ROLE OF MEDIA IN FUELING AFFECTIVE POLARIZATION

ANALYSIS OF ONLINE NEWS ABOUT TURKEY’S OPPOSITION PARTY

CEMALETTİN YILMAZ

BOĞAZİÇİ UNIVERSITY

2023

ROLE OF MEDIA IN FUELING AFFECTIVE POLARIZATION

ANALYSIS OF ONLINE NEWS ABOUT TURKEY'S OPPOSITION PARTY

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Political Science and International Relations

by

Cemalettin Yılmaz

Boğaziçi University

2023

DECLARATION OF ORIGINALITY

I, Cemalettin Yılmaz, certify that

- I am I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature: ________________

Date: ____________________

ABSTRACT

Role of Media in Fueling Affective Polarization Analysis of

Online News About Turkey's Opposition Party

Today, amid the absence of major election-day frauds and overtly violent repression, autocratizing incumbents depend on some level of genuine popular support for political survival. Affective polarization and negative partisanship are considered crucial factors for this support. This thesis attempts to contribute to this line of research by tracking the evolution of content in the media that could fuel this polarization in the Turkish context. It analyzes more than 45.000 online news articles about the main opposition party published since 2009 in two newspapers. Focusing on news from pro-government Sabah and centrist Hürriyet, machine learning algorithms predict the amount of language that portrays opposition as an enemy or rival within politics. Together with the change in ownership of centrist outlet to a pro-government conglomerate in 2018, and the establishment of hyper-presidential system in Turkey, the analysis establishes the long-term trajectory of such language, its changes during election periods, and how it is affected by these junctures across the newspapers. Findings show that vilifying language increases during most election periods for both newspapers, that the pro-government newspaper uses more vilifying language, and that both newspapers got more vilifying towards the opposition after Turkish regime change. In addition, with one newspaper increasingly vilifying while the other reducing the overall coverage of opposition after its change in ownership, it also shows that the influence of incumbent-control over different friendly news outlets may be heterogeneous, depending on the audience.

ÖZET

Duygusal Kutuplaşmayı Körüklemede Medyanın Rolü: Türkiye'de Muhalefet Partisi Hakkındaki İnternet Haberlerinin Analizi

Bugün, büyük çapta seçim hilelerinin ve alenen şiddetli baskıların olmadığı bir ortamda, otokratikleşen iktidarların siyasi olarak ayakta kalabilmesi en azından bir düzeyde gerçek halk desteğine bağlıdır. Duygusal kutuplaşma ve siyasi partilere karşı negatif kimliklenme bu destek için önemli faktörler olarak kabul edilir. Bu tez, Türkiye bağlamında bu kutuplaşmayı körükleyebilecek medyadaki içeriğin gelişimini izleyerek bu araştırma hattına katkıda bulunmayı amaçlamaktadır. Tez, ana muhalefet partisi hakkında 2009'dan bu yana iki gazetede yayınlanan 45.000'den fazla çevrimiçi haberi analiz etmektedir. Hükümet yanlısı Sabah ve merkezci Hürriyet haberlerine odaklanan makine öğrenimi algoritmaları, muhalefeti siyasette düşman veya rakip olarak gösteren dili ölçümlemiştir. 2018'de merkezci Hürriyet'in hükümet yanlısı bir holding tarafından satın alınması ve Türkiye'de hiper-başkanlık sisteminin kurulmasıyla birlikte, analiz, bu dilin uzun vadeli yörüngesini, seçim dönemlerindeki değişimlerini ortaya koymaktadır. Bulgular, her iki gazete için de çoğu seçim döneminde karalayıcı dilin arttığını, iktidar yanlısı gazetenin daha fazla karalayıcı dil kullandığını ve her iki gazetenin de rejim değişikliğinden sonra muhalefete yönelik aşağılayıcı söylemlerinin arttığını göstermektedir. Ayrıca, bir gazetenin sahibi değiştikten sonra muhalefet hakkındaki haberlerin azalması ve diğerinin muhalefete giderek düşmanlaşması, medyanın iktidara yakın sermaye grupları tarafından kontrolünün içeriğe etkisinin, izleyicilere bağlı olarak heterojen olabileceğini de göstermektedir.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my thesis advisor, Alper Yağcı, who contributed greatly to the design of the study by pushing me to novel methods and guiding me through the concepts and methods applied in this thesis. I am greatly indebted to him for the opportunity to turn this thesis into an actual academic paper with the help of others. It was his crucial methodological and conceptual interventions, and influential discussions that lead this study into something tangible.

I always felt like whenever I was writing an assignment, a term paper, or this thesis, narrowing down the question to manageable and studiable bits was the most difficult part. Mine Eder's insightful guidance proved crucial for the process and if not for our initial discussions with her, all the thoughts in my head would be nothing more than a mess.

It was my brief but influential first meeting with Ali Çarkoğlu, which was possible thanks to Alper Yağcı and Mine Eder, that transformed the initial ideas in my head into a media analysis that measures polarizing content. Başak Taraktaş lead me to learn computer coding, insisting this project was possible with no prior experience. She helped me acquire the skills which proved essential for this thesis and also referred me to Uzay Çetin. His guidance and expertise have led us to the algorithm we used in this thesis which would otherwise be much more challenging. I am also indebted to Ezgi Elçi who offered critical assistance and guidance when the work on this thesis was at a crossroads and did not know where to go.

I would also like to acknowledge the contributions of the POLS 509 course of Mert Arslanalp. The discussions there inspired the conceptual framework. If not for the invaluable discussions here, it would be impossible for me to bring together the thesis in this way. As someone who took all graduate courses over Zoom during COVID, this course I was following along face to face with friends was a blessing during my last term in the university.

The moment in which I am writing this acknowledgments page, represents for me much more than the completion of this thesis. For me, it puts an end to an era in my life and opens a new one for me and my family in a country far away. All the professors and friends I mentioned above, and countless others have also given me their much-needed support to even think about pursuing an academic career and find myself in this moment. Rabia Kutlu and Gül Çatır was immensely motivating during the graduate and scholarship application process through which they never hesitated to offer their generous help. Yeşim Arat and Şevket Pamuk, in addition to Mine Eder, whom I admire and look up to greatly, gave me exceptional support for continuing my academic journey abroad and provided excellent advice for navigating it.

To Aybala, Ceren, and Yusuf, I should express my heartfelt admiration and appreciation. In my eight years in this school, for the first time, it was them with their companionship during the past year who gave me a glimpse of what living as a Boğaziçi student must feel like. You have all affected me in profound ways with all your qualities which, if I am to admit, I have often envied you for. Enes and I joined the master's program and assistantship at the university at the same time and struggled together throughout the graduate applications and master's courses and I appreciate the times we vented off and complained together. Lastly, I would like to mention my colleagues and friends Sümeyye, Tayfun and Umut. My life and days at the office would be much less inspiring and much more boring without them. I thank my friends for being with me along this journey so far, enriching me personally and professionally at every step. And I hope they stay with me for the rest of it.

My dear parents, Mustafa and Necmiye - their love, support and encouragement are what made my journey possible. Without their unwavering belief and sacrifices, their endless support, it would be impossible for me to imagine myself as a future scholar and write this page. I hope you think you have raised a good kid.

I think I thank my wife Damla the most, the source of meaning for me with everything she brings to this otherwise empty life, including our pets, Bruce and Robin. If I have the slightest idea about what I want to do in this life, it is thanks to her. She has been an amazing partner, wonderfully right here, by my side, bringing so much passion and joy for the past eleven years. I could not imagine going abroad to a new and unknown life if she wasn't here with me. I'm incredibly fortunate to have you in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDIX TABLES

LIST OF FIGURES

LIST OF APPENDIX FIGURES

CHAPTER 1

INTRODUCTION

How democracy breaks down today is markedly different from previous periods. Not only its more conspicuous forms such as military coups, election-day frauds, or executive takeovers are less likely to be observed in the world today, but also the backsliding process unfolds gradually and over a long period of time, with the process itself being legitimized through democratic institutions (Bermeo, 2016). Although backsliding has been usually conceptualized as “incremental, within-regime change” (Waldner & Lust, 2018, p. 95), in some cases the extent of the incremental change is so large that it becomes difficult and misleading to categorize these regimes as democracies. Such regimes, often resulting from this sustained but gradual decline of democracy and demonstrating “hybrid” features blending a minimum of procedurally democratic institutions with authoritarian forms of governance, have been studied under concepts such as competitive authoritarianism, electoral autocracies, and others (Gandhi & Lust-Okar, 2009).

Today, these regimes constitute the most common form of authoritarianism, operating “behind the institutional facades of representative democracy” (Schedler, 2013, p.1). Given the “hybrid” nature of these regimes that rely on democratic-looking elections, popular support maintaining undemocratic incumbents in power is crucial. Amid the absence of major election-day frauds and overtly violent repression (Bermeo, 2016), these incumbents have to muster some level of genuine support from society to stay in power. Under the light of growing evidence that political polarization might be crucial in generating this support, this thesis attempts to measure negative rhetoric launched towards political opposition in the news media which could potentially polarize the electorate along affective lines (Iyengar, Lelkes, Levendusky, Malhotra, & Westwood, 2019) and contribute to popular support for authoritarian incumbents.

More specifically, we focus on the Turkish case, which is perhaps the country that experienced one of the biggest downfalls in quality of democratic governance in the recent years. Analyzing more than 45.000 online news articles about the opposition across two of the most popular newspapers, Sabah and Hürriyet, we attempt to gain insights into the content about opposition in news media during autocratization. Two types of negative rhetoric used in these articles against the main opposition *Cumhuriyet Halk Partisi* (Republican People's Party, CHP) in Turkey is measured: portrayal of CHP (a) as an “enemy,” an illegitimate and undemocratic player in the political game, or (b) as a “rival” of the incumbent party, an opponent whose competence and leadership credentials are put to question. By analyzing the above data in light of the Turkish media landscape and Turkey’s regime trajectory, this thesis argues that pro-government news media in electoral authoritarian regimes could fulfill two distinct but interrelated functions. On the one hand, it launches an “enemy” rhetoric against the political opposition which denies them legitimacy and morality, enabling the ground on which authoritarian actions can be justified to the public. On the other hand, pro-government news media in electoral autocracies can help deplatform the political opposition by allocating little coverage to it, without explicitly denouncing its legitimacy and declaring it an enemy in the political arena.

This thesis attempts to observe empirically both functions at work in Turkey’s two of the largest newspapers, Sabah and Hürriyet. Sabah has been an ardent supporter of the incumbent government of *Adalet ve Kalkınma Partisi* (Justice and Development Party, AKP) while Hürriyet was neutral -if not often critical of the government- until its acquisition by a pro-government conglomerate, Demirören Holding in 2018. With a machine-learning based text classification approach utilizing BERTurk (Schweter, 2020), this thesis analyses all of the national news articles about CHP published online on the websites of these two news outlets between 2009 and 2022. Operating at the sentence level, algorithms predict how many sentences depict CHP as an “enemy” or “rival” each day on Sabah and Hürriyet and compare their levels across time and across the news outlets.

Findings reveal novel patterns across incumbent-friendly media, pointing at divergent effects of incumbent influence over media. We show that while the pro-government newspaper Sabah increasingly depicts CHP as an “enemy” rather than “rival” since 2009, centrist outlet Hürriyet goes considerably silent in its coverage of the opposition which declines dramatically after its acquisition by pro-government Demirören Holding in 2018. In addition, findings show that news articles published during election periods contain generally more criticisms of the opposition compared to periods adjacent to election periods, and that the news articles on both outlets gets more critical of the opposition after the establishment of the hyper-presidential system in Turkey in 2017.

The research contributes to academic literature in three ways. First, complementing the studies that attempt to understand the relationship between polarization and autocratization, it implies that the relationship is cyclical. In addition to polarization creating the ground to enable authoritarian actions and policies, incumbents aiming to initiate or intensify autocratization may also deliberately fuel polarization through friendly news media, influenced and controlled through carrots and sticks.

Secondly, it generates a large dataset of around 45,000 news about the opposition party CHP and analyzes them in terms of how they portray opposition. This dataset contains all of the text-based news content published about CHP nationally on Sabah and Hürriyet, two of the largest newspapers in Turkey between 2009 and 2022. Third, this study also offers a new methodological approach to measure media content. It constructs a novel measure of measuring potentially polarizing rhetoric towards the political opposition which make possible tracking changes in the language media uses over large periods. For each day, machine learning algorithms predict how many sentences depict CHP as either a rival, or an enemy on news items from Sabah and Hürriyet. These enemy and rival categories of criticism are aggregated from depictions of CHP as an immoral, illegitimate, incompetent, or incohesive player of the political game. Former two depictions constitute the enemy dimension which portrays opposition as an "enemy" of the people and of the state, while the latter two dimensions constitute the rival dimension which criticizes the opposition, but nevertheless acknowledging it as a legitimate and valid actor of politics. The machine-learning process employed here can be extended to cover additional components of negative and positive rhetoric with relatively small effort.

Next chapter presents a survey of the existing literature on the conceptualization of political polarization, its relationship to democratic backsliding and autocratization, and how it can be affected by media content, in addition to measuring polarization on the news media. Chapters Three and Four detail the case, hypotheses, and the research design. Chapter Five presents and discusses the findings of the analysis conducted, while Chapter Six concludes the thesis.

CHAPTER 2

LITERATURE REVIEW

This chapter provides a survey of conceptualization of political polarization, the relationship of these conceptions to democratic backsliding and autocratization, the relationship between media and political polarization, and how polarization is measured in media content across a variety of research designs.

Following these conceptual and methodological elaborations, the chapter explains the interventions of the thesis to the existing literature on the relationship between autocratization and polarization, its contributions to it and to the measurement of polarization in news media. Then, following chapter details how these measurements are conceptualized and applied in this study.

2.1 Political Polarization and Democratic Backsliding

The literature on political polarization usually distinguishes ideological polarization from affective polarization. While ideological polarization refers to the diversity of policy preferences and different positions taken up in issue areas (Dalton, 1987), affective polarization operates as an extension of social identity (Iyengar et al., 2019), and influences political behavior via like and trust towards one's own party and distrust and dislike towards other parties (Iyengar, Sood, & Lelkes, 2012; Kubin & von Sikorski, 2021).

Both types of polarization have been associated with consequences harmful to democracies. In the US context, Svolik (2020) demonstrates that voters are more willing to let go off democratic preferences for partisan economic interests when society is polarized, therefore illustrating “why polarized democracies vulnerable to subversion by elected politicians” (p. 24). Gidengil et al. (2021) find evidence of a similar trade-off between ideological agendas on abortion and upholding democratic checks and balances in the US and Canada. In the Turkish context, Şaşmaz et al. (2022) find that affective polarization increases “elastic support” for executive aggrandizement, that is, voters support authoritarian changes because when they believe it will facilitate the party they support in enacting their partisan vision.

Not only differences in ideological and partisan positions like those above, but also affective ties to political leaders and parties have been found to be contributing to degradation of democratic qualities of governance. More particularly, negative partisanship has been shown to have detrimental consequences for democracy. Negative partisanship, or negative party identification with a party, can be defined as “an affective repulsion from that party” which is “more stable than a current dislike and more strongly held than a passing opinion” (Caruana, McGregor, & Stephenson, 2015, p. 772). Abramowitz and Webster (2016) find that Americans dislike and hold negative views of the opposing party at unprecedented levels, even though their like for own party has remained at similar levels. They note that the rise of negative partisanship has contributed to higher party loyalty, straight-ticket voting, politics of confrontation and gridlock, while also having deleterious consequences for governance, representation, and bipartisanship in policymaking.

Comparative evidence from advanced democracies suggests that the rise of negative partisanship, and its influence over political preferences are not unique to American or two-party system contexts. Ridge (2020), looking at forty-three elections in established democracies between 2006 and 2011, shows that negative partisans are less satisfied with democracy, and negative party identification thus has the potential to undermine democratic systems of government. Garzia and Ferreira da Silva (2021) find that negative identification with political leaders has been increasing over the past six decades and is influential for voting behavior in fourteen Western European democracies. They underline the possible consequences of voting primarily against a candidate in degrading the quality of democratic representation.

Although most studies explore affective polarization and negative partisanship in established democracies and especially in the United States, these factors have been receiving increasing attention in weakly or non-democratic contexts as well, including Turkey. Laebens and Öztürk (2021) find, similarly in democracies, partisanship affects political behavior and vote choice in Turkey. In electoral autocracies, even though clientelistic linkages may downplay the role of partisanship in influencing political action, partisanship nevertheless contributes to the reinforcing of threat perceptions to one's well-being from other party supporters (Laebens & Öztürk, 2021). Şaşmaz et al. (2022) also find in the Turkish context, partisans can care less about democracy and support authoritarian changes when they are affectively polarized and believe such changes will help their leader with their preferred policies.

Another related research agenda that focuses on the detrimental effects of polarization for democracy, comes from the work of McCoy et al. (2018) who argue that similar processes of polarization can be encountered in democratic and non-democratic contexts. They conceptualize political polarization more in line with the definition of affective polarization rather than ideological or policy-based: "the severe polarization we examine includes a significant affective dimension, when distance between groups moves beyond principled issue-based differences to a social identity" (p. 19). Severe polarization accompanies the increasing affective distance between groups, reinforcing Manichean perceptions of each other and dividing the society into “us versus them” camps with leaders mobilizing the public by increasing the salience of existing societal cleavages (McCoy et al., 2018).

2.2 Media and Political Polarization

Given the importance of affective polarization in influencing voter behavior and its potentially damaging consequences for democratic governance, a host of studies explore how it emerges. Iyengar et al. (2019) identify several in the US context. Partisan sorting, in which voters tend to “sort” themselves into parties according to their ideology, has been on the rise in the past 50 years, contributing to the social distance between partisans of different parties by declining cross-cutting identities (Levendusky, 2010; Levendusky & Malhotra, 2016; Mason, 2015, 2018; Roccas & Brewer, 2002). Iyengar et al. (2019) infer that these make inaccurate stereotypes about the other party supporters much more likely, and thus contribute to the reinforcing of social identities and increasing affective polarization. In addition, ideological polarization is found to be related to affective polarization (Iyengar et al., 2019), with ideological extremity increasing affective polarization (Bougher, 2017).

More relevant for this paper, however, is the influence of news media content on political polarization. Although the early works exploring the effects of media often argued that their effects on political choices were minimal, more recent, and more sophisticated studies point to the substantial and frequent effects of media over political behavior (Stecula, 2018). In the context of negative partisanship, as negative party identification is “resilient in part because it entails selective information gathering and processing that is capable of overriding rational updating" (Caruana et al., 2015, p. 772), media effects emerge as crucial. Kubin and von Sikorski's recent review (2021) reveals that most studies that investigate the effects of media on political polarization focus on the US context and mostly look at social media, especially Twitter, rather than traditional media. Although there are forms of exposure that depolarize or have no relationship to polarization (Kubin & von Sikorski, 2021), interactions with news media, in particular the consumption of like-minded rather than cross-cutting news content (Levendusky, 2013), negative political advertisements (Lau, Andersen, Ditonto, Kleinberg, & Redlawsk, 2017), incivility on out-party sources (Druckman, Gubitz, Levendusky, & Lloyd, 2019), news about in-party scandals (Rothschild, Keefer, & Hauri, 2021) increase affective polarization (Kubin & von Sikorski, 2021).

The role of media in contributing to public support for authoritarianism and backsliding is not limited only to its effects through increasing affective polarization. Diverse strands of works studying authoritarianism frequently emphasize the role of media in influencing the public opinion through various mechanisms and underline the importance of its independence from political interests for healthy democratic processes (e.g. Bermeo, 2016; Esen & Gumuscu, 2016; Guriev & Treisman, 2019; Levitsky & Way, 2020; Levitsky & Ziblatt, 2018; Mechkova et al., 2017; Waldner & Lust, 2018). When autocrats have control over the media, they can achieve dominance "by distorting information flows," and if successful, "many citizens do not realize that they are being dominated" (Guriev & Treisman, 2019, p. 123). These distortions can take up different forms. Presenting information with partisan cues is more effective in changing individual's opinion of about electoral rules, meaning that strong partisanship can be utilized to bend institutional rules in undemocratic ways (Ahlquist, Ichino, Wittenberg,

& Ziblatt, 2018). In addition, perceptions of economic performance can be biased by partisan links and media choices, and even when controlled for partisanship, pro-government news can significantly bias assessments of the economy positively, which translates into votes for the incumbent (Yagci & Oyvat, 2020).

2.3 Measuring Polarization on Media

Despite the emphasis on the importance of media during the autocratization process, and its significance in altering affect towards out-groups, perceptions about the economy, or institutional changes, research that explicitly and systematically investigates its content, especially on the traditional news sources, is relatively rare (Stecula, 2018). Those that do so usually report partisan bias in a given issue area such as COVID-19 (Hart, Chinn, & Soroka, 2020), climate change (Merkley & Stecula, 2018), or abortion (Carmines, Gerrity, & Wagner, 2010) and use manual or automated content analysis procedures to produce insights about how news content has evolved over time.

In the US context, other studies look at political news from different media outlets and compare patterns in terms of content categories and partisan bias. Budak et al. (2016) compare partisan slant and categories of political news across fifteen online news outlets via crowdsourced content analysis and find no substantial difference in descriptive news articles of major platforms. In addition, when there is partisan bias, it is reflected not in favorability towards the party ideologically closer to the news outlet, but in disproportionate criticism of the other party (Budak et al., 2016). Stecula (2018) looks at 600,000 news articles over a broad range of political topics through using dictionaries to detect tone, conflict and cooperation, and Linguistic Inquiry and Word Count (LIWC) software to measure affect and incivility. While most of these measures are found to be stable since 1980s, one significant change is that more extreme politicians receive increasingly disproportionate amount of attention over time, and Stecula concludes this has likely contributed to polarization in the US.

In the Turkish context, some studies focus on polarization on issue areas as well. Kurt (2016) investigates six national newspapers in Turkey for their coverage of the nation-wide Gezi protests for three months during which the protests were ongoing. Sources quoted in the news articles about the protests, and the article's tone (positive, negative, or neutral) towards the source are analyzed. She finds that there is a polarization in tone towards news sources which converges with the newspapers' political outlooks. Kalkar and Öcal (2020) look at the establishment of the presidential system in Turkey by comparing the first pages of neutral, pro-government and pro-opposition newspapers. Starting from three months before the referendum which approved the institutional change, they find pro-government newspapers do not provide any critical commentary on the issue, in addition to often using derogatory language and defamation against opposition elites. Opposition newspapers on the other hand, although mostly reporting viewpoints opposed to the presidential system, present some extent of pro-government arguments (Kalkar & Öcal, 2020). Şirvanli (2021), similarly, compares one mainstream and one alternative newspaper that could be defined as politically centrist in terms of their coverage of COVID-19. Looking at whether scientists or politicians are quoted, and whether the arguments presented in the articles resemble those made by the incumbent and opposition, he finds that politicians are disproportionately quoted in both outlets, and that while the mainstream newspaper more closely resemble the incumbent's attitude towards COVID-19 measures, the alternative one is more in line with opposition's arguments about the topic. He concludes, even when the issue is arguably non-political and even in newspapers that could be said to be at the ideological center, political polarization permeates coverage in both mainstream and alternative sources (ŞiRvanli, 2021).

A related line of work focuses on the coverage of political parties in Turkey which is the literature on press-party parallelism. Çarkoğlu et al. (2014) compare fifteen newspapers grouped under opposition, incumbent and mainstream outlets in terms of voice given to and favorability towards the incumbent and main opposition party for three months before the general elections of 2011. Again, polarization is found in terms of voice allocation and tone of coverage, in line with the political outlooks of these newspapers. As the election approaches, pro-incumbent newspapers favor the incumbent and disfavor opposition, and vice versa in opposition papers (Çarkoğlu et al., 2014). Yıldırım et al. (2021) following a similar research design, look at four general election periods in 2002, 2007, 2011 and 2015 to see the evolution of media coverage under competitive authoritarianism in Turkey. They manually code the front-pages of thirteen newspapers and track them for 60 - 90 days prior to the election day in each period. Yıldırım and colleagues find that negative mentions of opposition parties increase at each subsequent election while the incumbent party establishes dominance in terms of both visibility and favorability.

Studies surveyed above suggest that media is an influential factor both as a source of information and has the potential to induce negative affect towards political parties and leaders. Line of studies coming from the democratic backsliding and autocratization literature also suggest that this type of polarization has been important to understand the contemporary decline in democratic qualities of governance around the world. In addition, media's role has also been noted by this strand of works, which suggest control over media is an important tool from the authoritarian handbook.

When it comes to the studies that look at polarization in the media, most of the literature focuses on social media platforms such as Twitter due to easy access to data and widespread usage. Nevertheless, polarizing and negative content on traditional news media has also been studied, usually around selected topics, or within set timeframes, such as election periods. To help understand the relationship between polarization and autocratization, this thesis takes a broader view on such content in the media and analyzes all the news about CHP in Turkey on two most popular news websites in terms of negative content the published over a 13-year period since 2009, regardless of theme or topic, utilizing machine learning.

2.4 Contributions of the Study

Contributions of this study are arguably three-fold. First, its findings can be read as evidence of the cyclical nature of the relationship between polarization and autocratization contributing to the theoretical debates in the literature. Most studies emphasize and draw the causality between polarization and autocratization as the former providing the ground for the latter while some acknowledging the more nuanced and sometimes reciprocal relationship between the two (McCoy et al., 2018). The present study provides preliminary empirical evidence that suggests a more nuanced and reciprocal relationship, in that authoritarian elites also fuel polarization deliberately which can be observed through language used against the opposition in the media. It focuses on Sabah, an ardent supporter of the government since 2008, and on Hürriyet centrist if not opposed to the government until it is captured by AKP-friendly conglomerate Demirören Holding in 2018. It shows that after 2018, under the hyper-presidential system of Turkey, two newspapers controlled by different pro-government conglomerates diverge in their coverage of CHP hinting at two possible distinct functions of pro-government media in electoral authoritarian regimes. In the Turkish context, findings reveal that, one newspaper, Sabah, arguably helps declare CHP as an internal enemy to justify the autocratization undergoing in the hands of the incumbent. While the other, formerly centrist Hürriyet, deplatforms the opposition by lowering the amount of coverage it allocates to CHP after it is captured by the government-friendly conglomerate. Overall, for the readers of these newspapers, CHP either receives little interest or shown often as an enemy of the people and Turkey. Findings show an increasingly negative portrayal of CHP with increasing emphasis on the enemy dimension of coverage in Sabah. While in Hürriyet, the space allocated to CHP is significantly reduced, reducing the chances for opposition to defend itself against the rhetorical attacks of the incumbent. Taken together, while societal polarization enables the ground for the justification of authoritarian actions of the incumbent, this study provides evidence that incumbent-friendly actors also either deplatform or vilify the opposition, arguably to contribute to societal polarization and block outreach channels for democratic opposition which helps justify autocratization to the public.

Empirically, it generates a large dataset of around 45,000 news about the opposition party CHP and analyzes them in terms of how they portray the opposition. This dataset contains all the text-based news content published nationally on the websites of Sabah and Hürriyet, two of the largest newspapers in Turkey about CHP between 2009 and 2022. This dataset can be utilized through various textual data analysis methods as it contains important information about the opposition, such as the opposition’s response to political, social, and economic developments over a large period of time.

Methodologically, it proposes novel measures that can measure the amount of negative content about political opposition and differentiate between different types of it. Four machine learning algorithms working at sentence level, predict whether or not a given sentence portrays CHP as an immoral, illegitimate, incompetent, or incohesive player in politics. The number of sentences predicted as such are then aggregated daily to compare different periods across the same news outlet or normalized in terms of total content published by Hürriyet and Sabah to compare news outlets for the same time periods. Machine-learning approach allows for easy replication and performance measurement, and can be applied to other country cases, different textual content, or extended with additional dimensions of content that could induce positive or negative sentiments towards political parties, potentially offering a way to empirically complement the studies that look at the relationship between partisan attachments and regime change more broadly.

Analyses focusing on similar measurements of media content, usually has relied on samples from narrower timeframes using extensive resources such as Yıldırım et al. (2021) and Çarkoğlu et al., (2014) and Somer (2010). A machine-learning approach allows for building on and broadening the time scope of such works covering a longer period. Especially today, where outright censoring and suppression are less common than other, “behind the scenes” tactics (Mechkova et al., 2017, p. 166), we need novel measures capable of tracking changes in the media other than outright censoring. Given that in emblematic cases of electoral authoritarian regimes, or countries which have experienced tremendous declines in their democratic governance like Turkey and

Hungary, the autocratizing incumbent has significant influence over the media, these novel measures should be able to delve deeper into a substantial amount of content about the opposition to understand how the coverage of opposition changes during episodes of autocratization which is nowadays a gradual and long-term process. The measures proposed and tested in this study could potentially lead to the development of a more comprehensive measure which could undertake this task, in addition to revealing patterns in the media coverage of opposition in Turkey.

CHAPTER 3

CONCEPTUAL FRAMEWORK

Studies cited in the previous chapter and other works in the literature that examine the relationship between polarization and autocratization, generally put forward polarization as the leading force behind autocratization, while some noting reciprocal effects between the two processes. The data generated for this study, on its own, is not enough to present a decisive conclusion to this question. But when coupled with Turkey's regime trajectory and media ownership structure in Turkey, findings of the study strengthen the argument that the relationship between polarization and autocratization is reciprocal. Findings also show that pro-government media can fulfill roles that help vilify political opposition and reduce the channels it can contest autocratization publicly to tilt the playground in favor of the incumbent. Rather than positing a unidirectional relationship from polarization to autocratization, findings of this study suggest that the relationship is instead cyclical, and polarization may also be deliberately fueled by authoritarian incumbents or by businesses friendly to the leading political elite, possibly to enable the ground for and muster popular support behind authoritarian measures.

This study looks for two types of negative rhetoric launched against the opposition in news articles published on the main opposition party in Turkey, CHP, since 2009. We operationalize and quantify the amount of these rhetoric types from the information at the sentence level. Relying on machine-learning algorithms, we attempt to detect whether each sentence that mentions the main opposition party, CHP, belongs to either "enemy" or "rival" type of rhetoric towards it, or neither. The enemy type of rhetoric consists of two components: CHP's portrayal as an illegitimate or immoral political actor whereas rival type consists of another two components: CHP's portrayal as an incompetent or incohesive political actor, which is visualized in Figure 1.

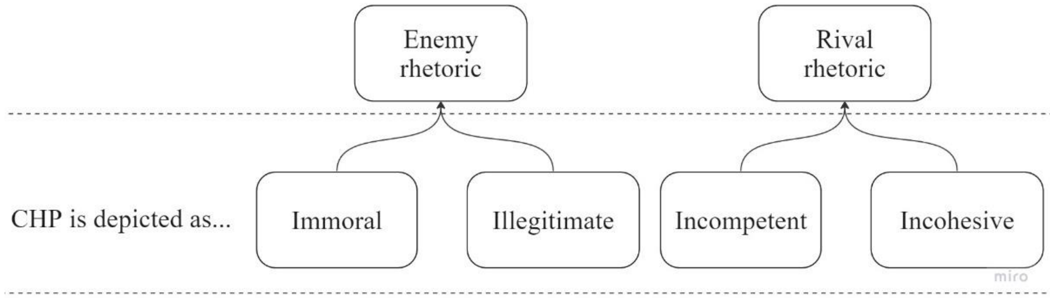

Figure 2. Conceptualization of negative rhetoric against CHP.

3.1 Opposition as an enemy versus a rival of the incumbent

Relying on the strands of literature on polarization and backsliding, this study suggests an approach to measure the level of polarization in news media and attempts the understand the findings in the light of the authoritarian transformation of Turkey and changes in the media ownership. It focuses on two types of negative rhetoric of the opposition in news media, one which portrays it as an enemy, and the other as a rival. The conceptual line between the two types is inspired by Schmitt's (2005) friend-enemy distinction, and what counts as a valid opponent in politics versus an illegitimate enemy that needs to be destroyed, and the emergency politics literature that builds upon Schmitt's thought (e.g., Agamben, 2005; Lührmann & Rooney, 2021).

Depiction of the political opposition as an enemy, without legitimacy within the political space, can enable and justify authoritarian rule through three mechanisms. First, portraying the political opposition as an enemy often works through characterizing the opposition politicians and voter base as enemies of the state that poses a danger to the existence of the society and the nation (Levitsky & Ziblatt, 2018, p. 24). Autocratizing incumbents can claim that their opposition is not a mere opponent and adversary, but an existential and ontological threat.

Second, this depiction of the political opposition contributes to the “us versus them” dynamic (McCoy et al., 2018). Portraying the legitimate opposition parties as hostile forces that seeks to undermine the nation and state, authoritarian incumbents build the image that they are the only legitimate representatives of the “people,” creating a sense of identity among their voters in electoral authoritarian regimes. By dividing the public into us versus them camps, authoritarian incumbents aim to consolidate their support by inducing feelings of loyalty to the nation and the state in their popular base, who may feel that they have a duty to uphold what they perceive to be the national interests, and muster behind authoritarian actions that aim to eliminate or neutralize the political “enemy.”

Finally, these processes arguably contribute to the sense of urgency and emergency. The perceived presence of internal or external enemies helps create the need for immediate action. The resulting emergency then can further prepare and strengthen the ground on which authoritarian actions can be taken (Arslanalp & Erkmen, 2020). The sense of urgency contributes to the justification of authoritarian measures which would have been perhaps opposed by the public under “normal” conditions, to be interpreted as necessary under emergency conditions.

Keeping these mechanisms in mind, this thesis focuses on and conceptualizes two types of content that could induce negative affect towards the opposition parties. One of these types portray opposition as a valid and legitimate “rival” within the political arena, while the other portrays the opposition as an “enemy” arguably contributing to the spiral of “pernicious polarization” as described by Somer (2019). These two main types of depictions of the opposition are aggregations of four ways in which the opposition could be portrayed in news media. These are portrayals of the opposition party in the media as an immoral, illegitimate, incompetent, or incohesive player of the political game. The former two representations form the portrayal of the opposition as an “enemy,” while the latter two form the “rival” dimension.

More specifically, this study compares the levels of depictions of CHP as an enemy and rival across Hürriyet and Sabah, before and after the transition of Turkey to a hyper-presidential system, and before and after the changes in the media ownership of Hürriyet. The results will show that while Sabah allocates significantly more visibility to CHP and portrays it as “enemy” significantly more as autocratization unfolds, Hürriyet significantly lowers its coverage of CHP after its capture by incumbent friendly Demirören Holding.

We explain below the coding rules used to create the training sets for the algorithms to predict such sentences in all news in Sabah and Hürriyet about the opposition since 2009, and then discuss the data collection and analysis process and findings. The code used for scraping the news articles and training the machine learning algorithm can be found in Appendix A.

3.2 Opposition as the enemy

3.2.1 Illegitimacy Component

For the illegitimacy component of polarizing rhetoric, we focus on sentences which portray the main opposition party, CHP, as an illegitimate political actor. More specifically, sentences containing similar claims to categories below, are coded as instances where CHP is shown as illegitimate.

CHP plotting or supporting coup plans, and/or undermining democracy: Such sentences claim that CHP is upholding the legacy of the military coups in Turkey, or as having the same mindset and actively supporting the past military interventions. In addition, implications of CHP’s cooperation in or support of coup plans against the AKP government are also coded as illegitimate mentions of CHP.

CHP working with terrorist groups and/or their supporters: These sentences claim that CHP is cooperating with or supporting groups recognized by Turkey as terrorist organizations, such as the Kurdish PKK, KCK, and Gülenist FETÖ. I also include claims that CHP is collaborating with or advocating for the Kurdish party in the parliament, HDP. AKP government and allies since the breaking down of peace talks on the Kurdish issue, have repeatedly called HDP as the “political wing” of the PKK, and blamed CHP as supporting terrorism (Kaya & Whiting, 2019).

CHP being against nation’s interest or working to undermine the “general will” of the public: These sentences claim that CHP is against the nation’s interests, that it does not have the interest of the nation in mind in its policies, actions, or ideas. In addition, sentences that say the opposition party is actively working against the interests of the Turkish nation or state, for example in areas of foreign policy, fight against terror, etc., are also included.

Implications of CHP cadres as taken over by or supportive of foreign interests: These claim that CHP administration has been bribed to, coerced to, or willingly partakes in working with international and domestic interests that are undermining the Turkish state or nation.

3.2.2 Immorality Component

For the immorality component of polarizing rhetoric, we focus on sentences which portray the main opposition party, CHP, as an immoral political actor. More specifically, sentences containing similar claims to categories below, are coded as instances where CHP is shown as immoral.

Implications of CHP and its politicians in general as lacking morals and principles: These sentences vilify the opposition as liars, hypocrites, etc., and/or without principles or morals.

Allegations of corruption, misconduct, and malpractice: These claim that CHP and its politicians engage in corrupt practices or abuse their power.

Interpersonal insults, threats, attacks, and slander: These sentences show politicians from CHP as insulting, slandering, or threatening parts of society, other politicians, public figures, including the president.

Not respecting public morals or religious values of society: These sentences either show CHP as violating or disrespecting supposedly established norms and values in society. These also include references to past occurrences of such instances.

3.3 Opposition as the rival

3.3.1 Incompetency Component

This component refers to the perceived lack of ability to create positive results (Bellucci, 2006). These sentences include allegations about CHP regarding incompetence, negligence, election losses, defeats, and failures.

Incompetence, inefficiency, negligence, election losses, defeats, and failures: These sentences claim that the party lacks the ability to achieve positive results due to factors such as inefficiency, negligence, election losses, defeats, and failures.

Mismanagement of the economy, including the accumulation of debt, bankruptcy: These sentences claim the opposition has ruled public institutions and municipalities inefficiently often leading to accumulation of debt, bankruptcies. Such mentions are also included when they refer to past periods.

Indecision, ambiguity, and failure to keep promises: These sentences claim that the party or party leaders are unable to take a clear stance on various issues. Mentions of indecision, ambiguity, going back on promises are included.

Protests against actions such as layoffs made by CHP administrations: These sentences mention protests against actions and layoffs made by CHP administrations, which can be seen as a sign of the party's inability to handle issues properly.

3.3.2 Incohesiveness Component

Internal disunity, indiscipline, resignations seen in political parties, and public criticisms to the party by party members can also signal incompetency to voters (Greene & Haber, 2015). Sentences with such claims are coded as incohesive as they discredit the party’s ability to govern and the effectiveness of its leadership.

Criticisms of CHP by CHP members: These sentences contain criticisms of the CHP made by its own members, indicating a lack of internal unity and a divided front.

Factionalism, internal conflicts, and leadership struggles: These sentences depict internal factions, conflicts, and leadership struggles within the party, which indicates a lack of coordination leading to effectiveness.

Resignations: These sentences portray resignations of party members, indicating a lack of confidence in the party's leadership and direction.

Confusion, differences of opinion and disciplinary actions within the party: These sentences depict opposition members as divided on issues and having differences of opinion. Mentions of disciplinary actions within the party are also counted, indicating a lack of internal unity and coherence.

Unfair treatment of party members: Such sentences refer to instances of unfair treatment of party members, which can lead to a sense of disenchantment and lack of loyalty among party members.

CHAPTER 4

DATA AND METHOD

As described in Chapter Three, this study analyzes sentences from news about CHP published online on Sabah and Hürriyet by looking at whether or not each sentence depicts CHP negatively according to certain criteria. These are depictions of CHP as an immoral or illegitimate actor which constitutes the “enemy” type of rhetoric, or as an incompetent or incohesive actor in the political arena, constituting the “rival” type. The rival type of rhetoric does not problematize the legitimacy of the political opposition and sees it as valid and legitimate opponent within the political arena. On the other hand, the enemy type of claims comes with arguably important consequences for democratic well-being. Parties, if seen as illegitimate by large segments of the population, can lead to distrust and dissatisfaction in democratic institutions (Ridge, 2020) and justify authoritarian action against that party. In addition, claiming that the opposition is immoral, in the sense that it is against the moral order of the society, or against established religious customs and practices, contributes to the strengthening of us versus them camps within the society, harming democracy through lowering interpersonal trust and enabling justifications of authoritarian action against the other camp.

4.1 Independent variables

Separate fine-tuned machine-learning algorithms utilizing the same pre-trained algorithm are used to detect each of these four types of sentences focused by this study. We measure the daily number of sentences belonging to one or more components through predictions made by algorithms dedicated to each component. Then, the daily number of sentences that depict CHP as immoral or illegitimate are summed to arrive at the measure of enemy type of rhetoric, and the number of sentences with incompetent and incohesive depictions are summed to construct the measure of rival type of rhetoric. From these daily numbers of enemy and rival types of criticisms, two additional variables are created to be used in hypotheses testing. Percentage versions of these two types of criticisms are constructed by dividing the number of sentences containing each type of criticism to the total number of sentences input to the algorithms for each day. Normalized versions of these two types of criticisms are constructed by aggregating the daily number of two types at quarterly level, which are then divided by the number of total news published from a representative sample for each quarter**.**

4.2 Data Collection

The news articles from Sabah and Hürriyet were collected in two steps. The first step was to create an inventory of main headings, URLs, and dates of news that appeared in keyword searches on Sabah and Hürriyet. The second step was to retrieve and store the main heading, lead paragraphs (a brief paragraph summarizing the article at the beginning of most articles) and the main bodies, exhausting all the news article URLs collected in the inventory.

For a comprehensive evaluation of the news about the main opposition party CHP, keyword searches on the websites of Sabah and Hürriyet are conducted. We searched for "CHP" on both websites and scraped the URLs, headers, and dates of the news that appear in the search results to generate an inventory of all the news that matched my keyword filter on each website. For both Sabah and Hürriyet, a news article appears in the search if it contains the keyword "CHP" somewhere in the article at least once.

These searches resulted in 40940 news articles from Hürriyet and 37688 articles from Sabah whose date, heading and URL were stored in csv files. I scraped the main heading, lead paragraphs, and the text body of each of these news stories with the Beautiful Soup 4 library for Python (Richardson, 2007). I stored this information in csv files separately for Sabah and Hürriyet with each row containing the date and URL of the news published, its heading, lead paragraph, and body in separate columns.

As mentioned, these news articles contain the keyword "CHP" at least once somewhere in their text, which means not all of them are about the opposition. To narrow down the scope further and make it more likely that a news story is actually about the opposition, I filtered out the news that do not contain keywords "Kılıçdaroğlu," "İmamoğlu," "Yavaş," "Sarıgül," "İhsanoğlu," "İnce," "CHP," and "muhalefet" in neither their heading nor the lead paragraph after I scraped the text of the news[*]. All the filter keywords were case sensitive to make sure they apply to proper names, except for the keyword "muhalefet." These narrowed down sets of news were further cleaned from duplicate URLs, photo-gallery and video news, opinion pieces, and local news. The resulting clean sets of news consisted of 25951 stories from Hürriyet and 20117 stories from Sabah.

4.3 Creating the training datasets

To train the machine learning algorithms used in this study, a total of 16,765 sentences were selected from a sample of the online news articles on these two newspapers

[*] These keywords were selected as they correspond to the surnames of the leader of CHP (Kılıçdaroğlu), mayoral candidates for Istanbul and Ankara in local elections of 2014 (Sarıgül and Yavaş) and 2019 (İmamoğlu and Yavaş), its presidential candidates in the elections of 2014 (Ihsanoğlu) and 2018 (İnce), party's abbreviation (CHP), and Turkish for opposition (muhalefet). Application of these keywords are not date sensitive, i.e., they are applied equally across the total time scope.

published about the opposition. We used constructed week sampling to sample these news by constructing two weeks per each quarter. These sentences are coded in terms of whether they portray CHP as illegitimate, immoral, incohesive, or incompetent. Due to the research design having been changed several times, sampling process of the news from these newspapers and coding process to create the training datasets includes some asymmetries, which are described below. Originally, research would include the newspaper Sabah only, with news from Hürriyet included in the scope of the research only later in the process. Therefore, the first versions of the training sets only included news that were sampled from Sabah with 11334 sentences in the training sets. After Hürriyet is also included in the research design, additional news are sampled from Hürriyet. Using exactly the same procedure on Hürriyet, namely generating the sample news from two constructed weeks per quarter, resulted in 43441 sentences in the sample, which proved too large to be coded manually with the available resources and time. To narrow down the sample, we selected every eighth sentence from the tokenized sentences of the sampled news, which resulted in 5431 sentences from Hürriyet. These sentences were manually coded along the four components mentioned above. These processes are visualized in Figures 2 and 3.

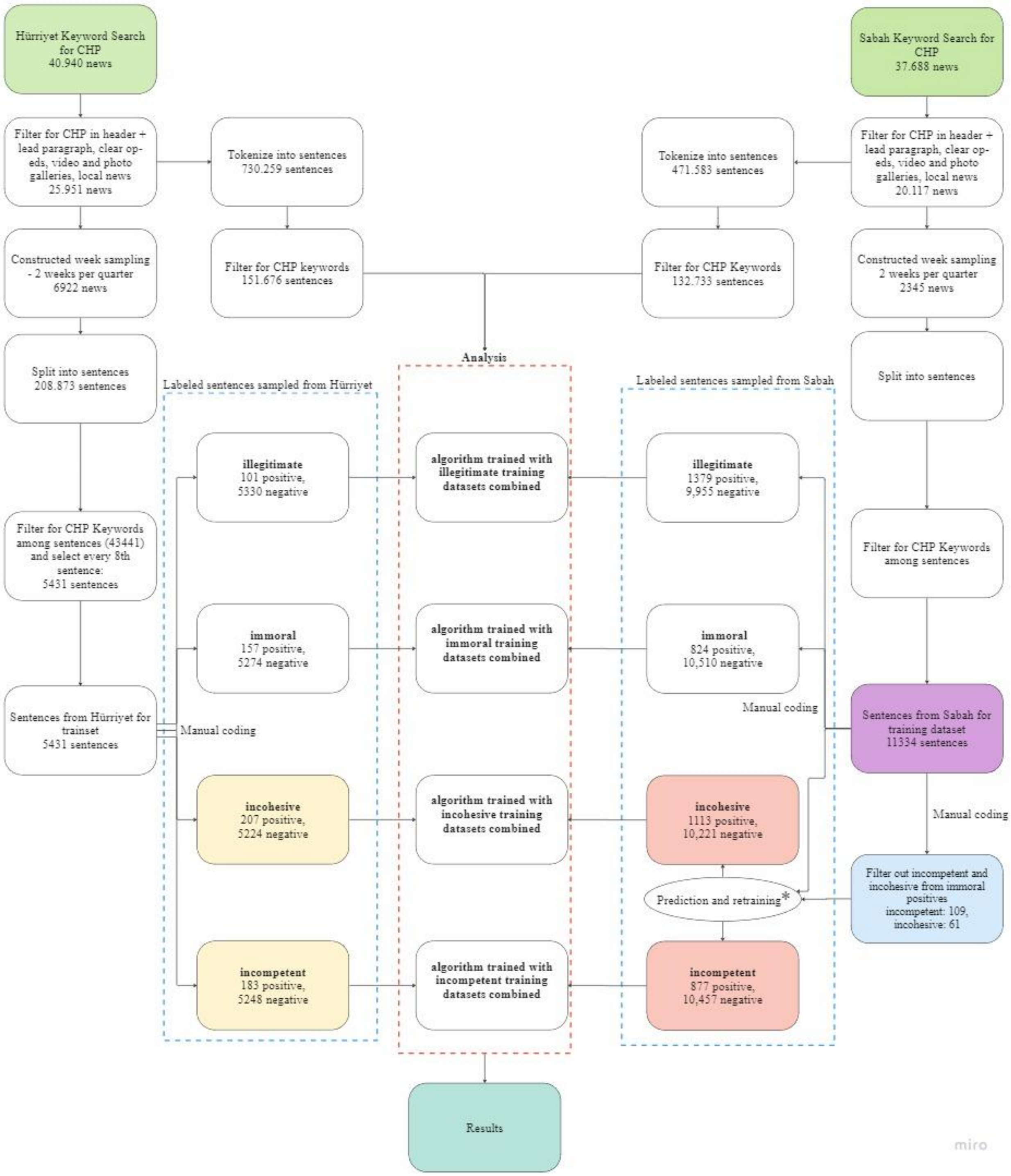

Figure 3. Flowchart describing the data collection process for analysis and training sets.

(*) Prediction and retraining process to create the incohesive and incompetent training sets is detailed in Figure 3. Shared background colors represent the same datasets across figures 2 and 3.

Figure 4. Flowchart describing the prediction and retraining process for incohesive and incompetent training datasets. Shared rectangle background colors represent the same datasets across Figures 2 and 3.

The rival conceptualization of the criticisms against the opposition, namely the sentences that depicted CHP as incompetent and incohesive, were also included in the research design at a later phase. For Hürriyet, training sets all four components including incompetent and incohesive were coded manually. On Sabah, while sentences with immorality and illegitimacy connotations about CHP were coded manually, the datasets for the other two components (incompetent and incohesive) forming the rival type were coded with the help of machine-learning. The training data for this retraining process was built by using the manually coded incohesive and incompetent sentences from the news sample of Hürriyet. This was supplemented by sentences coded as immoral at an early version of conceptualization of the components but were changed to either incompetent or incohesive after conceptualization had been subject to changes. Then, these labeled datasets for incompetent and incohesive algorithms, are used to predict the sampled sentences from Sabah. Then we manually went over the predicted positives, and corrected any misclassifications made by the algorithms, and used these as the training datasets to fine-tune the actual algorithm used to predict incompetent and incohesive sentences in all the sentences under consideration.

In total, out of the 11,334 sentences from the sampled news from Sabah, 828 were coded as depicting CHP as immoral, 1335 were coded as illegitimate, 734 as incompetent, 1133 as incohesive. Out of the 5,431 sentences from Hürriyet, 157 were coded as immoral, 101 as illegitimate, 183 as incompetent, and 207 as incohesive. These labeled datasets of each component are then merged together. Then, four machine-learning algorithms are fine-tuned using these labeled data to detect each component - illegitimacy, immorality, incompetency, and incohesiveness- separately. These fine-tuned algorithms are then applied across all the 45,000 news articles published by Sabah and Hürriyet since 2009 about the political opposition and around 300,000 sentences extracted from these to predict the existence of similar sentences in all the sentences from news about CHP.

Even though the creation of the training datasets contains some asymmetries which might translate into inaccuracies made in the predictions, we believe the impact of these is minimal. The main findings of the study are robust to even to the original training datasets, which include sentences sampled from Sabah only. This robustness we believe suggests Sabah sentences provide enough variation for the algorithms.

Differently from the studies reviewed Chapter 2, we do not differentiate between, or measure the mentions of different topics of the news articles. Rather, we focus on the amount and ratios of such sentences at a daily level, regardless of in which topic or context these sentences are used. We also do not differentiate between the sources using such sentences. Both editorial use of these sentences by those who write the articles themselves, and their use by sources quoted in the article, are counted equally when aggregating the amount and the ratio of sentences at the daily level. We only code a sentence positive for a given component if the coding criteria described in Chapter 3 is explicitly present in the sentence, without any additional reading of the context.

4.4 Analyzing the Data

I conduct the main analysis on sentence level with a fine-tuned BERT (Bidirectional Encoder Representations from Transformers) sequence classification algorithm for Turkish (Scheweter, 2020) to detect sentences that portray the opposition as immoral, illegitimate, incompetent, or incohesive. I then aggregate the former two into the "enemy" type of criticism measure and the latter two into the "rival" type. I measure the number of such sentences at the daily level and calculate their ratio to all sentences input to the algorithm on each day which gives us the percentage variable. In addition, we sample one constructed week per quarter within the time scope of this study and collect every news item published on Hürriyet and Sabah, to help get a sense of their total volume of news and how much relative coverage CHP receives at a quarterly level. These are the main variables on which the hypotheses are tested.

We rely on the Statistical Functions module from SciPy library (Virtanen et al., 2020) to conduct the hypotheses tests using Python. In particular, one-tailed, independent, two sample t-tests are used to compare the above variables on either same time periods across the two outlets, or on the same outlets across different time periods when testing the hypotheses. The Python module reports two values as a result of this test: the t-test statistic corresponding to the magnitude of difference of means, and p-value associated with this magnitude. We report the t-statistic in tables in Chapter 6, and present p-values as asterisks next to them.

4.4.1 BERT sequence classification algorithm

BERT is a pre-trained machine learning algorithm for language processing. It is a general-purpose language understanding model trained on a very large text corpora, which could then be "fine-tuned" to handle specific tasks like question answering, next sentence prediction, and text classification (Devlin et al., 2018). The particular BERT model used in this paper is BERTurk, which is pre-trained on Turkish corpora composed of 44 billion words (Scheweter, 2020).

Four different fine-tuned BERT algorithms are used to predict immorality, illegitimacy, incompetency, and incohesiveness components separately. This allows for an approach to these four problems as binary classification (e.g., whether a sentence is portraying opposition as illegitimate or not) rather than multilabel classification problem. Under multilabel classification, algorithm would have predicted only one class given a sentence, since each sentence can only be under either immoral or illegitimate dimension, not both. In this case, formulation of the problem as four different binary classification is preferred because a single sentence can have multiple negative connotations about the opposition.

4.4.2 Creating the training dataset used in fine-tuning BERT

Fine-tuning procedure for binary text classification task consists of inputting manually labeled training data into the model which consists of the sentences itself and their associated class-labels (positive, 1 or negative, 0). In this case, positive class (1) corresponds to the presence of illegitimate or immoral rhetoric about the opposition whereas negative class (0) corresponds to their absence. To make the training data as representative of the whole population as possible, we conduct constructed week sampling. Constructed week sampling outperforms random sampling in analyzing content with periodic cycles (e.g., newspapers, magazines) in terms of its ability to represent the population (Riffe, Aust, & Lacy, 1993). For online news content, (Hester & Dougall, 2007) suggest two to five constructed weeks for a representative sample from a population of six months of online news. For the training data, I chose at random two Mondays, two Tuesday, two Sundays, etc., from each quarter, and included all the news articles that were published on these days. I construct four weeks for each six-month period, which should be representative enough of the general population to avoid any biases due to sampling of the training data used in fine-tuning.

After sampling which news would belong to the training data, each news article in the sample is split into sentences using the Turkish sentence tokenizer module from the NLTK library for Python (Loper & Bird, 2002).We filter out the sentences that do not contain one of the eight keywords listed in section 4.2. and code the filtered sentences along four components of polarizing language discussed in Sections 3.2. and 3.3 In total, out of the 11334 sentences from the sampled news from Sabah, 828 were coded depicting CHP as immoral, 1335 were coded as illegitimate, 734 as incompetent, 1133 as incohesive. Out of the 5431 sentences from Hürriyet, 157 were coded as immoral, 101 as illegitimate, 183 as incompetent, and 207 as incohesive. These constituted the labeled training datasets used for fine-tuning algorithms to predict the existence of similar sentences in all the sentences from news about CHP.

4.4.3 Adjusting class weights for unbalanced classification

As mentioned, these data are used in fine-tuning four separate BERT algorithms, to predict sentences that depict CHP as immoral, illegitimate, incompetent, and incohesive. Classes in these datasets however are unbalanced (i.e., class distribution is not even across the positive and negative classes but skewed heavily in the negative class) and may cause issues, as most algorithms, including BERT, are designed with balanced datasets in mind by giving equal weights to each class while optimizing its performance (Luque, Carrasco, Martín, & de Las Heras, 2019). In this case, since most of the sentences from news items in training data belong to the negative class (i.e., they do not contain depictions of CHP as immoral, incompetent, etc.), assigning equal weights to each class biases the algorithm in favor of the negative class, omitting the less common, but more important positive class (Pereira & Saraiva, 2020). Although more sophisticated methods exist, one straightforward way to overcome this problem and adapt the algorithm to optimize itself on unbalanced data, is called cost sensitive learning (Fernández et al., 2018). Cost sensitive learning techniques manipulate the misclassification costs (i.e., class-weights) associated with the classes that algorithm takes into consideration while optimizing itself (Kaur, Pannu, & Malhi, 2019). We adjusted the class-weights to be inversely proportionate to size of the classes so that misclassifying the rare class has a higher cost (Kamaldeep Singh, 2020).

BERT also allows manipulating several parameters to change the prediction algorithm slightly (Devlin, Chang, Lee, & Toutanova, 2019). These parameters are optimized through trial-and-error in fine-tuning the algorithms. The values of each parameter from the final versions of these algorithms are presented in Table 1.

Table 3. Parameters of Final Versions of Algorithms to Predict Each Component

Algorithm	Learning rate	Batch Size	Evaluation set ratio	Epochs
Immoral	4E-05	16	0.2	2
Illegitimate	4E-05	32	0.2	3
Incompetent	4E-05	32	0.1	3
Incohesive	4E-05	32	0.1	3

4.5 Evaluating the Performance of the Analysis

Performance of machine-learning algorithms is measured by creating a confusion matrix on which most performance metrics are based. The structure of the confusion matrix for a binary classification problem is shown in Table 2.

Table 4. Confusion Matrix for Binary Classification (Source: Lipton, Elkan, & Naryanaswamy, 2014, p. 226)

	Actual Positive	Actual Negative
Predicted Positive	True Positive	False Positive
Predicted Negative	False Negative	True Negative

The confusion matrix has two dimensions, actual known values of the dataset fed into the algorithm to train and evaluate it, and the values of the same dataset as predicted by the algorithm. Metrics that evaluate the performance of the algorithm are calculated by using the true negative, true positive, etc. values inside the table. A summary of these metrics, which range between 0 to 1, can be found in Table 3.

Among these, accuracy is the most frequently employed as it captures the overall performance of the model by looking at the ratio of correct predictions of the algorithm over its total predictions (Luque et al., 2019). However, its use with unbalanced data is problematic as when the class unbalance is high, algorithm will learn to classify the more common class, and since classification is done mostly on the more common class, accuracy metric will not be able to evaluate the prediction performance of the smaller class (Tan, Tan, Dara, & Mayeux, 2015). For example, if only 1% of the sentences portrays opposition as illegitimate, then the classifier can achieve 99% accuracy score by classifying all sentences as not illegitimate. Therefore, I rely on F1 scores to measure the performance of the fine-tuned algorithms. F1 score is the geometric mean of recall and precision. These two metrics focus on the performance of identifying the positive class, with recall answering the question “What proportion of actual positives was identified correctly?”, and precision answering “What proportion of positive predictions was actually correct?” (“Classification,” 2022). Therefore, correct evaluations of the more common but less important negative class do not skew the metric as is the case with accuracy score. These performance metrics for immoral and illegitimate classification algorithms can be found in Table 4.

Table 5. Performance Metrics for Classification Algorithms (Source: Erickson & Kitamura, 2021)

Metric	Meaning	Synonyms	Mathematical Formula
Sensitivity	The fraction of positive cases predicted as positive	Recall, true-positive rate	$\text{Sensitivity} = \frac{TP}{TP + FN}$
Specificity	The fraction of negative cases predicted as negative	Selectivity, true-negative rate	$\text{Specificity} = \frac{TN}{TN + FP}$
False-positive rate	The fraction of cases predicted positive that were actually negative	Fall-out, probability of false alarm	$FPR = \frac{FP}{TN + FP}$
False-negative rate	The fraction of cases predicted negative that were actually positive	Miss rate	$FNR = \frac{FN}{TP + FN}$
Positive predictive value	The fraction of truly positive cases from all cases the model predicted positive	Precision	$PPV = \frac{TP}{TP + FP}$
Negative predictive value	The fraction of truly negative cases from all cases the model predicted negative	None	$NPV = \frac{TN}{TN + FN}$
Accuracy	The fraction of cases the model correctly predicted	None	$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP}$
F1 score	The harmonic mean of positive predictive value and sensitivity	F score, F measure, Dice similarity coefficient	$F1 = \frac{2TP}{2TP + FP + FN}$

Note.—FN = false negative, FNR = false-negative rate, FP = false positive, FPR = false-positive rate, NPV = negative predictive value, PPV = positive predictive value, TN = true negative, TP = true positive.

These algorithms are used to generate variables that measure the enemy and rival types of criticisms across Sabah and Hürriyet. Next chapter discusses the case and hypotheses which are tested with these data, and Chapter 6 presents the test results.

Table 6. Performance Metrics of Final Versions of algorithms to Predict Each Component

Algorithm	Precision	Recall	F1 - Score
Immoral	0.77940	0.73180	0.75485
Illegitimate	0.88600	0.80170	0.84174
Incompetent	0.84590	0.68930	0.75961
Incohesive	0.80230	0.88040	0.83954

CHAPTER 5

CASE AND HYPOTHESES

To track potentially polarizing rhetoric against the opposition in news articles, I focus on the two of the most popular newspapers in Turkey both in terms of online popularity and print circulation rates: Sabah and Hürriyet.

The Turkish media, in part due to its ownership structure, has always been sensitive to political demands (Yeşil, 2018). With the neoliberal transformation of the economy during 1980s, media outlets started to be acquired by a handful of large conglomerates (Tunç, 2018). These conglomerates with various investments in non-media sectors, became open to political influence through state advertisements and their dependence on business deals provided by the government (Tunç, 2018). As a result, during the 1990s, news media in Turkey mostly agreed on and reproduced the boundaries of politics backed by the secularist-militarist ruling elite, not problematizing the securitization of Kurdish and Islamist movements (Yeşil, 2016).

When AKP first came to power in Turkey in 2002, as an offshoot of the Islamist *Refah Partisi* (Welfare Party), it was seen as an agent of democratization both domestically and internationally, promising full membership in the EU, and ending of Turkey's secularist-militarist tutelage which excluded Kurdish and Islamist demands from the political arena (Insel, 2003). However, the initial momentum of fulfilling these promises was lost and began to be reversed after mid-2000s.

Starting perhaps with the mass-level and nation-wide Gezi protests and elite-level infighting between Gülen movement and AKP government in Turkey, the latter increasingly adopted a polarizing and exclusionary rhetoric towards broad groups in the society (Somer, 2019). Kurdish movement in Turkey, who were promised a more inclusionary political system under AKP and made substantial gains with the Kurdish opening under AKP-rule, was put squarely back into the previous exclusionary framework with their demands securitized as part of the authoritarian AKP's fight against terrorism with the breakdown of the peace talks in 2015 (Kaya & Whiting, 2019).

Since mid-2010s, Turkey has been experiencing a rapid and significant decline into authoritarianism. With the coup attempt against the AKP government in 2016, an emergency rule was declared which lasted for two years (Çalışkan, 2018). The new hyper-presidential system, approved by a referendum in 2017 under emergency rule, put an end to the parliamentary system in Turkey and consolidated unprecedented levels of executive power in the hands of Erdoğan, who was elected the first president of the new system in 2018 (Esen & Gümüşçü, 2017).

AKP during this process also benefited from the corporatist-clientelist media structure in Turkey and was able to establish hegemony over most major news outlets through friendly business conglomerates, including Sabah in 2008 and Hürriyet in 2018 (Yıldırım et al., 2021), on which this paper focuses. Below I briefly discuss their place in the current Turkish media landscape, and their ties to the ruling AKP government.

5.1 Sabah and Hürriyet in the Current Turkish Media Landscape

Sabah is the second-most popular printed newspaper, and its online version was the third-most popular news website in October 2022 in Turkey according to similarweb.com. Originally established in 1985, Sabah along with some other dailies and popular TV channels was acquired by a business group friendly to the AKP government, Çalık Holding, in 2008 (Media Ownership Monitor, 2021b). Çalık Holding's CEO at the time, Berat Albayrak, was the son-in-law of then-prime minister Erdoğan, and became the group's CEO at age 26 in 2007, three years after he married Erdoğan's daughter, Esra Erdoğan (OSC, 2008). Since 2013, Sabah, along with other dailies and TV channels operating under Turkuvaz Media Group, is owned by Kalyon Holding (Media Ownership Monitor, 2021b). There are claims that Erdogan himself arranged the purchase of Turkuvaz Media from Çalık Holding by Kalyon Holding, in return for public tenders for government projects (Media Ownership Monitor, 2021b). Kalyon Holding, managing megaprojects in construction, energy, and infrastructure sectors, is one of the top 10 companies in the world to receive public tenders (Media Ownership Monitor, 2021b).

Hürriyet has been the most popular printed and online news outlet in Turkey and was founded in 1948. In 1994, it was purchased by Doğan Holding that published and ran several other popular print and TV media outlets (Media Ownership Monitor, 2021a). Doğan Holding, with investments in sectors such as energy, manufacturing, tourism, and others, operated the largest major media group not affiliated with the AKP in Turkey (Reporters Without Borders, 2018). Starting with the first few years of AKP rule in Turkey, relationships between Doğan Holding and AKP turned increasingly sour with AKP government rejecting offers of Doğan Holding for a petroleum refinery and the purchase of a valuable public property in Turkey, and news published by Doğan Media about the corruption scandals related to AKP and Islamic charities in 2008 (Yeşil, 2018, p. 246). AKP government then fined Doğan media a total of four billion dollars on allegations of tax evasion in 2009, forcing the company to close down or sell several of its media outlets (Yeşil, 2016, 2018).

Demirören Holding, which acquired some of Doğan's outlets in the aftermath, also bought the remainder of Doğan Media in 2018 which consisted of a news agency, several high-rating TV channels, and popular newspapers, including the then- and still the most popular newspaper Hürriyet, financed by credits from the public agricultural bank, *Ziraat Bankası* (Reporters without Borders, 2018). The new owners of Hürriyet, Demirören Group, active in mining, energy, construction, and other sectors, have close ties to Erdoğan (Media Ownership Monitor, 2021b). In a phone call leaked in 2014, allegedly between the chairperson of Demirören Holding and then-prime minister Erdoğan, Erdoğan is heard scolding Demirören about news articles on the Kurdish issue, where Demirören sobbingly asks for forgiveness and promises to deal with those responsible for the news item (Reporters without Borders, 2018). In addition, daughter of the chairperson of Demirören, and the son of the founder of Kalyon Group, got married in 2019, where Erdoğan and his wife were the marriage witnesses (Media Ownership Monitor, 2021b), hinting at very close relations between two conglomerates and President Erdoğan's family.

5.2 Hypotheses

This study focuses on the levels of two distinct types of negative rhetoric launched against the main opposition party, CHP, in the online news articles published by these two newspapers. These types are depiction of CHP as an illegitimate enemy, and depiction of CHP as a valid and legitimate rival within the political arena. Given the Turkish autocratization recently and the media's ownership structure, which increasingly has closer ties to AKP elite, hypotheses put forward below attempt to establish and compare how the levels of these two types of depictions of opposition has varied and evolved across these two newspapers since 2009. We also look at how CHP coverage changed during these periods.

First, we expect election periods to contain more criticisms of the opposition, compared to non-election periods. We focus on election periods, defined as the 60 days prior to the election day, and adjacent periods which are defined as either the 60-day period just after the election period, or the 60-day period just before the election period. We expect that election periods should contain more of each type of rhetoric compared to adjacent periods.

> *H1a: News published in election periods will contain more "enemy" type of rhetoric about the opposition compared to news published in periods adjacent to the election period.*
>
> *H1b: News published in election periods will contain more "rival" type of rhetoric about the opposition compared to news published in periods adjacent to the election period.*

Moreover, given the ownership patterns of Sabah and Hürriyet and their target audiences, with Sabah firmly belonging to the pro-government camp and with sizable readership among incumbent voters (Yıldırım et al., 2021), and Hürriyet having a neutral -if not opposing- outlook to the AKP government until its acquisition by Demirören, and more centrist readership (Çarkoğlu et al., 2014), we expect each type of negative rhetoric to also be more common in Sabah compared to Hürriyet. In addition, we expect relative levels of "enemy" and "rival" types of rhetoric to be higher in Sabah compared to Hürriyet, given the latter's firmer place among pro-government news media and in light of the theoretical framework.

H2a: News published by Sabah will contain more "enemy" type of rhetoric about the opposition compared to news published by Hürriyet.

H2b: News published by Sabah will contain more "rival" type of rhetoric about the opposition compared to news published by Hürriyet.

H2c: News published by Sabah will have a higher ratio of "enemy" to "rival" type of rhetoric about the opposition compared to news published by Hürriyet.

While above hypotheses look at Sabah and Hürriyet, comparing all the news content from between the years 2009-2022 at once, we also compare narrower time periods, limited only to the election periods defined as 60-days prior to the election day, and limited to one-year intervals of all the timeframe within the scope of the study.

We also look at Hürriyet only, to hint at the possible effects of the change in its ownership. As mentioned, a pro-government business conglomerate, Demirören Holding acquired Hürriyet in 2018, and we expect this change in ownership to lead to each type of negative rhetoric to also be more frequent after the Demirören. In addition, we expect relative levels of "enemy" and "rival" types of rhetoric to be higher after the acquisition. We also test for CHP coverage, and whether it increased after Hürriyet was taken over by the pro-government Demirören Holding.

H3a: News published by Hürriyet will contain more "enemy" type of rhetoric about the opposition after its acquisition by Demirören Holding.

H3b: News published by Hürriyet will contain more "rival" type of rhetoric about the opposition after its acquisition by Demirören Holding.

H3c: News published by Hürriyet will have a greater ratio of enemy to rival rhetoric after its acquisition by Demirören Holding.

Moreover, Turkey has went through a regime change in recent years. After the failed coup attempt in 2016, AKP has brought constitutional changes regarding the establishment of a hyper-presidential system to a referendum in 2017. The changes, approved by a small margin (51,4% yes vote share) under emergency rule, entered into force with the general elections of 2018. We expect both newspapers to become more critical of the opposition, due to the institutional features of the new system turning political competition into a winner-takes-all framework (Esen & Gumuscu, 2018) and strengthening of "us versus them" camps in the country (Somer, 2019). We also check whether CHP coverage has increased after Turkey's switch to the hyper-presidential system.

H4a: News published by Sabah will contain more "enemy" type of criticism about the opposition after Turkey's switch to hyper-presidential system.

H4b: News published by Sabah will contain more "rival" type of criticism about the opposition after Turkey's switch to hyper-presidential system.

H4c: News published by Sabah will have a higher ratio of "enemy" to "rival" type of criticism about the opposition after Turkey's switch to hyper-presidential system.

H5a: News published by Hürriyet will contain more "enemy" type of criticism about the opposition after Turkey's switch to hyper-presidential system.

H5b: News published by Hürriyet will contain more "rival" type of criticism about the opposition after Turkey's switch to hyper-presidential system.

H5c: News published by Hürriyet will have a higher ratio of "enemy" to "rival" type of criticism about the opposition after Turkey's switch to hyper-presidential system.

The following chapter presents the results of the analysis. First, long term trends are considered and then the results for hypotheses tests are given.

CHAPTER 6

RESULTS

We conducted a content analysis of more than 45.000 online news articles (all the national level online news articles about CHP since 2009) to measure two types of criticism that has been launched against the main political opposition party, CHP, in Turkey across Sabah and Hürriyet. These are the two of the most popular newspapers in Turkey, former having a neutral -if not opposed- stance to the AKP government until its acquisition by a pro-government conglomerate in 2018, and the latter having close ties to the government since 2008. Tracking the evolution of these two types of criticisms, namely depiction of CHP as an enemy or rival, can provide important insights into the role of media in electoral autocracies. When systematically employed against opposition, such language can contribute to the affective polarization of the electorate which may translate into negative partisanship against opposition and support for undemocratic incumbents.

We operationalize enemy and rival types at the sentence level. We use two machine learning algorithms to classify each sentence in these articles in terms of whether or not they contain references to two components of the enemy type: opposition being cast as an illegitimate or immoral political actor. We use another two machine learning algorithms to classify each sentence in these articles in terms of whether or not they contain references to two components of the rival type: opposition being cast as an incompetent or incohesive political actor.

From a representative sample of these news, 16776 sentences are labeled in terms of these components to create training datasets to fine-tune these algorithms. These algorithms are then applied to all the news articles from the selected period. In this chapter, we first give an overview of the overall results predicted by these algorithms on these two online news outlets, and then focus on the hypotheses discussed previously.

6.1 Overview and Long-Term Trends

6.1.1 Number of News and CHP coverage on Sabah and Hürriyet

Before moving on to the main variables used to conduct the hypotheses testing, it makes sense to look at the total news volume and the coverage of CHP to see how they evolved during the 13-year period. Next figures show the amount of all the news from the constructed week samples for each quarter, total number of CHP news per quarter and the ratio of CHP news to the amount of all the news from the constructed week samples for each quarter, to gives an idea about the relative amount of CHP coverage.

As can be seen in Figure 4, Hürriyet overall publishes more news compared to Sabah for most quarters. However, long term trends are in opposite directions for Sabah and Hürriyet. Sabah gradually increases the amount of news published on its website while Hürriyet sees a gradual decrease in the long term. Peak number of total news for Sabah is around the coup attempt in July 2016, whereas for Hürriyet it is the quarter which witnessed the 2018 presidential elections in Turkey. Overall, numbers for Hürriyet start around a level that is three times higher compared to Sabah, but they gradually converge around a similar level with Hürriyet being slightly higher towards the end of the timeframe.

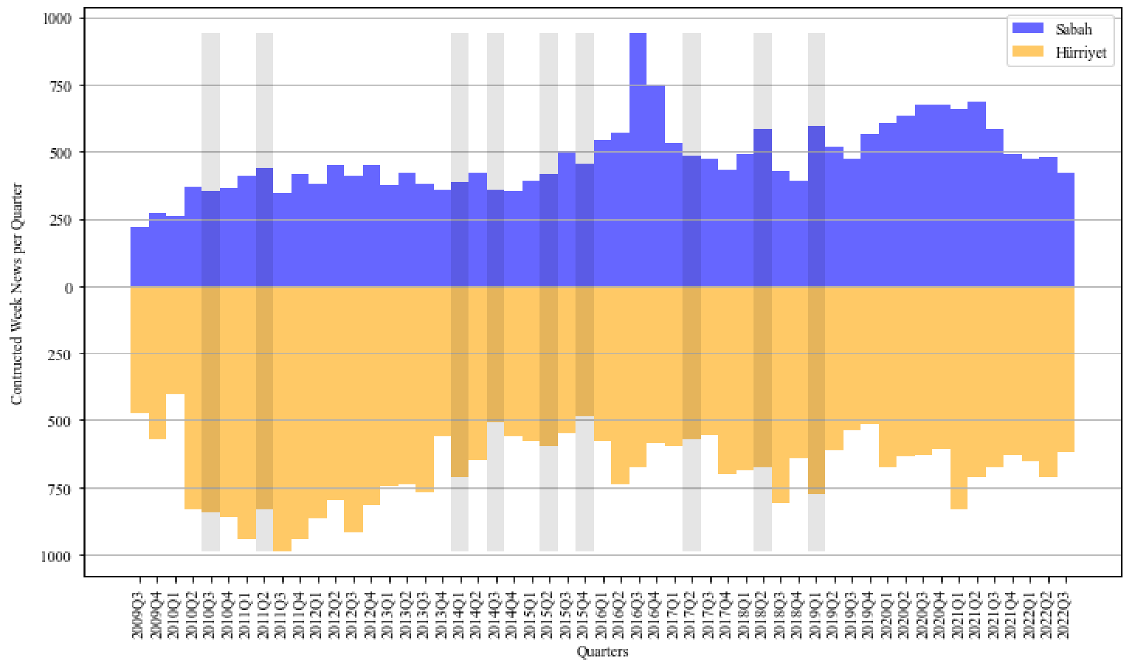

Figure 5. Quarterly number of news proxy. Number of news from one constructed week sample per quarter

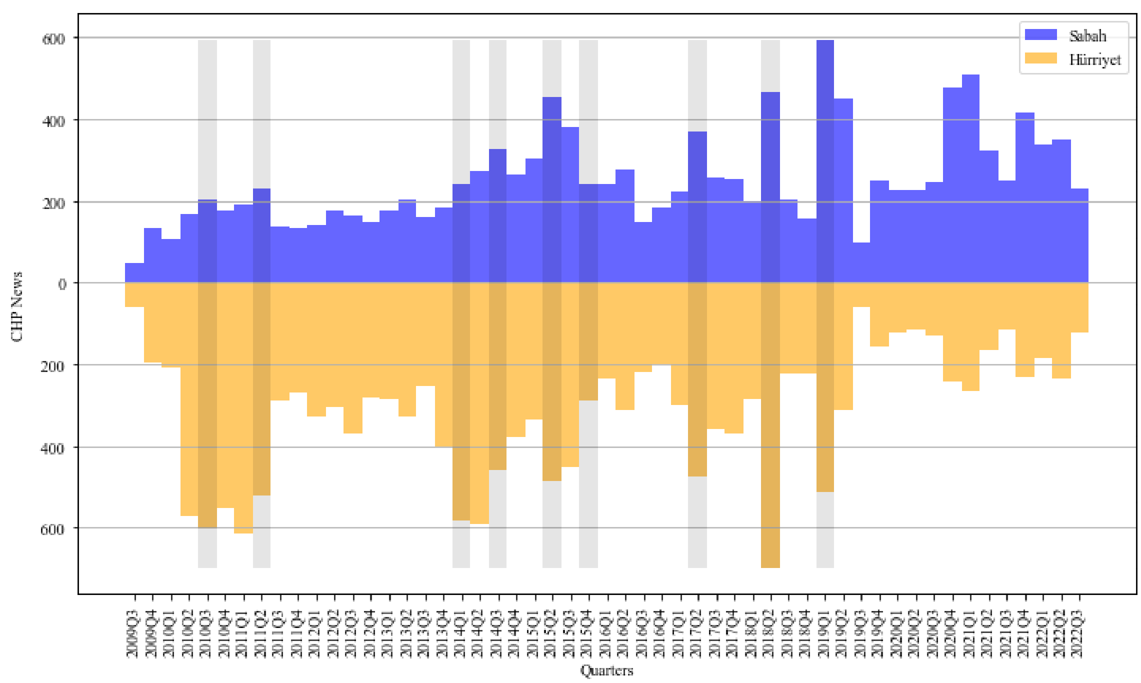

Figure 6. Quarterly number of CHP News

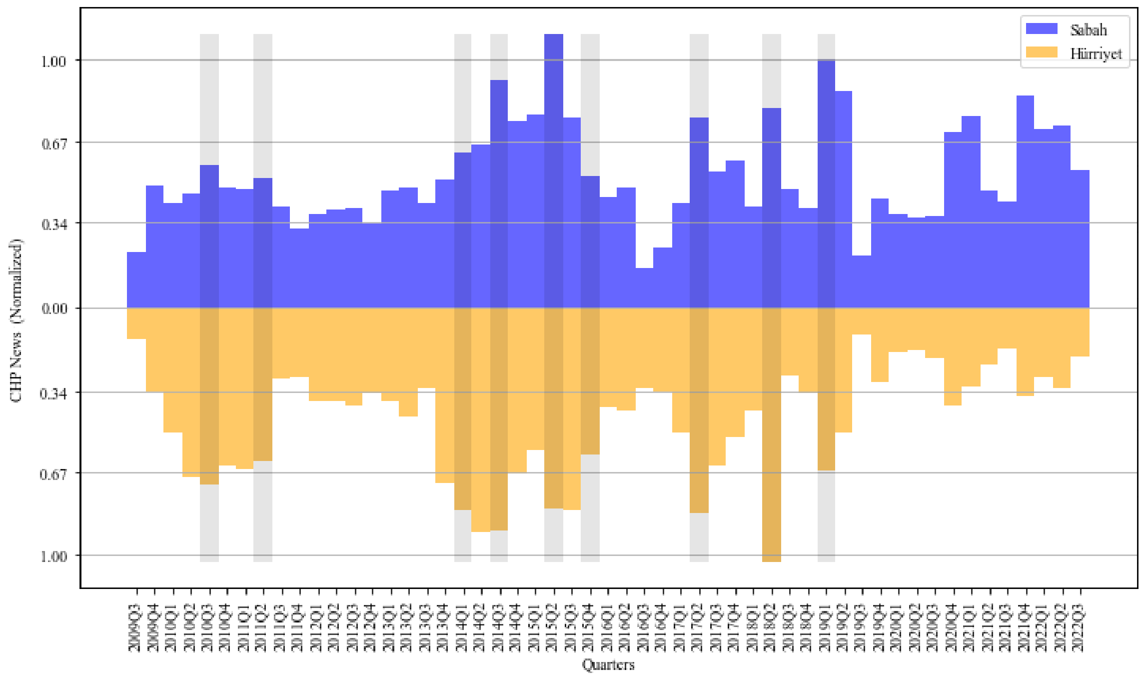

Figure 7. Quarterly number of CHP news normalized by total number of news proxy.

Figure 5 reports the numbers of CHP news we scraped and analyzed. Long term trends are similar to Figure 4, with Sabah again increasing over the long term and Hürriyet experiencing a more visible decline starting in 2019. Quarters with elections (highlighted gray in Figures 4, 5, and 6) usually correspond to local peaks in both newspapers. Another interesting period with unusually high number of news about CHP on Sabah are the last quarter of 2020 and the first quarter of 2021, where Turkey experienced several economic and political turmoil, such as the conflict between Azerbaijan and Armenia and Turkey's involvement and explicit support for Azerbaijan (Keddie, 2020), Turkey's military operations in country's Kurdish regions, cross-border operations into Syria (Aslan et al., 2021), and the peak of the currency crisis of the Turkish lira (Ewing, 2020).

Figure 6 presents the normalized version, suggesting relative to the total number of news published by Sabah and Hürriyet, CHP coverage proceeds at similar levels throughout the years until 2019 with Hürriyet allocating slightly more coverage to CHP, after which it gets as low as around 40 % of Sabah's coverage.

6.1.2 Enemy and Rival Type Sentences on Sabah and Hürriyet

Moving on to the variables measuring two types of criticisms, Figures 6 and 7 provide the visualization of the enemy and rival type of criticisms over the long term respectively, normalized for total news published. The number of sentences belonging to a given type is aggregated at the quarterly level, and then divided by the number of news published on the constructed week sampling of all the news from that quarter.

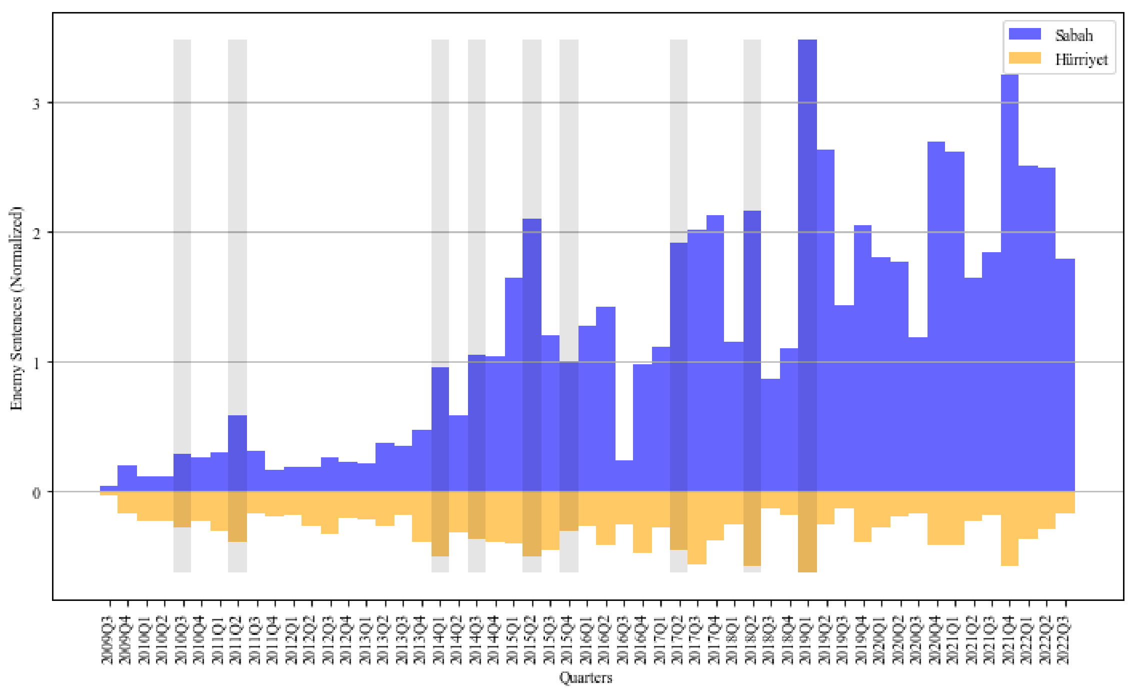

Figure 8. Quarterly number of enemy type of sentences about CHP (normalized)

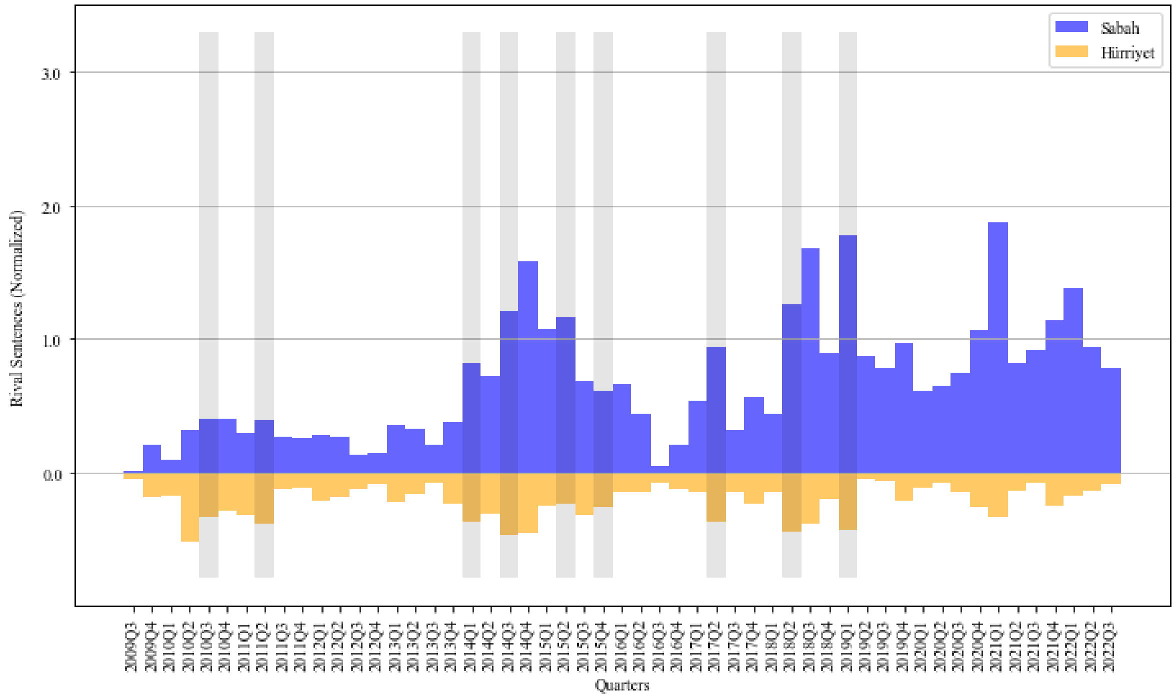

Figure 9. Quarterly number of rival type sentences about CHP (normalized)

The stark difference between Sabah and Hürriyet can be seen in Figures 6 and 7. While Hürriyet and Sabah start at similar levels of enemy and rival depictions of CHP, divergence begins around the end of 2013. Beginning from this date of divergence which coincides with the aftermath of Gezi Protests in Turkey (Göle & Eşkinat, 2013), such sentences on Sabah starts to increase dramatically compared to Hürriyet which stays around similar levels over the long term. The rise of enemy type is much higher in Sabah compared to the rival type. Quarters with elections also often witness local peaks of both types across both newspapers in addition to the late 2020 to early 2021 period described above as a period which also witnessed higher coverage of CHP in both newspapers. It is also interesting to note the outlier period Sabah's long-term trend of significant increase. The third quarter of 2016, which is just after the coup attempt in July 2016, saw very little enemy or rival depictions of CHP on Sabah.

Overall, these long-term overviews on the news coverage of and negative rhetoric about CHP by Sabah and Hürriyet over the 13-year period reveal interesting trends. While at the beginning Hürriyet publishes more news about CHP compared to Sabah, there is a gradual decrease over the long term, with Sabah taking over after 2018. In addition, figures representing the enemy and rival types of rhetoric show that both types are more common in Sabah, which experiences dramatic increases over the long term, compared to Hürriyet which stays relatively stable on both types. Interestingly, the divergence between Sabah and Hürriyet starts to be visible after the Gezi protests in Turkey.

Below, we test the hypotheses elaborated in Chapter 5. We use one-tailed, independent, two sample t-tests from SciPy Python library (Virtanen et al., 2020). For each hypothesis test results are given in the following section, with visualizations, and robustness checks with different operationalizations of the variables in t-tests are reported in Appendices B and C.

6.2 Testing the Hypotheses

6.2.1 H1: Coverage of election periods versus adjacent periods

H1 explores whether election periods have increased levels of negative rhetoric compared to previous or following periods. Election periods are defined as the 60-day period before the election day. Then the mean levels of enemy and rival type of rhetoric are compared between election periods, and adjacent 60-day periods (i.e., just before or after the election period) by one tailed, two sample independent t-tests.

Table 7. One Tailed t-test Results for H1, Election Periods.

	Election	2010 Referendum		2011 General Elections	
Website	Depiction	Election vs. Pre	Election vs. Post	Election vs. Pre	Election vs. Post
Sabah	Rival	-1.115	-1.829	0.972	-1.131
	Enemy	3.981***	2.529***	3.338***	2.945***
Hürriyet	Rival	-2.141	-0.023	0.114	-0.614
	Enemy	3.939***	6.537***	2.259**	3.290***
		2014 Local Elections		2014 Presidential Elections	
Sabah	Rival	0.972	-1.131	2.614***	0.470
	Enemy	3.338***	2.945***	3.318***	3.377***
Hürriyet	Rival	0.114	-0.614	0.641	0.836
	Enemy	2.259**	3.290***	1.869**	4.357***
		2015 June General Elections		2015 November General Elections	
Sabah	Rival	1.379*	0.859	-2.975	-2.266
	Enemy	1.567*	1.760**	-2.167	0.848
Hürriyet	Rival	-0.463	-2.342	-1.585	-1.178
	Enemy	0.030	-2.715	-2.173	2.284**
		2017 Referendum		2018 General Elections	
Sabah	Rival	1.195	-1.172	2.889***	-3.591
	Enemy	1.600*	1.813**	3.367***	3.896***
Hürriyet	Rival	0.000	-2.473	1.012	-4.126
	Enemy	1.029	0.063	3.153***	4.981***
		2019 Local Elections			
Sabah	Rival	4.509***	6.400***		
	Enemy	8.963***	6.696***		
Hürriyet	Rival	2.115**	7.398***		
	Enemy	5.481***	4.518***		

Note: Null hypothesis is period means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Table 5 presents the t-test results for comparing two 60-day period for each election for mean levels of rival and enemy type in Sabah and Hürriyet. The table contains results for the variable that counts the number of sentences that contain a given type for each day. Visualizations with daily numbers, tests for percentage and the normalized versions at the quarter level of the variables can be found in Appendix C.

For the enemy type of criticism on Sabah about CHP, we see that election periods almost always saw significantly higher mean levels of it except for the general elections in November 2015. Levels of statistical significance and the magnitude of mean differences, however, is also lower in June 2015 elections and 2017 referendum compared to other periods. On Hürriyet, most elections periods also witnessed significantly higher mean levels of enemy type of criticism, except for June 2015, pre test of November 2015 elections, and the 2017 referendum. Overall, enemy rhetoric shows more statistical significance for news from Sabah with 16 out of 18 tests, compared to news from Hürriyet with 13 out of 18 tests.

For rival rhetoric used against CHP on Sabah, fewer election periods have statistically significant differences compared to the enemy rhetoric. Elections in 2010, 2011, 2014, 2015, 2017, and 2018 did not see higher levels of rival type compared to the respective adjacent periods. Nevertheless, pre test for 2014 presidential, June 2015, and 2018 general elections show statistical significance, in addition to the 2019 local elections which has statistically significant mean difference compared to both previous and next period adjacent to the election. Whereas on Hürriyet, the only election that witnesses significantly higher levels of rival rhetoric is the 2019 local elections. Overall, election periods covered on Sabah saw more references to rival rhetoric with 5 out of 18 tests, compared to Hürriyet with 2 out of 18 tests.

We believe the lack of statistical significance in the elections in 2014, 2015, and 2017 points to a limitation of our chosen method of analysis. As can be seen in figures 3 and 4, this period corresponds to the timeframe where enemy and rival types of rhetoric are on the increase. Therefore, simple t-tests utilized in this study are not able to isolate the effect of this long-term increase, and -if any- seasonality in the data. More complex econometric and time-series models should be able to give a clearer picture, especially during such periods. However, due to time and skill constraints, simpler t-tests are employed in this study.

6.2.2 H2: Comparison of Sabah and Hürriyet

Next three hypotheses concern the comparison of enemy and rival types of rhetoric across Sabah and Hürriyet over the same time periods. H2 looks at the coverage within the whole period at once, while H3 looks at election periods only and H4 looks at one-year intervals of the 13-year period. These tests also include the relative ratios of enemy and rival type of rhetoric used against CHP. The expectation is that these levels should be higher in Sabah, due to its closer relationship to the incumbent AKP government in Turkey.

Since the amount of CHP news and CHP's relative coverage is different on Sabah and Hürriyet across the same timeframes, we conduct our main tests utilizing the percentage variable, defined as the predicted number of sentences for a given rhetoric type divided by the total number of sentences input to the algorithm for each day. We also present robustness checks with different versions of the variable (absolute number of sentences and normalized), in Appendix D.

Table 6 presents t-test results for the percentage value comparisons across all news sentences from Sabah and Hürriyet. As expected, more staunchly pro-government Sabah depicts CHP more intensely across all variables when looking at the news items across all the timeframe under consideration. Looking at the magnitude of mean differences reported by the t-tests, Sabah relies on both types of enemy and rival rhetoric at higher levels than Hürriyet. Also, ratio variable suggests that compared to Hürriyet, Sabah uses enemy rhetoric more heavily relative to the rival rhetoric, in line with our expectations about H2c.

Table 8. Sabah vs Hürriyet Comparison, Overall

Variable (Percentage)	t-test
Immoral	24.699***
Illegitimate	14.78***
Incompetent	32.223***
Incohesive	24.188***
Rival	26.823***
Enemy	38.129***
Enemy / Rival	18.435***

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is Sabah mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Next, we look at election periods and compare the variables only during election periods, defined as the 60-day period before the election day. Table 7 presents the results for each election. As can be seen, most election periods witnessed higher mean levels of variables measured in this study. Looking at the main variables, enemy rhetoric on Sabah is significantly higher from 2011 general elections onwards, whereas rival rhetoric starts to be significant after 2014 local elections. The ratio of these types is also significant for most periods, except for 2014 presidential elections and 2019 local elections. However, it should be noted that both enemy and rival rhetoric types separately are significantly higher in Sabah compared to Hürriyet in these elections. Overall, enemy rhetoric is statistically significantly higher in Sabah compared to Hürriyet 8 elections out of 9 total, rival rhetoric is statistically significantly higher for 7, and their ratio is higher for 7 elections out of 9.

Table 9. Sabah vs. Hürriyet Comparison, Election Periods

Election / Variable (Percentage)	Incom-petent	Incoh-esive	Immoral	Illegitimate	Rival	Enemy	Enemy / Rival
2010 Referendum	1.263	0.167	-0.565	1.739***	0.829	1.117	1.848**
2011 General	0.321	1.027	0.776	3.603***	1.087	3.398***	3.788***
2014 Local	2.510***	2.360***	1.474*	3.546***	3.048***	3.068***	1.625*
2014 Presidential	2.338**	4.110***	2.758***	1.736**	4.771***	2.799***	0.662
2015 June General	6.163***	4.007***	7.248***	5.004***	8.292***	7.764***	1.464*
2015 November General	3.324***	2.724***	4.609***	5.236***	3.671***	6.701***	4.876***
2017 Referendum	5.856***	1.428*	8.095***	4.842***	5.230***	8.123***	4.390***
2018 General	5.024***	4.113***	6.021***	6.676***	6.452***	8.476***	4.084***
2019 Local	6.619***	2.489***	7.305***	8.363***	5.653***	10.832***	1.111

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

When looked at year by year, all variables are at significantly higher levels on news items from Sabah compared to those from Hürriyet from 2014 and onwards as seen in Table 8. Looking at the one year intervals, enemy rhetoric on Sabah is higher than Hürriyet in 9 out of 14 years under consideration, rival rhetoric is higher on Sabah in 12 out of 14 years, their ratio on Sabah is significantly higher compared to Hürriyet 11 out of 14 years.

Table 10. Sabah vs. Hürriyet Comparison, Year by Year

Year / Variable (Percentage)	Incom-petent	Incoh-esive	Immoral	Illegitimate	Rival	Enemy	Enemy / Rival
2009	1.824**	-1.472	-1.602	0.488	0.234	-0.423	-0.360
2010	1.656**	0.281	-0.851	1.617*	1.119	0.599	1.814**
2011	0.134	1.911**	0.749	1.034	1.674**	1.268	2.081**
2012	0.654	2.537***	-0.122	-0.787	2.511***	-0.706	-0.095
2013	2.071**	2.751***	2.286**	-0.477	3.374***	1.146	0.474
2014	5.682***	6.688***	4.590***	5.591***	8.379***	7.091***	3.109***
2015	8.864***	7.224***	10.712***	9.450***	10.642***	13.845***	6.387***
2016	5.831***	4.088***	7.965***	10.157***	6.894***	12.813***	8.284***
2017	8.834***	1.925**	13.180***	11.766***	7.601***	17.135***	10.922***
2018	10.095***	6.553***	12.794***	10.517***	11.483***	15.759***	9.007***
2019	12.321***	4.539***	15.948***	11.133***	12.487***	19.973***	6.560***
2020	11.879***	3.959***	16.754***	8.470***	11.594***	18.989***	5.970***
2021	9.910***	5.788***	18.204***	7.259***	11.109***	19.774***	3.910***
2022	7.176***	5.564***	14.238***	6.613***	8.974***	15.470***	3.130***

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is Sabah mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

It is interesting to note that t-test magnitudes and their statistical significance gradually increase for most variables over time, with differences between Sabah and Hürriyet becoming significant from 2012-2013 onwards and beginning to somewhat decrease again after 2018. Another interesting point is that Sabah starts to diverge significantly from Hürriyet earlier in terms of rival rhetoric compared to enemy. Whereas rival rhetoric is significantly higher in Sabah from 2012 onwards, enemy rhetoric begins to diverge significantly only after 2014. Even though divergence starts two years later, the magnitude of differences gradually increases in enemy type of rhetoric more than rival rhetoric over the years. In addition, the magnitude of the differences between relative amount of enemy rhetoric to rival rhetoric on Sabah and Hürriyet increases each year until 2018, when Hürriyet was captured by Demirören Holding, and starts decreasing gradually until 2022. Nevertheless, the difference remains statistically significant after Hürriyet's acquisition.

6.2.3 H3: Hürriyet's acquisition by Demirören

Next set of hypotheses focuses on the change in ownership of Hürriyet and attempts to understand what kind of changes may have accompanied the news outlet when it was acquired by a pro-government business conglomerate in terms of the language it uses against the opposition.

For these tests, we group the variables under consideration before and after the date of acquisition (22 March 2018) and run a t-test to compare the means and to see whether the difference between the groups is significant. Results are presented in Table 9. Tests for the number of sentences and the normalized version can be found in Appendix E.

Looking at the hypotheses under H5, enemy, rival, and enemy / rival variables in terms of percentages are statistically significant. After Hürriyet was acquired by Demirören Holding, t-tests suggest that rhetoric against CHP got more intense relatively.

We also look at the number of news to assess the coverage of CHP. Looking at number of news published before Demirören purchase and comparing it with the period after the acquisition, we see a statistically significant decline (t-test: 21.623, p value <0.01).

Table 11. Hürriyet Comparison Before and After Demirören Acquisition

Variable (Percentage)	t-test
Incompetent	6.668***
Incohesive	0.807
Immoral	8.521***
Illegitimate	0.149
Enemy	4.968***
Rival	4.066***
Enemy / Rival	2.493***

One possible explanation behind why Hürriyet chose to publish much fewer news articles after it was bought by the pro-government conglomerate could relate to the readership base of Hürriyet. As opposed to Sabah with its readership concentrated mostly among incumbent AKP voters, Hürriyet still continues to have a more centrist readership base (Yıldırım et al., 2021). One can argue that the different nature of their readership base could be influential for this diverging pattern in the coverage of opposition. While on Sabah, opposition coverage increasing over time with more and more references to vilifying rhetoric, on Hürriyet, instead the coverage of opposition declines substantially, as the centrist readership base has the potential to stop reading news from Hürriyet if they notice blatantly vilifying attitude towards CHP. Hürriyet therefore might have chosen instead to go silent about the opposition, decreasing its overall coverage after its acquisition by the pro-government conglomerate Demirören.

6.2.4 H4 and H5: Sabah and Hürriyet after the hyper-presidential system

Next set of hypotheses consider the effect of regime change on Hürriyet and Sabah. The hyper-presidential system in Turkey was approved with popular vote by a small margin on 16 April 2017 under a state of emergency following the failed coup attempt nine months. The constitutional amendments entered into force on 9 July 2018 after the general and presidential elections held concurrently. H4 checks the changes in enemy and rival types of rhetoric and their ratio on Sabah and H5 checks the changes in Hürriyet. Daily percentage values before the regime change and after are grouped for Sabah and Hürriyet differently, then one-tailed, two sample independent t-tests are applied. Table 10 presents the results for these t-tests. Tests for the number of sentences and the normalized version can be found in Appendix F.

Table 12. Sabah and Hürriyet Comparison Before and After Regime Change

Variable (Percentage)	Sabah	Hürriyet
Incompetent	20.356***	6.635***
Incohesive	4.371***	0.939
Immoral	32.957***	9.218***
Illegitimate	5.921***	0.403
Enemy	24.551***	5.561***
Rival	16.346***	4.162***
Enemy / Rival	11.282***	3.544***

Note: Null hypothesis is period means are not different. Alternative hypothesis is period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

As can be seen, all variables in terms of percentages are statistically significantly higher on Sabah after the regime change, whereas on Hürriyet incohesive and illegitimate components are the only exceptions. The magnitudes of change are substantially higher on Sabah compared to Hürriyet for all variables as well.

However, another important limitation of the research design can be observed here as well. In the case of Hürriyet, effects of regime change on negative rhetoric cannot clearly be separated from the effects of its change in ownership to a pro-government conglomerate. However, the tests establish that there are significant differences in the coverage of CHP before and after Turkey’s switch to presidentialism.

Overall, findings show that election periods are associated with increased levels of negative language in both newspapers, particularly the enemy rhetoric on Sabah while rival rhetoric does not consistently increase during elections compared to adjacent periods. When comparing Sabah and Hürriyet, Sabah's emphasis on criticisms of CHP diverges starting from mid-2014, in the aftermath of Gezi protests from Hürriyet. From then on, Sabah consistently depicts CHP more intensely as an enemy and rival. This finding aligns with the expectation that Sabah, owned by the pro-incumbent business, vilifies the opposition more with heavier emphasis on showing CHP as an immoral and illegitimate enemy in the political system.

Moreover, acquisition of Hürriyet by another pro-government conglomerate, Demirören Holding, is correlated with a significant increase in enemy and rival depictions of CHP in terms of percentages. After being purchased by Demirören, Hürriyet depicts CHP more negatively as indicated by the statistically significant increases in enemy, rival, and enemy/rival variables. However, these levels are well below compared to those of Sabah, even after the acquisition. In fact, Hürriyet decreases its coverage of CHP significantly afterwards, and even though the relative levels of variables of negative rhetoric increase, in absolute terms, they decline significantly. This is possibly related to the readership base of Hürriyet, which has a more centrist and mixed audience, compared to Sabah which is read mostly by incumbent voters. The contrasting patterns between Sabah and Hürriyet suggest that the readership base and ownership changes play a role in shaping the coverage and language used towards the opposition.

Finally, the hypotheses tests concerned correlations between the regime change in Turkey and negative depictions of CHP. Both Hürriyet and Sabah gets more intense in their criticisms of CHP as a rival and an enemy after Turkey changes into the hyper-presidential system.

CHAPTER 7

CONCLUSION

The unique qualities of authoritarianism today, namely that incumbents undermine democracy gradually, and that elections are the primary source of legitimacy for the incumbent in this process, make partisan ties between the broader society and political elites, and the role of affective polarization in reinforcing these ties central mechanisms. Popular support not only can bring authoritarian leaders to power, but also sustain and legitimize their rule at each election throughout the gradual process of backsliding. As a result, some authoritarian leaders, such as Erdoğan in Turkey, Orban in Hungary, Modi in India enjoy genuine public support while they are paradoxically undermining the democratic institutions of their country (Svolik, 2020).

This study has put forward a novel measure of polarizing rhetoric in the media used against the political opposition, as an attempt to complement the line of studies that focus on affective polarization and negative partisanship to explain political preferences and behavior. In the Turkish case, we attempt to measure two crucial components of such language in all the online news published about the opposition in two largest newspapers since 2009. The goal of the hypotheses tested here were to hint at how the use of this language might have been affected by election periods, different media outlets, and critical junctures relating to the Turkish political system and ownership structures of the newspapers under consideration. Briefly, findings show that election periods witnessed more vilifying language against opposition in the news sources under consideration, that the pro-government media outlet used such language much more commonly than the formerly neutral one, and that they got more polarizing after Turkish authoritarianism consolidated with a hyper-presidential system in 2017. Findings also emphasized the influence of the incumbent over different friendly news outlets is not uniform, and that the readership structure might be influential in determining how the opposition coverage will evolve.

Overall, we observe that these two newspapers, after they are both in the hands of incumbent-friendly conglomerates, may fulfill two interrelated but distinct functions in preparing the ground for authoritarianism. One pro-government news outlet, Sabah, which is read mostly by pro-government voters, increasingly and dramatically casts opposition as immoral and illegitimate, declaring it as an enemy. This helps consolidate and mobilize the voter-base when needed and prevents voters from switching to the opposition. The other one with a more centrist readership base, Hürriyet, instead decreases its coverage of the opposition, giving less and less allocation to the opposition party among other news after its acquisition by another pro-government conglomerate. One implication of this is that decreased coverage helps deplatform and silence the opposition, and pro-democratic opposition loses an important channel of reaching the voter base, while its increasingly portrayed as an enemy of the nation to the incumbent's supporters.

However, there are also several limitations with the research design. The manually coded training data used to train algorithms which predicted the occurrences of immoral or illegitimate representations of the main political opposition in Turkey, was not validated by additional coders and relies, at the end, on the author's subjective assessment. In addition, the processes to create the labeled training datasets used for training the classification algorithms contained some asymmetries, whose impact should nevertheless be minimal. Due to resource constraints, it only focused on two components of negative language, where other components of negative rhetoric and potential changes in components of positive rhetoric remain out of the scope.

Future work can improve the design both methodologically and conceptually. More advanced analysis into the trajectory and amount of negative language over the years with time-series and econometric methods may uncover sudden changes and breakpoints, hinting at events and dates worthy of more scholarly interest. In addition, other dimensions which could instill negative partisanship and new components for the positive dimension of partisanship can be conceptualized and similarly measured. These same dimensions could be tracked also about the incumbent's portrayal in the media. Moreover, many arguably related variables, such as traffic information of the websites under consideration here, their revenue, changes in their size of employment and other patterns could complement the data generated here, to further clarify the evolution of vilifying language through more complex models instead of simple t-tests.

Such a measure with the potential to allow for mobility across different contexts, could be applied to news articles or speeches of political elite across different contexts. Given the rising interest in the rise of affective polarization and negative partisanship across democratic and non-democratic contexts alike, having a comparative measure of the language that could fuel this polarization could provide crucial insights about the support behind authoritarian leaders and what can pro-democratic forces do to claim it.

APPENDIX A: PYTHON SCRIPTS

Below is the code for scraping CHP keyword search results from Hürriyet. The page number parameter in the search URL is updated to reach new search results. Then titles, dates, categories, and descriptions of the news items are saved in a csv file. This constitutes the inventory file which is used to access news links to scrape the text content.

```
for x in range(3, 22):

    df = pd.read_csv("C:/Users/cemal/Desktop/Tez/Hurriyet/CHPNewsLinks-2009.csv", encoding="utf-8-sig")

    headers = list(df["Headers"])
    links = list(df["Links"])
    dates = list(df["Dates"])
    descs = list(df["Descriptions"])
    catgs = list(df["Categories"])
    page = list(df["Page"])
    pageNo = df["Page"].iat[-1]

    for i in range(pageNo+1, pageNo+11):
        options = ChromeOptions()
        options.headless=True

        url = "https://www.hurriyet.com.tr/arama/#/?page=" + str(i) + "&key=CHP&where=hurriyet&how=Article&and=CHP&startDate=01/09/2009&finishDate=31/03/2010&platform=hurriyet&isDetail=true"
        sleep(random.randint(2, 6))
        browser = webdriver.Chrome(chrome_options=options, executable_path=r"C:\Users\cemal\Desktop\Tez\chromedriver.exe")
        browser.minimize_window()
        browser.get(url)
        soup = BeautifulSoup(browser.page_source,"html.parser")
        datas = soup.find_all("div", {"class":"hs-cn-new asd"})
        data = ""

        for data in datas:
            page.append(i)

            headerdata = data.find("p", class_="hs-cnnc-title")
            header = headerdata.get_text()
            headers.append(header)

            datedata = data.find("p", class_="hs-cnncc-date")
            date = datedata.get_text()
            dates.append(date)

            linkblockdata = data.find("div", class_="hs-cnn-content")
            linkdata = linkblockdata.find("a")
            link = linkdata.get("href")
            links.append(link)
```

```
            descdata = data.find("p", class_="hs-cnnc-text")
            desc = descdata.get_text(" ", strip=True)
            descs.append(desc)

            catgsdata = data.find("p", class_="hs-cnncc-category")
            catg = catgsdata.get_text()
            catgs.append(catg)

        print(str(len(datas)) + " headers found on page #" + str(i))
        print(str(len(datas)) + " dates found on page #" + str(i))
        print(str(len(datas)) + " links found on page #" + str(i))
        print(str(len(datas)) + " descriptions found on page #" + str(i))
        print(str(len(datas)) + " categories found on page #" + str(i))

    dfFinal = pd.DataFrame({"Dates":dates, "Links": links, "Headers": headers, "Descriptions": descs, "Categories":catgs, "Page":page})
    print(dfFinal)
    dfFinal.to_csv("C:/Users/cemal/Desktop/Tez/Hurriyet/CHPNewsLinks-2009.csv", encoding="utf-8-sig")
    print("Dataframe is saved. Continuing with the next batch. (" + str(x+1) + "/22 done)")

print("All batches finished (25/25).")
```

Below is the code for scraping news items from Hürriyet. Links are stored in and read from a csv file, and the news texts are written and saved to another one.

```
import requests
from bs4 import BeautifulSoup
import random
from time import sleep
import pandas as pd

def getdata(url):
    r = requests.get(url, headers={"User-Agent": 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/79.0.3945.130 Safari/537.36'})
    return r.text

for x in range(1,200):
    print("At batch #" + str(x))

    df = pd.read_csv("C:/Users/cemal/Desktop/Tez/Hurriyet/OppoNewsLinks.csv", encoding="utf-8-sig")
    df_news = pd.read_csv("C:/Users/cemal/Desktop/Tez/Hurriyet/AllNewsText.csv", encoding="utf-8-sig")
    lastrow = df_news.index[-1]
    firstrow = lastrow + 1

    urlList = list(df_news["URL"]) + list(df["SLinks"][firstrow:])
    dates = list(df_news["Dates"]) + list(df["Dates"][firstrow:])
    h1 = list(df_news["H1"])
    h2 = list(df_news["H2"])
    news = list(df_news["Text"])

    for i in range(firstrow,firstrow+40):
        sleep(random.uniform(1.2, 3.2))
        url = urlList[i]
        htmldata = getdata(url)
        soup = BeautifulSoup(htmldata, "html.parser")
```

```
        data = ""
        data2= ""

        if soup.find("h1"):
            data = soup.find("h1").get_text(" ", strip=True)
            h1.append(data)
            print("H1 of News Story #" + str(i) + ": " + urlList[i][27:] + " is printed")
        if not soup.find("h1"):
            h1.append("None")
            print("No H1 found on News Story #" + str(i) + ": " + urlList[i][27:])

        if soup.find("h2"):
            data2 = soup.find("h2").get_text(" ", strip=True)
            h2.append(data2)
            print("H2 of News Story #" + str(i) + ": " + urlList[i][27:] + " is printed")
        if not soup.find("h2"):
            h2.append("None")
            print("No H2 found on News Story #" + str(i) + ": " + urlList[i][27:])

        if soup.find("div", {"class":"news-detail-text"}):
            newsDt = []
            for new in soup.find_all("div", {"class":"news-detail-text"}):
                newsDt.append(new.get_text(" ", strip=True))
            news.append(" ".join(newsDt))
            print("Text of News Story #" + str(i) + ": " + urlList[i][27:] + " is printed")
        if not soup.find("div", {"class":"news-detail-text"}):
            if soup.find("div", {"class":"news-content readingTime"}):
                newsDt2 = []
                for new in soup.find_all("div", {"class":"news-content readingTime"}):
                    newsDt2.append(new.get_text(" ", strip=True))
                news.append(" ".join(newsDt2))
                print("Text of News Story #" + str(i) + ": " + urlList[i][27:] + " is printed")
            if not soup.find("div", {"class":"news-content readingTime"}):
                news.append("None")
                print("No Text found on News Story #" + str(i) + ": " + urlList[i][27:])

    dfFinal = pd.DataFrame({"Dates":dates[:i], "URL":urlList[:i], "H1":h1[:i], "H2":h2[:i], "Text":news[:i]})
    print(dfFinal)
    dfFinal.to_csv("C:/Users/cemal/Desktop/Tez/Hurriyet/AllNewsText.csv", encoding="utf-8-sig")
    print("dfFinal is saved.")

```

Below is the code for scraping CHP keyword search results from Sabah. The page number parameter in the search URL is updated to reach new search results. Then titles, dates, categories, and descriptions of the news items are saved in a csv file. This constitutes the inventory file which is used to access news links to scrape the text content.

```
import requests
from bs4 import BeautifulSoup
import random
from time import sleep
import pandas as pd

def getdata(url):
    r = requests.get(url, headers={"User-Agent": 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/79.0.3945.130 Safari/537.36'})
    return r.text

df = pd.read_csv("C:/Users/cemal/Desktop/Tez/Sabah/OppoNewsv2Links.csv", encoding="utf-8-sig")
lastrow = df.index[-1]
firstrow = lastrow + 1

dates = list(df["Dates"])
headers = list(df["Headers"])
links = list(df["Links"])
pages = list(df["PageNo"])
pageNo = df["PageNo"].iat[-1]
first = pageNo+1

#specify date range & page range, run beautifulsoup to extract headers, date, and links of news items.
for i in range(first,10812):
    sleep(random.randint(3, 6))
    URL = "https://www.sabah.com.tr/get/arama?query=CHP&page=" + str(i)
    htmldata = getdata(URL)
    soup = BeautifulSoup(htmldata, "html.parser")
    data = ""
    datas = soup.find_all("div", {"class":"col-lg-4 col-md-6 col-sm-6 view20"})
    for data in datas:
        caption = data.find("figcaption")
        a = caption.find("a")
        headers.append(a.get_text(" ", strip=True))
        links.append(a.get("href"))

        info = data.find("span")
        dates.append(info.get_text(" ", strip=True))

        pages.append(i)

    print("Page #" + str(i) + " complete. " + str(len(datas)) + " news found." )
```

Below is the code for scraping news items from Sabah. Links are stored in and read from a csv file, and the news texts are written and saved to another one.

```python
import requests
from bs4 import BeautifulSoup
import random
from time import sleep
import pandas as pd

def getdata(url):
    r = requests.get(url, headers={"User-Agent": 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/105.0.0.0 Safari/537.36'})
    return r.text

for c in range(1761):

    df = pd.read_csv("C:/Users/cemal/Desktop/Tez/Sabah/AllNewsv3Links.csv", encoding="utf-8-sig")
    df_news = pd.read_csv("C:/Users/cemal/Desktop/Tez/Sabah/AllNewsv3Text.csv", encoding="utf-8-sig")

    lastrow = df_news.index[-1]
    firstrow = lastrow + 1
    urlList = list(df_news["URL"]) + list(df["Links"][firstrow:])
    dates = list(df_news["Dates"]) + list(df["Dates"][firstrow:])
    h1 = list(df_news["H1"])
    h2 = list(df_news["H2"])
    news = list(df_news["Text"])

    for i in range(firstrow,firstrow+11):
        sleep(random.uniform(1.0, 2.5))
        url = urlList[i]
        htmldata = getdata(url)
        soup = BeautifulSoup(htmldata, "html.parser")
        data = ""
        data2= ""
        if soup.find("h1"):
            data = soup.find("h1").get_text(" ", strip=True)
            h1.append(data)
            print("H1 of News Story #" + str(i) + ": " + urlList[i][43:] + " is printed")
        if not soup.find("h1"):
            h1.append("None")
            print("No H1 found on News Story #" + str(i) + ": " + urlList[i][43:])
        if soup.find("h2"):
            data2 = soup.find("h2").get_text(" ", strip=True)
            h2.append(data2)
            print("H2 of News Story #" + str(i) + ": " + urlList[i][43:] + " is printed")
        if not soup.find("h2"):
            h2.append("None")
            print("No H2 found on News Story #" + str(i) + ": " + urlList[i][43:])
        if soup.find("div", {"class":"newsDetailText"}):
            newsDt = []
            for new in soup.find_all("div", {"class":"newsDetailText"}):
                newsDt.append(new.get_text(" ", strip=True))
            news.append(" ".join(newsDt))
            print("Text of News Story #" + str(i) + ": " + urlList[i][43:] + " is printed")
        if not soup.find("div", {"class":"newsDetailText"}):
            news.append("None")
            print("No Text found on News Story #" + str(i) + ": " + urlList[i][43:])

    dfFinal = pd.DataFrame({"Dates":dates[:i], "URL":urlList[:i], "H1":h1[:i], "H2":h2[:i], "Text":news[:i]})
    print(dfFinal)
    dfFinal.to_csv("C:/Users/cemal/Desktop/Tez/Sabah/AllNewsv3Text.csv", encoding="utf-8-sig")
```

Below is the code to fine-tune the pre-trained BERTurk algorithm. Adapted from online tutorial by (Albanese, 2022). First the training set is read. Then the algorithm is trained on the training set and applied to the list of sentences stored in a csv file. Below code is for training on the illegitimacy component.

```
import logging
import torch
from torch.utils.data import TensorDataset, DataLoader, RandomSampler, SequentialSampler
from transformers import BertTokenizer, BertForSequenceClassification
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix
from transformers import AutoModel, AutoTokenizer
import torch.nn as nn
import torch.nn.functional as F
import pandas as pd
import numpy as np
from tabulate import tabulate
from tqdm import trange
import random
from cleantext import clean

df = pd.DataFrame({'label':int(), 'text':str()}, index = [])

path = "C:/Users/cemal/Desktop/Tez/Part 2/trainsets/"
trainset = "ml-train-illv1.csv"

params = "ml128-

logger = logging.getLogger(__name__)
logger.setLevel(logging.INFO)

df = df.append(pd.read_csv(path + trainset, encoding="utf-8-sig"))

df["text"] = df["text"].apply(lambda row: clean(row, no_emoji=True, to_ascii=False, lower=False))
df.tail(20)

text = df.text.values
labels = df.label.values

tokenizer = AutoTokenizer.from_pretrained("dbmdz/bert-base-turkish-cased", do_lower_case=False)

def print_rand_sentence():
  '''Displays the tokens and respective IDs of a random text sample'''
  index = random.randint(0, len(text)-1)
  table = np.array([tokenizer.tokenize(text[index]),
tokenizer.convert_tokens_to_ids(tokenizer.tokenize(text[index]))]).T
  print(tabulate(table,
                 headers = ['Tokens', 'Token IDs'],
                 tablefmt = 'fancy_grid'))

print_rand_sentence()

token_id = []
attention_masks = []
```

```

def preprocessing(input_text, tokenizer):
  '''
  Returns <class transformers.tokenization_utils_base.BatchEncoding> with the following fields:
    - input_ids: list of token ids
    - token_type_ids: list of token type ids
    - attention_mask: list of indices (0,1) specifying which tokens should considered by the model (return_attention_mask = True).
  '''
  return tokenizer.encode_plus(
                        input_text,
                        add_special_tokens = True,
                        max_length = 128,
                        pad_to_max_length = True,
                        return_attention_mask = True,
                        return_tensors = 'pt'
                   )

for sample in text:
  encoding_dict = preprocessing(sample, tokenizer)
  token_id.append(encoding_dict['input_ids'])
  attention_masks.append(encoding_dict['attention_mask'])

token_id = torch.cat(token_id, dim = 0)
attention_masks = torch.cat(attention_masks, dim = 0)
labels = torch.tensor(labels)

token_id

def print_rand_sentence_encoding():
  '''Displays tokens, token IDs and attention mask of a random text sample'''
  index = random.randint(0, len(text) - 1)
  tokens = tokenizer.tokenize(tokenizer.decode(token_id[index]))
  token_ids = [i.numpy() for i in token_id[index]]
  attention = [i.numpy() for i in attention_masks[index]]

  table = np.array([tokens, token_ids, attention]).T
  print(tabulate(table,
                 headers = ['Tokens', 'Token IDs', 'Attention Mask'],
                 tablefmt = 'fancy_grid'))

print_rand_sentence_encoding()

val_ratio = 0.2
# Recommended batch size: 16, 32. See: https://arxiv.org/pdf/1810.04805.pdf
batch_size = 32

# Indices of the train and validation splits stratified by labels

train_idx, val_idx = train_test_split(
    np.arange(len(labels)),
    test_size = val_ratio,
    shuffle = True,
    stratify = labels,
    )

# Train and validation sets
train_set = TensorDataset(token_id[train_idx],
                          attention_masks[train_idx],
                          labels[train_idx])

val_set = TensorDataset(token_id[val_idx],
                        attention_masks[val_idx],
                        labels[val_idx])
```

```

# Prepare DataLoader
train_dataloader = DataLoader(
            train_set,
            sampler = RandomSampler(train_set),
            batch_size = batch_size
        )

validation_dataloader = DataLoader(
            val_set,
            sampler = SequentialSampler(val_set),
            batch_size = batch_size
        )

def b_tp(preds, labels):
  '''Returns True Positives (TP): count of correct predictions of actual class 1'''
  return sum([preds == labels and preds == 1 for preds, labels in zip(preds, labels)])

def b_fp(preds, labels):
  '''Returns False Positives (FP): count of wrong predictions of actual class 1'''
  return sum([preds != labels and preds == 1 for preds, labels in zip(preds, labels)])

def b_tn(preds, labels):
  '''Returns True Negatives (TN): count of correct predictions of actual class 0'''
  return sum([preds == labels and preds == 0 for preds, labels in zip(preds, labels)])

def b_fn(preds, labels):
  '''Returns False Negatives (FN): count of wrong predictions of actual class 0'''
  return sum([preds != labels and preds == 0 for preds, labels in zip(preds, labels)])

def b_metrics(preds, labels):
  '''
  Returns the following metrics:
    - accuracy    = (TP + TN) / N
    - precision   = TP / (TP + FP)
    - recall      = TP / (TP + FN)
    - specificity = TN / (TN + FP)
  '''
  preds = np.argmax(preds, axis = 1).flatten()
  labels = labels.flatten()
  tp = b_tp(preds, labels)
  tn = b_tn(preds, labels)
  fp = b_fp(preds, labels)
  fn = b_fn(preds, labels)
  b_accuracy = (tp + tn) / len(labels)
  b_precision = tp / (tp + fp) if (tp + fp) > 0 else 'nan'
  b_recall = tp / (tp + fn) if (tp + fn) > 0 else 'nan'
  b_specificity = tn / (tn + fp) if (tn + fp) > 0 else 'nan'
  return b_accuracy, b_precision, b_recall, b_specificity

# Load the BertForSequenceClassification model
model = BertForSequenceClassification.from_pretrained(
    'dbmdz/bert-base-turkish-cased',
    num_labels = 2,
    output_attentions = False,
    output_hidden_states = False,
)
"""
model = AutoModel.from_pretrained("dbmdz/bert-base-turkish-cased")
"""
# Recommended learning rates (Adam): 5e-5, 3e-5, 2e-5. See: https://arxiv.org/pdf/1810.04805.pdf
optimizer = torch.optim.AdamW(model.parameters(),
                              lr = 2e-5,
                              eps = 1e-08
```

```
                               )

# Run on GPU
model.cuda()

device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

# Recommended number of epochs: 2, 3, 4. See: https://arxiv.org/pdf/1810.04805.pdf
epochs = 3
total_preds=[]
train_losses = []
num_mb_train = len(train_dataloader)

# Create a file handler and add it to the logger
handler = logging.FileHandler(path + params + "lr" +
str(optimizer.param_groups[0]['lr']) + "-bs" + str(batch_size) + "-valr" + str(val_ratio) +
"-eps" + str(epochs) + trainset + '.log')
handler.setLevel(logging.INFO)
formatter = logging.Formatter('%(asctime)s - %(levelname)s - %(message)s')
handler.setFormatter(formatter)
logger.addHandler(handler)

for _ in trange(epochs, desc = 'Epoch'):

    # ========== Training ==========

    # Set model to training mode

    model.train()

    # Tracking variables
    tr_loss = 0
    nb_tr_examples, nb_tr_steps = 0, 0
    w0 = df.label.count() / (2*(df.label.count() - df.label.sum()))
    w1 = (df.label.count() / (2*df.label.sum()))
    weights = torch.FloatTensor([w0,w1]).cuda()

    for step, batch in enumerate(train_dataloader):
        batch = tuple(t.to(device) for t in batch)
        b_input_ids, b_input_mask, b_labels = batch



        # Forward pass
        train_output = model(b_input_ids,
                             token_type_ids = None,
                             attention_mask = b_input_mask,
                             labels = b_labels
)
        criterion = torch.nn.CrossEntropyLoss(weight=weights,reduction='mean')
        loss = criterion(train_output['logits'], b_labels)

        loss.backward()
        torch.nn.utils.clip_grad_norm_(model.parameters(), 1.0)
        optimizer.step()

        # Backward pass
        #train_output.loss.backward()
        #optimizer.step()
        # Update tracking variables
        tr_loss += train_output.loss.item()
        nb_tr_examples += b_input_ids.size(0)
        nb_tr_steps += 1
        optimizer.zero_grad()
        train_losses.append(tr_loss)
```

```
    # ========== Validation ==========

    # Set model to evaluation mode
    model.eval()

    # Tracking variables
    val_accuracy = []
    val_precision = []
    val_recall = []
    val_specificity = []

    for batch in validation_dataloader:
        batch = tuple(t.to(device) for t in batch)
        b_input_ids, b_input_mask, b_labels = batch
        with torch.no_grad():
          # Forward pass
          eval_output = model(b_input_ids,
                              token_type_ids = None,
                              attention_mask = b_input_mask)
        logits = eval_output.logits.detach().cpu().numpy()
        label_ids = b_labels.to('cpu').numpy()
        # Calculate validation metrics
        b_accuracy, b_precision, b_recall, b_specificity = b_metrics(logits, label_ids)
        val_accuracy.append(b_accuracy)
        # Update precision only when (tp + fp) !=0; ignore nan
        if b_precision != 'nan': val_precision.append(b_precision)
        # Update recall only when (tp + fn) !=0; ignore nan
        if b_recall != 'nan': val_recall.append(b_recall)
        # Update specificity only when (tn + fp) !=0; ignore nan
        if b_specificity != 'nan': val_specificity.append(b_specificity)

    los = tr_loss / nb_tr_steps
    accu = sum(val_accuracy)/len(val_accuracy)
    pres = sum(val_precision)/len(val_precision) if len(val_precision)>0 else '\t - Validation Precision: NaN'
    rec = sum(val_recall)/len(val_recall) if len(val_recall)>0 else '\t - Validation Recall: NaN'
    spec = sum(val_specificity)/len(val_specificity) if len(val_specificity)>0 else '\t - Validation Specificity: NaN'

    print('\n\t - Train loss: {:.4f}'.format(tr_loss / nb_tr_steps))
    print('\t - Validation Accuracy: {:.4f}'.format(sum(val_accuracy)/len(val_accuracy)))
    print('\t - Validation Precision: {:.4f}'.format(sum(val_precision)/len(val_precision)) if len(val_precision)>0 else '\t - Validation Precision: NaN')
    print('\t - Validation Recall: {:.4f}'.format(sum(val_recall)/len(val_recall)) if len(val_recall)>0 else '\t - Validation Recall: NaN')
    print('\t - Validation Specificity: {:.4f}\n'.format(sum(val_specificity)/len(val_specificity)) if len(val_specificity)>0 else '\t - Validation Specificity: NaN')
    print("Learning rate:" + str(optimizer.param_groups[0]['lr']))
    print("Batch Size:" +str(batch_size))

    logger.info('\n\t - Train loss: {:.4f}'.format(tr_loss / nb_tr_steps))
    logger.info('\t - Validation Accuracy: {:.4f}'.format(sum(val_accuracy)/len(val_accuracy)))
    logger.info('\t - Validation Precision: {:.4f}'.format(sum(val_precision)/len(val_precision)) if len(val_precision)>0 else '\t - Validation Precision: NaN')
    logger.info('\t - Validation Recall: {:.4f}'.format(sum(val_recall)/len(val_recall)) if len(val_recall)>0 else '\t - Validation Recall: NaN')
```

```
    logger.info('\t - Validation Specificity: {:.4f}\n'.format(sum(val_specificity)/len(val_specificity)) if len(val_specificity)>0 else '\t - Validation Specificity: NaN')
    logger.info("Learning Rate: " + str(optimizer.param_groups[0]['lr']))
    logger.info("Batch Size: " +str(batch_size))
    logger.info("Val Raito: " + str(val_ratio))
    logger.info("Epochs: " + str(epochs))
    logger.info("Training Set: " + trainset)

df_sentsH =  pd.read_csv("C:/Users/cemal/Desktop/Tez/Part 2/Sabah/sabah-oppo-sents.csv", encoding="utf-8-sig")

sents = []
predictions = []
counter = 0

for sent in df_sentsH["Sentences"]:
  new_sentence = sent

  # We need Token IDs and Attention Mask for inference on the new sentence
  test_ids = []
  test_attention_mask = []

  # Apply the tokenizer
  encoding = preprocessing(new_sentence, tokenizer)

  # Extract IDs and Attention Mask
  test_ids.append(encoding['input_ids'])
  test_attention_mask.append(encoding['attention_mask'])
  test_ids = torch.cat(test_ids, dim = 0)
  test_attention_mask = torch.cat(test_attention_mask, dim = 0)

  # Forward pass, calculate logit predictions
  with torch.no_grad():
    output = model(test_ids.to(device), token_type_ids = None, attention_mask = test_attention_mask.to(device))

  prediction = 1 if np.argmax(output.logits.cpu().numpy()).flatten().item() == 1 else 0
  sents.append(new_sentence)
  predictions.append(prediction)
  counter = counter + 1
  if counter % 1000 == 0:
      print("At sentence #" + str(counter) + " out of total " + str(len(df_sentsH["Sentences"])))

df_predH = pd.DataFrame({ trainset: predictions, "text":sents})

df_finalH = pd.concat([df_sentsH, df_predH], axis=1)
df_finalH.to_csv('C:/Users/cemal/Desktop/Tez/Part 2/Hurriyet/hur-res' + "lr"+ str(optimizer.param_groups[0]['lr']) + "-bs" + str(batch_size) + "-valr" + str(val_ratio) + "-eps" + str(epochs) +"-"+ trainset, encoding="utf-8-sig")

```

Below code are for the statistical tests reported in the results section.

```
# Election periods normalized quarterly

from scipy import stats
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from datetime import timedelta

df = pd.read_csv("C:/Users/cemal/Desktop/Tez/Part 2/daily-vars.csv", encoding="utf-8-sig")

df.Dates = pd.to_datetime(df.Dates)

df = df.set_index(df.Dates)
df['Dates'] = df['Dates'].dt.strftime('%Y-%m-%d')
#df = df.reset_index()
df = df.drop(labels="Dates", axis=1)


df["henemy"] = df.hill + df.himm
df["senemy"] = df.sill + df.simm

df["hrival"] = df.hinc + df.hinf
df["srival"] = df.sinc + df.sinf

df["snem"] = df.senemy.div(df.sall).fillna(0)
df["hnem"] = df.henemy.div(df.hall).fillna(0)

df["snri"] = df.srival.div(df.sall).fillna(0)
df["hnri"] = df.hrival.div(df.hall).fillna(0)

df["sdiff"] = df.senemy - df.srival
df["hdiff"] = df.henemy - df.hrival

df["sndiff"] = df.sdiff.div(df.sall).fillna(0)
df["hndiff"] = df.hdiff.div(df.hall).fillna(0)

df["snnews"] = df.snews.div(df.sall).fillna(0)
df["hnnews"] = df.hnews.div(df.hall).fillna(0)

df['sratio'] = np.where(df['srival'] == 0, df['senemy'], df['senemy'] / df['srival'])
df['hratio'] = np.where(df['hrival'] == 0, df['henemy'], df['henemy'] / df['hrival'])

df["sratio-p"] = np.where(df['stotal'] == 0, df['sratio'], df['sratio'] / df['stotal'])
df["hratio-p"] = np.where(df['htotal'] == 0, df['hratio'], df['hratio'] / df['htotal'])

"""
pre = "2018Q4"


election = "2019Q1"


post = "2019Q2"


df1 = df[df['Quarter'] == election]
df2 = df[df["Quarter"] == pre]
df3 = df[df["Quarter"] == post ]

```

```
v=["hnem","snem","hnri","snri","hnnews","snnews","hndiff","sndiff"]

for var in v:

    var1 = df1[var]
    var2 = df2[var]
    var3 = df3[var]

    varss = [var2, var3]
    strs = ["Pre", "Post"]
    dfs = [pre, post]

    for i in range(len(varss)):

        res = stats.ttest_ind(var1,varss[i],alternative="greater")
        print("Election "+ election + " vs. " + strs[i] + " " + dfs[i] +" var  :" + var + " t-stat:" + str(res[0])+":" + str(res[1]))




"""
"""
# REGIME & DEMİRÖREN COMPARISON

from scipy import stats
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from datetime import timedelta

df = pd.read_csv("C:/Users/cemal/Desktop/Tez/Part 2/daily-together-q.csv", encoding="utf-8-sig")

df.Dates = pd.to_datetime(df.Dates)

df = df.set_index(df.Dates)
df['Dates'] = df['Dates'].dt.strftime('%Y-%m-%d')

df["henemy"] = df.hill + df.himm
df["senemy"] = df.sill + df.simm

df["hrival"] = df.hinc + df.hinf
df["srival"] = df.sinc + df.sinf

df["snem"] = df.senemy.div(df.sall).fillna(0)
df["hnem"] = df.henemy.div(df.hall).fillna(0)

df["snri"] = df.srival.div(df.sall).fillna(0)
df["hnri"] = df.hrival.div(df.hall).fillna(0)

df["sdiff"] = df.senemy - df.srival
df["hdiff"] = df.henemy - df.hrival

df["sndiff"] = df.sdiff.div(df.sall).fillna(0)
df["hndiff"] = df.hdiff.div(df.hall).fillna(0)

df["snnews"] = df.snews.div(df.sall).fillna(0)
df["hnnews"] = df.hnews.div(df.hall).fillna(0)

after = "2018-10-01"
before = "2018-06-30"

df1 = df[:before]
df2 = df[after:]
```

```
var = "snnews"
var1 = df1[var]
var2 = df2[var]
"""
"""
res = stats.ttest_ind(var2,var1,alternative="greater")
print("After "+ after + " vs. Before " + before + ": " + "snnews " + " t-stat:" + str(res[0])+":" + str(res[1]))
"""
"""
# ELECTION QUARTERS COMPARISON
from scipy import stats
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from datetime import timedelta
df = pd.read_csv("C:/Users/cemal/Desktop/Tez/Part 2/daily-together-q.csv", encoding="utf-8-sig")
df.Dates = pd.to_datetime(df.Dates)
df = df.set_index(df.Dates)
df['Dates'] = df['Dates'].dt.strftime('%Y-%m-%d')
df["henemy"] = df.hill + df.himm
df["senemy"] = df.sill + df.simm
df["hrival"] = df.hinc + df.hinf
df["srival"] = df.sinc + df.sinf
df["snem"] = df.senemy.div(df.sall).fillna(0)
df["hnem"] = df.henemy.div(df.hall).fillna(0)
df["snri"] = df.srival.div(df.sall).fillna(0)
df["hnri"] = df.hrival.div(df.hall).fillna(0)
df["sdiff"] = df.senemy - df.srival
df["hdiff"] = df.henemy - df.hrival
df["sndiff"] = df.sdiff.div(df.sall).fillna(0)
df["hndiff"] = df.hdiff.div(df.hall).fillna(0)
quarters = ["2010Q3",
"2011Q2",
"2014Q1",
"2014Q2",
"2014Q3",
"2015Q2",
"2015Q3",
"2015Q4",
"2017Q1",
"2017Q2",
"2018Q2",
"2019Q1"]
for q in quarters:
```

```
    dfx = df[df['Quarter'] == q]



    var = dfx["snri"]
    var1 = dfx["hnri"]


    res = stats.ttest_ind(var,var1,alternative="greater")
    print(dfx["Quarter"][0] + ": " + var.name + "t-stat: " + str(res[0]) + ": pvalue: " + str(res[1]))


"""

"""
Tek tek variablelar ve seçim dönemleri



df10refpre = df["2010-05-16":"2010-07-14"]
df10ref = df["2010-07-15":"2010-09-12"]
df10refpost = df ["2010-09-13":"2010-11-11"]

df11genpre = df["2011-02-13":"2011-04-13"]
df11gen = df["2011-04-14":"2011-06-12"]
df11genpost = df ["2011-06-13":"2011-08-11"]

df14locpre = df["2013-12-01":"2014-01-29"]
df14loc = df["2014-01-30":"2014-03-30"]
df14locpost = df ["2014-03-31":"2014-05-29"]

df14prepre = df["2014-04-13":"2014-06-11"]
df14pre = df["2014-06-12":"2014-08-10"]
df14prepost = df["2014-08-11":"2014-10-09"]

df15junpre = df["2015-02-08":"2015-04-08"]
df15jun = df["2015-04-09":"2015-06-07"]
df15junpost = df["2015-06-08":"2015-08-06"]

df15novpre = df["2015-07-05":"2015-09-02"]
df15nov = df ["2015-09-03":"2015-11-01"]
df15novpost = df["2015-11-02":"2015-12-31"]

df17refpre = df["2016-12-18":"2017-02-15"]
df17ref = df["2017-02-16":"2017-04-16"]
df17refpost = df["2017-04-17":"2017-06-15"]

df18genpre = df["2018-02-25":"2018-04-25"]
df18gen = df["2018-04-26":"2018-06-24"]
df18genpost = df["2018-06-25":"2018-08-23"]

df19locpre = df["2018-12-02":"2019-01-30"]
df19loc = df["2019-01-31":"2019-03-31"]
df19locpost = df["2019-04-01":"2019-05-30"]

dfs_dict = {'df10refpre': df10refpre,
            'df10ref': df10ref,
            'df10refpost': df10refpost,
            'df11genpre': df11genpre,
            'df11gen': df11gen,
            'df11genpost': df11genpost,
            'df14locpre': df14locpre,
            'df14loc': df14loc,
```

```
            'df14locpost': df14locpost,
            'df14prepre': df14prepre,
            'df14pre': df14pre,
            'df14prepost': df14prepost,
            'df15junpre': df15junpre,
            'df15jun': df15jun,
            'df15junpost': df15junpost,
            'df15novpre': df15novpre,
            'df15nov': df15nov,
            'df15novpost': df15novpost,
            'df17refpre': df17refpre,
            'df17ref': df17ref,
            'df17refpost': df17refpost,
            'df18genpre': df18genpre,
            'df18gen': df18gen,
            'df18genpost': df18genpost,
            'df19locpre': df19locpre,
            'df19loc': df19loc,
            'df19locpost': df19locpost}

keys = list(dfs_dict.keys())

dfs = [df10refpre,df10ref, df10refpost, df11genpre, df11gen, df11genpost, df14locpre,df14loc, df14locpost, df14prepre, df14pre,df14prepost, df15junpre,df15jun, df15junpost, df15novpre,df15nov, df15novpost, df17refpre,df17ref, df17refpost, df18genpre, df18gen, df18genpost, df19locpre,df19loc, df19locpost]

variables = [ 'sinc',
 'sinf',
 'simm',
 'sill',
 'hinc',
 'hinf',
 'himm',
 'hill',
 'srival',
 'senemy',
 'hrival',
 'henemy',
 'sneg',
 'hneg',
 "sdiff",
 "hdiff"]



for i in range(len(dfs)):
    if i % 3 == 1:
        for var in variables:
            res = stats.ttest_ind(dfs[i][var], dfs[i-1][var],alternative="greater")
            res1 = stats.ttest_ind(dfs[i][var], dfs[i+1][var],alternative="greater")
            print(var + " : " + keys[i] + "-pre-post : " +  str(res[0]) + " : " +str(res[1]) + " : " + str(res1[0]) + " : " + str(res1[1]))
    else:
        continue

"""
"""

#Tek tek variablelar ve Hürriyet vs. Sabah overall



variables_s = [ 'sinc-p',
 'sinf-p',
```

```
 'simm-p',
 'sill-p',
 'srival-p',
 'senemy-p',
 'snegp',
 "sdiff-p",
 "sratio-p"]

variables_h = [  'hinc-p',
 'hinf-p',
 'himm-p',
 'hill-p',
 'hrival-p',
 'henemy-p',
 'hnegp',
 "hdiff-p",
 "hratio-p"]



for i in range(len(variables_s)):
    res = stats.ttest_ind(df[variables_s[i]], df[variables_h[i]], alternative="greater")
    print(str(variables_s[i]) + " : " +  str(res[0]) + " : " +str(res[1]))

"""
"""
#Tek tek variablelar ve Hürriyet vs. Sabah year by year
"""
"""
df09 = df["2009-09-01":"2009-12-31"]
df10 = df["2010-01-01":"2010-12-31"]
df11 = df["2011-01-01":"2011-12-31"]
df12 = df["2012-01-01":"2012-12-31"]
df13 = df["2013-01-01":"2013-12-31"]
df14 = df["2014-01-01":"2014-12-31"]
df15 = df["2015-01-01":"2015-12-31"]
df16 = df["2016-01-01":"2016-12-31"]
df17 = df["2017-01-01":"2017-12-31"]
df18 = df["2018-01-01":"2018-12-31"]
df19 = df["2019-01-01":"2019-12-31"]
df20 = df["2020-01-01":"2020-12-31"]
df21 = df["2021-01-01":"2021-12-31"]
df22 = df["2022-01-01":"2022-08-31"]

dfs = [df09,df10,df11,df12,df13,df14,df15,df16,df17,df18,df19,df20,df21,df22]

dfs_dict = {'df09': df09,
            'df10': df10,
            'df11': df11,
            'df12': df12,
            'df13': df13,
            'df14': df14,
            'df15': df15,
            'df16': df16,
            'df17': df17,
            'df18': df18,
            'df19': df19,
            'df20': df20,
            'df21': df21,
            'df22': df22}

keys = list(dfs_dict.keys())

variables_s = [ 'sinc-p',
```

```
 'sinf-p',
 'simm-p',
 'sill-p',
 'srival-p',
 'senemy-p',
 'snegp',
 "sdiff-p",
 "sratio-p"]

variables_h = [  'hinc-p',
 'hinf-p',
 'himm-p',
 'hill-p',
 'hrival-p',
 'henemy-p',
 'hnegp',
 "hdiff-p",
 "hratio-p"]

for x in range(len(dfs)):

    for i in range(len(variables_s)):
        res = stats.ttest_ind(dfs[x][variables_s[i]], dfs[x][variables_h[i]], alternative="greater")
        print(keys[x] + " : " +  str(variables_s[i]) + " : " +  str(res[0]) + " : " +str(res[1]))


"""
"""
Tek tek variablelar ve Hürriyet vs. Sabah election periods
"""
"""


df10ref = df["2010-07-14":"2010-09-12"]
df11gen = df["2011-04-13":"2011-06-12"]
df14loc = df["2014-01-29":"2014-03-30"]
df14pre = df["2014-06-11":"2014-08-10"]
df15jun = df["2015-04-08":"2015-06-07"]
df15nov = df["2015-09-02":"2015-11-01"]
df17ref = df["2017-02-15":"2017-04-16"]
df18gen = df["2018-04-25":"2018-06-24"]
df19loc = df["2019-01-30":"2019-03-31"]

dfs = [df10ref,df11gen,df14loc,df14pre,df15jun,df15nov,df17ref,df18gen,df19loc]

dfs_dict = {'df10ref': df10ref,
            'df11gen': df11gen,
            'df14loc': df14loc,
            'df14pre': df14pre,
            'df15jun': df15jun,
            'df15nov': df15nov,
            'df17ref': df17ref,
            'df18gen': df18gen,
            'df19loc': df19loc}

keys = list(dfs_dict.keys())

variables_s = [ 'sinc-p',
 'sinf-p',
 'simm-p',
 'sill-p',
 'srival-p',
 'senemy-p',
```

```
 'snegp',
 "sdiff-p",
 "sratio-p"]

variables_h = [  'hinc-p',
 'hinf-p',
 'himm-p',
 'hill-p',
 'hrival-p',
 'henemy-p',
 'hnegp',
 "hdiff-p",
 "hratio-p"]

for x in range(len(dfs)):

    for i in range(len(variables_s)):
        res = stats.ttest_ind(dfs[x][variables_s[i]], dfs[x][variables_h[i]], alternative="greater")
        print(keys[x] + " : " +  str(variables_s[i]) + " : " +  str(res[0]) + " : " +str(res[1]))

"""
"""
#Demirören Test for Hürriyet

#Daily Totals
df = df.reset_index()
df["Demirören"] = df.apply(lambda row: 1 if row.Dates > pd.to_datetime("2018-03-22") else 0, axis= 1)
df = df.set_index(["Dates"])

dfD = df.query("Demirören == 1")
dfH = df.query("Demirören == 0")

variables_hp = [  'hinc-p',
 'hinf-p',
 'himm-p',
 'hill-p',
 'hrival-p',
 'henemy-p',
 'hnegp',
 "hdiff-p",
 "hratio-p",
 "hnews"]

variables_h = [  'hinc',
 'hinf',
 'himm',
 'hill',
 'hrival',
 'henemy',
 'hneg',
 "hdiff",
 "hratio-p"]

for i in range(len(variables_hp)):
    res = stats.ttest_ind(dfD[variables_hp[i]], dfH[variables_hp[i]], alternative="greater")
    print("Percents: " +str(variables_hp[i]) + " : " +  str(res[0]) + " : " + str(res[1]))

for i in range(len(variables_h)):
    res = stats.ttest_ind(dfD[variables_h[i]], dfH[variables_h[i]], alternative="greater")
```

```
    print("Numbers: " +str(variables_h[i]) + " : " +  str(res[0]) + " : " + str(res[1]))

"""

#Hürriyet and Sabah Before & After Regime Change (regime change = 16.04.2017)


df = df.reset_index()
df["Regime1"] = df.apply(lambda row: 1 if row.Dates > pd.to_datetime("2017-04-16") else 0, axis= 1)
df = df.set_index(["Dates"])

dfR = df.query("Regime1 == 1")
dfP = df.query("Regime1 == 0")

variables_h = [  'hinc',
 'hinf',
 'himm',
 'hill',
 'hrival',
 'henemy',
 'hneg',
 "hdiff",
 'hinc-p',
 'hinf-p',
 'himm-p',
 'hill-p',
 'hrival-p',
 'henemy-p',
 'hnegp',
 "hdiff-p",
 "hratio-p"]

variables_s = [  'sinc',
 'sinf',
 'simm',
 'sill',
 'srival',
 'senemy',
 'sneg',
 "sdiff",
 'sinc-p',
 'sinf-p',
 'simm-p',
 'sill-p',
 'srival-p',
 'senemy-p',
 'snegp',
 "sdiff-p",
"sratio-p" ]

#Sabah
for i in range(len(variables_s)):
    res = stats.ttest_ind(dfR[variables_s[i]], dfP[variables_s[i]], alternative="greater")
    print( "Sabah : " +  str(variables_s[i]) + " : " +  str(res[0]) + " : " +str(res[1]))


#Hurriyet
for i in range(len(variables_h)):
    res = stats.ttest_ind(dfR[variables_h[i]], dfP[variables_h[i]], alternative="greater")
```

```
    print( "Hurriyet : " +  str(variables_h[i]) + " : " +  str(res[0]) + " : " +str(res[1]))

"""

Hürriyet and Sabah Before & After Regime Change (regime change = 09.07.2018)


df = df.reset_index()
df["Regime2"] = df.apply(lambda row: 1 if row.Dates > pd.to_datetime("2018-07-09") else 0, axis= 1)
df = df.set_index(["Dates"])

dfR = df.query("Regime2 == 1")
dfP = df.query("Regime2 == 0")

variables_h = [  'hinc',
 'hinf',
 'himm',
 'hill',
 'hrival',
 'henemy',
 'hneg',
 "hdiff",
 'hinc-p',
 'hinf-p',
 'himm-p',
 'hill-p',
 'hrival-p',
 'henemy-p',
 'hnegp',
 "hdiff-p"]

variables_s = [  'sinc',
 'sinf',
 'simm',
 'sill',
 'srival',
 'senemy',
 'sneg',
 "sdiff",
 'sinc-p',
 'sinf-p',
 'simm-p',
 'sill-p',
 'srival-p',
 'senemy-p',
 'snegp',
 "sdiff-p" ]

#Sabah
for i in range(len(variables_s)):
    res = stats.ttest_ind(dfR[variables_s[i]], dfP[variables_s[i]], alternative="greater")
    print( "Sabah : " +  str(variables_s[i]) + " : " +  str(res[0]) + " : " +str(res[1]))


#Hurriyet
for i in range(len(variables_h)):
    res = stats.ttest_ind(dfR[variables_h[i]], dfP[variables_h[i]], alternative="greater")
    print( "Hurriyet : " +  str(variables_h[i]) + " : " +  str(res[0]) + " : " +str(res[1]))
```

APPENDIX B: FIGURES FOR ELECTION PERIODS

Below figures illustrate the 180 period for each election day. Vertical lines delineate the 60-day marks, which are the boundaries of the periods compared in the t-tests. On the left axis, a 60-day moving average of the variable under consideration is given which smooths the area for visual clarity. When taking the average, 60 days before are counted and averaged from a given value's date. On the right axis, dots represent the number of given rhetoric type for each day. Darker colored diamonds visualize the t-tests conducted, with diamonds corresponding to respective period means and error bars representing 1.5 times the standard error.

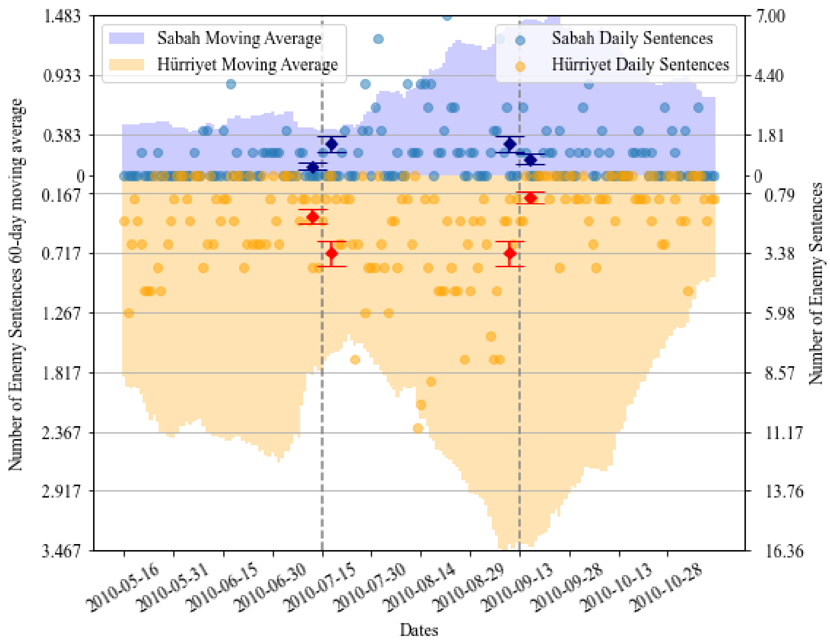

Figure B1. Enemy sentences, 2010 referendum

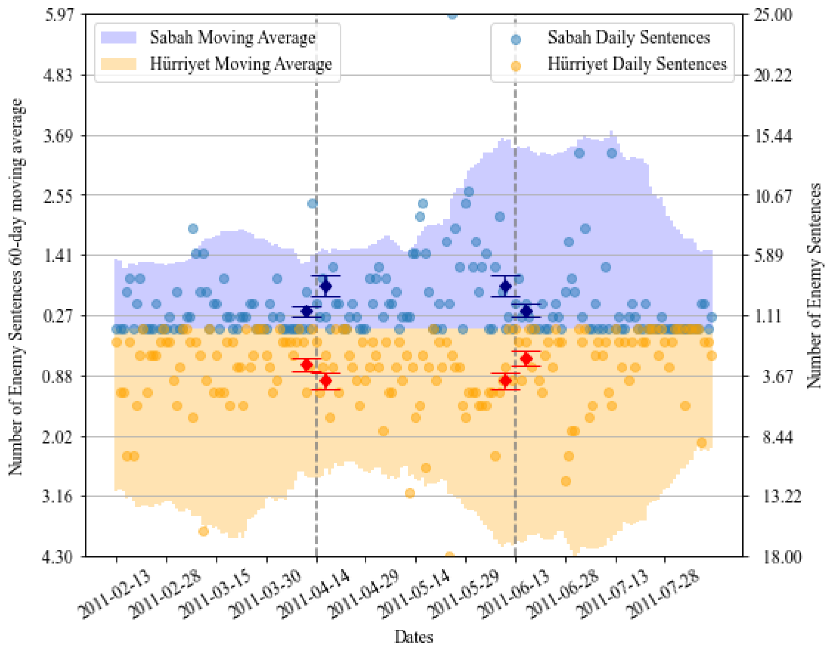

Figure B2. Enemy sentences, 2011 general elections

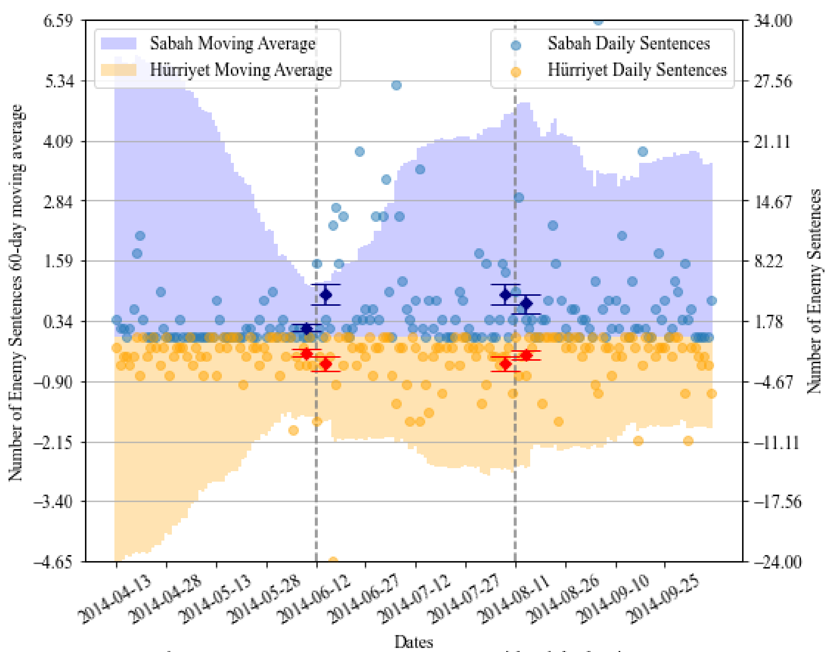

Figure B3. Enemy sentences, 2014 presidential elections

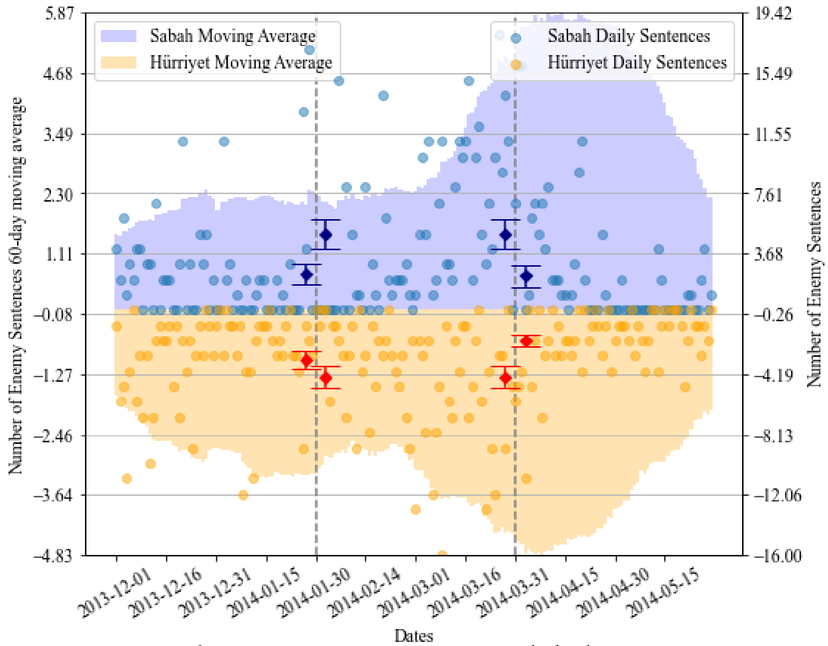

Figure B4. Enemy sentences, 2014 local elections

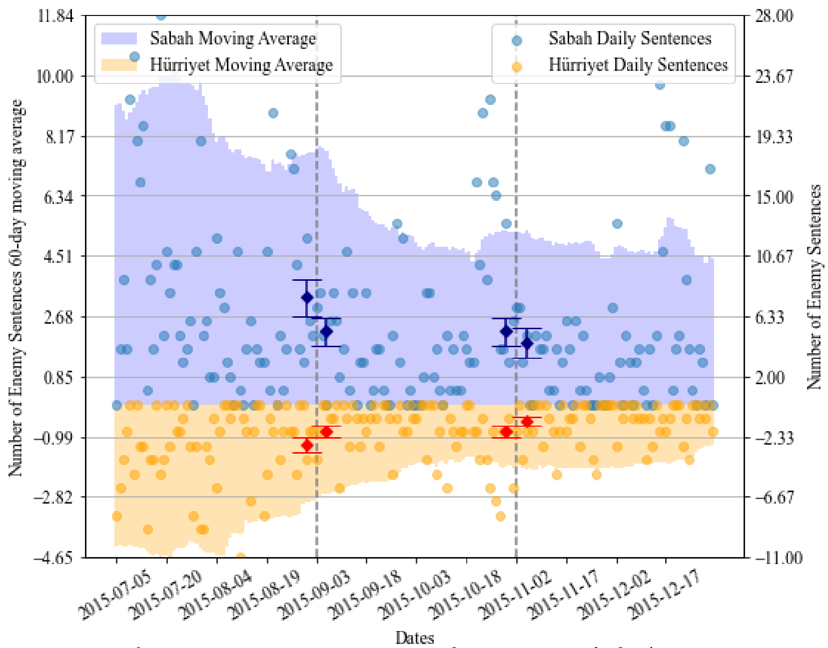

Figure B5. Enemy sentences, November 2015 general elections

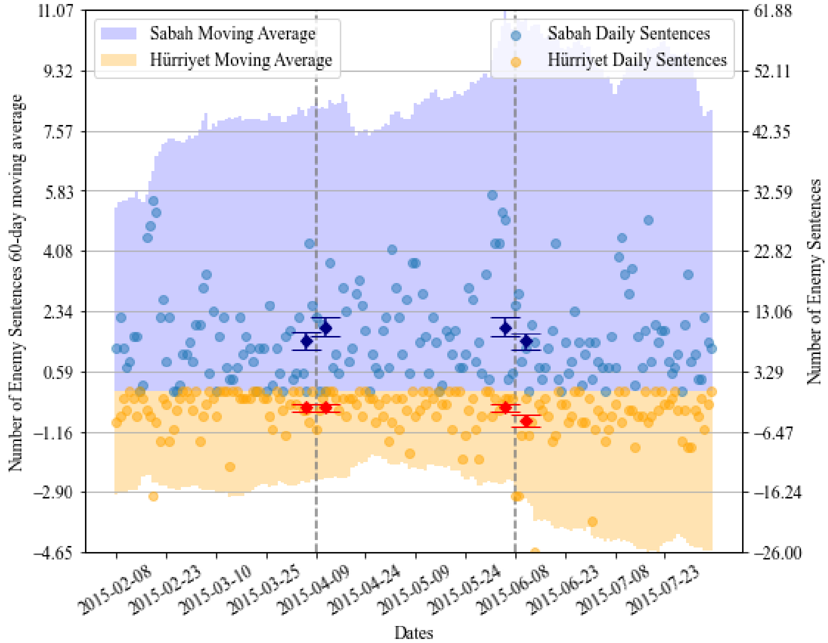

Figure B6. Enemy sentences, June 2015 general elections

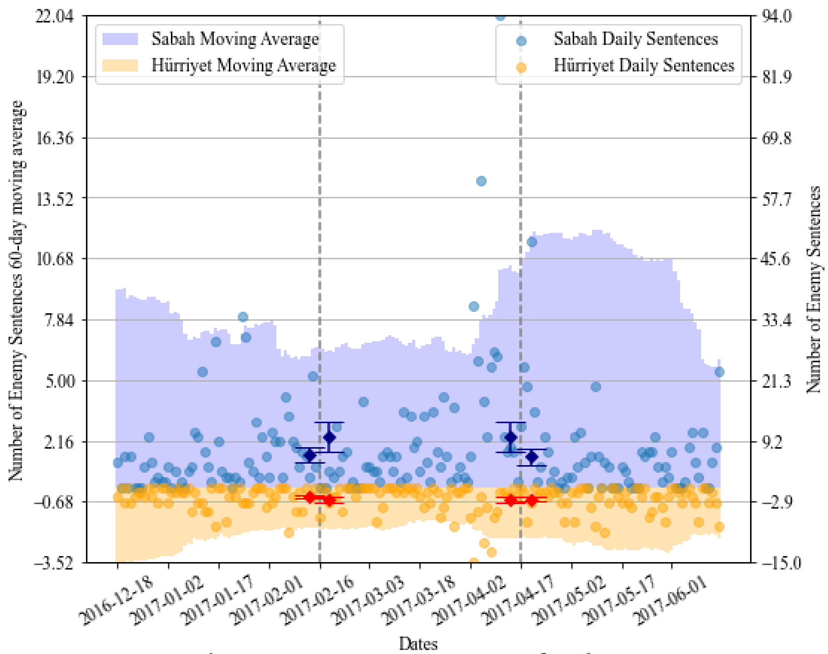

Figure B710. Enemy sentences, 2017 referendum

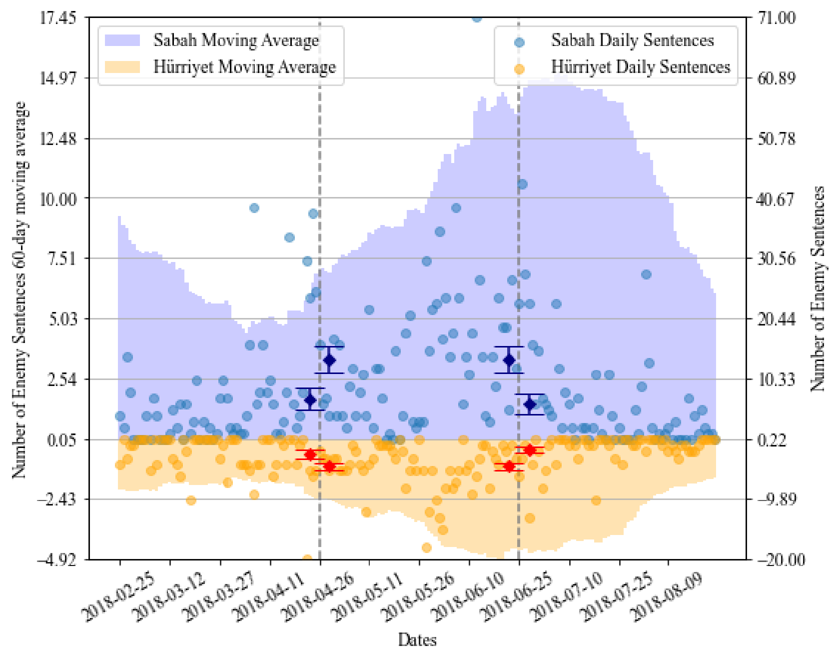

Figure B8. Enemy sentences, 2018 presidential elections

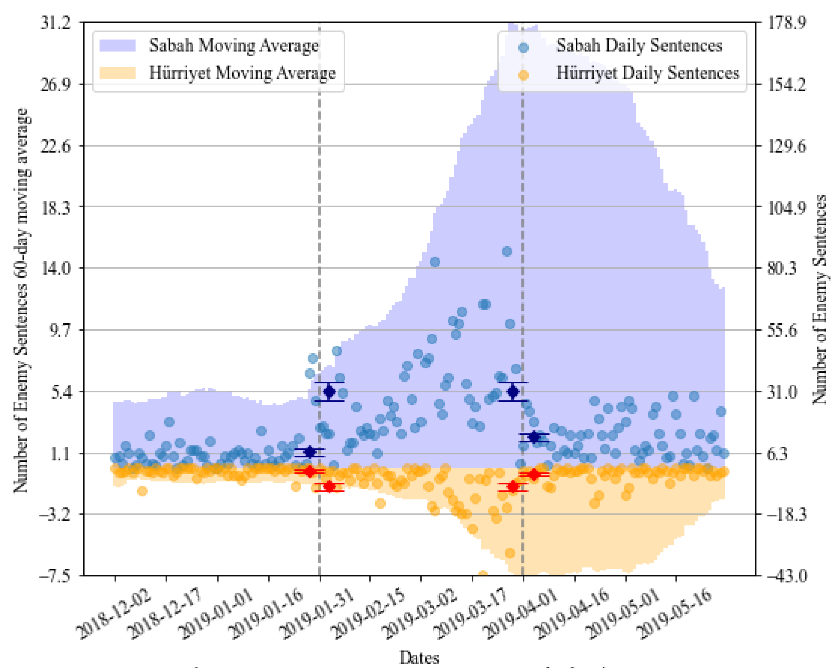

Figure B9. Enemy sentences, 2019 local elections

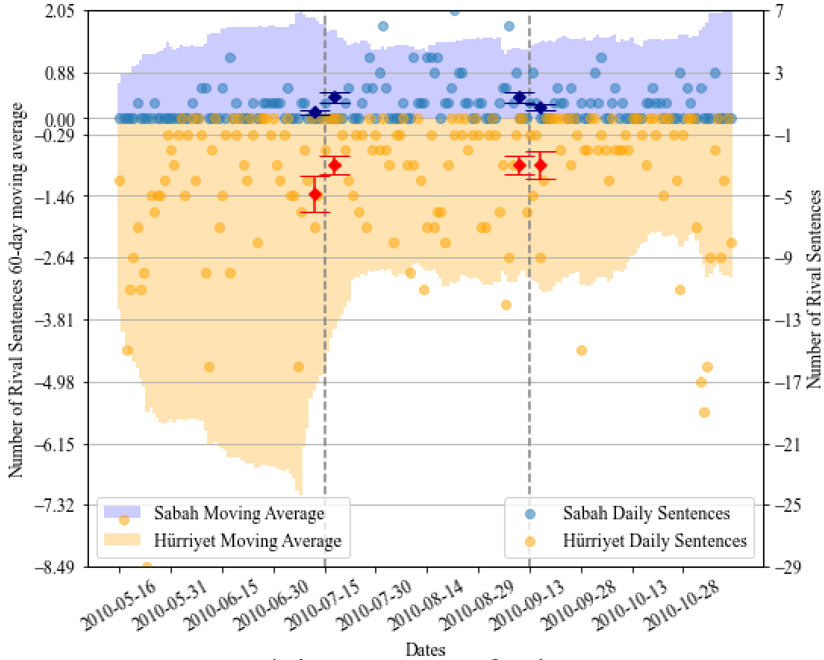

Figure B10. Rival sentences, 2010 referendum

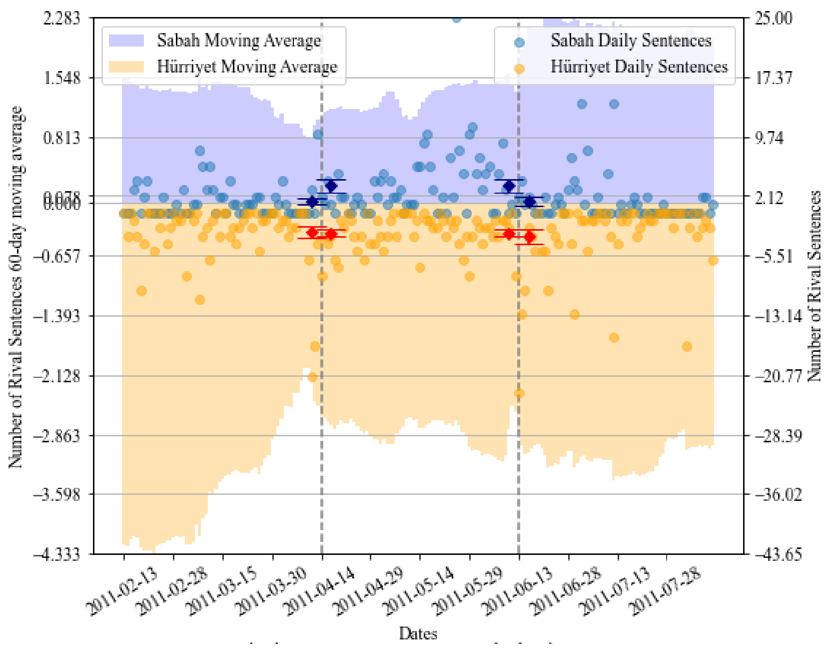

Figure B11. Rival sentences, 2011 general elections

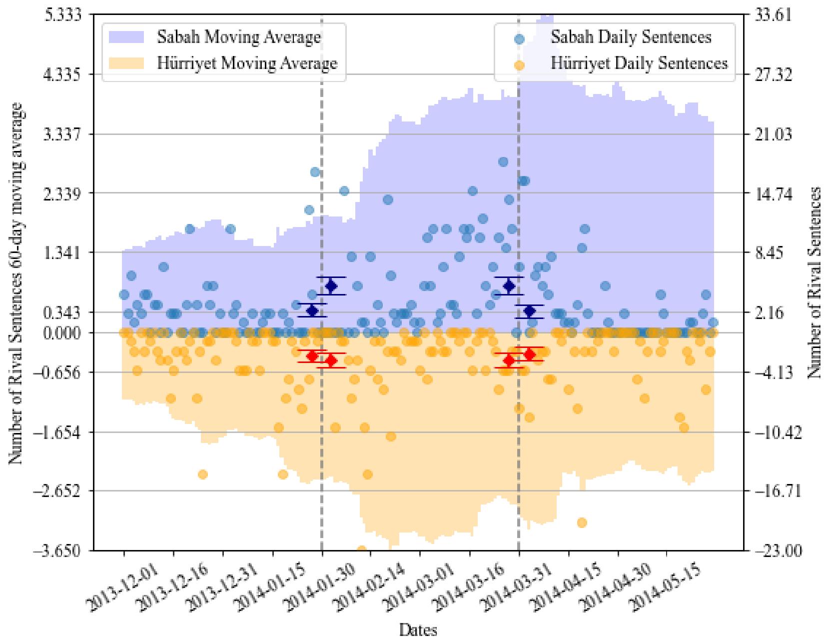

Figure B13. Rival sentences, 2014 local elections

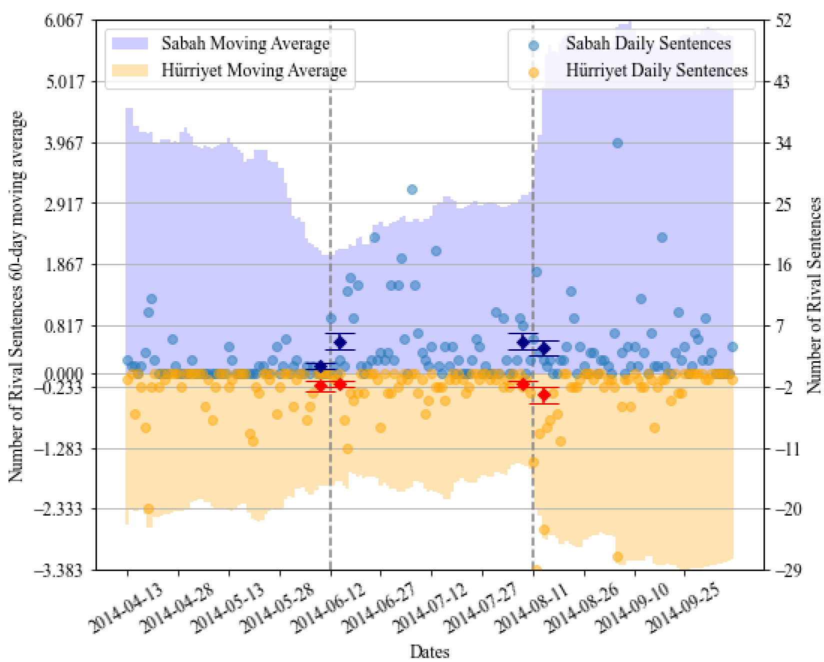

Figure B14. Rival sentences, 2014 presidential elections

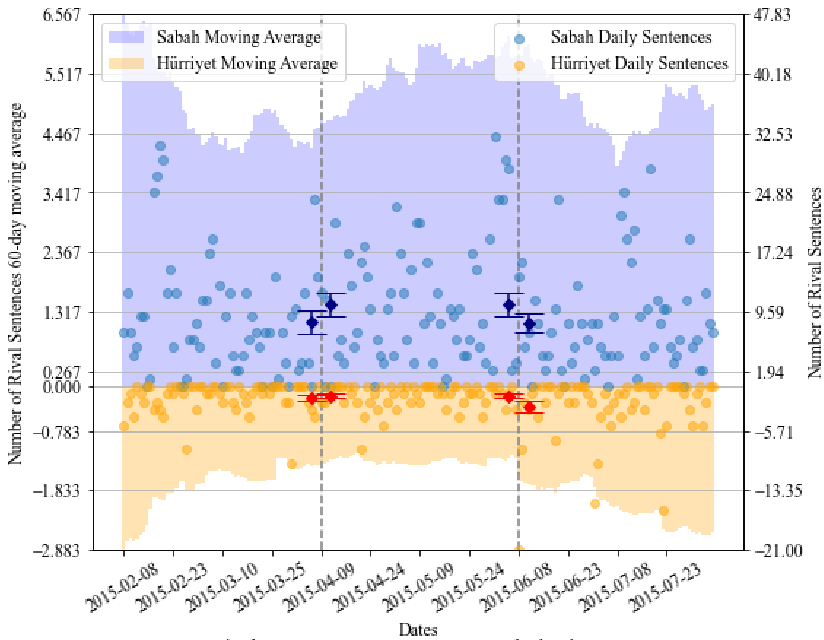

Figure B15. Rival sentences, June 2015 general elections

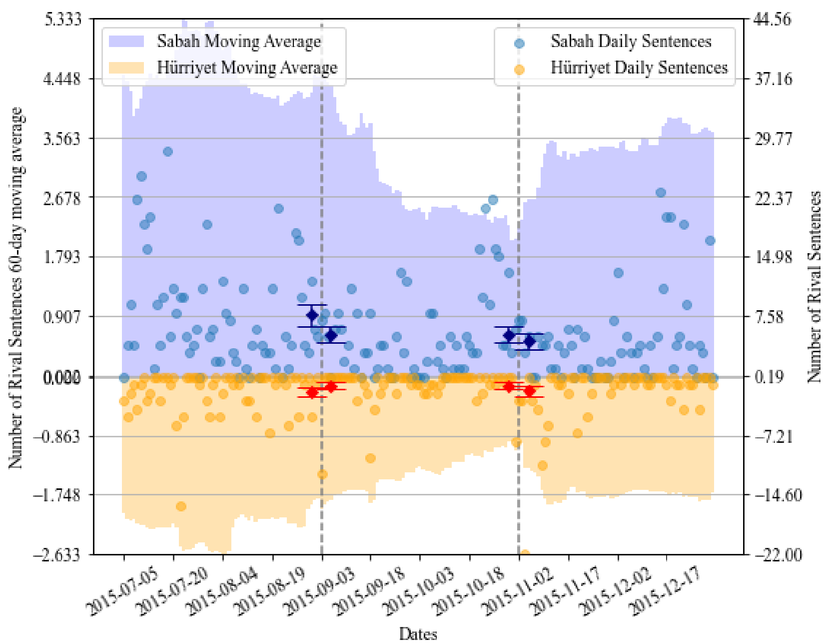

Figure B16. Rival sentences, November 2015 general elections

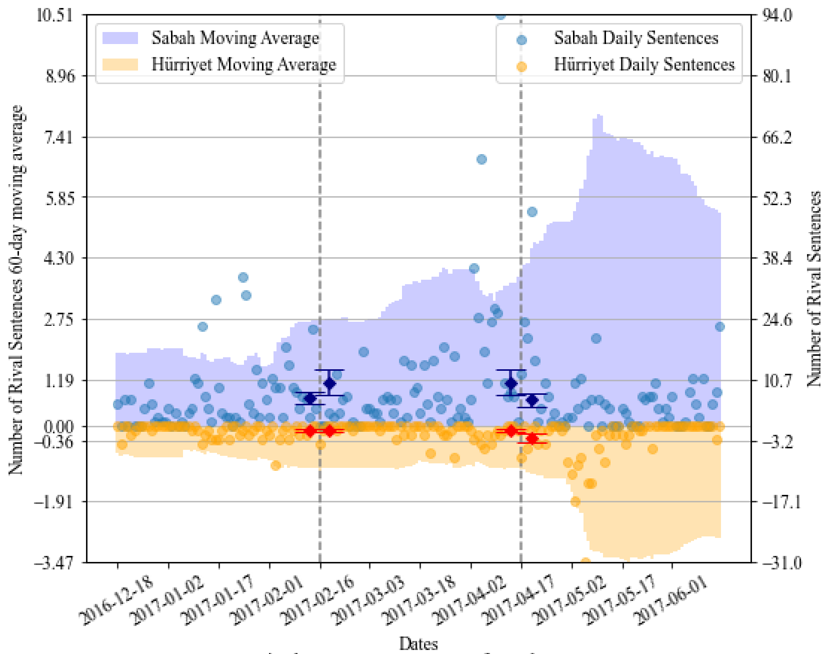

Figure B17. Rival sentences, 2017 referendum

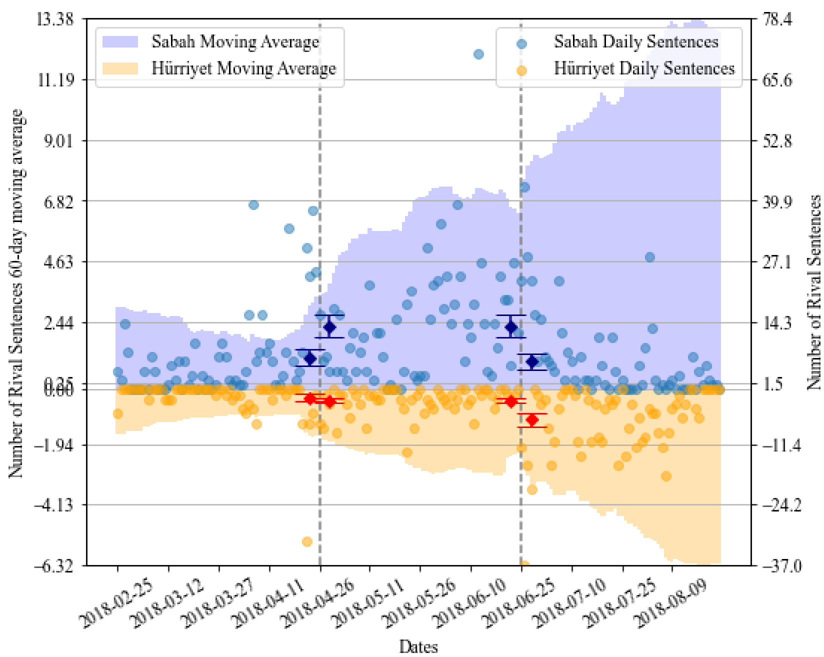

Figure B18. Rival sentences, 2018 presidential elections

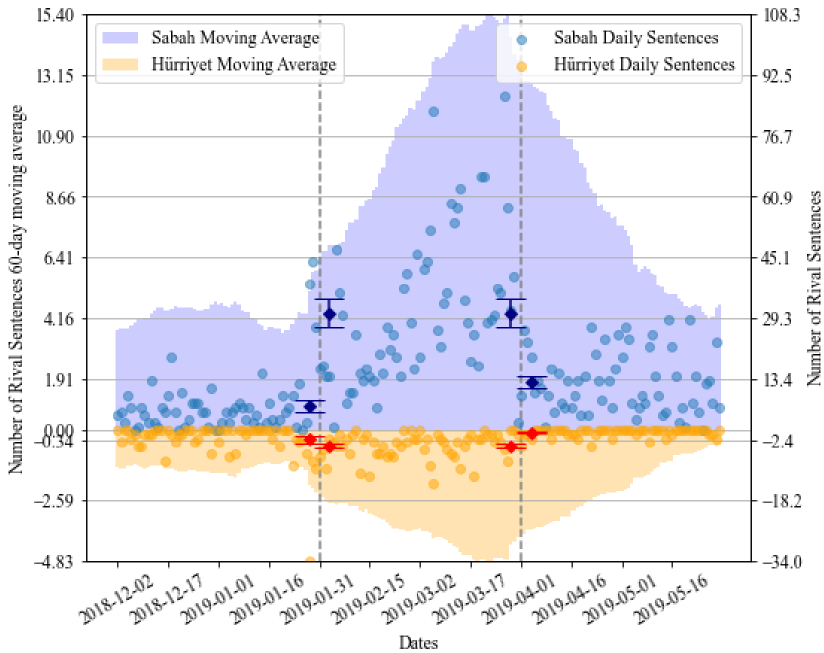

Figure B19. Rival sentences, 2019 local elections

APPENDIX C: ROBUSTNESS CHECKS FOR H1.

Table 13. One Tailed t-test Results for Election Periods, Percentage Values

	Election	2010 Referendum		2011 General Elections	
Website	Variable (Percentage)	Election vs. Pre	Election vs. Post	Election vs. Pre	Election vs. Post
Sabah	Rival	-1.214	-2.286	0.328	-0.972
	Enemy	1.958 **	1.847 **	1.613 *	2.345 **
Hürriyet	Rival	-2.459	-0.415	-0.955	-2.737
	Enemy	2.057 **	5.218 ***	0.648	0.067
	Election	2014 Local Elections		2014 Presidential Elections	
Sabah	Rival	1.130	0.454	-0.123	-1.898
	Enemy	1.541 *	3.205 ***	1.924 **	-1.009
Hürriyet	Rival	0.497	0.960	-1.041	-2.753
	Enemy	-0.307	2.587 ***	0.327	-0.844
	Election	2015 June General Elections		2015 November General Elections	
Sabah	Rival	0.287	1.205	-0.786	-2.422
	Enemy	0.633	2.153 **	1.974 **	1.712 **
Hürriyet	Rival	-0.828	-1.714	-1.709	-3.165
	Enemy	0.560	-2.011	-0.856	-1.084
	Election	2017 Referendum		2018 General Elections	
Sabah	Rival	1.974 **	0.471	-1.035	-6.704
	Enemy	-0.588	0.533	1.499 *	4.489 ***
Hürriyet	Rival	-1.944	-2.966	-1.130	-6.544
	Enemy	-1.626	-1.726	-0.243	0.939
	Election	2019 Local Elections			
Sabah	Rival	-0.102	3.741 ***		
	Enemy	4.901 ***	3.366 ***		
Hürriyet	Rival	0.688	6.103 ***		
	Enemy	5.123 ***	2.646 ***		

Note: Null hypothesis is period means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Table 14. One Tailed T-Test Results For Election Periods, Normalized For Total News Proxy

	Election	2010 Referendum		2011 General Elections	
Website	Variable (Normalized)	Election vs. Pre	Election vs. Post	Election vs. Pre	Election vs. Post
Sabah	Rival	0.903	0.007	1.103	1.379 *
	Enemy	3.364 ***	0.437	2.915 ***	2.539 ***
Hürriyet	Rival	-1.758	0.939	0.799	4.486 ***
	Enemy	1.280	1.172	1.576 *	5.124 ***
	Election	2014 Local Elections		2014 Presidential Elections	
Sabah	Rival	3.366 ***	0.692	2.009 **	-1.236
	Enemy	3.546 ***	2.438 ***	2.453 ***	0.072
Hürriyet	Rival	2.165 **	0.952	1.503 *	0.170
	Enemy	1.678 **	2.832 ***	0.800	-0.314
	Election	2015 June General Elections		2015 November General Elections	
Sabah	Rival	0.499	3.203 ***	-0.581	-0.423
	Enemy	1.824 **	4.353 ***	-1.167	-1.547
Hürriyet	Rival	-0.317	-1.170	-0.651	1.866 **
	Enemy	1.251	0.611	-2.207	0.700
	Election	2017 Referendum		2018 General Elections	
Sabah	Rival	2.063 **	3.679 ***	4.413 ***	-1.613
	Enemy	2.459 ***	-0.219	4.027 ***	5.611 ***
Hürriyet	Rival	2.656 ***	2.650 ***	3.414 ***	0.633
	Enemy	2.846 ***	-0.955	4.653 ***	7.347 ***
	Election	2019 Local Elections			
Sabah	Rival	4.591 ***	4.866 ***		
	Enemy	7.137 ***	2.318 **		
Hürriyet	Rival	3.647 ***	6.368 ***		
	Enemy	4.556 ***	3.767 ***		

Note: Null hypothesis is period means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels. Quarters with the election day are taken as the election quarter and compared with the next and previous quarters.

APPENDIX D: ROBUSTNESS CHECKS FOR H2

Below tables present robustness checks for H2 by comparing absolute number of sentences and their normalized versions by the quarterly total number of news proxy.

Table D1. Sabah vs Hürriyet Comparison, Numbers Overall

Variable (Number of sentences)	t-test
Immoral	27.321***
Illegitimate	21.488***
Incompetent	23.486***
Incohesive	11.971***
Rival	20.506***
Enemy	29.737***
Enemy / Rival	14.380***

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is Sabah mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Table D2. Sabah vs Hürriyet Comparison, Normalized Overall

Variable (number of sentences, normalized)	t-test
Immoral	31.759***
Illegitimate	26.386***
Incompetent	28.631***
Incohesive	18.330***
Rival	35.468***
Enemy	27.918***
Enemy / Rival	22.089***

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is Sabah mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Below tables present robustness checks for H2, focused on the election periods by comparing absolute number of sentences and their normalized versions by the quarterly total number of news proxy.

Table D3. Sabah vs. Hürriyet Comparison, Numbers Election Periods

Election / Variable (Number)	Incompetent	Incohesive	Immoral	Illegitimate	Rival	Enemy	Enemy / Rival
2010 Referendum	-2.280	-2.905	-4.440	-2.585	-3.501	-5.092	-3.306
2011 General	-2.441	-2.079	-2.534	0.758	-3.386	-0.749	1.047
2014 Local	1.277	0.919	-0.099	1.172	1.222	0.595	-0.693
2014 Presidential	0.766	3.144***	2.136**	1.149	2.814***	1.981**	-0.430
2015 June General	5.719***	3.125***	6.062***	6.005***	6.579***	7.461***	1.431**
2015 November General	1.808**	1.339*	3.475***	3.688***	1.898**	4.597***	3.016***
2017 Referendum	5.426***	0.958	4.663***	2.784***	4.306***	3.825***	2.791***
2018 General	3.750***	3.104***	4.559***	4.889***	4.590***	5.724***	2.794***
2019 Local	6.020***	4.343***	6.764***	7.919***	7.166***	8.649***	1.382*

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Table D4. Sabah vs. Hürriyet Comparison, Normalized Election Periods

Election / Variable (Normalized)	Incompetent	Incohesive	Immoral	Illegitimate	Rival	Enemy	Enemy / Rival
2010 Referendum	1.357*	0.283	0.336	-0.101	0.871	0.132	1.209
2011 General	0.150	0.104	2.385***	0.623	0.1460	2.119**	2.863***
2014 Local	2.545***	2.666***	3.464***	2.605***	3.374***	3.534***	1.640*
2014 Presidential	1.828**	3.323***	3.317***	3.840***	3.115***	4.117***	2.388***
2015 June General	6.306***	4.147***	6.587***	7.530***	7.260***	8.541***	3.600***
2015 November General	2.621***	2.941***	4.379***	5.040***	3.325***	5.486***	3.228***
2017 Referendum	4.737***	1.502*	4.476***	4.311***	3.141***	4.945***	3.321***
2018 General	4.716***	2.818***	6.532***	5.859***	4.164***	7.604***	3.921***
2019 Local	6.836***	4.572***	7.755***	7.643***	7.333***	8.599***	2.890***

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

.

Table D5. Sabah vs. Hürriyet Comparison, Numbers and Normalized Year by Year

Year / Variable (Numbers)	Incompetent	Incohesive	Immoral	Illegitimate	Rival	Enemy	Enemy / Rival
2009	0.000	-2.562	-3.364	-0.293	-2.361	-2.197	-1.777
2010	-3.901	-5.041	-6.932	-4.404	-5.628	-7.882	-5.241
2011	-3.500	-3.532	-4.847	-3.614	-4.438	-5.137	-2.776
2012	-3.192	-1.732	-4.974	-4.563	-3.127	-5.549	-5.258
2013	0.358	-0.277	-0.640	-3.572	-0.102	-2.725	-3.427
2014	2.493***	4.203***	2.176**	3.784***	4.348***	3.567***	-1.235
2015	7.377***	6.210***	9.883***	8.529***	8.639***	11.454***	5.014***
2016	5.314***	4.489***	5.950***	7.360***	6.054***	8.411***	5.401***
2017	8.190***	1.150	8.799***	8.630***	4.819***	10.190***	7.534***
2018	7.668***	4.827***	9.607***	8.311***	7.138***	10.356***	6.439***
2019	13.329***	5.932***	14.004***	10.600***	11.806***	14.937***	8.322***
2020	14.730***	9.010***	14.005***	10.614***	14.766***	16.500***	10.361***
2021	13.395***	8.205***	18.978***	9.837***	12.329***	18.580***	8.019***
2022	10.915***	9.106***	13.408***	8.310***	12.656***	14.292***	6.465***

Note: Null hypothesis is period means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Table D6. Sabah vs. Hürriyet Comparison, Normalized Year by Year

Year / Variable (Normalized)	Incompetent	Incohesive	Immoral	Illegitimate	Rival	Enemy	Enemy / Rival
2009	2.021**	-1.391	-0.777	1.841**	-0.495	0.952	1.074
2010	1.193	-1.213	-2.292	-0.127	-0.588	-1.763	0.144
2011	1.362*	1.367*	1.066	2.297**	1.730**	2.236**	3.150***
2012	0.627	2.119**	0.037	-1.156	2.050**	-0.850	-0.017
2013	2.767***	3.095***	3.151***	0.682	3.937***	2.498***	1.775**
2014	5.297***	6.499***	5.844***	6.960***	7.250***	7.706***	3.240***
2015	8.547***	7.521***	11.098***	10.132***	10.358***	13.185***	6.919***
2016	5.404***	4.821***	6.030***	7.360***	6.219***	8.416***	5.144***
2017	9.163***	2.119**	9.534***	9.559***	6.139***	11.144***	8.977***
2018	9.596***	7.388***	11.836***	10.118***	10.061***	12.865***	9.130***
2019	14.234***	6.735***	14.692***	11.758***	13.251***	16.397***	9.730***
2020	14.547***	8.908***	14.104***	10.490***	14.638***	16.464***	10.219***
2021	14.386***	9.069***	19.913***	9.806***	13.775***	18.661***	9.732***
2022	12.447***	10.062***	14.046***	9.566***	14.216***	15.595***	9.191***

Note: Null hypothesis is newspaper means are not different. Alternative hypothesis is election period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

APPENDIX E: ROBUSTNESS CHECKS FOR H3

Table E1. Hürriyet Comparison Before and After Demirören Acquisition Numbers

Comparison period	Variable (Numbers)	t-test
After 2018-03-22 vs. Before 2018-03-21	Incompeten	1.464**
After 2018-03-22 vs. Before 2018-03-21	Immoral	2.186**
After 2018-03-22 vs. Before 2018-03-21	Enemy	-0.324
Before 2018-03-21 vs. After 2018-03-22	Incohesive	3.661***
Before 2018-03-21 vs. After 2018-03-22	Illegitimate	2.173**
Before 2018-03-21 vs. After 2018-03-22	Rival	2.352***
Before 2018-03-21 vs. After 2018-03-22	Enemy / Rival	2.992**
Before 2018-03-21 vs. After 2018-03-22	CHP News	21.624***

Note: Null hypothesis is period means are not different. Alternative hypothesis is first period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Table E2. Hürriyet Comparison Before and After Demirören Acquisition Normalized

Comparison period	Variable (Normalized)	t-test
After 2018-03-22 vs. Before 2018-03-21	Incompetent	1.792**
After 2018-03-22 vs. Before 2018-03-21	Immoral	2.745**
After 2018-03-22 vs. Before 2018-03-21	Enemy	-0.158
Before 2018-03-21 vs. After 2018-03-22	Incohesive	4.043***
Before 2018-03-21 vs. After 2018-03-22	Illegitimate	2.413**
Before 2018-03-21 vs. After 2018-03-22	Rival	2.42***
Before 2018-03-21 vs. After 2018-03-22	Enemy / Rival	3.003***
Before 2018-03-21 vs. After 2018-03-22	CHP News	21.178***

Note: Null hypothesis is period means are not different. Alternative hypothesis is first period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

APPENDIX F: ROBUSTNESS CHECKS FOR H4 & H5

Below tables present robustness tests for hypotheses tested in H4 and H5, focusing on the periods before and after Turkey's regime change into hyper-presidentialism.

Table F1. Sabah and Hürriyet Comparison, Numbers Before and After Regime Change

Variable (Numbers)	Sabah	Hürriyet
Incompetent	26.476***	-0.216
Incohesive	12.407***	-4.7576
Immoral	34.682***	1.425*
Illegitimate	14.557***	-3.632
Enemy	30.214***	-1.7515
Rival	22.531***	-3.978
Enemy / Rival	11.507***	-3.936

Note: Null hypothesis is period means are not different. Alternative hypothesis is period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

Table F2. Sabah and Hürriyet Comparison, Normalized Before and After Regime Change

Variable (Normalized)	Sabah	Hürriyet
Incompetent	21.825***	0.139
Incohesive	8.027***	-5.153
Immoral	30.960***	2.022**
Illegitimate	11.740***	-3.856
Enemy	25.998***	-1.557
Rival	16.895***	-4.053
Enemy / Rival	7.803***	-3.899

Note: Null hypothesis is period means are not different. Alternative hypothesis is period mean is higher. Asterisks indicate statistical significance of mean differences at 1% (***), 5% (**), 10% (*) levels.

REFERENCES

Abramowitz, A. I., & Webster, S. (2016). The rise of negative partisanship and the nationalization of U.S. elections in the 21st century. *Electoral Studies*, *41*, 12–22. https://doi.org/10.1016/j.electstud.2015.11.001

Agamben, G. (2005). *State of exception*. Chicago: University of Chicago Press.

Ahlquist, J. S., Ichino, N., Wittenberg, J., & Ziblatt, D. (2018). How do voters perceive changes to the rules of the game? Evidence from the 2014 Hungarian elections. *Journal of Comparative Economics*, *46*(4), 906–919. https://doi.org/10.1016/j.jce.2018.01.001

Albanese, N. C. (2022, May 16). Fine-tuning BERT for text classification. Retrieved May 27, 2023, from Medium website: https://towardsdatascience.com/fine-tuning-bert-for-text-classification-54e7df642894

Arslanalp, M., & Erkmen, T. D. (2020). Repression without exception: A study of protest bans during Turkey's state of emergency (2016-2018). *South European Society and Politics*, *25*(1), 99–125. https://doi.org/10.1080/13608746.2020.1748353

Bellucci, P. (2006). Tracing the cognitive and affective roots of 'party competence': Italy and Britain, 2001. *Electoral Studies, 25*(3), 548–569. https://doi.org/10.1016/j.electstud.2005.06.014

Bermeo, N. (2016). On democratic backsliding. *Journal of Democracy*, *27*(1), 5–19. https://doi.org/10.1353/jod.2016.0012

Bougher, L. D. (2017). The correlates of discord: Identity, issue alignment, and political hostility in polarized America. *Political Behavior, 39*(3), 731–762. https://doi.org/10.1007/s11109-016-9377-1

Budak, C., Goel, S., & Rao, J. M. (2016). Fair and balanced? Quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, *80*(S1), 250–271. https://doi.org/10.1093/poq/nfw007

Çalışkan, K. (2018). Toward a new political regime in Turkey: From competitive toward full authoritarianism. *New Perspectives on Turkey*, *58*, 5–33. https://doi.org/10.1017/npt.2018.10

Çarkoğlu, A., Baruh, L., & Yıldırım, K. (2014). Press-party parallelism and polarization of news media during an election campaign: The case of the 2011 Turkish elections. *The International Journal of Press/Politics*, *19*(3), 295–317. https://doi.org/10.1177/1940161214528994

Carmines, E. G., Gerrity, J. C., & Wagner, M. W. (2010). How abortion became a partisan issue: Media coverage of the interest group-political party connection: carmines et al. / how abortion became a partisan issue. *Politics & Policy*, *38*(6), 1135–1158. https://doi.org/10.1111/j.1747-1346.2010.00272.x

Caruana, N. J., McGregor, R. M., & Stephenson, L. B. (2015). The power of the dark side: Negative partisanship and political behaviour in Canada. *Canadian Journal of Political Science*, *48*(4), 771–789. https://doi.org/10.1017/S0008423914000882

Classification: Precision and recall | machine learning. (2022). Retrieved May 27, 2023, from Google for Developers website: https://developers.google.com/machine-learning/crash-course/classification/precision-and-recall

Dalton, R. J. (1987). Generational change in elite political beliefs: The growth of ideological polarization. *The Journal of Politics*, *49*(4), 976–997. https://doi.org/10.2307/2130780

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019, May 24). *BERT: Pre-training of deep bidirectional transformers for language understanding*. arXiv. https://doi.org/10.48550/arXiv.1810.04805

Druckman, J. N., Gubitz, S. R., Levendusky, M. S., & Lloyd, A. M. (2019). How incivility on partisan media (de)polarizes the electorate. *The Journal of Politics*, *81*(1), 291–295. https://doi.org/10.1086/699912

Erickson, B. J., & Kitamura, F. (2021). Magician's corner: 9. performance metrics for machine learning models. *Radiology. Artificial Intelligence*, *3*(3), e200126. https://doi.org/10.1148/ryai.2021200126

Esen, B., & Gumuscu, S. (2016). Rising competitive authoritarianism in Turkey. *Third World Quarterly*, *37*(9), 1581–1606. https://doi.org/10.1080/01436597.2015.1135732

Esen, B., & Gümüşçü, Ş. (2017). A small yes for presidentialism: The Turkish constitutional referendum of april 2017. *South European Society and Politics*, *22*(3), 303–326. https://doi.org/10.1080/13608746.2017.1384341

Esen, B., & Gumuscu, S. (2018). The perils of "Turkish presidentialism." *Review of Middle East Studies*, *52*(1), 43–53. https://doi.org/10.1017/rms.2018.10

Ewing, J. (2020, August 27). Turkey braces for yet another currency crisis. *The New York Times*. Retrieved from https://www.nytimes.com/2020/08/27/business/turkey-currency-crisis.html

Fernández, A., García, S., Galar, M., Prati, R., Krawczyk, B., & Herrera, F. (2018). *Cost-sensitive learning*. https://doi.org/10.1007/978-3-319-98074-4_4

Gandhi, J., & Lust-Okar, E. (2009). Elections under authoritarianism. *Annual Review of Political Science*, *12*(1), 403–422. https://doi.org/10.1146/annurev.polisci.11.060106.095434

Garzia, D., & Ferreira da Silva, F. (2021). Negative personalization and voting behavior in 14 parliamentary democracies, 1961–2018. *Electoral Studies*, *71*, 102300. https://doi.org/10.1016/j.electstud.2021.102300

Gidengil, E., Stolle, D., & Bergeron-Boutin, O. (2021). The partisan nature of support for democratic backsliding: A comparative perspective. *European Journal of Political Research*, 1475-6765.12502. https://doi.org/10.1111/1475-6765.12502

Göle, N., & Eşkinat, D. (2013). Gezi – anatomy of a public square movement. *Insight Turkey, 15(3), 7-14.*

Greene, Z. D., & Haber, M. (2015). The consequences of appearing divided: An analysis of party evaluations and vote choice. *Electoral Studies*, *37*, 15–27. https://doi.org/10.1016/j.electstud.2014.11.002

Guriev, S., & Treisman, D. (2019). Informational autocrats. *Journal of Economic Perspectives*, *33*(4), 100–127. https://doi.org/10.1257/jep.33.4.100

Hart, P. S., Chinn, S., & Soroka, S. (2020). Politicization and polarization in COVID-19 news coverage. *Science Communication*, *42*(5), 679–697. https://doi.org/10.1177/1075547020950735

Hester, J. B., & Dougall, E. (2007). The efficiency of constructed week sampling for content analysis of online news. *Journalism & Mass Communication Quarterly*, *84*(4), 811–824. https://doi.org/10.1177/107769900708400410

Insel, A. (2003). The AKP and normalizing democracy in Turkey. *South Atlantic Quarterly*, *102*(2–3), 293–308. https://doi.org/10.1215/00382876-102-2-3-293

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology. *Public Opinion Quarterly*, *76*(3), 405–431. https://doi.org/10.1093/poq/nfs038

Kalkar, Ö. Ş., & Öcal, A. (2020). Bir içerik analizi: Cumhurbaşkanliği hükümet sistemi. *Balıkesir Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *1*(2), 1–30.

Kamaldeep Singh. (2020). How to improve class imbalance using class weights in machine learning. Retrieved May 27, 2023, from Analytics Vidhya website: https://www.analyticsvidhya.com/blog/2020/10/improve-class-imbalance-class-weights/

Kaur, H., Pannu, H., & Malhi, A. (2019). A systematic review on imbalanced data challenges in machine learning: Applications and solutions. *ACM Computing Surveys*, *52*, 1–36. https://doi.org/10.1145/3343440

Kaya, Z. N., & Whiting, M. (2019). The HDP, the AKP and the battle for Turkish democracy. *Ethnopolitics*, *18*(1), 92–106. https://doi.org/10.1080/17449057.2018.1525168

Keddie, P. (2020, October 30). What's Turkey's role in the Nagorno-Karabakh conflict? Retrieved May 14, 2023, from https://www.aljazeera.com/features/2020/10/30/whats-turkeys-role-in-the-nagorno-karabakh-conflict

Kubin, E., & von Sikorski, C. (2021). The role of (social) media in political polarization: A systematic review. *Annals of the International Communication Association*, *45*(3), 188–206. https://doi.org/10.1080/23808985.2021.1976070

Kurt, E. V. (2016). Medya ve kutuplaşma: Kutuplaşmanın medyaya yansımalarının, 'Gezi Parkı eylemleri' üzerinden değerlendirilmesi. *Toplum ve Demokrasi Dergisi*, *9*(19), 193–223.

Laebens, M. G., & Öztürk, A. (2021). Partisanship and autocratization: Polarization, power asymmetry, and partisan social identities in Turkey. *Comparative Political Studies*, *54*(2), 245–279. https://doi.org/10.1177/0010414020926199

Lau, R. R., Andersen, D. J., Ditonto, T. M., Kleinberg, M. S., & Redlawsk, D. P. (2017). Effect of media environment diversity and advertising tone on information search, selective exposure, and affective polarization. *Political Behavior*, *39*(1), 231–255. https://doi.org/10.1007/s11109-016-9354-8

Leonard Richardson. (2007). *Beautiful soup documentation*. Retrieved from https://www.crummy.com/software/BeautifulSoup/bs4/doc/index.html

Levendusky, M. (2010). *The partisan sort: How liberals became Democrats and conservatives became Republicans*. Chicago: University of Chicago Press.

Levendusky, M. (2013). Partisan media exposure and attitudes toward the opposition. *Political Communication*, *30*(4), 565–581. https://doi.org/10.1080/10584609.2012.737435

Levendusky, M. S., & Malhotra, N. (2016). (Mis)perceptions of partisan polarization in the American public. *Public Opinion Quarterly*, *80*(S1), 378–391. https://doi.org/10.1093/poq/nfv045

Levitsky, S., & Way, L. (2020). The new competitive authoritarianism. *Journal of Democracy*, *31*(1), 51–65. https://doi.org/10.1353/jod.2020.0004

Levitsky, S., & Ziblatt, D. (2018). *How democracies die* (First edition). New York: Crown.

Lipton, Z. C., Elkan, C., & Naryanaswamy, B. (2014). Optimal thresholding of classifiers to maximize F1 measure. In T. Calders, F. Esposito, E. Hüllermeier, & R. Meo (Eds.), *Machine Learning and Knowledge Discovery in Databases* (pp. 225–239). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-44851-9_15

Loper, E., & Bird, S. (2002). NLTK: The natural language toolkit. *CoRR*, *cs.CL/0205028*. https://doi.org/10.3115/1118108.1118117

Lührmann, A., & Rooney, B. (2021). Autocratization by decree: States of emergency and democratic decline. *Comparative Politics*, *53*(4), 617–649. https://doi.org/10.5129/001041521X16004520146485

Luque, A., Carrasco, A., Martín, A., & de Las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, *91*, 6829. https://doi.org/10.1016/j.patcog.2019.02.023

Mason, L. (2015). “I disrespectfully agree”: The differential effects of partisan sorting on social and issue polarization: partisan sorting and polarization. *American Journal of Political Science*, *59*(1), 128–145. https://doi.org/10.1111/ajps.12089

Mason, L. (2018). Ideologues without issues: The polarizing consequences of ideological identities. *Public Opinion Quarterly*, *82*(S1), 866–887. https://doi.org/10.1093/poq/nfy005

McCoy, J., Rahman, T., & Somer, M. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic polities. *American Behavioral Scientist*, *62*(1), 16–42. https://doi.org/10.1177/0002764218759576

Mechkova, V., Lührmann, A., & Lindberg, S. I. (2017). How much democratic backsliding? *Journal of Democracy*, *28*(4), 162–169. https://doi.org/10.1353/jod.2017.0075

Media Ownership Monitor. (2021a). Hürriyet. Retrieved May 27, 2023, from http://turkey.mom-rsf.org/en/media/detail/outlet/1076/

Media Ownership Monitor. (2021b). Sabah. Retrieved May 27, 2023, from http://turkey.mom-rsf.org/en/media/detail/outlet/1077/

Merkley, E., & Stecula, D. A. (2018). Party elites or manufactured doubt? The informational context of climate change polarization. *Science Communication*, *40*(2), 258–274. https://doi.org/10.1177/1075547018760334

Murat Aslan, Ferhat Pirinççi, Bilgehan Öztürk, Sibel Koru, Ömer Özkizilcik, Rıfat Öncel, … Batu Coşkun. (2021). *Seta security radar Turkey's security landscape in 2021* (No. 181). SETA. Retrieved from SETA website: https://setav.org/en/assets/uploads/2021/01/R181En.pdf

OSC. (2008). *Turkey—Guide to major Turkish daily newspapers 2008*. Open Source Center. Retrieved from Open Source Center website: https://irp.fas.org/dni/osc/turkish-news.pdf

Pereira, J., & Saraiva, F. (2020, July 1). *A comparative analysis of unbalanced data handling techniques for machine learning algorithms to electricity theft detection*. 1–8. https://doi.org/10.1109/CEC48606.2020.9185822

Reporters Without Borders. (2018, March 22). Doğan media group sale completes government control of Turkish media | RSF. Retrieved May 27, 2023, from https://rsf.org/en/do%C4%9Fan-media-group-sale-completes-government-control-turkish-media

Ridge, H. M. (2020). Enemy mine: Negative partisanship and satisfaction with democracy. *Political Behavior*. https://doi.org/10.1007/s11109-020-09658-7

Riffe, D., Aust, C. F., & Lacy, S. R. (1993). The effectiveness of random, consecutive day and constructed week sampling in newspaper content analysis. *Journalism Quarterly*, *70*(1), 133–139. https://doi.org/10.1177/107769909307000115

Roccas, S., & Brewer, M. B. (2002). Social identity complexity. *Personality and Social Psychology Review*, *6*(2), 88–106. https://doi.org/10.1207/S15327957PSPR0602_01

Rothschild, Z. K., Keefer, L. A., & Hauri, J. (2021). Defensive partisanship? Evidence that in-party scandals increase out-party hostility. *Political Psychology*, *42*(1), 3–21. https://doi.org/10.1111/pops.12680

Şaşmaz, A., Yagci, A. H., & Ziblatt, D. (2022). How voters respond to presidential assaults on checks and balances: Evidence from a survey experiment in Turkey. *Comparative Political Studies*, 001041402110662. https://doi.org/10.1177/00104140211066216

Schedler, A. (2013). *The politics of uncertainty: Sustaining and subverting electoral authoritarianism* (1st ed). Oxford : New York, NY: Oxford University Press.

Schmitt, C. (2005). *Political theology: Four chapters on the concept of sovereignty* (University of Chicago Press ed). Chicago: University of Chicago Press.

Schweter, S. (2020). *BERTurk—BERT models for Turkish*. Zenodo. https://doi.org/10.5281/ZENODO.3770924

Şirvanli, T. (2021). Covid-19 haberlerinde siyasi kutuplaşmanın yansımaları: Hürriyet ve gazete pencere üzerine bir inceleme. *SELÇUK ÜNİVERSİTESİ İLETİŞİM FAKÜLTESİ AKADEMİK DERGİSİ*. https://doi.org/10.18094/josc.883597

Somer, M. (2010). Media values and democratization: What unites and what divides religious-conservative and pro-secular elites? *Turkish Studies*, *11*(4), 555–577. https://doi.org/10.1080/14683849.2010.540114

Somer, M. (2019). Turkey: The slippery slope from reformist to revolutionary polarization and democratic breakdown. *The ANNALS of the American Academy of Political and Social Science*, *681*(1), 42–61. https://doi.org/10.1177/0002716218818056

Stecula, D. A. (2018). *Mass media and political polarization in the United States* (The University of British Columbia). The University of British Columbia, Vancouver. Retrieved from https://open.library.ubc.ca/media/stream/pdf/24/1.0368990/4

Svolik, M. w. (2020). When polarization trumps civic virtue: Partisan conflict and the subversion of democracy by incumbents. *Quarterly Journal of Political Science*, *15*(1), 3–31. https://doi.org/10.1561/100.00018132

Tan, M., Tan, L., Dara, S., & Mayeux, C. (2015). Online defect prediction for imbalanced data. *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, *2*, 99–108. https://doi.org/10.1109/ICSE.2015.139

Tunç, A. (2018). All is flux: A hybrid media approach to macro-analysis of the Turkish media. *Middle East Critique*, *27*(2), 141–159. https://doi.org/10.1080/19436149.2018.1433581

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Waldner, D., & Lust, E. (2018). Unwelcome change: Coming to terms with democratic backsliding. *Annual Review of Political Science*, *21*(1), 93–113. https://doi.org/10.1146/annurev-polisci-050517-114628

Yagci, A. H., & Oyvat, C. (2020). Partisanship, media and the objective economy: Sources of individual-level economic assessments. *Electoral Studies*, *66*, 102135. https://doi.org/10.1016/j.electstud.2020.102135

Yeşil, B. (2016). *Media in new Turkey: The origins of an authoritarian neoliberal state*. Urbana: University of Illinois Press.

Yeşil, B. (2018). Authoritarian turn or continuity? Governance of media through capture and discipline in the AKP era. *South European Society and Politics*, *23*(2), 239–257. https://doi.org/10.1080/13608746.2018.1487137

Yıldırım, K., Baruh, L., & Çarkoğlu, A. (2021). Dynamics of campaign reporting and press-party parallelism: Rise of competitive authoritarianism and the media system in Turkey. *Political Communication*, *38*(3), 326–349. https://doi.org/10.1080/10584609.2020.1765913