

Table 16: A correlation matrix for the non-media and news media factors used in the regression models predicting electoral success.

	1. Vote Share (Rep)	2. Conservative Rating	3. Midterm (=yes)	4. Funding (Rep)	5. No Exp. (Rep)	6. Congressman (Rep)	7. Governor (Rep)	8. Other Exp. (Rep)	9. Senator (Rep)	10. # Articles (Rep)	11. Doc length (Rep)	12. Tone (Rep)	13. Funding (Dem)	14. No Exp. (Dem)	15. Congressman (Dem)	16. Governor (Dem)	17. Other Exp. (Dem)	18. Senator (Dem)	19. # Articles (Dem)	20. Doc length (Dem)	21. Tone (Dem)	22. TM Diversity (Rep)	23. SNA Diversity (Rep)	24. RSNA Diversity (Rep)	25. TM Diversity (Dem)	26. SNA Diversity (Dem)	27. RSNA Diversity (Dem)
1	1																										
2	.71	1																									
3	.18	-.01	1																								
4	.2	.2	.03	1																							
5	-.47	-.45	.13	-.3	1																						
6	.01	-.09	.1	.13	-.22	1																					
7	.02	-.02	-.1	-.06	-.1	-.08	1																				
8	-.28	-.21	0	-.08	-.25	-.2	-.09	1																			
9	.59	.61	-.15	.24	-.45	-.36	-.16	-.39	1																		
10	.19	.3	.08	.59	-.27	-.04	-.07	-.1	.36	1																	
11	.24	.22	.17	.04	-.14	.01	-.13	-.12	.26	.06	1																
12	.05	.2	-.11	.04	-.07	-.05	.06	-.05	.12	-.03	.1	1															
13	-.43	-.41	-.03	.49	.11	.13	-.03	.09	-.25	.33	-.13	-.17	1														
14	.54	.53	.01	.03	-.21	-.14	-.02	-.21	.45	0	.17	.02	-.31	1													
15	-.01	-.03	-.06	.13	-.16	.04	.22	.01	.01	.19	-.07	-.06	.04	-.2	1												
16	-.08	-.03	-.13	.07	-.05	-.1	.11	-.03	.1	.02	-.03	-.01	.09	-.12	-.09	1											
17	.25	.23	.06	.15	-.18	.1	-.03	-.14	.2	.18	.02	.17	-.18	-.29	-.21	-.12	1										
18	-.64	-.61	.04	-.27	.47	.05	-.15	.3	-.61	-.29	-.1	-.12	.35	-.41	-.3	-.18	-.42	1									
19	-.37	-.33	.1	.38	.16	.09	-.05	.18	-.32	.23	-.09	-.04	.57	-.28	-.01	.05	-.2	.4	1								
20	-.05	-.11	-.07	-.07	.11	.02	-.1	-.01	-.06	-.12	-.02	-.12	.03	.02	.03	-.1	-.22	.2	.02	1							
21	-.14	-.14	.06	-.06	.04	-.09	.01	.12	-.06	-.09	.11	.33	-.06	-.11	0	.05	-.03	.09	0	0	1						
22	-.44	-.34	-.14	-.38	.28	-.05	.06	.12	-.31	-.35	.05	.14	-.05	-.19	-.09	-.07	-.17	.4	.08	.03	.18	1					
23	-.32	-.35	-.11	-.58	.36	-.23	.03	.1	-.22	-.55	-.13	-.03	-.19	-.06	-.16	-.1	-.23	.4	-.19	.2	.11	.41	1				
24	-.09	-.15	.02	-.38	.25	-.27	-.07	.05	-.03	-.33	-.05	-.11	-.18	.04	-.14	-.02	-.13	.19	-.23	.18	.03	.02	.82	1			
25	.43	.29	-.16	-.09	-.24	-.05	.14	-.15	.3	-.07	.19	.13	-.36	.31	-.02	-.04	.1	-.33	-.3	-.06	.11	.03	-.05	-.08	1		
26	.52	.43	-.11	-.29	-.14	-.19	-.05	-.21	.44	-.23	.16	.05	-.56	.47	-.18	-.08	.14	-.36	-.61	-.05	-.08	-.12	.29	.33	.36	1	
27	.32	.23	0	-.22	.01	-.19	-.18	-.15	.32	-.21	.16	-.08	-.31	.36	-.25	-.01	.04	-.16	-.38	-.03	-.14	-.14	.3	.39	.05	.87	1



To better focus on the main variables of interest, Table 17 shows the correlation matrix for a smaller set of these predictors, including vote share, and the various thematic diversity estimations.

Table 17: A correlation matrix for the news media thematic diversity factors used to predict electoral success.

	1. Vote Share (Rep)	22. TM Diversity (Rep)	23. SNA Diversity (Rep)	24. RSNA Diversity (Rep)	25. TM Diversity (Dem)	26. SNA Diversity (Dem)	27. RSNA Diversity (Dem)
1. Vote Share (Rep)	1						
22. TM Diversity (Rep)	-.44	1					
23. SNA Diversity (Rep)	-.32	.41	1				
24. RSNA Diversity (Rep)	-.09	.02	.82	1			
25. TM Diversity (Dem)	.43	.03	-.05	-.08	1		
26. SNA Diversity (Dem)	.52	-.12	.29	.33	.36	1	
27. RSNA Diversity (Dem)	.32	-.14	.3	.39	.05	.87	1

Highlighted in orange are the between-party correlations. As Table 17 illustrates, these correlations are, at best, weak to moderate. This finding is supported by the low variation inflation factors for the various models, including both parties’ predictors in tandem. Further, the correlation between the diversity scores and vote share are similar to their performance in the regression models, with a negative relationship between candidates’ vote share and their level of thematic diversity, and a positive relationship between candidates’ vote share and their opponents’ diversity.

A similar analysis was performed for the social media data as well, although the results differ. Table 18 shows the full correlation matrix for the non-media and social-media factors, and Table 19 a collated matrix to highlight the most important factors:

Table 18: A correlation matrix for the non-media and social-media factors used in the regression models predicting electoral success.

	1. Vote Share (Rep)	2. Conservative Rating	3. Midterm (1=yes)	4. Funding (Rep)	5. No Exp. (Rep)	6. Congressman (Rep)	7. Governor (Rep)	8. Other Exp. (Rep)	9. Senator (Rep)	10. # Tweets (Rep)	11. Doc length (Rep)	12. Tone (Rep)	13. Funding (Dem)	14. No Exp. (Dem)	15. Congressman (Dem)	16. Governor (Dem)	17. Other Exp. (Dem)	18. Senator (Dem)	19. # Tweets (Dem)	20. Doc length (Dem)	21. Tone (Dem)	22. RSNA Diversity (Rep)	23. SNA Diversity (Rep)	24. TM Diversity (Rep)	25. RSNA Diversity (Dem)	26. SNA Diversity (Dem)	27. TM Diversity (Dem)		
1	1																												
2	.76	1																											
3	.15	.06	1																										
4	.16	.18	-.08	1																									
5	-.4	-.38	.15	-.25	1																								
6	.05	-.09	.09	.06	-.21	1																							
7	-.13	-.02	-.14	0	-.09	-.07	1																						
8	-.27	-.24	.07	-.16	-.25	-.19	-.08	1																					
9	.55	.57	-.19	.29	-.46	-.36	-.15	-.42	1																				
10	-.06	-.03	-.11	.33	-.06	-.08	.27	-.26	.22	1																			
11	.2	.22	-.16	.11	-.16	.01	-.07	-.03	.18	.01	1																		
12	-.02	-.17	-.05	-.1	0	.07	.16	.11	-.19	-.25	-.06	1																	
13	-.42	-.38	-.02	.59	.06	0	.02	.06	-.11	.15	-.05	.01	1																
14	.51	.54	.12	.07	-.11	-.22	-.09	-.26	.48	-.14	.14	-.1	-.25	1															
15	-.19	-.19	-.13	.14	-.09	-.03	.45	.04	-.08	.19	-.07	.09	.11	-.21	1														
16	.01	-.11	-.17	.29	-.11	-.09	-.04	-.1	.24	.39	.18	.03	.17	-.11	-.08	1													
17	.32	.32	-.11	-.03	-.26	.1	-.08	-.05	.21	-.06	0	.11	-.22	-.27	-.19	-.1	1												
18	-.57	-.54	.15	-.25	.41	.16	-.14	.27	-.62	-.12	-.14	-.08	.25	-.45	-.32	-.17	-.41	1											
19	-.19	-.25	.09	.2	.07	.02	.07	.06	-.15	.09	-.12	.03	.43	-.3	.15	.24	-.15	.17	1										
20	-.2	-.17	-.22	.09	.1	-.01	0	.01	-.08	.11	.13	.07	.16	-.19	0	.12	.03	.09	.03	1									
21	-.04	-.12	.33	-.15	.11	.1	-.06	-.1	-.06	-.07	-.26	.3	.02	-.02	-.09	-.05	.05	.06	.13	-.31	1								
22	.17	.19	-.13	-.09	-.14	0	-.06	.16	.01	-.53	.21	.06	-.11	.18	-.04	-.19	.01	-.05	-.07	-.09	-.08	1							
23	.24	.22	-.09	-.14	-.14	.02	-.12	.09	.07	-.41	.18	-.02	-.21	.2	-.08	-.24	.01	-.03	-.13	-.17	-.05	.83	1						
24	-.18	-.18	-.19	.28	-.04	-.04	.05	-.01	.05	.6	.13	-.21	.24	-.24	-.02	.64	-.09	.04	.26	.15	-.23	-.16	-.15	1					
25	.27	.15	.13	-.15	-.05	.04	-.08	-.02	.06	-.1	-.33	-.03	-.32	.22	-.2	-.14	.06	-.05	-.41	-.3	.09	.02	.12	-.11	1				
26	.31	.21	.11	-.2	-.07	.03	-.02	-.09	.12	-.1	-.27	-.08	-.42	.23	-.13	-.17	.05	-.08	-.51	-.37	-.03	.12	.25	-.15	.91	1			
27	.19	.03	-.23	.37	-.17	-.12	-.05	-.14	.34	.48	.18	-.12	.09	.02	-.07	.72	-.08	-.2	.2	.11	-.27	-.1	-.06	.77	.04	.01	1		

Table 19: A correlation matrix for the social media thematic diversity factors used to predict electoral success.

	1. Vote Share (Rep)	22. RSNA Diversity (Rep)	23. SNA Diversity (Rep)	24. TM Diversity (Rep)	25. RSNA Diversity (Dem)	26. SNA Diversity (Dem)	27. TM Diversity (Dem)
1. Vote Share (Rep)	1						
22. RSNA Diversity (Rep)	.17	1					
23. SNA Diversity (Rep)	.24	.83	1				
24. TM Diversity (Rep)	-.18	-.16	-.15	1			
25. RSNA Diversity (Dem)	.27	.02	.12	-.11	1		
26. SNA Diversity (Dem)	.31	.12	.25	-.15	.91	1	
27. TM Diversity (Dem)	.19	-.1	-.06	.77	.04	.01	1

Looking at the correlations between vote share and thematic diversity in social media activity using topic modeling (highlighted in blue), the results match the relationships found in the previous several regression models (reflecting the broader trend of semantic network models’ performance for social media throughout this dissertation). A negative relationship was found between candidates’ vote share and their own thematic diversity, and a positive relationship between candidates’ vote share and their opponents’ diversity.

Although the Variation Inflation Factor scores for the multiple regression models were within reasonable limits, there was a high correlation between the diversity scores for Republican and Democratic candidates estimated with topic modeling. I offer three different solutions for this issue. First, as the diversity estimations for both candidates and

their opponents was found to be correlated, Table 20 present the results of the regression models estimated independently for the Democratic and for the Republican candidates. As can be seen in these models, thematic diversity was found to be a significant predictor of electoral success for Democratic candidates (in the same negative direction found in previous models). However, thematic diversity was not found to be a significant predictor in the context of Republican candidates.

Table 20: Summary of multiple regression analysis of social media factors and non-media factors on candidate vote share, separated by party

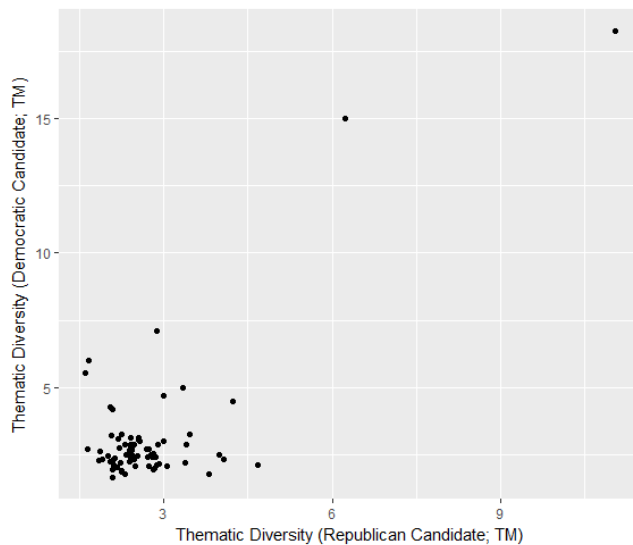
Variable	Model (Rep)		Model (Dem)	
	Beta	p	Beta	p
Constant	.00	.014*	.00	<.001***
Midterm (1=yes)	.16	.049*	-.18	.027*
Conservative Rating	.56	<.001***	-.49	<.001***
Funding	.21	.409	.57	.01*
Funding^2	-.24	.333	-.37	.063•
Senator	.34	.009**	.19	.181
Governor	-.04	.657	.10	.366
Congressman	.17	.095•	.04	.748
Other Exp.	.00	.994	-.05	.577
# Tweets	-.06	.561	-.07	.383
Doc length	.05	.538	-.03	.689
Tone	.12	.131	-.01	.943
TM Diversity	.01	.884	-.28	.01*
Adj R <sup>2</sup>	.63		.68	

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=71).

These results raise two questions. First, is thematic diversity in candidates' social media activity related to the thematic diversity of their opponents? As shown earlier, while non-media factors shape thematic diversity in the context of the news coverage, such explanations were found to be weaker in the context of social media. Therefore, while some moderate correlation was expected for the news media corpus (with both

candidates in a state being influenced by similar non-media variables), the same correlation was expected to be weaker for social media. Second, why is there such a large difference in the relationship between thematic diversity and electoral success for Republican and Democratic candidates?

I therefore examined more closely the correlations between the diversity scores of both candidates.



*Figure 37: Scatter plot of the thematic diversity for the Democratic and Republican candidates*

As can be seen in Figure 37, there are several extreme observations in the data that might be responsible for the high correlation. I therefore transformed the two variables (thematic diversity for Republican and for Democratic candidates) using a natural log to reduce their skewness. Using the natural logarithm of these values, the correlation between the variables was reduced from  $r=.77$  to  $r=.46$ . The regression results remained similar to previous results, or even improved, as can be seen in Table 21.

Table 21: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share (using natural log transformation for the topic modeling diversity estimation)

Variable	model 9 (RSE)		model 10 (RSE)	
	Beta	p	Beta	p
Constant	.00	.143	.00	.881
Midterm (1=yes)			.18	.017*
Conservative Rating			.28	.015*
Funding (Rep)			.66	.02*
Funding (Dem)			-.65	.005**
Funding^2 (Rep)			-.61	.029*
Funding^2 (Dem)			.51	.014*
Senator (Rep)			.03	.81
Governor (Rep)			-.01	.927
Congressman (Rep)			.09	.284
Other Exp. (Rep)			.01	.865
Senator (Dem)			-.11	.4
Governor (Dem)			.00	.974
Congressman (Dem)			-.11	.3
Other Exp. (Dem)			.08	.328
# Tweets (Rep)	.04	.699	.00	.981
# Tweets (Dem)	-.10	.331	.08	.287
Doc length (Rep)	.14	.156	.03	.671
Doc length (Dem)	-.17	.101	-.06	.396
Tone (Rep)	-.08	.423	.00	.95
Tone (Dem)	.04	.733	-.03	.706
TM Diversity (Rep; log)	-.60	<.001***	-.27	.007**
TM Diversity (Dem; log)	.59	<.001***	.37	<.001***
Adj R^2	.40		.77	

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=71).

This result also raises the possibility that the extreme observations are influencing the regression results in an unwanted way (especially in terms of increasing type I errors). Therefore, aside from separating the models along party lines and using a log transformation on the skewed variables, I also revisited the models but with the suspected outliers removed. To do that, I removed observations for any candidates whose

thematic diversity measures were larger than two standard deviations over the mean thematic diversity in the whole sample. This resulted in three candidates being removed in two races (as one race had two candidates with extreme scores). The reduced correlation matrix for this data set can be seen in Table 7 (n=69, in contrast with n=71 in the earlier models).

Table 22: A correlation matrix for the social media thematic diversity factors used to predict electoral success, with outlier observations removed

	1. Vote Share (Rep)	24.RSNA Diversity (Rep)	23.SNA Diversity (Rep)	22.TM Diversity (Rep)	25.RSNA Diversity (Dem)	26.SNA Diversity (Dem)	27.TM Diversity (Dem)
1. Vote Share (Rep)	1						
22.RSNA Diversity (Rep)	.19	1					
23.SNA Diversity (Rep)	.25	.83	1				
24.TM Diversity (Rep)	-.5	-.01	-.04	1			
25.RSNA Diversity (Dem)	.27	.02	.12	-.23	1		
26.SNA Diversity (Dem)	.31	.11	.24	-.21	.91	1	
27.TM Diversity (Dem)	.45	.23	.3	-.08	.16	.22	1

First, as can be seen highlighted in orange in Table 22, no relationship was found between the thematic diversity of candidates' and their opponents' coverage or social media activity after the extreme observations were removed. Moreover, as can be seen in blue, the correlations between diversity and electoral success become stronger, not weaker, after removing these observations. This correlation translates also into the results for the new regression models as well. Table 23 presents the results of the regression models run for the two parties separately. As can be seen from these results, thematic

diversity was again found to be a significant negative predictor of electoral success, for candidates from both parties. The regression models presented in Table 24, in which the factors for both the candidates' and their opponents are used concurrently, show similar results.

Table 23: Summary of multiple regression analysis of social media factors and non-media factors on candidate vote share, separated by party

Variable	Model (Rep)		Model (Dem)	
	Beta	p	Beta	p
Constant	.00	<.001***	.00	<.001***
Midterm (1=yes)	.10	.143	-.23	.005**
Conservative Rating	.55	<.001***	-.44	<.001***
Funding	.29	.184	.64	.004**
Funding^2	-.28	.188	-.41	.033*
Senator	.27	.018*	.14	.322
Governor	.03	.668	-.04	.56
Congressman	.15	.074•	.00	.986
Other Exp.	.00	.966	-.10	.285
# of Tweets	-.17	.044*	-.07	.399
Doc length	-.04	.545	-.01	.889
Tone	.04	.56	-.01	.938
TM Diversity	-.25	.001***	-.27	.001***
Adj R^2	.73		.71	

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=71).

In summary, examining the correlation matrices for the independent variables used across the models, a potentially problematic correlation was detected when topic modeling was used to estimate diversity, between candidates' thematic diversity and their opponents'. I used three different methods to address this issue. First, I ran separate regression models for the Republican and Democratic candidates. Second, I used a log transformation of the thematic diversity scores as independent variables in the models. Third, I removed three candidates (competing in two races) whose thematic diversity values were higher than two standard deviations over the mean thematic diversity for the



whole sample. The three different methods offer support for the original conclusion that there is a significant negative relationship between candidates' thematic diversity on social media and their candidates electoral success.

Table 24: Summary of multiple regression analysis of social media factors and non-media factors on Republicans vote share (using topic modeling for diversity estimation)

Variable	model 9 (RSE)		model 10 (RSE)	
	Beta	p	Beta	p
Constant	.00	.003**	.00	<.001***
Midterm (1=yes)			.18	.017*
Conservative Rating			.31	.007**
Funding (Rep)			.76	.006**
Funding (Dem)			-.61	.01**
Funding^2 (Rep)			-.78	.005**
Funding^2 (Dem)			.56	.008**
Senator (Rep)			.03	.822
Governor (Rep)			.00	.951
Congressman (Rep)			.06	.472
Other Exp. (Rep)			-.03	.723
Senator (Dem)			-.07	.568
Governor (Dem)			.03	.7
Congressman (Dem)			-.09	.415
Other Exp. (Dem)			.10	.229
# Tweets (Rep)	-.03	.758	-.09	.283
# Tweets (Dem)	-.10	.303	.05	.499
Doc length (Rep)	.12	.239	.01	.91
Doc length (Dem)	-.18	.084•	-.08	.215
Tone (Rep)	-.06	.56	.00	.953
Tone (Dem)	.00	.971	-.05	.544
TM Diversity (Rep)	-.43	<.001***	-.18	.019*
TM Diversity (Dem)	.38	.001***	.27	.002**
Adj R^2		.37		.79

Note: •p < .1. \*p < .05. \*\*p < .01. \*\*\*p < .001 (two-sided). Models marked as (RSE) computed with Robust Standard Errors (N=71).

## 6. SUMMARY AND DISCUSSION

I began this study with several questions and hypotheses about the role of media in the electoral process. I explored the relationship between electoral success, traditionally studied media features, such volume and tone, and, most centrally, thematic diversity. Using data from U.S. Senate races, I examined these relationships in two distinct contexts: news coverage of political candidates and candidates' direct communication with potential voters via social media.

### 6.1 Volume and Tone: News

First, I hypothesized that the volume of coverage in the news media would be positively related to electoral success. I found that in the models including only media features, candidates who received more news coverage also won a larger share of the votes. This result is in line with most of the current research and can be explained by various theories reviewed in the theoretical framework (Section 2.1.1), including simple exposure and agenda setting theory. Similarly, I hypothesized that the tone of the news coverage that candidates receive would be positively correlated with their electoral success. This hypothesis was also supported, but to a lesser extent. The tone of coverage that a candidate's opponent received was found to be a significant predictor of electoral success ( $p < 0.05$ ), with candidates whose opponents received more positive coverage receiving a lower share of the total votes. This, again, is in line with existing research, including theories such as affective priming and second-level agenda setting, and in accordance with accepted wisdom regarding the benefits of positive coverage for candidates' chances of winning elections.

The influence of the volume and tone of coverage, however, became less clear in models that included thematic diversity as well as in the full model, which controlled for a wide array of non-media factors. Indeed, at times, the relationship with volume was found to be in the opposite direction than was predicted. However, in further analyses, in which the *relative* volume of news coverage was used rather than the absolute number of articles mentioning a candidate, volume again emerged as a significant and positive predictor of electoral success. While more research is needed, it seems likely that this result owes to the conflicting tendencies of more competitive races attracting a higher volume of news coverage and more successful candidates receiving more news coverage. Thus, while there is a general positive relationship between the number of articles a candidate appears in and their electoral success, this trend line curves around the 50% vote share line, as more competitive races tend to receive more coverage as a whole for both the winning and the losing candidates. Regarding tone, the full model, including non-media controls, detected no significant relationship for tone (whereas relative volume was significant at the non-traditional threshold of  $p < 0.1$  in the full models).

In addition, the overall explanatory power of models that included only volume and tone as predictors was much smaller than in the models that incorporated thematic diversity. This suggests that the two traditional factors play a smaller role in the relationship between media coverage and electoral success than was previously believed, both in absolute terms and in comparison, when using a more nuanced method for textual analysis. The explanatory power of the full model that included non-media factors was much larger than the model that included only volume and tone, and larger than the models containing the three media factors of volume, tone, *and* thematic diversity.

I draw two tentative but potentially important conclusions from this admittedly complicated picture. First, as suggested previously by Belanger and Soroka (2010), the impact of the media on political success is likely not independent. In other words, the diminished power of volume and tone as predictors when controlling for non-media factors seems to indicate that these factors might either be spuriously related to electoral success, or that they mediate the impact of non-media factors, at least to some extent, on electoral success. Incumbency, for example, might impact the volume of news coverage that candidates receive, as well as their chances of winning elections (thus indicating a spurious relationship). On the other hand, incumbency (and the advantages of candidate name-recognition and access to the media that this entails) might contribute to the volume of coverage, which, in turn, increases the likelihood of electoral success (thus indicating a mediated relationship). I elaborate more on this issue when discussing a similar debate on the role of thematic diversity in the news.

Second, even if media factors such as the volume, tone, and thematic diversity of news coverage demonstrate mediating and even independent effects on electoral coverage, these effects are small relative to those of more critical non-media explanations. Factors such as the conservative or liberal leaning of a state, the political arena context (such as midterm elections), and the candidate's experience (especially in Senate elections, where incumbency is known to be extremely beneficial to a candidate's success), all seem to be more important (taken together) than the impact of media on the electoral process.

In summary, this study provides some support for existing research on the positive relationship between volume of coverage, tone of coverage, and electoral

success. However, it also lends support to arguments that criticize naïve or overstated interpretations of this relationship. The impact of media tone and volume on electoral success is relatively small compared with the non-media factors that shape larger political processes; to the extent that such relationships exist, they are likely the result of some combination of independent effects, mediating effects, and spuriousness. (Bélanger & Soroka, 2012).

## **6.2 Volume and Tone: Social Media**

Similar to news coverage, I hypothesized that the volume of candidates' social media activity will be positively related to electoral success. This hypothesis was not supported by the data. Candidates who were more active on social media were not found to perform significantly better in terms of vote share. I hypothesized that candidates' voicing a more negative tone on social media would relate to greater electoral success. As no published studies have examined this relationship in the context of social media, I drew on previous literature on the impact of negative tone in televised political ads (Krupnikov, 2011). However, and similar to the volume of social media activity, no significant relationship was found. This could be in part due to the fact that my analyses did not distinguish between, on one hand, positive or negative tones in personal attacks on a candidate's opponents, rivals, or political nemeses, and, on the other, tone associated with support or opposition to specific policies and political issues. Regardless of the reason, these findings suggest that at least in the context of social media, no clear benefits can be discerned for "going negative."

Previous research has provided mixed results for the relationship between volume of social media activity and electoral success, across a wide range of contexts and

countries. This study can be added to the growing line of evidence illustrating that mere volume of activity is not indicative of candidate performance. In addition, no previous studies have addressed the relationship between the tone of candidates' social media activity and their electoral success. This study, which offers the first such examination, found no evidence for a relationship between negativity or positivity in candidates' online rhetoric and their electoral success. These results do seem to be in line with a different yet related area of research—the relationship between volume and sentiment in general Twitter chatter and electoral success. Echoing Beauchamp's (2017) emphasis on statistical testing of these relationships and the inclusion of adequate controls (Beauchamp, 2017), and as several researchers have already argued (Gayo-Avello, 2012; 2013; Metaxas et al., 2011), at the macro-level, volume and tone may simply be too crude as measures to correctly characterize candidates' social media activity. However, as I elaborate in the limitations section, this evidence should be considered in context: while volume of activity was not found to be impactful in the case of Senate candidates' Twitter activity for election cycles between 2012-2016, these results might be different in other contexts, platforms, or countries.

### **6.3 Thematic Diversity: News**

This study's main contribution is that it goes beyond previous explorations of volume and tone to examine the relationship between thematic diversity, in both candidates' news coverage and their social media activity, and electoral success. Based on previous literature, as well as common wisdom on the subject, I offered a set of competing hypotheses for both news and social media. One hypothesis predicted a negative relationship between thematic diversity and electoral success, and the other

predicted a positive relationship. The results indicate strong support for the first hypothesis for both media types.

In the context of political candidates' news media coverage, thematic diversity was found to be a significant negative predictor of electoral success. That is, more successful candidates generally received less diverse coverage in the news media. In the media-only models, the inclusion of thematic diversity more than doubled the models' predictive capacity, thus highlighting its importance compared with other common predictors used in such research, such as volume and tone. Moreover, although the inclusion of additional non-media factors reduced the impact of thematic diversity, it was still found to be a significant and independent predictor of electoral success, with the slope of its relationship with candidate vote share having a similar magnitude to candidates' experience (though with a smaller magnitude than states' political leaning or candidates' funding). This relationship was detected when using two very different methods of analysis, topic modeling and semantic network analysis, and when using random semantic networks as a benchmark for diversity, thus further increasing the validity of these findings.

However, it should also be noted that improvements to the predictive performance of models containing non-media factors was found to be modest at best. In both topic modeling and semantic network analysis, the inclusion of thematic diversity improved the adjusted R-squared scores of the models from a 0.7 for the non-media predictors only model to a score of 0.74 for models containing both non-media predictors and thematic diversity. Even when both the topic modeling and semantic network analysis estimations

of diversity were included concurrently, the model's adjusted R-squared score only improved to 0.76. These findings point toward several conclusions.

First, from a methodological perspective, while both thematic diversity estimations correlate well, each captures a slightly different aspect of diversity and should therefore be viewed as complementary rather than competing explanations. Second, and most importantly, even when controlling for a wide array of non-media factors, thematic diversity was found to have an independent relationship with electoral success. In line with theories such as issue ownership, and in support of the common wisdom of “staying on message,” it seems that a monothematic media image is more beneficial to political candidates. This conclusion, of course, is limited in terms of the causal direction of this relationship, as I discuss further in Sections 6.5 and 6.7. However, the wide set of controls used in this study, as well as the subsequent analysis of the factors shaping thematic diversity in news coverage, seem to increase the validity of this argument.

Third, as has been argued in the past and similar to the discussion of volume of news coverage in electoral success, while the media can be related to electoral success, its role is much smaller than other “real-world” factors. If thematic diversity (or, rather, non-diversity) is indeed important, then the funding that candidates receive, the state they compete in, and/or their level of previous experience in office (and their opponents' experience) can all play a much larger role.

Fourth, it has previously been argued that media factors often thought to be correlated with electoral success might not be independent predictors of electoral success (Bélanger & Soroka, 2012). Rather, their relationship with electoral success might be



spurious, with non-media factors influencing both media coverage and electoral success concurrently, or mediating, as non-media factors influence media coverage and as a result, via changes to media coverage, candidates' electoral success. While thematic diversity remains a significant predictor of electoral success, even when controlling for a wide array of non-media factors (in contrast with volume and tone), arguments for the non-independence of media predictors for electoral success nonetheless receive some support from the findings presented here. While media factors—and chief among them, thematic diversity—can explain a large share of the variability in electoral success, when these are incorporated into models that control for non-media factors, their relationship with electoral success diminishes. Thus, it seems that these variables are related to some extent to the non-media factors presented in earlier models.

In a further exploration of the factors influencing thematic diversity in news coverage, it was found that thematic diversity can be significantly predicted by a wide array of non-media factors. These relationships have not previously been addressed in the literature. However, some initial explanations can be provided based on theories such as issue ownership, as well as the logic of media coverage of election campaigns.

First, midterm elections were found to have a negative impact on thematic diversity. This might be explained by the tendency of midterm elections to serve as a referendum on the sitting president. In midterm elections, campaign topics tend to gravitate toward issues and topics strongly connected with the sitting president (Barack Obama), limiting the importance of other secondary themes and, as a result, the variety of themes in news coverage of the campaign.

Second, the extent to which the state was liberal or conservative had a significant impact on the thematic diversity of news coverage of both Republican and Democratic candidates, but in opposite directions. The more conservative the state was, the less diverse the coverage of the Republican candidates was, and the more diverse the news coverage of the Democratic candidates was. An explanation for this might be the relevancy of core Republican issues to more conservative states and the relevancy of core Democratic issues to more liberal states. Issue ownership dictates that some candidates or parties “own” specific issues on which they are awarded higher credibility and for which they are considered more capable. With issue ownership being deeply rooted in a country’s history and culture, for countries as diverse as the US, issue ownership could vary dramatically across states. More conservative states might attribute issue ownership for specific issues to candidates more in line with the state’s dominant ideology, whereas in liberal states, other issues would be owned by the Democratic Party. As a consequence, Republican candidates in conservative states would be more strongly connected to the “owned” issues and use them to leverage their candidacy. By contrast, in these same states, Democratic candidates would not be able to attach themselves to such prominent issues as effectively. While a Republican candidate would be constantly connected to an “owned” issue like immigration on which they can campaign effectively, the Democratic candidates would be more likely to emphasize (or be connected to) a wider range of other issues, thus increasing thematic diversity.

Third, a similar argument can be used to explain why incumbents were found to have less thematically diverse news coverage than challengers. Incumbents are more likely to have specific issues on which they ran their previous campaign—issues that are

based on the candidates' defining achievements and successes (Sellers, 1998). However, challengers can run on multiple platforms and attack various issues, decisions, and policies that the incumbent executed during his or her political past, thereby targeting both negative topics and positive topics, framed negatively (Sigelman & Buell, 2004).

Fourth, using both methods, campaign funding was found to have a negative relationship with thematic diversity for both Republican and Democratic candidates. Candidates that received more contributions were found to have received less thematically diverse news coverage. One possibility for this is that greater funding is an indication that a candidate has a better-run and more professional campaign. As political strategists often advocate for candidates to stay on message, a more professional campaign might have a more concentrated message strategy, thus exhibiting lower thematic diversity in general.

Finally, more competitive races appeared to have greatly reduced the thematic diversity of news coverage of political candidates. However, this impact was found to be more pronounced when diversity was measured with semantic network analysis and less so when topic modeling was used. Existing literature does not seem to offer a plausible explanation for this finding. This might be the result of non-competitive races becoming competitive due to specific events or scandals, such as the special Senate election in Alabama in 2018, which was defined by the sexual misconduct of Republican candidate Roy Moore—a scandal that was the main talking point for news coverage of the race. Such defining events are likely to limit the range of topics discussed in the media for both candidates. Alternatively, for other more “purple states,” it may be that these findings reflect defining “wedge issues,” on which campaign coverage tends to converge.

There are two important caveats that should be considered. First, the explanations that are suggested above are, in a number of cases, provided post-hoc. Further research is needed to test the new hypotheses generated from the data. Moreover, the cross-sectional nature of this study limits the ability to determine causality. Incumbency, for example, might be influenced by thematic diversity rather than influencing it. It might also be that incumbency indicates that a candidate had previously run a self-disciplined monothematic campaign and is therefore more experienced in and/or has greater ability to shape news coverage that is less thematically diverse. To take another example, campaign funding might be connected to diversity and electoral success in various ways. Funding could be connected to both electoral success and thematic diversity concurrently (i.e., a spurious relationship). Or it might be that funding influences media coverage and thematic diversity, and thematic diversity only then influences electoral success (i.e., thematic diversity plays a mediating role). I revisit this critical issue when discussing the underlying mechanisms of these relationships and in the limitations section, focusing on the cross-sectional nature of the study.

All in all, while being cautious about overstating the importance of media factors, it should be noted that the set of non-media predictors used as controls is quite comprehensive, as indicated by the large adjusted R-squared for the model with only non-media factors (0.7). An improvement of 6% over this benchmark can be seen as at least suggestive of the relative independence of thematic diversity, as well as the strength of thematic diversity's relationship with electoral success. Second, and from a practical point of view, even small effects such as that attributed to thematic diversity here can be very important in a winner-takes-all political system, which is the case in the U.S. With

more than a few races being decided by razor-thin margins, thematic diversity can play a decisive, even if modest, role. Thus, in the context of news coverage, it can be argued that limiting thematic diversity can indeed be beneficial to candidates.

#### **6.4 Thematic Diversity: Social Media**

In the context of political candidates' social media activity, thematic diversity was also found to be a significant and negative predictor of electoral success. In the media-only models, the inclusion of thematic diversity improved the models' predictive capacity greatly (from an adjusted R-squared score of .04 to a score of 0.31). However, a much more modest improvement was found when using semantic network analysis to estimate thematic diversity (with an adjusted R-squared score of 0.1—although the relationship was still found to be significant, especially when using randomly generated networks as a benchmark for diversity. I discuss the possible reasons for this difference in the limitations and future directions section of this chapter.

Although the inclusion of additional non-media factors reduced its impact, thematic diversity was still found to be a significant predictor (when estimated using topic modeling and, to a lesser extent, when using semantic network analysis), with the slope of its relationship with candidate vote share having a similar magnitude to state political leaning, though a smaller magnitude than with funding. These results support the logic of issue ownership, as well as of “staying on message” as a campaigning strategy. Generally, candidates who discussed fewer issues, or alternatively, issues that were more closely related to each other, often performed better in terms of vote share. This conclusion is especially relevant for social media, considering that the analysis focused

on candidates' own activity on Twitter, which campaigns have much greater and more direct control than news media coverage.

When exploring the factors that influence or explain the degree of thematic diversity, some results were similar to those found for news coverage. Using topic modeling to estimate diversity, candidates in the midterm elections were found to have significantly less thematic diversity than candidates in quadrennial elections. Consistent with observations from previous midterm elections (Grofman et al., 1998), the 2014 midterm elections included in my analyses might be viewed as a referendum on then-sitting president Barack Obama, thus focusing the attention on a more limited set of topics and reducing thematic diversity. Similarly, a significant relationship was found between the competitiveness of the race and thematic diversity, with candidates in more competitive races exhibiting lower thematic diversity in their social media activity—a finding consistent with the relationship between competitiveness and thematic diversity in news coverage.

In other instances, the results differed greatly between media channels. News coverage of incumbents differed significantly from that of challengers, the coverage of well-funded candidates differed from that of less well-funded candidates, and candidates in midterm elections, tight races, or conservative states received less diverse news coverage. However, with the partial exception of incumbency and a race's competitiveness, candidates' social media activity was found to be a much more independent phenomenon. In the context of the news media, the adjusted R-scores for the semantic network diversity estimation for Republican (0.45) and Democratic (0.48) candidates, as well as the models estimating thematic diversity using topic modeling for

Republican (0.25) and Democratic (0.2) candidates, showed very strong results. Compare these scores with the predictive capacity of the corresponding models in the context of social media, in which the adjusted R-squared scores were only 0.13, 0.18, 0.11, and 0.03, respectively.

Such dramatic differences in explained variance may reflect the fact that candidates' ability to stay on message, offer a more monothematic message strategy, and connect all issues to one specific talking point are all under their control when using social media. Consequently, social media activity should be less influenced by the macro-level factors that shape thematic diversity in news coverage. Nonetheless, these findings raise new and important questions about the factors that influence thematic diversity in social media use—questions that should be answered in future research.

All in all, these results suggest that thematic diversity is an important component in election campaigns. Thematic diversity in both news coverage and candidates' social media activity was found to be closely and negatively related to their electoral success, as measured by the share of the votes gathered by a candidate. This holds true even when accounting for a wide range of non-media predictors. In addition, in the case of the news media, non-media factors were also shown to impact thematic diversity, raising the possibility that at least part of the relationship between thematic diversity and electoral success can be attributed to non-media factors, which are either mediated through media coverage or which spuriously affect media coverage and electoral success concurrently. Finally, although both topic modeling and semantic network analysis performed well as estimations of diversity in the context of news coverage, in the context of social media, the topic modeling-based diversity estimations performed much better than the semantic

network analysis-based estimations, leading to some methodological considerations that will be addressed in the limitations and future directions section of this chapter.

### **6.5 Thematic Diversity and Electoral Success – Possible Mechanisms**

While the results presented in this dissertation strongly support a relationship between thematic diversity in candidates' news media coverage and social media activity and their electoral success, several possible explanations for these relationships can be considered based on different factors and the suggested causal direction of the effects.

The first set of explanations rests largely on the premise of setting and agenda diversity—that is, that the range of issues to which the media affords attention can impact the range of issues that voters consider when making political decisions. For example, in accordance with issue ownership theory, focusing on a single message can benefit candidates, as each party “owns” only a limited set of issues. In another example, because audiences are not always attentive, during moments when they are paying attention, the message delivered needs to be the strongest message that the campaign has to offer (Benoit et al., 2011). Given that campaign messages need only be either a rationale for choosing the candidate or a rejection of his or her opponent's message, the range of messages that can be seen as beneficial is somewhat limited. A more limited set of themes can create a more coherent campaign, thus increasing its effectiveness. Finally, as repetition and reinforcement are critical for message effect (Allport & Lepkin, 1945; Henkel & Mattson, 2011), it makes sense that campaigns should focus only on a small set of messages with as little variation as possible.



However, another group of explanations for the relationship between thematic diversity and electoral success focuses on the opposite causal direction, in which electoral success impacts thematic diversity—or in which both are spuriously affected by a third variable concurrently. While the results of this study show that thematic diversity’s relationship is independent of a wide set of non-media-related third variables, such as incumbency, state-leaning, funding, and experience, these are in no way a complete list of all possible third variables. Thus, one could reasonably argue that media practices or text generation circumstances impact both electoral success and thematic diversity concurrently.

Different media practices can impact the thematic diversity of candidates’ news coverage and social media activity. For example, less important Senate races might draw fewer resources for coverage, thereby limiting the diversity of themes, issues, and topics discussed in coverage of the race. However, considering the linear relationship found in this study, this possibility seems unlikely. The level of a campaign’s organization could also be a third variable influencing both thematic diversity and electoral success. A well-coordinated campaign can be more successful in keeping its theme or message coherent, thus reducing diversity for more successful candidates while at the same time contributing independently to electoral success. It could also be argued, however, that the degree to which a campaign is well coordinated or organized is likely related, at least to some degree, to the amount of funding it receives and the political experience of the candidate, variables which are controlled for in the regression models.

As shown in Section 4.4.2.3, the randomized networks created using the configuration models offered higher diversity than the observed networks in over 90% of

the cases. It thus seems likely that more coherent and structured media activity, in both news and social media, will result in networks that are less diverse. This relates to the linguistic characteristics and processes by which texts are generated, which can be extended to candidates' vocabulary. Candidates with a more limited vocabulary will tend to have lower thematic diversity, at least in their social media activity, as the density of their semantic networks are expected to be larger. However, to argue that limited vocabulary contributes to candidates' success will likely requires additional evidence before it can be supported. Nonetheless, the repetition of specific phrases, sound-bites, and slogans can have a similar effect on thematic structure. Mentioning keywords from the campaign slogan every time an issue is discussed, for various issues, can help connect candidates with a coherent theme. To use an example from outside this study's sample, if Donald Trump's campaign slogan, "Make America Great Again," is connected to economic, foreign policy, and security issues simultaneously, then disparate concrete actions, such as reducing unemployment, renegotiating treaties, and adding funds to the military, can all be connected to one thematic framework. As semantic network analysis and topic modeling address word co-occurrence, this relationship between the impact of topical or linguistic diversity and electoral success remains an open question for future research.

As I discuss in the limitations section, the larger question about the mechanisms underlying the relationship between electoral success and thematic diversity also remains open. While an extensive set of control variables was used in the regression models to examine the independent—or semi-independent—nature of thematic diversity, the study's cross-sectional design does not allow for a definite conclusion regarding the

causal order of this relationship. The question of whether this relationship results from advantages to the monothematic campaign message strategy, the result of media practices related to successful candidates' social media activity, the media practices of newsrooms when covering successful candidates, or factors related to text generation and candidates' vocabulary, will need to be addressed using alternative research designs.

## **6.6 Summary of Contributions**

This dissertation offers several contributions to current knowledge, both theoretically and methodologically. First, this study contributes to our understanding of the relationship between the volume and tone of political candidates' news coverage and their electoral success (Bélanger & Soroka, 2012; De Vreese, 2010; Hopmann et al., 2010; Norris et al., 1999). More specifically, this study shows that the volume and tone of coverage (positive or negative) can impact candidates' electoral success (Balmas & Sheafer, 2010; Boomgaarden et al., 2012; Coleman & Wu, 2010; Eberl et al., 2017; Geers & Bos, 2017; Geiß & Schäfer, 2017; Hopmann et al., 2010; Johann et al., 2017; Kiousis et al., 2006; Lengauer & Johann, 2013; M. McCombs et al., 1997; Norris et al., 1999; Oegema & Kleinnijenhuis, 2000). However, it also lends support to recent criticisms that call into question the extent to which these relationships are truly independent (Bélanger & Soroka, 2012), rather than spurious or mediated ones.

Second, this study offers evidence for the impact of volume and tone in the much less researched context of political candidates' activity in social media. In line with various studies offering conflicting findings regarding the impact of volume and tone in general Twitter "chatter" about political candidates (Beauchamp, 2017; Gayo-Avello, 2013; Jungherr, 2016; Tumasjan et al., 2010), and adding to the relatively few studies that

explore the relationship between candidates' social media activity and their electoral success (Bright et al., 2018; LaMarre & Suzuki-Lambrech, 2013; Vergeer, Hermans, & Sams, 2011), this study concludes that volume and tone of social media activity cannot be shown to be consistently related to electoral success. This contribution is especially important for the tone of candidates' social media activity, a topic that has not been previously addressed by researchers (Jungherr, 2016).

Third and most importantly, from a theoretical perspective, this study presents a systematic and empirical analysis of the role of thematic diversity in candidates' news coverage and social media activity. While some previous research on the role of thematic diversity does exist, it suffers from several limitations that were addressed in Chapter 2 of this dissertation. Much of the research on monothematic campaign strategies—i.e., the “It’s the economy stupid” approach, or “staying on message,” as it is often referred to in practice (Benoit et al., 2011—relies on case studies and anecdotal evidence. Additionally, existing research is largely limited to national as opposed to state-level election campaigns and focuses on more traditional forms of political advertising. Much less attention has been paid to the role of thematic diversity, either in news coverage of campaigns or in new forms of direct communication, such as candidates' social media activity. Using unsupervised machine learning to estimate thematic diversity in large corpora, this study shows that thematic diversity is strongly and negatively related to electoral success, both in candidates' coverage in the news media as well as their social media activity. This relationship remains significant even when controlling for a large host of non-media factors and is found to significantly improve predictions for models

based solely on volume and tone, and modestly improve models on volume, tone, and a wide array of non-media factors.

The lack of previous systematic evidence on the relationship between thematic diversity and electoral success, I argue, is due in part to the complex nature of thematic diversity. This complexity makes its estimation challenging, especially when addressing it in the context of large and varied corpora or when no reliable a priori topic lists are available for a given corpus' specific context. Thus, the fourth contribution of this study is that it addresses challenges in the operationalization, conceptualization, and measurement of thematic diversity. In this study, I conceptualize and estimate thematic diversity using two different unsupervised machine learning methods. Neither of these requires a priori assumptions of possible themes, issues, or topics in a given corpus prior to analysis. Both also allow for comparisons across discourses using an identical procedure over all corpora. Used in tandem, the two methods enable the researcher to account not only for the number of categories but also the interconnectivity and distribution of those categories. Finally, these methods offer a relatively cost-effective way to estimate thematic diversity without having to rely on human coders—a process that is costly for analyzing large corpora and impossible for extremely large datasets, such as those used in this study.

Fifth, using semantic network analysis to estimate thematic diversity not only benefits existing research, but also ongoing research on semantic network analysis as a method for textual analysis in political communication. While research on semantic networks has grown considerably in the last decade, and while researchers often use semantic network analysis as a tool for analyzing various discourses and corpora,

relatively few researchers have extended their analysis from the one-network perspective to a multiple or between-network perspective (Baden, 2010; Carley & Palmquist, 1992; Danowski, 2012a; Doerfel & Connaughton, 2009a; Eberl et al., 2014a; Qin, 2015a; Shim et al., 2015). Moreover, even the studies that do tend to be limited to a small number of graphs and to a more basic set of qualitative methods for comparison. While applicable to small scale comparative analytics, such methods are inadequate for comparing larger sets of semantic networks. Therefore, this dissertation follows in the footsteps of studies such as those conducted by Eberl et al. (2014) and Doerfel and Connaughton (2009a), which extend semantic network analysis scholarship through the use of prominent network graph-level indicators to conduct large-scale comparisons of multiple semantic networks and their impact. This study likewise advances research on thematic diversity in semantic network analysis (Eberl et al., 2014a) by focusing on a prominent set of measures related to network cohesion and partitioning, by taking advantage of the role of sub-graphs, and by providing a novel method for estimating network diversity. The potential for this method in analyzing discourse structure can also be extended beyond estimations of thematic diversity, as I discuss in the following section.

## **6.7 Limitations and Future Directions**

This study suffers from several limitations that highlight potential future directions for further elaboration on the theories and methods presented in this study. First, the main drawback of this study is its cross-sectional design. While attempting to strengthen the validity of the findings using a wide array of non-media controls, the causal and independent nature of the relationship between thematic diversity and electoral success remains to be determined. While this study identifies a significant

negative relationship between electoral success and thematic diversity, and while arguments are provided in support of thematic diversity's effect on electoral success, the causal ordering of these two phenomena requires further validation. While it is likely that thematic diversity in media coverage influences electoral success directly, it is also likely that this relationship is more complex than that. Evidence for this stance can be found in the relationship between thematic diversity in candidates' news coverage and various non-media factors. On the one hand, this relationship can be viewed as somewhat independent. There is evidence to suggest that thematic diversity affects electoral success, given the wide range of control variables included in the regression models and the semi-independent relationship that these models detected. On the other hand, the relationship can be accounted for, at least partially, by non-media factors, such as incumbency status, funding, or the type of election cycle. Thus, as discussed extensively in Section 6.4, questions about the causal direction and the mechanisms underlying the relationship remain open.

In a related limitation, the causal direction of the relationship between thematic diversity and non-media factors cannot be determined with the data used in this study. For example, although incumbency is more likely to influence thematic diversity, the relationship could also go in the opposite direction. Campaign funding might connect to both electoral success and thematic diversity concurrently, thus suggesting that the relationship is spurious; it could also be the case that funding influences media coverage and thematic diversity and that thematic diversity, in turn, influences electoral success. One possibility for addressing these issues is to design an experimental study that alters thematic diversity in various corpora and then measures the impact of these changes on

individual responses. Such an experiment would shed light on causality in the relationship between thematic diversity and electoral success, and whether there is a direct effect from one to the other. Unfortunately, not all of the study's findings can be validated through experimentation. As an obvious example, it would be near impossible to create a truly randomized experiment in which candidates received different levels of funding to observe its effect on candidates' thematic diversity choices on social media or in their news coverage.

A second limitation is that while the various non-media factors were found to shape thematic diversity in candidates' news coverage, their performance in predicting thematic diversity in candidates' social media activity was modest at best. Thus, questions as to what factors shape how candidates choose topics to discuss on social media, and the diversity of these topics, remains to be answered in future research.

Third, this study is limited in scope. It explores a very specific political arena (U.S. Senate elections) over a limited time frame (three to five election cycles) and using two specific platforms (Twitter and mainstream print news coverage). For example, the role of thematic diversity could be different in presidential elections, during which candidates likely need to connect with a wider set of topics to relate to a much larger public. It might also be different in other countries, where political rhetoric and the rules of the media differ from those in the U.S. While Twitter is an important platform through which candidates can communicate directly with potential voters and has received much scholarly attention in previous election cycles, other forms of political advertising and social media, both traditional and novel, might offer unique benefits and/or drawbacks in terms of studying the effects of political rhetoric's structure, tone, and volume on



electoral outcomes. In terms of news media in general, this study has explored only print media, and it might be that other media types, such as television or online news, require different theoretical considerations. Thus, future research should apply the schema developed in this study in other contexts, from direct emailing and televised advertising, to other political systems and political roles. Moreover, both channels examined in this study, print news and Twitter, are mass broadcast channels through which candidates connect with the public as a whole. With the rise of targeted advertising and personalized messages, the arguments offered here might be inapplicable or outdated. For targeted advertising, it might be that candidates might exhibit higher overall thematic diversity but that this is limited by media-type, indicating that different constituencies are addressed through different messages with greater efficiency—and that candidates who do this type of “narrowcasting” are those that are ultimately more successful at the polls.

Fourth, during the exploration of thematic diversity in candidates’ social media activity, the estimation based on topic modeling was found to perform better than that derived using semantic network analysis. I hypothesize that this difference was the result of a smaller corpus used to build the semantic networks within the social media context. However, to better understand the limitations of this method, especially in terms of minimal size, further research is needed. It is clear from the results that both topic modeling and semantic network analysis are potent tools for textual analysis. Topic modeling is the more advanced of the two methods, and has a distinct advantage in that it can cope with smaller amounts of data. However, this is contingent on the number of small cases it is applied to, as it is applied to several corpora at once. This proves both a strength and a weakness of the method. When analyzing a smaller number of corpora,

semantic networks might be the better approach—that is, assuming the corpora in question are large enough for a reliable network to be created. Moreover, once a study has already begun, semantic network analysis can handle the addition of new data and compare it to existing data far more easily. This requires only that the exact process for creating and analyzing the network is carried out for every new data point (such as new candidates). For topic modeling, however, estimating a new model, including all of the decisions about the adequate number of topics, must be carried out over all corpora, both new and old, again—that is, unless we explicitly assume that no new topics occur in the new data.

Another difference between the two approaches is how theme membership is defined for each word in the corpus. Topic modeling assumes a multiple membership model, where each word can be applied to several topics with differing magnitudes. By contrast, the community detection algorithm used to separate the semantic networks into communities in this study assumes single membership for each node. That is, each word can only be associated with one theme. This assumption is, of course, unrealistic, as words can be related to two topics—a word like “compete” is relevant to both politics and sports, for example. It should be noted that this limitation is not inherent to semantic network analysis in general, but rather to existing community detection algorithms. With more advanced algorithms being developed, such advances in network analysis can be directly applied to theme extraction in semantic networks (Bai, Yang, & Shi, 2017). This is perhaps one of the major advantages of the network-based approach, as well as a possible avenue for future research.

One the main contributions of this study is that it exemplifies the applicability of semantic network analysis for comparing disparate discourse structures. While this study focused solely on thematic diversity, additional network structures can and should be examined as well. The utility of semantic network analysis as a method is driven by the ever-growing literature, theory, methods, and measures associated with the approach. Just as this study took advantage of developments in community detection processes in large networks (Blondel et al., 2008; Fortunato & Hric, 2016), other network features might be applicable to semantic network analysis in other contexts. However, these features also pose a greater challenge in the form of “translating” general and social network features to discourse features. As this study shows, this endeavor might prove to be quite fruitful.

This study’s methods are also limited in terms of comparing the diversity of texts drawn from disparate media types. The fifth limitation of this study thus stems from the necessary usage of different methods and models to analyze social media data and candidates’ coverage in the news media. From a topic modeling perspective, differences in the size and nature of the two corpora required constructing different topic models for each media type—models that differed in the number of topics determined to be optimal and in terms of their interconnectivity. Similarly, from the perspective of semantic network analysis, differences between the size of the databases for the news media and social media activity required different definitions for word co-occurrence and dictated different network sizes across the two media channels. As such, while these methods enabled a direct comparison between different corpora drawn from the same media type, they were unable to offer a direct comparison between thematic diversity in social media activity and in candidates’ news media coverage.

It is also possible that there are differences between the two media types based on how the texts were generated. For example, candidates' social media activity is directed and created by fewer actors than the news media. It is therefore likely that thematic diversity will be found to be lower in the case of social media. This limitation is further exacerbated by the limited performance of semantic network analysis in the context of social media. However, taken to its extreme, this conceptual difference might ultimately limit the logic of comparison altogether.

Lastly, as mentioned in Section 2.4 on thematic diversity, the concept of thematic diversity can be applicable to numerous contexts and research questions beyond election campaigns and candidate news coverage. The importance of conceptualizing thematic diversity was underscored both by normative arguments related to the role of thematic diversity in the media and by empirical arguments about the effects of thematic diversity on public opinion. In the context of political communication, researchers have explored the relationship between the public's thematic diversity, thematic diversity in the media, and the causes and effects of media agenda diversity (Huang, 2010; Lee et al., 2014; McCombs & Zhu, 1995; Peter & De Vreese, 2003). Similarly, a line of research originating in psychology has explored the "real world" factors, such as crises, as possible antecedents to media diversity, the role of thematic diversity in the context of wars (Stewart & Suedfeld, 2012; Suedfeld & Tetlock, 1977) and revolutions (Suedfeld & Rank, 1976), and the role of diversity in more day-to-day settings, such as the creative and professional success of bi-cultural individuals living abroad (Tadmor, Galinsky, & Maddux, 2012), scientists' thinking on research and teaching (Feist, 1994), and the impact of positive and negative life-events (Suedfeld & Bluck, 1993). Each of these areas

could benefit from incorporating the thematic diversity measurement developed in this study (by accounting for variety, balance, and disparity), and could do so in a resource-efficient way by aiding, or even replacing, human coders with unsupervised machine learning methods.

## **6.8 It's the Structure**

I began this dissertation by pointing to a prominent example of political messaging strategy drawn from the 1992 U.S. Presidential elections between then-Governor of Arkansas, Bill Clinton, and incumbent George H. W. Bush. Repeated over and over, the adage that strategist James Carville coined during the Clinton campaign in 1992, “it’s the economy, stupid,” as well as his mission to connect every possible message opportunity to this theme, is perhaps one of the most prominent examples of “staying on message” and monothematic campaign strategies.

Of course, this was neither the first nor last election campaign defined by a strong central theme. In the 2008 Presidential elections, a theme similar to another of Carville’s three foci in the 1992 campaign, “change vs. more of the same,” was at the center of then-candidate Barack Obama’s campaign. Examples of this theme can be found in the memorable “change” poster that became a cultural phenomenon, ads that promoted McCain and Bush as “the same,” and even the nickname attributed to McCain as “McSame” (Kenski, Hardy, & Jamieson, 2010). In the 2016 presidential elections “make America great again” was a slogan repeated over and over by Trump, his supporters, and surrogates, in contrast with arguments regarding the lack of central theme for the Clinton campaign being raised. The same strategy can also be found in non-presidential races. For example, at the time of writing, the race for the 2018 Republican nomination for

governor of Ohio is in its initial stages. Candidate Mary Taylor is lobbing accusations that candidate Mike DeWite (also a Republican) is too liberal. Repeating keywords, such as “Obama,” “Hillary,” “liberal,” and “962” (the number of times DeWite allegedly voted with the Democrats in the last 6 years), through a series of televised ads, Taylor is attempting to connect key Republican issues, such as immigration, abortions and Obamacare, to the central theme of party loyalty. While still in its early stages, pundits argue that other campaigns also show signs of this central messaging strategy.<sup>7</sup>

Thus, while the economy might not always be the core issue around which all campaigns revolve, political strategists, consultants, and researchers remain committed to the notion that candidates must promote a specific and unified theme throughout the campaign—that in order to win, candidates need to keep the campaign message coherent, succinct, and as unidimensional as possible (Benoit et al., 2011; Bradshaw, 2004; Conway III et al., 2012). This study offered a systematic and empirical examination of this common wisdom. Various theories and arguments connected to this debate were reviewed, from agenda setting to issue ownership and issue convergence. I measured the level of thematic diversity in both news coverage of Senate candidates in five election cycles and in candidates’ direct communication with voters via social media in three recent election cycles.

I controlled for a wide array of non-media factors that previous scholarship suggests shape electoral results and used two different unsupervised machine learning methods to model discourse structure. The results show that structural features of

---

<sup>7</sup> <https://www.npr.org/2018/05/04/608193538/gop-primaries-focus-on-candidates-loyalty-to-president-trump>  
<http://www.chicagotribune.com/news/nationworld/politics/ct-trump-senate-races-20180415-story.html>

political discourse indicative of thematic diversity are significantly, strongly, and negatively related to electoral success. To rephrase Carville's famous adage, I argue that when it comes to candidate and news media political discourse during election campaigns, it might not necessarily be "the economy," "immigration," or "healthcare" that matter—but it is often "the structure" that does.

## REFERENCES

- Abe, J. A. A. (2012). Cognitive–affective styles associated with position on war. *Journal of Language and Social Psychology, 31*(2), 212–222.  
<https://doi.org/10.1177/0261927X12438532>
- Aizawa, A. (2003). An information-theoretic perspective of tf–idf measures. *Information Processing & Management, 39*(1), 45–65.
- Allport, F. H., & Lepkin, M. (1945). Wartime rumors of waste and special privilege: why some people believe them. *The Journal of Abnormal and Social Psychology, 40*(1), 3.
- Althaus, S. L., Nardulli, P. F., & Shaw, D. R. (2002). Candidate appearances in presidential elections, 1972–2000. *Political Communication, 19*(1), 49–72.  
<https://doi.org/10.1080/105846002317246489>
- Anholt, S. (2006). Anholt nation brands index: how does the world see America? *Journal of Advertising Research, 45*(03), 296.  
<https://doi.org/10.1017/S0021849905050336>
- Bache, K., Newman, D., & Smyth, P. (2013). Text-based measures of document diversity. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 23–31). ACM.
- Baden, C. (2010). *Communication, Contextualization & Cognition: Patterns & Processes of Frames' Influence on People's Interpretations of the EU Constitution* (Doctoral Dissertation). University of Amsterdam, Amsterdam School of Communication Research.



- Bai, X., Yang, P., & Shi, X. (2017). An overlapping community detection algorithm based on density peaks. *Neurocomputing*, *226*, 7–15.  
<https://doi.org/10.1016/j.neucom.2016.11.019>
- Baker-Brown, G., Ballard, E. J., Bluck, S., de Vries, B., Suedfeld, P., & Tetlock, P. E. (1992). The conceptual/integrative complexity scoring manual. In C. P. Smith (Ed.), *Motivation and Personality: Handbook of Thematic Content Analysis* (pp. 401–418). Cambridge, UK: Cambridge University Press.
- Balmas, M., & Sheafer, T. (2010). Candidate image in election campaigns: Attribute agenda setting, affective priming, and voting intentions. *International Journal of Public Opinion Research*, *22*(2), 204–229. <https://doi.org/10.1093/ijpor/edq009>
- Ban, P., Llaudet, E., & Snyder, J. M. (2016). Challenger quality and the incumbency advantage: Incumbency advantage. *Legislative Studies Quarterly*, *41*(1), 153–179.  
<https://doi.org/10.1111/lsq.12109>
- Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences*, *17*(7), 348–360. <https://doi.org/10.1016/j.tics.2013.04.010>
- Beauchamp, N. (2017). Predicting and interpolating state-level polls using Twitter textual data: Predicting polls with Twitter. *American Journal of Political Science*, *61*(2), 490–503. <https://doi.org/10.1111/ajps.12274>
- Bélangier, É., & Soroka, S. (2012). Campaigns and the prediction of election outcomes: Can historical and campaign-period prediction models be combined? *Electoral Studies*, *31*(4), 702–714. <https://doi.org/10.1016/j.electstud.2012.07.003>

- Benford, R. D., & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, 611–639.
- Benoit, W. L., Glantz, M. J., Phillips, A. L., Rill, L. A., Davis, C. B., Henson, J. R., & Sudbrock, L. A. (2011). Staying “on message”: Consistency in content of presidential primary campaign messages across media. *American Behavioral Scientist*, 55(4), 457–468. <https://doi.org/10.1177/0002764211398072>
- Benoit, W. L., Leshner, G. M., & Chattopadhyay, S. (2007). A meta-analysis of political advertising. *Human Communication*, 10, 507-522.
- Bernays, E. L. (1928). *Propaganda*. New York, NY: Horace Liveright.
- Bird, S. (2009). *Natural language processing with Python* (1st ed). Cambridge [Mass.]: O’Reilly.
- Blei, D. M. (2011). Introduction to probabilistic topic models. In *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993–1022.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 1–12. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Bode, L., & Dalrymple, K. E. (2016). Politics in 140 Characters or Less: Campaign Communication, Network Interaction, and Political Participation on Twitter. *Journal of Political Marketing*, 15(4), 311–332. <https://doi.org/10.1080/15377857.2014.959686>

- Bode, L., Lassen, D. S., Kim, Y. M., Shah, D. V., Fowler, E. F., Ridout, T., & Franz, M. (2016). Coherent campaigns? Campaign broadcast and social messaging. *Online Information Review*, 40(5), 580–594. <https://doi.org/10.1108/OIR-11-2015-0348>
- Boomgaarden, H. G., Vliegenthart, R., & de Vreese, C. H. (2012). A worldwide presidential election: The impact of the media on candidate and campaign evaluations. *International Journal of Public Opinion Research*, 24(1), 42–61. <https://doi.org/10.1093/ijpor/edr041>
- Borah, P. (2016). Political Facebook use: Campaign strategies used in 2008 and 2012 presidential elections. *Journal of Information Technology & Politics*, 13(4), 326–338. <https://doi.org/10.1080/19331681.2016.1163519>
- Bradshaw, J. C. (2004). Who will vote for you and why: Designing campaign strategy and message. In J. A. Thurber & C. J. Nelson (Eds.), *Campaigns and Elections American style* (pp. 37–56).
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318. <https://doi.org/10.1073/pnas.1618923114>
- Bright, J., Hale, S. A., Ganesh, B., Bulovsky, A., Margetts, H., & Howard, P. (2018). Does Campaigning on Social Media Make a Difference? Evidence from candidate use of Twitter during the 2015 and 2017 UK Elections.
- Broersma, M., & Graham, T. (2012). Social media as beat: Tweets as a news source during the 2010 British and Dutch elections. *Journalism Practice*, 6(3), 403–419. <https://doi.org/10.1080/17512786.2012.663626>

- Brosius, H.-B., & Weimann, G. (1996). Who sets the agenda: Agenda-setting as a two-step flow. *Communication Research*, 23(5), 561–580.  
<https://doi.org/10.1177/009365096023005002>
- Cacciatore, M. A., Scheufele, D. A., & Iyengar, S. (2016). The end of framing as we know it ... and the future of media effects. *Mass Communication and Society*, 19(1), 7–23. <https://doi.org/10.1080/15205436.2015.1068811>
- Caldarelli, G., Chessa, A., Pammolli, F., Pompa, G., Puliga, M., Riccaboni, M., & Riotta, G. (2014). A multi-level geographical study of Italian political elections from twitter data. *PLoS ONE*, 9(5), e95809.  
<https://doi.org/10.1371/journal.pone.0095809>
- Cappella, J. N., & Jamieson, K. H. (1997). *Spiral of Cynicism: The Press and the Public Good*. New York: Oxford University Press.
- Carley, K., & Palmquist, M. (1992). Extracting, representing, and analyzing mental models. *Social Forces*, 70(3), 601. <https://doi.org/10.2307/2579746>
- Carpenter, S. (2010). A study of content diversity in online citizen journalism and online newspaper articles. *New Media & Society*, 12(7), 1064–1084.  
<https://doi.org/10.1177/1461444809348772>
- Chaffee, S. H., & Wilson, D. (1976). Media rich, media poor: Two studies of diversity in agenda-holding. *Journalism Quarterly*, 54(3), 466–476.
- Chewning, L. V. (2015). Multiple voices and multiple media: Co-constructing BP's crisis response. *Public Relations Review*, 41(1), 72–79.  
<https://doi.org/10.1016/j.pubrev.2014.10.012>

- Choi, S. (2015). The two-step flow of communication in twitter-based public forums. *Social Science Computer Review*, 33(6), 696–711.  
<https://doi.org/10.1177/0894439314556599>
- Cohen, B. C. (1965). *The Press and Foreign Policy*. Princeton University Press.
- Coleman, R., & Wu, H. D. (2010). Proposing emotion as a dimension of affective agenda setting: Separating affect into two components and comparing their second-level effects. *Journalism & Mass Communication Quarterly*, 87(2), 315–327.
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Memory and Language*, 8(2), 240–247.
- Conway III, L. G., Gornick, L. J., Burfeind, C., Mandella, P., Kuenzli, A., Houck, S. C., & Fullerton, D. T. (2012). Does complex or simple rhetoric win elections? An integrative complexity analysis of U.S. presidential campaigns. *Political Psychology*, 33(5), 599–618. <https://doi.org/10.1111/j.1467-9221.2012.00910.x>
- Conway, L. G., Conway, K. R., Gornick, L. J., & Houck, S. C. (2014). Automated integrative complexity. *Political Psychology*, 35(5), 603–624.  
<https://doi.org/10.1111/pops.12021>
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5), 1–9.
- Danowski, J. A. (2012a). Semantic network analysis of Islamist sources using time slices as nodes and semantic similarity as link strengths: Some implications for propaganda analysis about jihad (pp. 164–171). IEEE.  
<https://doi.org/10.1109/EISIC.2012.53>

- Danowski, J. A., & Park, H. W. (2014). Arab Spring effects on meanings for Islamist web terms and on web hyperlink networks among Muslim-majority nations: A naturalistic field experiment. *Journal of Contemporary Eastern Asia*, 13(2), 15–39. <https://doi.org/10.17477/jcea.2014.13.2.015>
- De Vreese, C. H. (2010). Campaign communication and media. In L. LeDuc, R. G. Niemi, & P. Norris (Eds.), *Comparing Democracies 3: Elections and Voting in the 21st Century* (3rd ed.). London: SAGE.
- De Vreese, C. H., & Semetko, H. A. (2002). Cynical and engaged: Strategic campaign coverage, public opinion, and mobilization in a referendum. *Communication Research*, 29(6), 615–641. <https://doi.org/10.1177/009365002237829>
- Diesner, J. (2012). *Uncovering and managing the impact of methodological choices for the computational construction of socio-technical networks from texts* (Doctoral Dissertation). Carnegie Mellon University, Pittsburgh, PA.
- Doerfel, M. L., & Barnett, G. A. (1999). A semantic network analysis of the International Communication Association. *Human Communication Research*, 25(4), 589–603. <https://doi.org/10.1111/j.1468-2958.1999.tb00463>
- Doerfel, M. L., & Connaughton, S. L. (2009a). Semantic networks and competition: Election year winners and losers in U.S. televised presidential debates, 1960–2004. *Journal of the American Society for Information Science and Technology*, 60(1), 201–218. <https://doi.org/10.1002/asi.20950>
- Druckman, J. N. (2004). Priming the vote: Campaign effects in a US Senate election. *Political Psychology*, 577–594.

- Eberl, J.-M., Boomgaarden, H. G., & Wagner, M. (2017). One bias fits all? Three types of media bias and their effects on party preferences. *Communication Research*, 44(8), 1125–1148.
- Eberl, J.-M., Jacobi, C., & Schlögl, S. (2014). Measuring News media complexity using network analysis. Paper Presented at the International Communication Association (ICA), Seattle.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58.
- Evans, H. K., Cordova, V., & Sipole, S. (2014). Twitter style: An analysis of how house candidates used Twitter in their 2012 campaigns. *PS: Political Science & Politics*, 47(02), 454–462. <https://doi.org/10.1017/S1049096514000389>
- Feinerer, I. (2017). *Introduction to the TM Package Text Mining in R*.
- Feist, G. J. (1994). Personality and working style predictors of integrative complexity: A study of scientists' thinking about research and teaching. *Journal of Personality and Social Psychology*, 67(3), 474–484. <https://doi.org/10.1037/0022-3514.67.3.474>
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221–233. <https://doi.org/10.1037/h0057532>
- Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *ArXiv 1608.00163*. Retrieved from <http://arxiv.org/abs/1608.00163>
- Fournier, P., Nadeau, R., Blais, A., Gidengil, E., & Nevitte, N. (2004). Time-of-voting decision and susceptibility to campaign effects. *Electoral Studies*, 23(4), 661–681. <https://doi.org/10.1016/j.electstud.2003.09.001>

- Fridkin, K. L., & Kenney, P. J. (2008). The dimensions of negative messages. *American Politics Research*, 36(5), 694–723. <https://doi.org/10.1177/1532673X08316448>
- Fridkin, K. L., & Kenney, P. J. (2012). The impact of negative campaigning on citizens' actions and attitudes. In H. Semetko & M. Scammell, *The SAGE Handbook of Political Communication* (pp. 173–185). United Kingdom: SAGE Publications.
- Funkhouser, R. (1973). The issues of the Sixties: An exploratory study in the dynamics of public opinion. *The Public Opinion Quarterly*, 37(1), 62–75.
- Gamson, W. A., Croteau, D., Hoynes, W., & Sasson, T. (1992). Media images and the social construction of reality. *Annual Review of Sociology*, 373–393.
- Gayo-Avello, D. (2012). I wanted to predict elections with Twitter and all I got was this lousy paper: A balanced survey on election prediction using Twitter data. *ArXiv*, 13.
- Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Social Science Computer Review*, 31(6), 649–679. <https://doi.org/10.1177/0894439313493979>
- Geers, S., & Bos, L. (2017). Priming Issues, party visibility, and party evaluations: The impact on vote switching. *Political Communication*, 34(3), 344–366. <https://doi.org/10.1080/10584609.2016.1201179>
- Geiß, S., & Schäfer, S. (2017). Any publicity or good publicity? A competitive test of visibility- and tonality-based media effects on voting behavior. *Political Communication*, 34(3), 444–467. <https://doi.org/10.1080/10584609.2016.1271068>



- Gilmore, J. (2012). Ditching the pack: Digital media in the 2010 Brazilian congressional campaigns. *New Media & Society*, 14(4), 617–633.  
<https://doi.org/10.1177/1461444811422429>
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Boston: Northeastern University Press.
- Golan, G. J. (2006). Inter-media agenda setting and global news coverage. *Journalism Studies*, 7(2), 323–333. <https://doi.org/10.1080/14616700500533643>
- González-Bailón, S., & Paltoglou, G. (2015). Signals of public opinion in online communication: A comparison of methods and data sources. *The ANNALS of the American Academy of Political and Social Science*, 659(1), 95–107.  
<https://doi.org/10.1177/0002716215569192>
- Gosnell, H. F. (1926). An experiment in the stimulation of voting. *The American Political Science Review*, 20(4), 869. <https://doi.org/10.2307/1945435>
- Gottfried, J., Barthel, M., Shearer, E., & Mitchell, A. (2016). The 2016 presidential campaign – a news event that’s hard to miss. Retrieved March 8, 2018, from <http://www.journalism.org/2016/02/04/the-2016-presidential-campaign-a-news-event-thats-hard-to-miss/>
- Graber, D. A. (2003). The rocky road to new paradigms: Modernizing news and citizenship standards. *Political Communication*, 20(2), 145–148.  
<https://doi.org/10.1080/10584600390211163>
- Green, D. P., & Zelizer, A. (2017). How much GOTV mail is too much? Results from a large-scale field experiment. *Journal of Experimental Political Science*, 4(02), 107–118. <https://doi.org/10.1017/XPS.2017.5>

- Greene, Z. (2016). Competing on the issues: How experience in government and economic conditions influence the scope of parties' policy messages. *Party Politics*, 22(6), 809–822. <https://doi.org/10.1177/1354068814567026>
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297. <https://doi.org/10.1093/pan/mps028>
- Grofman, B., Brunell, T. L., & Koetzle, W. (1998). Why gain in the Senate but midterm loss in the house? Evidence from a natural experiment. *Legislative Studies Quarterly*, 23(1), 79–89. <https://doi.org/10.2307/440215>
- Gruenfeld, D. H. (1995). Status, ideology, and integrative complexity on the U.S. Supreme Court: Rethinking the politics of political decision making. *Journal of Personality and Social Psychology*, 68(1), 5–20. <https://doi.org/10.1037/0022-3514.68.1.5>
- Guo, L., & Vargo, C. (2015). The power of message networks: A big-data analysis of the network agenda setting model and issue ownership. *Mass Communication and Society*, 18(5), 557–576. <https://doi.org/10.1080/15205436.2015.1045300>
- Hänggli, R., & Kriesi, H. (2012). Frame construction and frame promotion (strategic framing choices). *American Behavioral Scientist*, 56(3), 260–278. <https://doi.org/10.1177/0002764211426325>
- Hansen, K. M., & Pedersen, R. T. (2014). Campaigns matter: How voters become knowledgeable and efficacious during election campaigns. *Political Communication*, 31(2), 303–324. <https://doi.org/10.1080/10584609.2013.815296>

- Hayes, D., & Lawless, J. L. (2015). As Local news goes, so goes citizen engagement: media, knowledge, and participation in US House elections. *The Journal of Politics*, 77(2), 447–462. <https://doi.org/10.1086/679749>
- Henkel, L. A., & Mattson, M. E. (2011). Reading is believing: The truth effect and source credibility. *Consciousness and Cognition*, 20(4), 1705–1721. <https://doi.org/10.1016/j.concog.2011.08.018>
- Hong, S. (2013). Who benefits from Twitter? Social media and political competition in the U.S. House of Representatives. *Government Information Quarterly*, 30(4), 464–472. <https://doi.org/10.1016/j.giq.2013.05.009>
- Hopmann, D. N., Vliegenthart, R., De Vreese, C., & Albæk, E. (2010). Effects of election news coverage: How visibility and tone influence party choice. *Political Communication*, 27(4), 389–405. <https://doi.org/10.1080/10584609.2010.516798>
- Huang, H. (2010). Frame-rich, frame-poor: An investigation of the contingent effects of media frame diversity and individual differences on audience frame diversity. *International Journal of Public Opinion Research*, 22(1), 47–73. <https://doi.org/10.1093/ijpor/edp024>
- Hummel, P., & Rothschild, D. (2014). Fundamental models for forecasting elections at the state level. *Electoral Studies*, 35, 123–139. <https://doi.org/10.1016/j.electstud.2014.05.002>
- Iyengar, S. (1987). *News That Matters: Television and American Opinion*. Chicago: University of Chicago Press.

- Iyengar, S. (1990). The accessibility bias in politics: Television news and public opinion. *International Journal of Public Opinion Research*, 2(1), 1–15.  
<https://doi.org/10.1093/ijpor/2.1.1>
- Iyengar, S., Peters, M. D., & Kinder, D. R. (1982). Experimental demonstrations of the “not-so-minimal” consequences of television news programs. *The American Political Science Review*, 76(4), 848. <https://doi.org/10.2307/1962976>
- Jackson, N., & Lilleker, D. (2011). Microblogging, constituency service and impression management: UK MPs and the use of Twitter. *The Journal of Legislative Studies*, 17(1), 86–105. <https://doi.org/10.1080/13572334.2011.545181>
- Jacobi, C. A. (2016). *The Quality of Political News in a Changing Media Environment* (Doctoral Dissertation). University of Amsterdam.
- Jacobs, L. R., & Shapiro, R. Y. (2005). Polling politics, media, and election campaigns. *Public Opinion Quarterly*, 69(5), 635–641. <https://doi.org/10.1093/poq/nfi068>
- Jamieson, K. H. (1996). *Packaging the Presidency: A History and Criticism of Presidential Campaign Advertising*. Oxford University Press.
- Jennings, W., Bevan, S., Timmermans, A., Breeman, G., Brouard, S., Chaques-Bonafont, L., Palau, A. M. (2011). Effects of the core functions of government on the diversity of executive agendas. *Comparative Political Studies*, 44(8), 1001–1030.  
<https://doi.org/10.1177/0010414011405165>
- Johann, D., Kleinen-von Königslöw, K., Kritzinger, S., & Thomas, K. (2017). Intra-campaign changes in voting preferences: The impact of media and party communication. *Political Communication*, 1–26.  
<https://doi.org/10.1080/10584609.2017.1339222>

- Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, 13(1), 72–91.  
<https://doi.org/10.1080/19331681.2015.1132401>
- Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the Pirate Party won the German Election of 2009 or the trouble with predictions: A response to Tumasjan. *Social Science Computer Review*, 30(2), 229–234.  
<https://doi.org/10.1177/0894439311404119>
- Katz, E. (1957). The two-step flow of communication: An up-to-date report on an hypothesis. *Public Opinion Quarterly*, 21(1), 61–78.  
<https://doi.org/10.1086/266687>
- Katz, E., & Lazarsfeld, P. F. (1955). *Personal Influence, the Part Played by People in the Flow of Mass Communications*. Transaction Publishers.
- Kenski, K., Hardy, B. W., & Jamieson, K. H. (2010). *The Obama victory: How media, money, and message shaped the 2008 election*. Oxford University Press.
- Kiousis, S., Mitrook, M., Wu, X., & Seltzer, T. (2006). First- and second-level agenda-building and agenda-setting effects: Exploring the linkages among candidate news releases, media coverage, and public opinion during the 2002 Florida gubernatorial election. *Journal of Public Relations Research*, 18(3), 265–285.  
[https://doi.org/10.1207/s1532754xjpr1803\\_4](https://doi.org/10.1207/s1532754xjpr1803_4)
- Kleinnijenhuis, J. (1991). Newspaper complexity and the knowledge gap. *European Journal of Communication*, 6(4), 499–522.  
<https://doi.org/10.1177/0267323191006004006>

- Kleinnijenhuis, J., Schultz, F., & Oegema, D. (2015). Frame Complexity and the financial crisis: A comparison of the United States, the United Kingdom, and Germany in the period 2007-2012. *Journal of Communication*, 65(1), 1–23. <https://doi.org/10.1111/jcom.12141>
- Kleinnijenhuis, J., van Hoof, A. M. J., Oegema, D., & de Ridder, J. A. (2007). A test of rivaling approaches to explain news effects: News on issue positions of parties, real-world developments, support and criticism, and success and failure. *Journal of Communication*, 57(2), 366–384. <https://doi.org/10.1111/j.1460-2466.2007.00347.x>
- Koo, J., & Han, J. (2007). Integrative complexity in correspondence. *Journal of Pacific Rim Psychology*, 1(02), 60–69.
- Kruikemeier, S. (2014). How political candidates use Twitter and the impact on votes. *Computers in Human Behavior*, 34, 131–139. <https://doi.org/10.1016/j.chb.2014.01.025>
- Krupnikov, Y. (2011). When does negativity demobilize? Tracing the conditional effect of negative campaigning on voter turnout. *American Journal of Political Science*, 55(4), 797–813.
- LaMarre, H. L., & Suzuki-Lambrech, Y. (2013). Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'. *Public Relations Review*, 39(4), 360–368. <https://doi.org/10.1016/j.pubrev.2013.07.009>
- Lancichinetti, A., & Fortunato, S. (2009). Community detection algorithms: A comparative analysis. *Physical Review E*, 80(5). <https://doi.org/10.1103/PhysRevE.80.056117>

- Larsson, A. O., & Kalsnes, B. (2014). 'Of course we are on Facebook': Use and non-use of social media among Swedish and Norwegian politicians. *European Journal of Communication, 29*(6), 653–667.
- Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: A Meta-analytic reassessment. *The Journal of Politics, 69*(4), 1176–1209. <https://doi.org/10.1111/j.1468-2508.2007.00618.x>
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1968). *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. Columbia University Press.
- Lee, E.-J., & Oh, S. Y. (2012). To Personalize or depersonalize? When and how politicians' personalized tweets affect the public's reactions. *Journal of Communication, 62*(6), 932–949. <https://doi.org/10.1111/j.1460-2466.2012.01681.x>
- Lee, E.-J., & Shin, S. Y. (2012). Are they talking to me? Cognitive and affective effects of interactivity in politicians' Twitter communication. *Cyberpsychology, Behavior, and Social Networking, 15*(10), 515–520. <https://doi.org/10.1089/cyber.2012.0228>
- Lee, J. K., Choi, J., & Kim, S. T. (2014). All things considered? Investigating the diversity of public affairs issues that individuals think about in the Internet age. *Computers in Human Behavior, 32*, 112–122. <https://doi.org/10.1016/j.chb.2013.11.009>
- Lee, M. (2014). *Structural and cultural processes in the development of the Frankfurt School*. University of Chicago, Illinois.

- Lengauer, G., & Johann, D. (2013). Candidate and party bias in the news and its effects on party choice: Evidence from Austria. *Studies in Communication Sciences*, 13(1), 41–49. <https://doi.org/10.1016/j.scoms.2013.04.011>
- Liht, J., Suedfeld, P., & Krawczyk, A. (2005). Integrative complexity in face-to-face negotiations between the Chiapas Guerrillas and the Mexican government. *Political Psychology*, 26(4), 543–552.
- Lippmann, W. (1922). *Public opinion*. Transaction Publishers.
- Magee, C. S. P. (2012). The incumbent spending puzzle. *Social Science Quarterly*, 93(4), 932–949. <https://doi.org/10.1111/j.1540-6237.2012.00845.x>
- Maoz, Z., & Astorino, A. (1992). The cognitive structure of peacemaking: Egypt and Israel, 1970-1978. *Political Psychology*, 13(4), 647.  
<https://doi.org/10.2307/3791495>
- Maslov, S., Sneppen, K., & Zaliznyak, A. (2004). Detection of topological patterns in complex networks: Correlation profile of the internet. *Physica A: Statistical Mechanics and Its Applications*, 333, 529–540.  
<https://doi.org/10.1016/j.physa.2003.06.002>
- Matthes, J. (2009). What's in a frame? A content analysis of media framing studies in the world's leading communication journals, 1990-2005. *Journalism & Mass Communication Quarterly*, 86(2), 349–367.
- McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), 176–187.



- McCombs, M., Llamas, J. P., Lopez-Escobar, E., & Rey, F. (1997). Candidate images in Spanish elections: Second-level agenda-setting effects. *Journalism & Mass Communication Quarterly*, 74(4), 703–717.
- McCombs, M. E., & Zhu, J.-H. (1995). Capacity, diversity, and volatility of the public agenda trends from 1954 to 1994. *Public Opinion Quarterly*, 59(4), 495–525.
- Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011). How (not) to predict elections (pp. 165–171). IEEE. <https://doi.org/10.1109/PASSAT/SocialCom.2011.98>
- Mickel, A. (2016). Tracing teams, texts, and topics: Applying social network analysis to understand archaeological knowledge production at Çatalhöyük. *Journal of Archaeological Method and Theory*, 23(4), 1095–1126.  
<https://doi.org/10.1007/s10816-015-9261-z>
- Murthy, D. (2015). Twitter and elections: Are tweets, predictive, reactive, or a form of buzz? *Information, Communication & Society*, 18(7), 816–831.  
<https://doi.org/10.1080/1369118X.2015.1006659>
- Mutz, D. C. (2006). *Hearing the Other Side: Deliberative Versus Participatory Democracy*. Cambridge University Press.
- Nenkova, A., & McKeown, K. (2012). A survey of text summarization techniques. In C. C. Aggarwal & C. Zhai (Eds.), *Mining Text Data* (pp. 43–76). Boston, MA: Springer US.
- Niven, D. (2006). A field experiment on the effects of negative campaign mail on voter turnout in a municipal election. *Political Research Quarterly*, 59(2), 203–210.

- Norpoth, H., & Buchanan, B. (1992). Wanted: The education president issue trespassing by political candidates. *Public Opinion Quarterly*, 56(1), 87–99.  
<https://doi.org/10.1086/269297>
- Norris, P., Curtice, J., Sanders, D., Scammell, M., & Semetko, H. A. (1999). *On Message: Communicating the Campaign*. London ; Thousand Oaks, Calif: SAGE.
- Oegema, D., & Kleinnijenhuis, J. (2000). Personalization in political television news: A 13-wave survey study to assess effects of text and footage. *Communications*, 25(1), 43–60. <https://doi.org/10.1515/comm.2000.25.1.43>
- Pan, Z., & Kosicki, G. M. (1997). Priming and media impact on the evaluations of the president's performance. *Communication Research*, 24(1), 3–30.  
<https://doi.org/10.1177/009365097024001001>
- Park, C. S. (2013). Does Twitter motivate involvement in politics? Tweeting, opinion leadership, and political engagement. *Computers in Human Behavior*, 29(4), 1641–1648. <https://doi.org/10.1016/j.chb.2013.01.044>
- Parmelee, J. H. (2014). The agenda-building function of political tweets. *New Media & Society*, 16(3), 434–450. <https://doi.org/10.1177/1461444813487955>
- Petchler, R., & Gonzalez-Bailon, S. (2013). *Automated Content Analysis of Online Political Communication* (SSRN Scholarly Paper No. ID 2306710). Rochester, NY: Social Science Research Network.
- Peter, J., & De Vreese, C. H. (2003). Agenda-rich, agenda-poor: A cross-national comparative investigation of nominal and thematic public agenda diversity. *International Journal of Public Opinion Research*, 15(1), 44–64.  
<https://doi.org/10.1093/ijpor/15.1.44>

- Peterson, R. D. (2012). To tweet or not to tweet: Exploring the determinants of early adoption of Twitter by House members in the 111th Congress. *The Social Science Journal*, 49(4), 430–438. <https://doi.org/10.1016/j.soscij.2012.07.002>
- Petrocik, J. R. (1996). Issue ownership in Presidential Elections, with a 1980 case study. *American Journal of Political Science*, 40(3), 825. <https://doi.org/10.2307/2111797>
- Petrovic, S., Osborne, M., Mccreadie, R., Macdonald, C., & Ounis, I. (2013). Can Twitter replace newswire for breaking news? [Conference Proceedings]. Retrieved October 28, 2016, from <http://www.icwsm.org/2013/>
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205.
- Pingree, R. J., & Stoycheff, E. (2013). Differentiating cueing from reasoning in agenda-setting effects: Agenda cueing and reasoning. *Journal of Communication*, n/a-n/a. <https://doi.org/10.1111/jcom.12051>
- Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, 16(1), 101–127. <https://doi.org/10.1146/annurev-polisci-100711-135242>
- Qin, J. (2015). Hero on Twitter, traitor on news: How social media and legacy news frame Snowden. *The International Journal of Press/Politics*, 20(2), 166–184. <https://doi.org/10.1177/1940161214566709>
- Quinn, K., & Powers, R. M. (2016). Revisiting the concept of ‘sharing’ for digital spaces: An analysis of reader comments to online news. *Information, Communication & Society*, 19(4), 442–460. <https://doi.org/10.1080/1369118X.2015.1092565>

- Ramage, D., Hall, D., Nallapati, R., & Manning, C. D. (2009). Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora (Vol. 1, p. 248). Association for Computational Linguistics.  
<https://doi.org/10.3115/1699510.1699543>
- Rao, C. R. (1982). Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology*, 21(1), 24–43. [https://doi.org/10.1016/0040-5809\(82\)90004-1](https://doi.org/10.1016/0040-5809(82)90004-1)
- Roberts, M., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... Rand, D. G. (2014). Structural topic models for open-ended survey responses *American Journal of Political Science*, 58(4), 1064–1082.  
<https://doi.org/10.1111/ajps.12103>
- Roberts, M., Stewart, B., & Tingley, D. (2014). *Navigating the local modes of big data: The case of topic models*. Cambridge University Press, New York.
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3), 1059–1069.  
<https://doi.org/10.1016/j.neuroimage.2009.10.003>
- Scheufele, D. A. (2000). Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication. *Mass Communication and Society*, 3(2–3), 297–316. [https://doi.org/10.1207/S15327825MCS0323\\_07](https://doi.org/10.1207/S15327825MCS0323_07)
- Scheufele, D. A., & Iyengar, S. (2012). The state of framing research: A call for new directions. *The Oxford Handbook of Political Communication Theories*. New York: Oxford University Press.

- Scheufele, D. A., & Tewksbury, D. (2007). Framing, agenda setting, and priming: the evolution of three media effects models. *Journal of Communication*, 57(1), 9–20. <https://doi.org/10.1111/j.1460-2466.2006.00326.x>
- Schuck, A. R. T., Boomgaarden, H. G., & de Vreese, C. H. (2013). Cynics all around? The impact of election news on political cynicism in comparative perspective. *Journal of Communication*, 63(2), 287–311. <https://doi.org/10.1111/jcom.12023>
- Sellers, P. J. (1998). Strategy and background in congressional campaigns. *American Political Science Review*, 92(01), 159–171. <https://doi.org/10.2307/2585935>
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Champaign, IL, US: University of Illinois Press.
- Sheafer, T. (2007). How to evaluate it: The role of story-evaluative tone in agenda setting and priming. *Journal of Communication*, 57(1), 21–39. <https://doi.org/10.1111/j.1460-2466.2006.00327.x>
- Sheafer, T., & Weimann, G. (2005). Agenda building, agenda setting, priming, individual voting intentions, and the aggregate results: An analysis of four Israeli elections. *Journal of Communication*, 55(2), 347–365. <https://doi.org/10.1111/j.1460-2466.2005.tb02676.x>
- Shim, J., Park, C., & Wilding, M. (2015). Identifying policy frames through semantic network analysis: An examination of nuclear energy policy across six countries. *Policy Sciences*, 48(1), 51–83. <https://doi.org/10.1007/s11077-015-9211-3>
- Shoemaker, P. J., & Reese, S. D. (2014). *Mediating the Message in the 21st Century: A Media Sociology Perspective* (Third edition). New York: Routledge/Taylor & Francis Group.

- Sigelman, L., & Buell, E. H. (2004). Avoidance or engagement? Issue convergence in US presidential campaigns, 1960–2000. *American Journal of Political Science*, 48(4), 650–661.
- Smith, A., Suedfeld, P., Conway, L., & Winter, D. (2008). The language of violence: distinguishing terrorist from non-terrorist groups by thematic content analysis. *Dynamics of Asymmetric Conflict*, 1(2), 142–163.  
<https://doi.org/10.1080/17467580802590449>
- Solow, A., Polasky, S., & Broadus, J. (1993). On the measurement of biological diversity. *Journal of Environmental Economics and Management*, 24(1), 60–68.  
<https://doi.org/10.1006/jeem.1993.1004>
- Soroka, S., Bodet, M. A., Young, L., & Andrew, B. (2009). Campaign news and vote intentions. *Journal of Elections, Public Opinion & Parties*, 19(4), 359–376.  
<https://doi.org/10.1080/17457280903275030>
- Sparck-Jones, K. (1972). A Statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1), 11–21.  
<https://doi.org/10.1108/eb026526>
- Squartini, T., & Garlaschelli, D. (2011). Analytical maximum-likelihood method to detect patterns in real networks. *New Journal of Physics*, 13(8), 083001.  
<https://doi.org/10.1088/1367-2630/13/8/083001>
- Stewart, M. R., & Suedfeld, P. (2012). Psychological concomitants and predictors of violence during political upheaval in Zimbabwe: A content analysis study. *Dynamics of Asymmetric Conflict*, 5(2), 77–95.  
<https://doi.org/10.1080/17467586.2012.740680>

- Steyvers, M., & Tenenbaum, J. B. (2005). The Large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1), 41–78. [https://doi.org/10.1207/s15516709cog2901\\_3](https://doi.org/10.1207/s15516709cog2901_3)
- Stirling, A. (2007). A general framework for analyzing diversity in science, technology and society. *Journal of The Royal Society Interface*, 4(15), 707–719. <https://doi.org/10.1098/rsif.2007.0213>
- Stouffer, D. B., Camacho, J., Jiang, W., & Nunes Amaral, L. A. (2007). Evidence for the existence of a robust pattern of prey selection in food webs. *Proceedings of the Royal Society B: Biological Sciences*, 274(1621), 1931–1940. <https://doi.org/10.1098/rspb.2007.0571>
- Stromer-Galley, J. (2014). *Presidential Campaigning in the Internet Age*. Oxford University Press.
- Suedfeld, P. (1992). Bilateral relations between countries and the complexity of newspaper editorials. *Political Psychology*, 13(4), 601. <https://doi.org/10.2307/3791492>
- Suedfeld, P. (2010). The cognitive processing of politics and politicians: Archival studies of conceptual and integrative complexity. *Journal of Personality*, 78(6), 1669–1702. <https://doi.org/10.1111/j.1467-6494.2010.00666.x>
- Suedfeld, P., & Bluck, S. (1993). Changes in integrative complexity accompanying significant life events: Historical evidence. *Journal of Personality and Social Psychology*, 64(1), 124–130. <https://doi.org/10.1037/0022-3514.64.1.124>

- Suedfeld, P., Cross, R. W., & Brcic, J. (2011). Two years of ups and downs: Barack Obama's patterns of integrative complexity, motive imagery, and values. *Political Psychology, 32*(6), 1007–1033. <https://doi.org/10.1111/j.1467-9221.2011.00850.x>
- Suedfeld, P., & Jhangiani, R. (2009). Cognitive management in an enduring international rivalry: The case of India and Pakistan. *Political Psychology, 30*(6), 937–951. <https://doi.org/10.1111/j.1467-9221.2009.00736.x>
- Suedfeld, P., & Rank, A. D. (1976). Revolutionary leaders: Long-term success as a function of changes in conceptual complexity. *Journal of Personality and Social Psychology, 34*(2), 169–178. <https://doi.org/10.1037/0022-3514.34.2.169>
- Suedfeld, P., & Tetlock, P. (1977). Integrative complexity of communications in international crises. *Journal of Conflict Resolution, 21*(1), 169–184.
- Suedfeld, P., Wallace, M. D., & Thachuk, K. L. (1993). Changes in integrative complexity among Middle East leaders during the Persian Gulf crisis. *Journal of Social Issues, 49*(4), 183–199.
- Tadmor, C. T., Galinsky, A. D., & Maddux, W. W. (2012). Getting the most out of living abroad: Biculturalism and integrative complexity as key drivers of creative and professional success. *Journal of Personality and Social Psychology, 103*(3), 520–542. <https://doi.org/10.1037/a0029360>
- Tan, Y., & Weaver, D. H. (2013). Agenda diversity and agenda setting from 1956 to 2004: What are the trends over time? *Journalism Studies, 14*(6), 773–789. <https://doi.org/10.1080/1461670X.2012.748516>



- Tetlock, P. E. (1988). Monitoring the integrative complexity of American and Soviet policy rhetoric: What can be learned? *Journal of Social Issues*, 44(2), 101–131.  
<https://doi.org/10.1111/j.1540-4560.1988.tb02065.x>
- Tetlock, P. E., Metz, S. E., Scott, S. E., & Suedfeld, P. (2014). Integrative complexity coding raises integratively complex issues: Complexities of integrative complexity coding. *Political Psychology*, 35(5), 625–634.  
<https://doi.org/10.1111/pops.12207>
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544–2558.  
<https://doi.org/10.1002/asi.21416>
- Thoemmes, F. J., & Conway, L. G. (2007). Integrative complexity of 41 U.S. presidents. *Political Psychology*, 28(2), 193–226. <https://doi.org/10.1111/j.1467-9221.2007.00562.x>
- Tolochko, P. (2016). Determining political text complexity: Measurements and application. Paper Presented at the International Communication Association Conference, Fukuoka, Japan.
- Tromble, R., Storz, A., & Stockmann, D. (2017). We don't know what we don't know: When and how the use of Twitter's public APIs biases scientific inference. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3079927>
- Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. *ICWSM*, 10(1), 178–185.

- Turney, P. D., Pantel, P., & others. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37(1), 141–188.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458.
- van Atteveldt, W., Kleinnijenhuis, J., & Ruigrok, N. (2008). Parsing, semantic networks, and political authority using syntactic analysis to extract semantic relations from Dutch newspaper articles. *Political Analysis*, 16(4), 428–446.  
<https://doi.org/10.1093/pan/mpn006>
- van Hoof, A. M., Jacobi, C., Ruigrok, N., & van Atteveldt, W. (2014). Diverse politics, diverse news coverage? A longitudinal study of diversity in Dutch political news during two decades of election campaigns. *European Journal of Communication*, 29(6), 668–686.
- van Wijk, B. C. M., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS ONE*, 5(10), e13701. <https://doi.org/10.1371/journal.pone.0013701>
- Vergeer, M., & Hermans, L. (2013). Campaigning on Twitter: Microblogging and online social networking as campaign tools in the 2010 general elections in the Netherlands. *Journal of Computer-Mediated Communication*, 18(4), 399–419.  
<https://doi.org/10.1111/jcc4.12023>
- Vergeer, M., Hermans, L., & Sams, S. (2011). Is the voter only a tweet away? Microblogging during the 2009 European Parliament election campaign in the Netherlands. *First Monday*, 16(8).

- Vergeer, M., Hermans, L., & Sams, S. (2013). Online social networks and micro-blogging in political campaigning: The exploration of a new campaign tool and a new campaign style. *Party Politics*, *19*(3), 477–501.  
<https://doi.org/10.1177/1354068811407580>
- Verweij, P. (2012). Twitter links between politicians and journalists. *Journalism Practice*, *6*(5–6), 680–691. <https://doi.org/10.1080/17512786.2012.667272>
- Voakes, P. S., Kapfer, J., Kurpius, D., & Chern, D. S.-Y. (1996). Diversity in the news: A conceptual and methodological framework. *Journalism & Mass Communication Quarterly*, *73*(3), 582–593. <https://doi.org/10.1177/107769909607300306>
- Walgrave, S., Lefevere, J., & Nuytemans, M. (2009). Issue ownership stability and change: How political parties claim and maintain issues through media appearances. *Political Communication*, *26*(2), 153–172.  
<https://doi.org/10.1080/10584600902850718>
- Walker, S. G., & Watson, G. L. (1994). Integrative complexity and British decisions during the Munich and Polish crises. *Journal of Conflict Resolution*, *38*(1), 3–23.  
<https://doi.org/10.1177/0022002794038001001>
- Wallace, M. D., Suedfeld, P., & Thachuk, K. (1993). Political rhetoric of leaders under stress in the Gulf Crisis. *Journal of Conflict Resolution*, *37*(1), 94–107.  
<https://doi.org/10.1177/0022002793037001004>
- Wang, M., & Chen, Y. (2006). Age differences in attitude change: Influences of cognitive resources and motivation on responses to argument quantity. *Psychology and Aging*, *21*(3), 581–589. <https://doi.org/10.1037/0882-7974.21.3.581>

- Wanta, W., King, P., & McCombs, M. E. (1995). A comparison of factors influencing issue diversity in the U.S. and Taiwan. *International Journal of Public Opinion Research*, 7(4), 353–365. <https://doi.org/10.1093/ijpor/7.4.353>
- Williams, C. B., & Gulati, G. J. (2012). Social networks in political campaigns: Facebook and the congressional elections of 2006 and 2008. *New Media & Society*, 14(6), 1461–1481. [1461444812457332](https://doi.org/10.1177/1461444812457332).
- Yang, S., & González-Bailón, S. (2015). Semantic networks and analysis of public opinion. In J. N. Victor, M. Lubell, & A. H. Montgomery (Eds.), *Oxford Handbook of Political Networks*. Oxford: Oxford University Press.
- Yuan, E. J., Feng, M., & Danowski, J. A. (2013). “Privacy” in Semantic networks on Chinese social media: The case of Sina Weibo. *Journal of Communication*, 63(6), 1011–1031. <https://doi.org/10.1111/jcom.12058>
- Zaller, J. (1992). *The Nature and Origins of Mass Opinion*. Cambridge university press.
- Zhou, Q., & Leydesdorff, L. (2016). The normalization of occurrence and co-occurrence matrices in bibliometrics using Cosine similarities and Ochiai coefficients. *Journal of the Association for Information Science and Technology*, 67(11), 2805–2814.
- Zywica, J., & Danowski, J. (2008). The faces of Facebookers: Investigating social enhancement and social compensation hypotheses; Predicting Facebook and offline popularity from sociability and self-esteem, and mapping the meanings of popularity with semantic networks. *Journal of Computer-Mediated Communication*, 14(1), 1–34. <https://doi.org/10.1111/j.1083-6101.2008.01429.x>