Computational Framing of Abortion Rights: Comparing Reddit with Liberal and Conservative News Coverage Before and After Roe v. Wade's Reversal

By

Sara Baniamerian

Submitted to
Central European University
Department of Network and Data Science

In partial fulfillment of the requirements for the degree of
Master of Science in Social Data Science

Supervisors: Prof. Eva Fodor
Associated Supervisor: Prof. Petra Kralj Novak

Vienna, Austria

2025

Author's Declaration

I, the undersigned, **Sara Baniamerian**, candidate for the MSc degree in Social Data Science declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright. I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 22 05 2025

Sara Baniamerian

Copyright Notice

---

[1]Icon by Font Awesome: https://fontawesome.com/

Abstract

The Dobbs draft leak in 2022 marked a major turning point in U.S. abortion policy, overturning the constitutional right established by *Roe v. Wade* and sparking heated debate around reproductive rights and abortion access. This research investigates how the issue was framed across traditional media—represented by CNN and Fox News, reflecting two sides of the U.S. political spectrum—and how the public framed it through Reddit posts and comments. The study also explores whether and how emotional tone and framing shifted following the Dobbs leak, and whether political affiliation or platform contributed to increased polarization.

To conduct this analysis, I used televised news transcripts and Reddit data from before and after the Dobbs draft leak. The study applies transformer-based Natural Language Processing (NLP) models for sentiment analysis, emotion classification, and embedding-driven topic modeling. Compared to traditional lexicon-based or qualitative methods, these models offer both scale and nuance—enabling the analysis of thousands of posts and transcripts while capturing subtle, context-dependent shifts in framing and emotional tone.

The results show clear differences in sentiment and framing across sources. CNN maintained a more positive tone and emphasized healthcare access, reproductive rights, and electoral implications. Fox News presented abortion in relation to political unrest and social critique, while omitting rights-based framing. Reddit displayed more emotionally diverse and morally charged discussions, with frames including bodily autonomy, personal experience, and ideological division. Statistical analysis using Chi-Square and Jensen–Shannon Divergence confirmed that polarization between CNN and Fox News significantly increased after the leak. Traditional media emphasized legal framing, while Reddit users expressed stronger emotional tones.

This thesis contributes by demonstrating how political ideology and platform shape discourse on abortion, especially during periods of legal and political change. It also highlights the value of applying transformer-based NLP to capture complex patterns of framing, sentiment, and emotion across diverse platforms on a highly contested and politically sensitive issue.

iii

Acknowledgements

I would like to express my deepest gratitude to my parents for their unwavering support and love—not only throughout my academic journey but throughout my life.

I am also thankful to my friends, whose companionship, understanding, and thoughtful conversations helped me stay grounded and inspired during the most challenging phases of this research.

I owe special thanks to my two supervisors, Professor Eva Fodor and Professor Petra Kralj Novak, for their invaluable guidance, constructive feedback, and consistent support. Their insights and mentorship played a crucial role in shaping this thesis and enhancing its academic rigor.

This work would not have been possible without the collective support and kindness of those around me. Thank you all.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Abortion laws have long been a highly controversial topic in the United States, dividing the two major political parties. The 2022 Supreme Court decision in Dobbs v. Jackson Women's Health Organization[1] marked a historic shift, overturning Roe v. Wade and ending nearly 50 years of federally protected abortion rights. This ruling gives the power to regulate abortion to individual states [1], triggering widespread political, legal, and social debates. The decision was largely made possible by the appointment of three conservative justices—Neil Gorsuch, Brett Kavanaugh, and Amy Coney Barrett—by President Donald Trump, who had pledged to appoint justices likely to overturn Roe during his 2016 campaign.

The Dobbs ruling set off immediate policy shifts across the United States. Fourteen states swiftly enacted full abortion bans, while some others imposed severe restrictions, significantly limiting access to reproductive healthcare. Meanwhile, public opinion remains largely in favor of abortion rights, with polling indicating that nearly two-thirds of Americans support legal access [2]. The decision also intensified partisan debates, with many conservatives pushing for a national abortion ban, while pro-choice advocates continue to fight for legal protections at the state level.

The outcome of the 2024 U.S. presidential election has raised concerns about reproductive rights and the potential for further restrictions. Examining the period around significant legal and political changes can help us understand how different groups justified their positions on the change in this case, abortion, and explore their attitudes toward it. Such an analysis offers insights into diverse perspectives and can identify the framing strategies used by media outlets with distinct political affiliations as well as by the public in online discussions. The overturning of the historical landmark Roe v. Wade provides a unique opportunity, as it not only intensified public debates on abortion but also allows for investigating whether legal changes influenced framing patterns, attitudes, and language usage in political discussions. Although the Dobbs decision officially occurred in June 2022, the leak of the draft opinion on May 2nd had already

---

[1]On June 24, 2022, the U.S. Supreme Court ruled on Dobbs vs. Jackson's Women's Health Organization, stating that the U.S. Constitution does not confer a woman's right to abortion. The ruling essentially overturned Roe v. Wade, stating the right to an abortion was no longer federally protected, but up to each individual state [1].

substantially heightened debates around abortion.

In this research, I examine how media coverage framed the abortion debate before and after the Dobbs v. Jackson draft leak, and I similarly analyze social media content to observe divergences between traditional and digital media. Media has consistently played a critical role in shaping public opinion, through both traditional news reporting and the growing influence of social media. Social media platforms, in particular, amplify ideological divides, mobilize activists, and reflect shifts in public sentiment. Comparing the framing patterns of traditional and digital media allows us to determine whether they follow similar or distinct trends. Analyzing how traditional and digital media framed abortion during this period provides deeper insights into the interplay between media narratives, policy changes, and public attitudes.

This research seeks to explore the implications of this legal shift by analyzing media framing, political rhetoric, and public response in the wake of Dobbs. Understanding these dynamics is crucial for assessing how policy changes can provoke public reactions, as well as for exploring the debates, opinions, and attitudes surrounding the issue—insights that are essential for predicting future developments in abortion rights and reproductive healthcare in the United States.

In this study, framing theory is applied using computational methods such as sentiment analysis, emotion recognition, and topic modeling. The data used in this research comes from two major media outlets with opposing political affiliations, covering the six months before and after the Dobbs draft leak on May 2nd, 2022. Additionally, user comments and posts from the Reddit platform during the same period have been collected. This study examines how different groups with varying political leanings frame their opinions. Furthermore, comparing media outlets and social media enables an analysis of their differences and their potential influence on each other.

To effectively analyze how media and public discussions framed the abortion debate during this period of intense political change, this thesis first engages with an extensive literature review on framing theory and then focus on its implications on abortion-related discourse. Previous studies have provided foundational insights into how framing shapes public understanding of abortion, typically emphasizing moral, legal, and health-related frames [3]. Moreover, recent computational analyses have begun to investigate sentiment and framing in online and traditional media separately, but often with methodological limitations such as reliance on lexicon-based sentiment detection [4, 5]. Recognizing these gaps and drawing on foundational theories and contemporary computational advances, this research introduces methodological innovations that allow for a deeper and more nuanced investigation of framing dynamics across both media platforms and public forums.

This thesis contributes methodologically to framing research by integrating advanced computational techniques, notably transformer-based sentiment analysis and contextual embedding-based topic modeling, significantly improving upon traditional lexicon-based and bag-of-words approaches [4,5]. Earlier computational framing studies predominantly employed simpler topic

2

modeling methods such as latent Dirichlet allocation (LDA) or keyword frequency analyses, limiting their capacity to accurately capture nuanced semantic contexts and emotional dimensions in politically sensitive discourse [4]. Additionally, prior analyses often relied exclusively on short texts such as news headlines or excerpts, thus neglecting deeper narrative and argumentative structures crucial for understanding frames [6]. By contrast, this thesis leverages full-text data from news articles and Reddit posts and comments, enhancing the granularity and interpretability of framing detection.

Moreover, this research addresses the methodological gap highlighted by scholars who analyzed either mainstream news media or social media platforms separately, without direct comparison [1, 7, 8]. Unlike these one sided analyses, this thesis explicitly compares framing across ideologically opposed media outlets (CNN and Fox News) and Reddit communities, thus answering recent calls for integrated analyses of institutional and public discourses [9,10]. Furthermore, by employing transformer-based models such as SiEBERT [11] instead of traditional lexicon-based sentiment tools [5], this study accurately captures emotional nuances within abortion-related framing. Finally, the inclusion of a temporal comparison of frames before and after the Roe v. Wade reversal can provide new insight in studying the intersection of policy and communication tools like news, and social media.

This research aims to answer the following questions:

Q1 – How do media outlets with different political affiliations and users on social media frame the abortion debate, and how did these framings change before and after the Dobbs v. Jackson decision?

Q2 – How do the attitudes and emotional expressions of media outlets and individuals with varying political affiliations differ, and how did they change in response to the Dobbs ruling?

Chapter 2

Literature Review

2.1 Literature Review

To begin investigating my research questions, I first reviewed the literature across several key areas. I started with framing theory, which provided insight into how an issue can be presented in different ways—highlighting certain aspects while omitting others—ultimately shaping the audience's interpretation of the same piece of information. After exploring the general foundations of framing theory, as discussed in Section 2.1.1, I focused specifically on its application to abortion and reproductive rights, as outlined in Section 2.1.2. This section offers an overview of the existing research, identifies what has already been done in the field, and establishes the foundation for understanding the unique contribution of this thesis.

2.1.1 Framing Theory

Framing has become a crucial concept in media research, widely applied in communication studies, sociology, and political science. It refers to how events and issues are organized and interpreted, particularly by media professionals and their audiences [12]. As one of the most widely used theories in the social sciences, framing remains central to analyzing how public discourse is shaped across different domains [13].

Entman [14] defines framing as the process of selecting specific aspects of a perceived reality and making them more salient within a communicative text to promote a particular problem definition, causal interpretation, moral evaluation, or policy solution. Similarly, Gitlin [15] describes frames as "persistent patterns of cognition, interpretation, and presentation, of selection, emphasis, and exclusion, by which symbol-handlers routinely organize discourse, whether verbal or visual". These principles demonstrate how framing structures public discourse by selectively highlighting certain aspects of an issue while omitting others.

Pan and Kosicki [10] conceptualize framing as a strategic action within discursive processes, where political actors construct meaning and engage in public deliberation. This view aligns with Tankard Jr. [16], who emphasizes that framing subtly defines the terms of political

debates, influencing public perception without audiences consciously recognizing its impact. This is particularly relevant to controversial issues such as abortion, where media framing often lacks neutrality. Even in medical coverage, abortion is frequently portrayed not as a routine healthcare service but as a contentious political issue. For instance, a study analyzing U.S. newspaper coverage found that abortion is predominantly presented through political and moral lenses, with minimal emphasis on its medical aspects [17]. Such framing can significantly influence public perception, leading audiences to view abortion primarily as a political debate rather than a standard component of healthcare.

Tuchman's [18] study, *The News Net*, explores how news organizations impose structure on reality through reporting practices, emphasizing that news is not merely a reflection of events but a constructed reality. Institutional routines, editorial decisions, and economic pressures determine what aspects of reality are emphasized or omitted, demonstrating that framing is deeply tied to power structures. Similarly, Reese [12] expands on framing theory within a hierarchy-of-influences framework, arguing that framing is not an individual journalistic choice but a structured process shaped by institutional, organizational, and ideological forces . His analysis aligns with Gitlin in demonstrating how framing reinforces dominant ideologies, privileging certain narratives while marginalizing dissenting voices. The media does not neutrally report events but actively constructs reality, reflecting the interests of political, economic, and social elites. Institutional constraints—such as editorial policies, corporate ownership, and relationships with state actors—further shape framing, determining which perspectives are amplified or suppressed. In this way, framing functions as a mechanism of symbolic power, controlling the visibility of competing viewpoints [15].

Entman [19] introduces *The Cascade Model*, which explains how political issues are framed through a hierarchical flow from government officials to the media and then to the public. However, framing is not solely a top-down process. Competing frames can emerge at different levels, and those that align with pre-existing cultural beliefs spread more effectively.

Framing theory has evolved alongside transformations in mass communication. The rise of digital platforms has shifted framing from a one-to-many broadcast model to a more interactive many-to-many form, where audiences engage with and reshape narratives in real time [13]. While early research on framing focused on how traditional news organizations structured reality through editorial decisions and economic constraints [12, 15, 18], the digital era has introduced new storytelling tools that change how frames are constructed, disseminated, and contested.

With the advent of social media, framing is no longer controlled exclusively by traditional gatekeepers. Digital platforms allow users to produce and distribute messages widely, sometimes even faster than mainstream media. Citizen-generated content has emerged as an alternative news source, often bypassing geographic and political constraints [9]. The emergence of mass media initially reinforced elite control over framing, monopolizing narratives along class and ideological lines. However, today's fragmented digital environment disrupts these

structures, enabling a more participatory form of framing [20].

The internet and social media have revolutionized public discourse by embedding audiences into the media production loop. This transformation has blurred the distinction between mass media and interpersonal communication, making the negotiation of social reality more complex. While some scholars predicted the internet would democratize framing, organized economic and political interests continue to exert significant influence online. Nevertheless, digital media remain a disruptive force, challenging traditional framing hierarchies [20].

Interactive platforms have fundamentally altered news framing by enabling multidimensional storytelling, where narratives are dynamic, interconnected, and audience-driven [21]. Traditional media followed a top-down framing model, with journalists and editors controlling narratives. In contrast, digital forums, comment sections, and social media discussions facilitate bottom-up framing, where public discourse actively reshapes dominant narratives [22].

Advancements in computational methods have further expanded the scope of framing research. While early studies relied on qualitative discourse and content analysis [23, 24], modern approaches incorporate computational techniques such as natural language processing (NLP), structural topic modeling, and machine learning. These methods allow researchers to analyze vast datasets, offering insights into media narratives, political communication, and social movements [4].

Framing remains a critical mechanism in shaping political and social discourse, influencing how societies define issues, assign responsibility, and advocate for solutions. Whether through traditional media, political rhetoric, or digital activism, framing plays a central role in structuring public debates, particularly on contested topics such as abortion policy.

2.1.2 Framing and Reproductive Rights

Policymaking on reproductive rights operates through both coercive measures—such as legal restrictions and criminalization—and ideological influence, which shapes public attitudes through political discourse. The media play a key role in legitimizing or criticizing policy decisions by constructing narratives around abortion access, morality, and individual rights. Social media platforms provide real-time insights into public attitudes and ideological shifts. Research demonstrates how digital platforms both reflect and shape public opinion. For instance, Kalabikhina et al. [25] examined pro-natalist and anti-natalist sentiment on Russian social media using data from the Russian social network VKontakte. The researchers conducted sentiment analysis on user comments from 314 pro-natalist groups (which support childbearing and reproductive policies) and eight anti-natalist groups (which advocate for child-free lifestyles) to assess public sentiment on demographic topics. Ujah et al. [8] analyzed Twitter reactions to the reversal of *Roe v. Wade*, finding widespread negative sentiment, particularly among marginalized racial and ethnic groups. These studies highlight the importance of monitoring digital discourse to understand evolving attitudes toward abortion.

The framing of abortion has evolved over time. Roberti [26] analyzed 1,706 anti-abortion bills across U.S. states from 2008 to 2017, finding that 70% employed a "pro-woman" frame—that is, framing abortion as harmful to women—while 38% emphasized fetal personhood. Anti-abortion rhetoric increasingly presents abortion as a public health risk, citing mental health consequences, coercion, and unsafe procedures to justify restrictions. Many laws combine protectionist rhetoric, framing the state as a guardian against the harm of abortion, with fetal personhood arguments to strengthen their position. Similarly, Moyer [3] found that while pro-choice groups rely heavily on scientific framing in Supreme Court briefs, anti-abortion organizations have increasingly adopted scientific rhetoric to support restrictions, though moral and emotional appeals remain dominant.

Legislative debates frequently frame abortion as a moral issue. Valdez and Goodson [5] observed consistently negative sentiment in legislative discourse, while Vossen et al. [27] examined Belgium's Chamber Committee on Justice debates, identifying competing morality frames [27]. Conservative opponents emphasized fetal sanctity, depicted abortion as harmful to women, and framed policy changes as threats to democratic procedures. In contrast, progressive advocates positioned abortion as a human rights issue, defended legislative reforms as essential for social justice, and framed women as autonomous decision-makers.

Siekiera & Szews [7] analyzed Polish media coverage of the 2020 abortion protests, showing how both conservative and liberal outlets reinforced ideological divisions. The debate was framed as a conflict between opposing ideological camps, with media narratives emphasizing political responsibility, religious tensions, and emotional responses. The findings illustrate how media actively shape public discourse, amplifying ideological divides rather than fostering dialogue.

Pan et al. [6] conducted a large-scale comparative analysis of framing differences between traditional news media and social media, focusing on topics like *Dobbs v. Jackson Women's Health Organization*. Their study applied computational methods, such as stance detection and topic modeling, to examine news headlines and Reddit comments. The findings highlighted that social media is more polarized than traditional media, with Reddit discussions on abortion often framing the issue through public sentiment and cultural identity, while news media focused more on policy and legal evaluations. Political polarization was also evident in news coverage, with left- and right-leaning outlets framing abortion differently. This study provides a foundation for understanding how media environments shape public discourse, reinforcing ideological divides.

Building on this work, the present research analyzes how abortion framing evolved before and after the *Dobbs* draft leak on may 2nd, 2022, incorporating political affiliations. Unlike previous studies that focused on static comparisons or news headlines, this study examines full-text articles, providing deeper insights into how framing is constructed throughout the content, not just in the headlines.

Using computational methods such as sentiment analysis, emotion recognition, and topic

7

modeling, this study examines attitudes toward abortion, identifies dominant frames through topic modeling—which serves as a framing analysis method in computational approaches—and analyzes the rhetorical strategies and justifications used by different groups in the abortion debate. By focusing on these justifications, the research provides a more nuanced understanding of how various actors construct and communicate their positions on abortion rights.

Additionally, this study employs context-aware transformer models—unlike earlier works that used Latent Dirichlet Allocation or lexicon-based tools like VADER—thereby better capturing subtle framing, emotional tone, and contextual nuances in abortion discourse. For instance, Pinto [28] used VADER and LDA for short-term Twitter analysis. Stanier and Shin [29] applied moral foundations theory and used term frequency and LDA to extract frames from Reddit comments in pro-life and pro-choice subreddits. Meanwhile, Rao et al. [30] employed transformer models to identify hostile framing patterns but focused exclusively on Twitter. This thesis expands on their findings by analyzing both Reddit and news data over a longer time frame and by integrating framing, sentiment, and emotion in a unified framework.

This research not only explores how sentiments and framing have evolved over time but also offers insights into the interaction between traditional media and social media in shaping public attitudes toward abortion policy.

CEU eTD Collection

8

Chapter 3

Data and Methodology

3.1 Data

In this section, I describe the data used to address my research question. My aim is to examine how news outlets frame abortion-related news, their attitudes toward abortion, and how these frames differ across sources with varying political leanings over time. Additionally, I analyze the same issue from the perspective of public opinion.

This research relies on two primary data sources to capture both sides of the issue and to explore how different outlets frame the topic of abortion. The first source consists of mainstream news media outlets, while the second focuses on public opinion. The dataset details are described in Sections 3.1.1 and 3.1.2. The process of preparing the data to address the research questions and to make it suitable for computational analysis is described in Section 3.1.3.

3.1.1 Data Sources

To answer the question of how the two major political groups frame the issue of abortion, I used two datasets collected from news channels that represent these groups. The goal was to select two outlets that serve as representatives of the prominent ideological divisions in the United States.

The data used in this study come from televised news programs and a public online platform. These sources were selected to capture a broad, ideologically diverse perspective on the abortion debate, including coverage from both Democratic-leaning and Republican-leaning media outlets, as well as public opinion expressed on social media.

The news data were obtained from the Harvard Dataverse, which hosts a collection of datasets related to media coverage and political discourse. Two datasets were used for this purpose: one containing transcripts from televised news programs on CNN [31], and the other from FOX News [32]. According to the AllSides website [33, 34], these outlets are considered reliable representations of two prominent ideological perspectives in the U.S.—FOX News representing conservative views and CNN representing liberal views.

To analyze how the public frames the issue of abortion, I used Reddit data, including both comments and posts. The dataset was obtained from Academic Torrents [35], a platform that hosts large-scale research data. I downloaded content from subreddits related to politics, American political parties, and abortion. The list of subreddits used is provided in Appendix A.1.

3.1.2 Dataset Overview

The data collected from FOX News, CNN, and Reddit form a large dataset enriched with comprehensive metadata. For mainstream media sources, only the transcribed text of televised news and the broadcast date were required. In contrast, the Reddit dataset includes the date, the text of posts (from the `selftext` field), and comments (from the `body` field). To capture potential signs of political affiliation, the fields `author_flair_text` and `author` were retained. Additionally, both the `date` and the `subreddit` were preserved to examine differences across communities and to analyze temporal patterns in framing.

Since the timeline of interest is the year 2022, the next step focuses specifically on data from that year. This period includes two key events that significantly intensified debates around abortion: the leak of the *Dobbs v. Jackson* draft opinion [36] and the official Supreme Court decision overturning *Roe v. Wade* [37]. These moments sparked intense media framing and online engagement, as reflected in the distribution of texts shown in Figures 3.1 and 3.3.

After filtering for data from 2022, the resulting dataset included 3,080 documents from FOX News, 8,444 from CNN, 557,649 Reddit comments, and 27,427 Reddit posts.

3.1.3 Data Preprocessing

After filtering the data by date, I applied a dictionary-based method to identify texts related to abortion. The keyword dictionary used in this study is adapted from previous work by Stanier and Shin [29], and further refined based on Pinto's thesis [28]. These sources categorize keywords by stance—neutral, pro-choice, and pro-life—based on their frequency and relevance in abortion-related discourse. The final dictionary used in this study is presented in Table 3.1.

For FOX News, the initial dataset contained 3,080 transcripts. After applying the dictionary, 221 transcripts were identified as relevant. For CNN, there were 8,444 transcripts, of which 2,985 were related to the subject of study.

From Reddit, the dataset of user comments initially included 557,649 documents after cleaning. After applying the dictionary filter, 482,999 comments remained. The dataset of Reddit posts, after removing missing values, included 27,427 posts. After filtering, 17,408 of those were related to abortion.

Table 3.2 shows a summary of the number of documents before and after filtering for each source.

The distribution of texts over time in Figures 3.1 and 3.2 shows how the document leak and the Supreme Court decision sparked intense debate around the topic of abortion. The normal-

Table 3.1: The dictionary for detecting related texts

Category	Terms
Core Abortion Terms	abortion, pregnancy, planned parenthood, fetus, embryo, miscarriage
Legal & Institutional	roe, wade, dobbs v. jackson, roe v. wade, roe v wade, abortion rights, abortion access, abortion law, abortion ban, pro-choice, pro-life, pro-choice movement, pro-life movement, reproductive rights, reproductive justice, reproductive health, dobbs, heartbeat bill, planned parenthood v. casey
Health-Related	reproductive health, contraception, birth control, ectopic pregnancy
Framing Terms	pro-life, prolife, pro-choice, prochoice, my body my choice, forced birth, abortion is murder
Rights and Identity	reproductive rights
Additional Terms	crisis pregnancy, family planning, pregnancy loss, post-abortive
Political Context Terms	roe, dobbs, planned parenthood
Rights Frame Terms	forced pregnancy

Table 3.2: Number of documents before and after filtering by dictionary

Source	Before Filtering	After Filtering
FOX News Transcripts	3,080	221
CNN Transcripts	8,444	2,985
Reddit Comments	557,649	482,999
Reddit Posts	27,427	17,408

ized distribution in Figure 3.2 illustrates that before and immediately after the leak, both FOX News and CNN gave relatively equal attention to the issue. However, while CNN—typically considered left-leaning—continued to cover the topic consistently throughout the year, FOX News gradually reduced its coverage of abortion-related news.

Furthermore, the Reddit activity shown in Figure 3.3 reveals significant spikes in discussion following both the leak and the official Supreme Court ruling. Over time, the intensity of the discussions returned to a level similar to that before the policy change.

These spikes in distribution support the choice of the Dobbs draft leak as a critical point dividing the timeline.
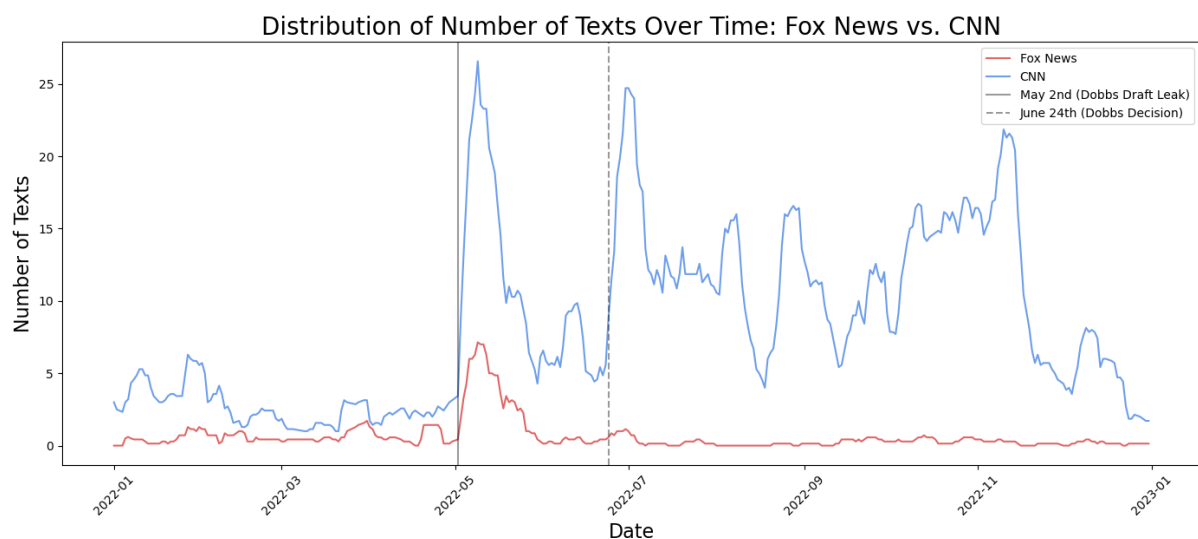
Figure 3.1: Distribution of abortion-related texts on CNN and Fox News over time. Coverage increases sharply around two major events: the leak of the *Dobbs v. Jackson Women's Health Organization* draft opinion on May 2nd and the official Supreme Court decision on June 24th.

In summary, CNN maintained steady coverage of abortion throughout 2022. FOX News, on the other hand, reduced its focus on the issue over time. Public engagement on Reddit shows strong initial reactions to the events, but the level of discussion eventually stabilized.

I used the spaCy library in Python for text preprocessing. This process included converting all text to lowercase, performing lemmatization, and removing special characters. In this study, each Reddit post or comment was treated as a single document. For news data, each transcript was also treated as one document; however, only a specific portion of the text—centered around relevant keywords—was used for analysis.

To ensure that the analysis focused specifically on abortion-related content within the news transcripts, I extracted a 40-word window before and after each occurrence of the keywords listed in Table 3.1. This approach was necessary because televised news segments often shift topics rapidly, and entire transcripts may include unrelated material.

The decision to use a 40-word window was based on iterative trials and qualitative evaluations of the resulting topic coherence. It was selected by the author as a balance between
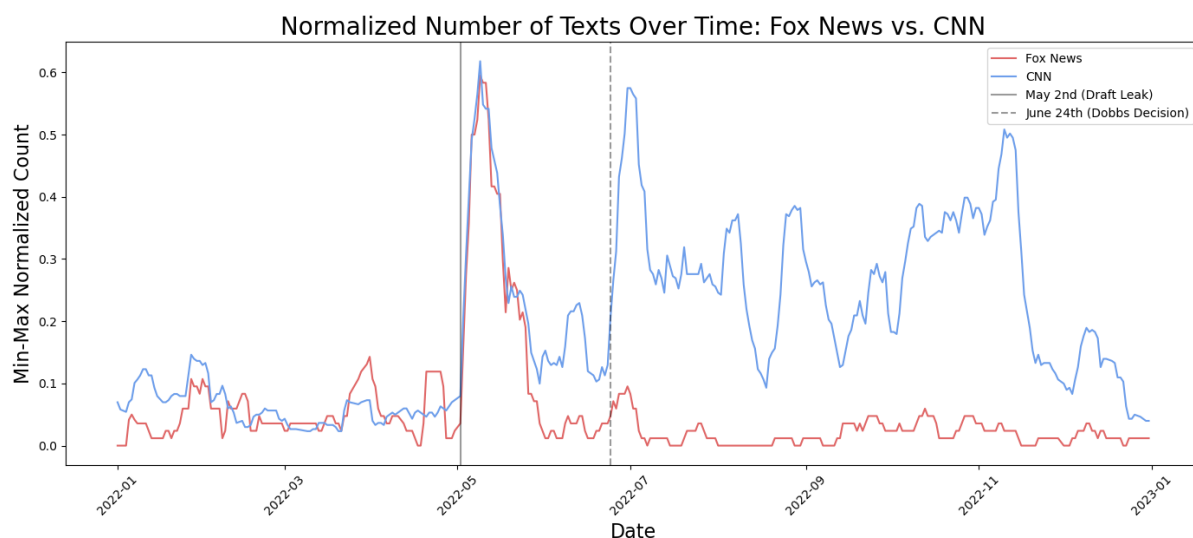
Figure 3.2: Min-Max normalized count of abortion-related texts on CNN and Fox News over time. The plot illustrates shifts in relative attention surrounding major events: the leak of the *Dobbs v. Jackson Women's Health Organization* draft opinion on May 2nd and the official Supreme Court decision on June 24th.
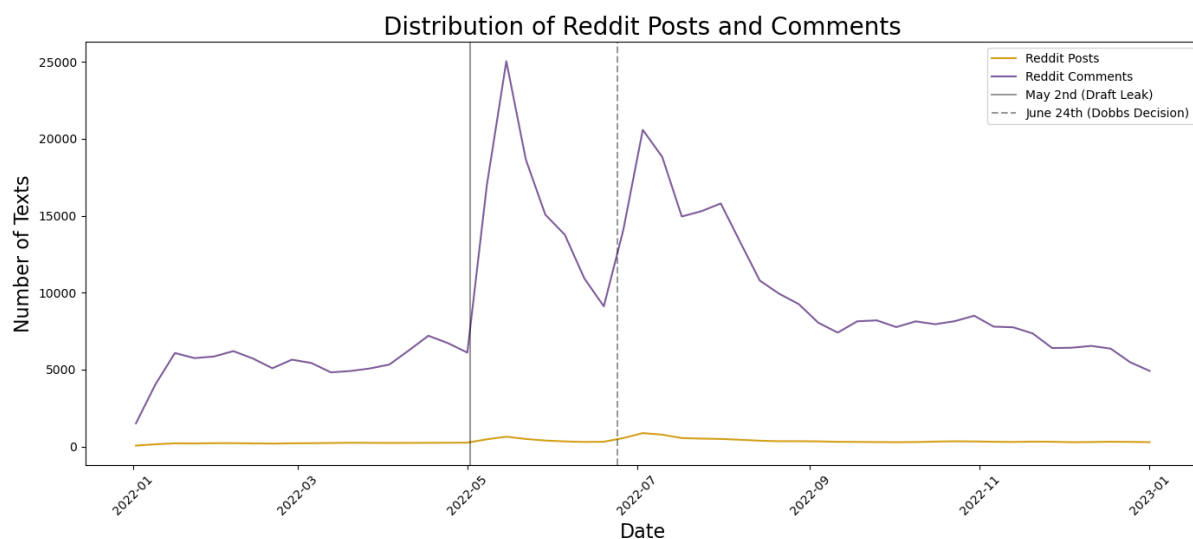
Figure 3.3: Distribution of Reddit posts and comments across 2022. Spikes are visible around key events: the leak of the *Dobbs v. Jackson Women's Health Organization* draft opinion on May 2nd and the official Supreme Court decision on June 24th.

13

capturing sufficient context and avoiding irrelevant content. This technique allowed for a more targeted and meaningful analysis of how abortion was framed in the media.

3.2 Methodology

The tools of Natural Language Processing (NLP) have opened new avenues for social science research, allowing for the analysis of large-scale textual data in a more systematic and quantitative manner. Many recent studies apply sentiment analysis, emotion recognition, stance detection, and topic modeling to examine attitudes and narratives expressed in news media and social platforms concerning various social, economic, and political issues [4, 8, 38].

Building on previous work, I also applied computational methods—including sentiment analysis, emotion recognition, and topic modeling—to address the central questions of my research: how discussions about abortion are framed, what attitudes are expressed, and which emotions are most commonly involved.

As a first step in analyzing how different sources express their attitudes toward abortion, I employed sentiment analysis. This method not only enables the detection of the general tone (positive or negative) toward abortion, but also allows for tracking how sentiment shifts over time in response to policy changes. In particular, it provides insight into how events such as the *Dobbs v. Jackson Women's Health Organization* leak and the official decision on June 24, 2022, may have triggered emotional reactions or shifts in discourse.

To carry out the sentiment analysis, I used a transformer-based model, SiEBERT (siebert/ sentiment-roberta-large-english) , which is based on RoBERTa-large and fine-tuned for binary sentiment classification across multiple domains including news, social media, and reviews [11]. This model was chosen due to its ability to interpret context and reduce false positives or false negatives, which are often caused by emotionally charged statements—particularly in abortion discourse, where some words may appear strongly positive or negative in isolation but carry different meanings in context. SiEBERT performs well across diverse textual genres and does not require further fine-tuning for the purposes of this study.

Sentiment scores were generated for each document from three sources: CNN, Fox News, and Reddit. This enabled a comparative analysis across different media and platforms, highlighting variations in sentiment by political orientation and over time. CNN and Fox News represent two major political ideologies in the U.S. In the case of Reddit, I utilized the `author` `_flair_text` and `author` metadata fields to infer users' political affiliations when available, and I also categorized data by subreddit (e.g., `r/prolife` vs. `r/prochoice`, and Democrat- or Republican-leaning subreddits) to examine variations in sentiment among ideologically distinct communities. This approach helps assess the extent of polarization across platforms and groups.

To infer the political affiliation of Reddit users, I employed a dictionary-based keyword matching approach using regular expressions. The fields `author` and `author_flair_text`—

14

which often contain self-identifying terms or ideologically relevant expressions—were scanned using predefined keyword sets representing three political orientations: Left-Leaning, Right-Leaning, and Center/Independent. Each set includes common political labels, as shown in Table 3.3. Documents that did not match any keywords were marked as "Unidentified" and excluded from this part of the analysis.

Table 3.3: Keyword Lists and Distribution for Political Affiliation Detection

Affiliation	Terms	Document Count
Left-Leaning	democrat, liberal, progressive, leftist, socialist, dems, blue wave, aoc, bernie, biden, kamala, elizabeth warren, green party, anarchist, antifa, communist	2,796
Right-Leaning	republican, conservative, right-wing, maga, gop, tea party, trump, desantis, fox news, patriot, rightist, red wave, marjorie, capitalist, alt-right, neoliberal, neoconservative	1,775
Center/Independent	centrist, moderate, independent, libertarian	5,923

To assess sentiment differences between ideologically distinct online communities, I categorized Reddit data based on the political leaning of the subreddits in which posts and comments appeared. Two lists of subreddit names were used to define Democrat-leaning and Republican-leaning groups. The list of subreddits is provided in 3.4. Each post or comment was assigned an affiliation label based on its subreddit tag.

Any subreddit not included in either list was labeled as *Other* and excluded from the final sentiment analysis in order to focus the comparison on explicitly partisan communities. This step helps isolate sentiment trends in clearly defined ideological spaces.

After labeling, I aggregated sentiment labels (positive, negative, etc.) within each group and calculated the proportion of each sentiment class. This approach allowed for a direct comparison of emotional tone between Democrat-leaning and Republican-leaning subreddits.

Table 3.4: Subreddits Used for Political Affiliation Classification

Affiliation	Subreddits
Democrat-Leaning Subreddits	Democrat, DemocratsforDiversity, DemocraticSocialism, DemocraticUnderground, Liberal, LeftistGamersUnion, LeftvsRightDebate, Leftyguitarists, Leftist, Political_Revolution
Republican-Leaning Subreddits	republicans, conservatives, conservativeterrorism, rightistvexillology

I also followed the same step for pro-life and pro-choice subreddits.

To evaluate sentiment differences, I employed Chi-Square tests and Jensen-Shannon Divergence (JSD) across several dimensions: traditional media outlets (CNN and FOX News), Reddit user political affiliations (inferred from metadata), and subreddit-level communities (e.g., Democrat-leaning vs. Republican-leaning). These methods were also applied temporally to assess shifts in sentiment before and after the leak of the *Dobbs v. Jackson Women's Health Organization* draft decision.

Chi-Square tests assessed whether differences in the proportions of positive versus negative sentiments across groups or time periods were statistically significant, indicating whether observed variations were likely due to chance or meaningful shifts in discourse. JSD, a symmetric and bounded measure of distributional divergence, quantified how sentiment profiles differed between groups. With values ranging from 0 (identical) to 1 (completely distinct), JSD enabled interpretation of polarization even when sentiment proportions were similar. Together, these methods captured both the significance and degree of sentiment polarization, offering a nuanced view of how political alignment and timing influenced abortion-related discourse [39].

This approach was applied not only across different sources and communities but also temporally, to examine how sentiment evolved before and after the Dobbs draft leak.

Finally, sentiment analysis was also used to measure the overall tone within each topic identified through topic modeling, providing an additional layer for interpreting the overall sentiment associated with specific themes.

To analyze the emotional tone of abortion-related news in news media and debates in social media, I employed a transformer-based emotion classification model, specifically the `j-hartmann/emotion-english-distilroberta-base` model [11]. This model is a fine-tuned version of DistilRoBERTa, trained on a cleaned subset of the GoEmotions dataset originally developed by Google. The GoEmotions corpus consists of over 58,000 English Reddit comments.

In this study, I used the model's 6-class configuration, which outputs probabilistic predictions across the following discrete emotions: anger, fear, joy, sadness, surprise, and neutral. This method is used to understand comparative differences across media sources (i.e., CNN vs. Fox News) and Reddit and to identify shifts in emotions around the Dobbs leak and decision).

In final step I conducted frame analysis. Framing analysis traditionally involves a mix of qualitative and quantitative approaches, often relying on manual content coding. However, advancements in computational text analysis offer new tools for extracting frames in an automated and scalable manner [4]. In this study, I employed an embedding-based topic modeling pipeline to analyze the thematic structure of abortion-related texts.

First, sentence-level embeddings were generated using the all-MiniLM-L6-v2 model from the Sentence-BERT framework [40]. This model preserves semantic similarity and is computationally efficient, making it well-suited for large corpora. These high-dimensional embeddings were then reduced using Uniform Manifold Approximation and Projection (UMAP) [41], a dimensionality reduction technique that facilitates HDBSCAN (Hierarchical Density-Based Spa-

tial Clustering of Applications with Noise). I chose $N = 7$ as the target number of dimensions.

Next, I applied HDBSCAN to perform unsupervised clustering [42]. HDBSCAN was chosen for its ability to handle uneven, noisy data—common in social and political discourse—without requiring a pre-specified number of clusters. It also naturally identifies and filters out outliers. Separate topic models were trained for each source (CNN, Fox News, Reddit) and for distinct time periods (pre- and post-Dobbs leak).

Each resulting cluster was interpreted based on its most representative terms, identified using class-based Term Frequency–Inverse Document Frequency (TF-IDF). This allowed for extraction of top unigrams, bigrams, and trigrams that distinguished each cluster. Non-informative or overly frequent terms (e.g., "say," "know," "cnn") were manually removed.

Clusters were manually labeled based on the salient keywords, and thematically similar clusters were merged to produce a coherent set of topics. These topics were then analyzed to understand how abortion-related frames vary across time, media platforms, and ideological groups.

Chapter 4

Results

4.1 Sentiment Analysis

Sentiment analysis is an established method for studying attitudes toward a subject in data-driven research within computational social sciences [43]. As explained earlier in Section 3.2, I used a transformer-based model, SiEBERT, to detect sentiments in text. Unlike lexicon-based models, this approach allows for better handling of sentimentally negative words in abortion-related texts, since the model considers the context in which the words appear. I applied the model to media sources (CNN and Fox News) to examine how these outlets reflect attitudes toward the issue. Additionally, I applied it to Reddit posts and comments to explore public attitudes.

The sentiment results presented in Figure 4.1 show that the two news sources—each representing opposing political leanings in the U.S.—exhibit different attitudes toward abortion. CNN, in comparison, demonstrates a higher level of positivity on the issue overall, which aligns with the Democratic Party's political stance on reproductive rights.

This result was further examined through statistical testing shown in Table 4.1. The Chi-Square test comparing sentiment distributions between CNN and Fox reveals a statistically significant difference ($\chi^2 = 166.68$, $p < 0.00001$), indicating that the proportions of positive and negative sentiments differ meaningfully between the two sources. CNN shows a much higher proportion of positive sentiment (2162 positive vs. 823 negative), while Fox exhibits more negative than positive sentiment (153 negative vs. 68 positive). The Chi-Square test is appropriate in this context because it does not assume a normal distribution and is suitable for analyzing relationships between categorical variables—such as sentiment labels across distinct media sources.

To complement this, Jensen–Shannon Divergence (JSD) was calculated to quantify the degree of distributional difference between CNN and Fox. The resulting JSD value of 0.3595 (on a scale from 0 to 1) indicates a moderate divergence in sentiment profiles, providing additional evidence that the two outlets framed the abortion debate with notably distinct emotional tones

during the analyzed period (2022).
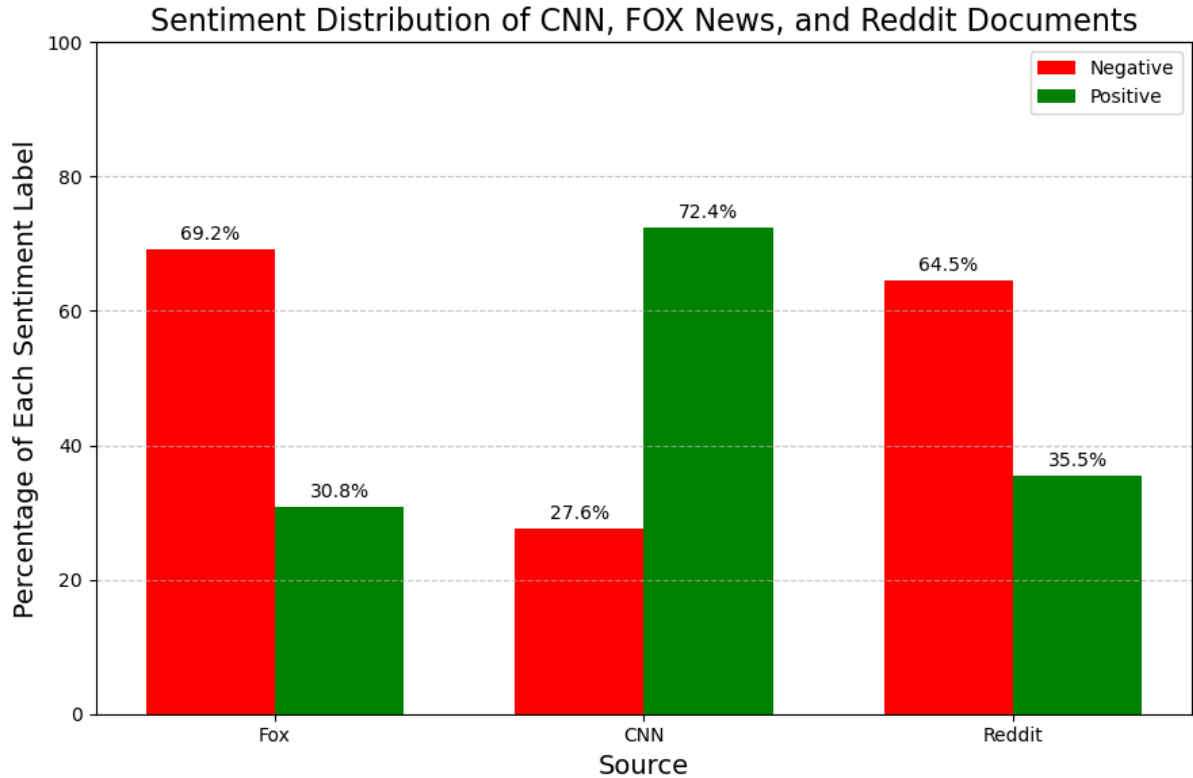
Figure 4.1: Distribution of Sentiments Across Different Data Sources.

Table 4.1: Sentiment distribution and Chi-Square test results comparing CNN and Fox News

Source	Positive	Negative
CNN	2162	823
Fox	68	153

Statistical Test Results
$\chi^2 = 166.68$, degrees of freedom = 1, $p < 0.00001$
Jensen–Shannon Divergence (JSD) = 0.3595

To check for evidence of whether sentiments correspond to the Dobbs v. Jackson Women's Health Organization Draft Opinion Leak on May 2, 2022, I used the Chi-Square test to assess whether sentiment distributions changed significantly after dobbs draft leak. The results are shown in Table 4.2

For CNN, the test revealed a statistically significant change in sentiment distribution ($\chi^2 = 9.13$, $p = 0.00251$). The proportion of positive sentiments increased notably after the leak (from 64.7% to 73.3%), indicating a shift toward a more favorable tone in CNN's abortion-related coverage. This change reflects CNN's broader editorial alignment with pro-choice perspectives.

19

For Fox News, the Chi-Square test yielded $\chi^2 = 3.48$, $p = 0.06202$, which is not statistically significant at the conventional 0.05 level. Although there was an increase in negative sentiment after the leak (from 59.4% to 73.3%), the evidence is not strong enough to conclude a meaningful change in Fox's overall sentiment.

In summary, CNN has a significantly more positive tone toward abortion than FOX News, and for CNN, the legal change coincided with a more positive tone, while FOX News maintained its general sentiment stance.

Table 4.2: Sentiment distribution (counts with percentages) before and after the Dobbs leak (May 2, 2022) for each news outlet, with Chi-Square test results

Source	Sentiment	Before	After	Chi-Square (χ^2), p-value
CNN	Positive	187 (64.7%)	1975 (73.3%)	$\chi^2 = 9.13$, $p = 0.00251$
	Negative	102 (35.3%)	721 (26.7%)	
Fox	Positive	26 (40.6%)	42 (26.7%)	$\chi^2 = 3.48$, $p = 0.06202$
	Negative	38 (59.4%)	115 (73.3%)	

To evaluate whether polarization between CNN and Fox increased following the Dobbs v. Jackson Women's Health Organization draft opinion leak, Chi-Square tests were conducted separately for the periods before and after the event. Resluts are shown in Table4.3.

Before the leak, there was already a statistically significant difference in sentiment distributions between CNN and Fox ($\chi^2 = 11.71$, $p = 0.00062$), with CNN showing a more positive tone and Fox exhibiting more negative sentiment. However, after the leak, the divergence grew substantially ($\chi^2 = 152.65$, $p < 0.00001$), indicating an increase in polarization. CNN's sentiment became overwhelmingly positive, while Fox remained largely negative in its tone. The dramatic rise in the Chi-Square statistic suggests that the Dobbs leak coincided with an intensification of sentiment-based polarization in media coverage between the two news outlets.

Before the Dobbs leak, the JSD between CNN and Fox was relatively moderate at 0.2058 (Table4.3), reflecting some baseline differences in tone. However, after the leak, the JSD rose sharply to 0.4027, indicating a much larger divergence in sentiment distributions. This confirms that not only were the sentiment shifts statistically significant, but the degree of polarization also intensified. These results suggest that the legal change coincided with increased polarization in how the issue was framed by the two media outlets.

The Chi-Square tests presented in Table 4.4 examine whether sentiment distributions differ significantly between Reddit and traditional media sources, and across time in Reddit posts and comments surrounding the Dobbs v. Jackson Women's Health Organization draft leak on May 2, 2022.

The comparison between Reddit and CNN revealed a statistically significant difference ($\chi^2 = 1758.00$, $p < 0.00001$), indicating that CNN's sentiment distribution was markedly more positive than Reddit's, which was predominantly negative. Conversely, the comparison be-

Table 4.3: Sentiment distribution comparison between CNN and Fox before and after the Dobbs leak, with Chi-Square and JSD results

Time Period	Source	Positive	Negative	χ^2, p-value	JSD
Before Dobbs	CNN	187 (64.7%)	102 (35.3%)	$\chi^2 = 11.71, p = 0.00062$	0.2058
	Fox	26 (40.6%)	38 (59.4%)		
After Dobbs	CNN	1975 (73.3%)	721 (26.7%)	$\chi^2 = 152.65, p < 0.00001$	0.4027
	Fox	42 (26.7%)	115 (73.3%)		

tween Reddit and Fox News ($\chi^2 = 1.98$, $p = 0.15961$) showed no statistically significant difference, suggesting that Reddit's sentiment profile is more similar to Fox's coverage in terms of sentiments.

When comparing Reddit sentiment before and after the Dobbs leak, the test also yielded no significant difference ($\chi^2 = 2.05$, $p = 0.15228$), indicating that while the volume of discussion increased, the proportion of positive to negative sentiment remained relatively stable over time. This stability is also shown in Figures 4.2 and 4.3 These results suggest that Reddit's overall tone did not shift in response to the event, while media responses—particularly CNN's—displayed clearer differences in sentiment framing.
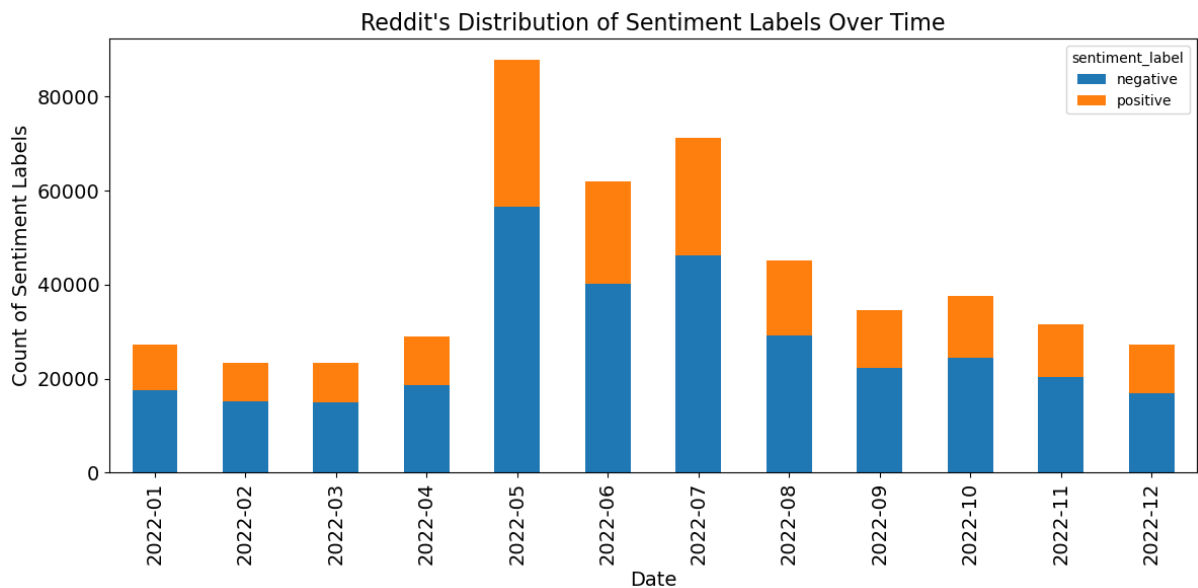
Figure 4.2: Distribution of Sentiments in Reddit Posts and Comments over Time.

Table 4.5 presents sentiment distributions across Reddit users and subreddits, categorized by political affiliation and abortion stance, along with Chi-Square test results for distributional differences.

For political affiliation inferred from usernames or flairs, the sentiment distributions were nearly identical across Center/Independent, Left-Leaning, and Right-Leaning users. The Chi-Square test confirmed no statistically significant difference ($\chi^2 = 0.03$, $p = 0.985$), suggesting
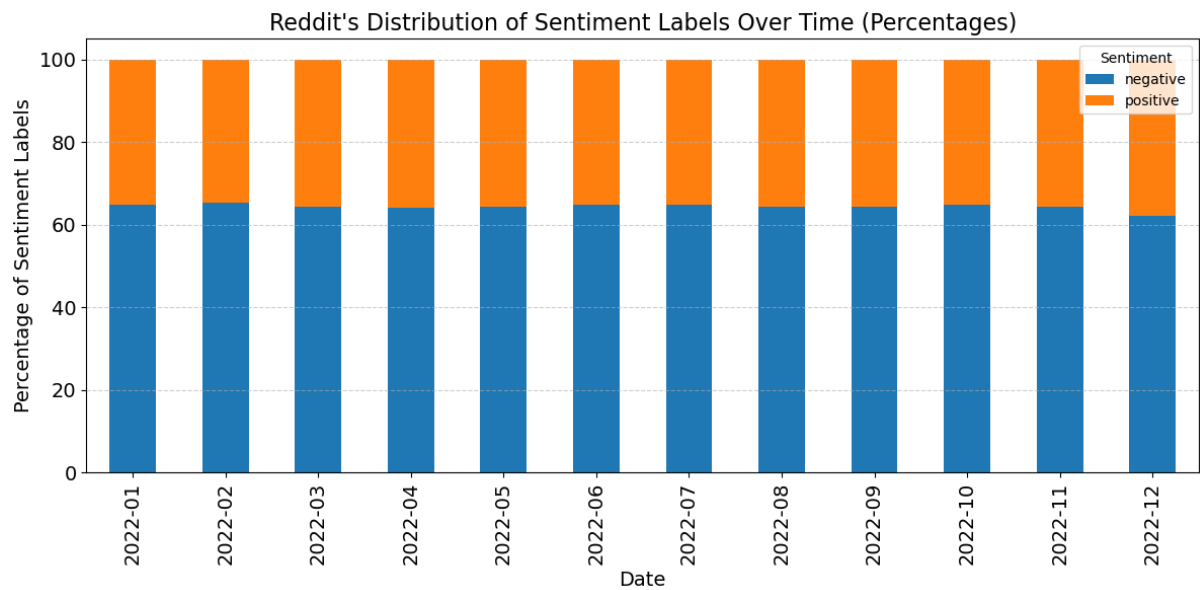
Figure 4.3: Distribution of Sentiments in Reddit Posts and Comments over Time (Percentage).

Table 4.4: Chi-Square test results comparing sentiment distributions across sources and time periods

Comparison	Reddit (Pos/Neg)	Media (Pos/Neg)	χ^2	p-value
Reddit vs. CNN	177,765 / 322,642	2,162 / 823	1758.00	<0.00001
Reddit vs. Fox	177,765 / 322,642	68 / 153	1.98	0.15961
Reddit Before vs. After Leak	36,705 / 67,174	141,060 / 255,468	2.05	0.15228

that user-declared or implied political leanings did not correlate with different sentiment expressions on abortion.

In contrast, a significant difference was observed between Democrat- and Republican-leaning subreddits ($\chi^2 = 5.24$, $p = 0.022$). Democrat-leaning subreddits exhibited a higher proportion of negative sentiment (69.13%) compared to Republican-leaning ones (63.25%).

The most pronounced difference appeared in abortion-specific subreddits. Pro-Choice and Pro-Life communities differed significantly in sentiment distribution ($\chi^2 = 77.12$, $p < 0.00001$), with Pro-Life subreddits showing slightly more positivity (37.13%) compared to Pro-Choice ones (34.46%).

These findings suggest that sentiment tone on Reddit is more strongly associated with community context (subreddit) than with individual user affiliation cues.

Table 4.6 presents sentiment distributions across Reddit users and communities before and after the Dobbs v. Jackson Women's Health Organization draft leak, along with Chi-Square test results for each subgroup.

Among user political affiliations, only Center/Independent users showed a statistically significant change in sentiment ($\chi^2 = 9.01$, $p = 0.00269$), with an increase in both negative and positive sentiments after the leak. Right-Leaning and Left-Leaning users did not display statis-

Table 4.5: Sentiment distribution across Reddit user and subreddit groups, with Chi-Square test results

Group	Subgroup	Negative (%)	Positive (%)	χ^2, p-value
User Political Affiliation	Center/Independent	67.20	32.80	
	Left-Leaning	67.20	32.80	$\chi^2 = 0.03, p = 0.985$
	Right-Leaning	66.99	33.01	
Subreddit Affiliation	Democrat-Leaning Subreddit	69.13	30.87	$\chi^2 = 5.24, p = 0.022$
	Republican-Leaning Subreddit	63.25	36.75	
Abortion Stance Subreddit	Pro-Choice Subreddit	65.54	34.46	$\chi^2 = 77.12, p < 0.00001$
	Pro-Life Subreddit	62.87	37.13	
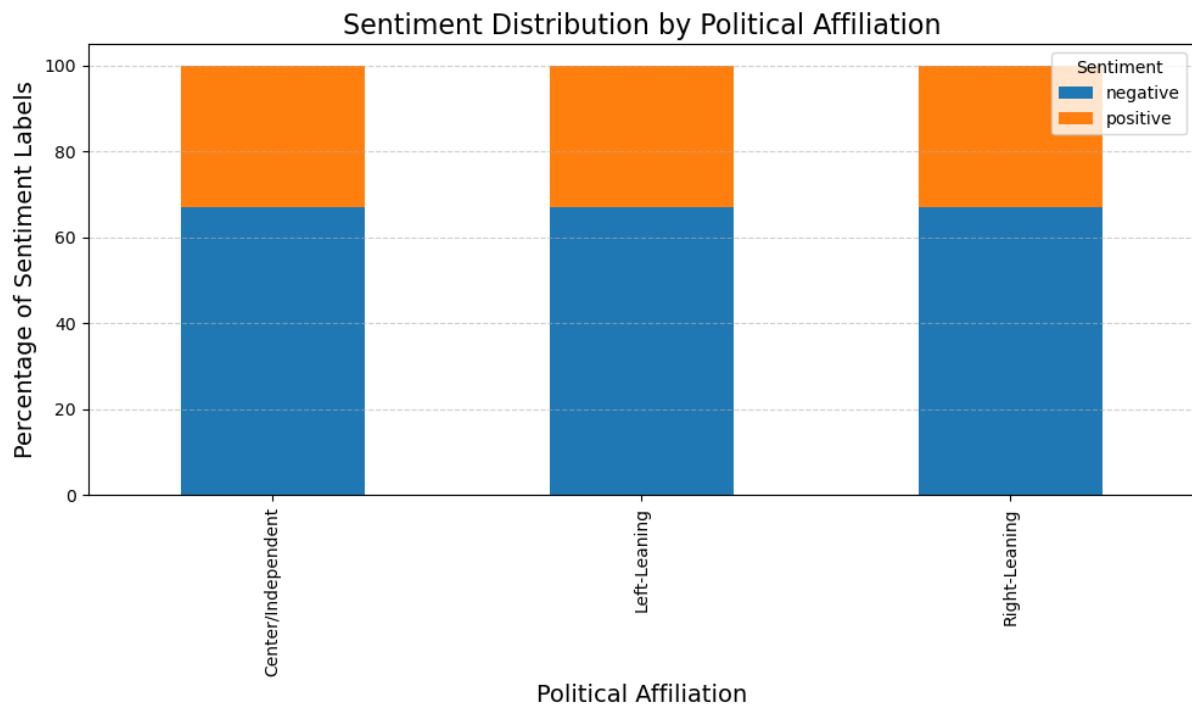
Figure 4.4: Distribution of Sentiments in Reddit Across Different Political Affiliation.

tically significant shifts in sentiment tone.

For subreddit-level affiliation, Democrat- and Republican-leaning communities both exhibited increases in sentiment volume post-leak. However, changes in sentiment distribution were not statistically significant, with p-values just above the 0.05 threshold ($p = 0.08383$ for Democrat-Leaning, $p = 0.09007$ for Republican-Leaning).

Finally, within abortion-specific subreddits, neither the Pro-Choice nor the Pro-Life communities exhibited statistically significant changes in sentiment tone before vs. after the leak. This suggests that these highly focused communities maintained consistent sentiment framing on the issue, even as broader discourse evolved.

Overall, the findings indicate that Center/Independent users were the only group to exhibit a statistically significant shift in sentiment, while most politically affiliated communities showed consistency in tone before and after the Dobbs event.

Table 4.6: Percentage of sentiment before and after the Dobbs leak (May 2, 2022) in Reddit based on political affiliation of user or community, with Chi-Square test results

Group	Subgroup	Period	Negative (%)	Positive (%)	χ^2	p-value
User Political Affiliation	Right-Leaning	Before	69.86	30.14	1.94	0.16382
		After	66.08	33.92		
	Center/Independent	Before	69.71	30.29	9.01	0.00269
		After	65.84	34.16		
	Left-Leaning	Before	68.38	31.62	0.49	0.48181
		After	66.82	33.18		
Subreddit Affiliation	Democrat-Leaning	Before	68.39	31.61	2.99	0.08383
		After	72.03	27.97		
	Republican-Leaning	Before	68.89	31.11	2.87	0.09007
		After	74.99	25.01		
Abortion Subreddits	Pro-Choice Subreddit	Before	65.00	35.00	0.78	0.37644
		After	65.63	34.37		
	Pro-Life Subreddit	Before	63.27	36.73	1.91	0.16705
		After	62.77	37.23		

Table 4.7 compares sentiment distributions between Reddit groups before and after the Dobbs v. Jackson Women's Health Organization draft leak, using Chi-Square tests and Jensen-Shannon divergence to assess polarization.

For user political affiliation inferred from usernames and flair, no significant difference in sentiment distribution was found between Left-Leaning and Right-Leaning users either before or after the leak. Jensen-Shannon divergence remained constant (0.0016), and Chi-Square p-values were well above 0.05, indicating no meaningful change in polarization at the individual user level.

In contrast, subreddit-level affiliation showed a modest increase in polarization. While Democrat- and Republican-leaning subreddits had nearly identical sentiment distributions before the leak ($p = 0.975$), the difference became statistically significant after the leak ($p = 0.0189$), even though the Jensen-Shannon divergence remained the same. This suggests that subreddit-level framing diverged more strongly after Dobbs.

The clearest evidence of increased polarization was found in abortion-specific subreddits. Although a significant difference already existed between Pro-Choice and Pro-Life communities before the leak ($p = 0.0207$), it became dramatically more significant afterward ($p < 0.00001$), reflecting a sharpening of differences in sentiment tone between the two sides. However, the Jensen-Shannon divergence did not increase, indicating that the volume of the difference, not the proportional distribution, grew.

Overall, polarization remained stable at the user level but increased at the community level, particularly around abortion-focused subreddits.

Table 4.7: Percentage of sentiment between Reddit communities before and after the Dobbs leak, with Chi-Square test results and Jensen-Shannon divergence

Comparison	Period	Group 1 (Neg/Pos %)	Group 2 (Neg/Pos %)	χ^2, p-value	Jensen-Shannon
User Political Affiliation (Left vs. Right)	Before	68.38 / 31.62	69.86 / 30.14	$\chi^2 = 0.21, p = 0.650$	0.0016
	After	66.82 / 33.18	66.08 / 33.92	$\chi^2 = 0.18, p = 0.675$	0.0016
Subreddit Affiliation (Dem vs. Rep)	Before	68.39 / 31.61	68.89 / 31.11	$\chi^2 = 0.00, p = 0.975$	0.0219
	After	72.03 / 27.97	74.99 / 25.01	$\chi^2 = 5.51, p = 0.0189$	0.0219
Abortion Subreddits (Pro-Choice vs. Pro-Life)	Before	65.00 / 35.00	63.27 / 36.73	$\chi^2 = 5.35, p = 0.0207$	0.0197
	After	65.63 / 34.37	62.77 / 37.23	$\chi^2 = 74.35, p < 0.00001$	0.0197

4.2 Emotion Recognition

The emotion analysis reveals substantial differences in how abortion is emotionally framed across mainstream media outlets and Reddit communities. In both CNN and Fox News, fear emerges as the dominant emotion. CNN displays the highest proportion, with over 50% of abortion-related content reflecting fear, compared to 45.6% in Fox News. This suggests that both networks frequently frame abortion in terms of threat, danger, or concern. However, Fox exhibits significantly more anger (25.8%) compared to CNN (17.4%), indicating a more confrontational and aggressive emotional tone. CNN shows a higher level of joy (14.1%) relative to Fox (11.4%), implying occasional positive narratives. Both outlets reflect relatively low levels of neutral, sadness, and surprise emotions, underscoring the emotionally charged nature of abortion coverage in traditional news media.

In contrast, Reddit presents a more balanced emotional distribution, with anger remaining prominent across all user groups, ranging from approximately 26% to 34%. The highest levels of anger are found in left-leaning users (33.7%) and Republican subreddits (30.9%) which may be a sign of heated and emotionally charged debate around the subject. Reddit also features higher levels of sadness (10–14%) and neutral sentiment (up to 25%) compared to mainstream media. These trends suggest that user-generated discourse allows for a broader affective spectrum, reflecting not only outrage but also expressions of grief, concern, and resignation—particularly among pro-choice and centrist subreddits. This can be explained by the natural difference between news media and public forum.

Fear, which dominates in media coverage, is notably less prominent on Reddit, where it ranges between 16–22% across communities. This suggests that Reddit users are less likely to frame abortion discourse in terms of threat and more inclined to express anger or social criticism. Furthermore, joy remains relatively consistent across Reddit groups (around 10–11%), and surprise remains marginal.

Overall, the findings suggest that while both media and Reddit reflect polarized emotional responses to abortion, their emotional structures differ. Mainstream media frames are dominated by fear and anger, while Reddit communities express a more nuanced and distributed emotional engagement with the issue.

The distribution of emotions over time shown in Figure 4.5 demonstrates an interesting

Figure 4.5: Line Plot of Distribution of Emotions in CNN and FOX News.
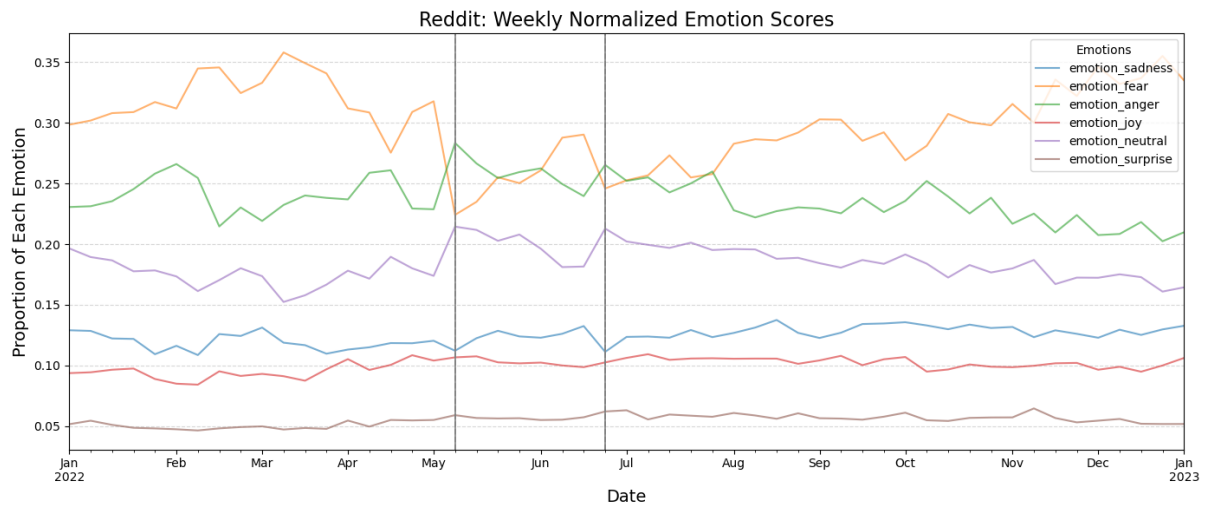
Figure 4.6: Line Plot of Distribution of Emotions in Reddit.

26

Table 4.8: Average normalized emotion distribution (%) across different media and Reddit groups

Source	Sadness	Fear	Anger	Joy	Neutral	Surprise
CNN	6.92	**50.53**	17.41	14.05	4.12	6.96
Fox	5.97	45.55	**25.80**	11.39	5.90	5.38
Reddit Combined	12.42	28.42	**24.37**	10.18	18.97	5.63
Reddit Pro-Choice	14.06	22.41	**26.71**	11.24	18.80	6.79
Reddit Pro-Life	13.35	21.34	**26.19**	11.60	21.17	6.35
Reddit Democrat Subreddits	10.36	16.23	**29.66**	11.26	24.15	8.34
Reddit Republican Subreddits	11.32	16.35	**30.90**	10.20	23.74	7.50
Reddit Left-Leaning Users	10.34	19.23	**33.70**	10.47	20.30	5.97
Reddit Right-Leaning Users	12.60	17.27	**31.40**	10.16	22.67	5.89
Reddit Center Users	12.94	17.31	28.11	10.48	**24.77**	6.40

difference between CNN and FOX News. While for both media dominant emotions are fear and anger, after draft opinion leak, CNN has small increase in fear, while FOX News' emotion of fear decreased. But in general CNN has more stable emotional tone, while FOX News has some jumps in fear, anger, and surprise.

Reddit shows more stable emotional tone, showing stability in public discussions. Although there are more emotional diversity, but fear and anger are still prominant emotions.

4.3 Topic Modeling

Identifying key topics in texts is an effective way to uncover the main themes being discussed. It is also one of the computational approaches used in frame analysis [4]. Therefore, for the next stage of the analysis, I applied topic modeling to examine how the framing of abortion differed across various levels of sources.

In political communication research, contextual embedding models such as BERT are increasingly used to generate context-sensitive semantic representations of text [4, 6, 38, 44]. In this study, I employed the Sentence-BERT model (all-MiniLM-L6-v2) [40].

After applying model, dimensionality reduction, I applied HDBSCAN [42], a density-based clustering algorithm. The resulting clusters represent topically coherent and semantically meaningful groups of texts based on their contextual similarity.

In the following section, I present the results of the topic modeling applied to news media data. This analysis was conducted to explore how the framing of abortion differs across media sources and how it changed over time, with particular attention to changes before and after the leak of the Dobbs draft. Topic modeling was applied to the corpus of news outlet texts to enable a comparative analysis of dominant frames before and after the leak. Additionally, I used the same approach on Reddit data to examine how the general public framed the issue of abortion over the same period (2022).

To improve the interpretability of the topic model, I removed words that were repeated across all documents or contributed little semantic value, such as technical or platform-specific terms (I provide the list of removed words in Appendix A.1). I also excluded documents labeled as -1, which were classified as noise by HDBSCAN, as well as clusters that contained fewer documents than a defined threshold for each dataset. This threshold varied by dataset size; for example, I retained all clusters for Fox News, while filtering the clusters with fewer than 20 documents for CNN and fewer than 1,000 for each topic on Reddit.

Topic modeling was applied to Fox News and CNN separately, and then to their combined dataset. In each case, I applied topic modeling on documents from before the Dobbs draft leak on May 2nd, 2022, and after. Again, the model was run on the entire Reddit dataset and then separately on documents from before and after the leak. Each step provided results that contributed to answering the research questions.

Finally, I performed a manual refinement step. Some clusters were found to be irrelevant, while others showed substantial overlap in their top keywords. I therefore removed the irrelevant clusters and merged those with significant keyword overlap. The final set of topics reported reflects both the original clustering model and this interpretive refinement.

4.3.1 News Outlet Frames

In this section, I present the results of topic modeling. I identified topics represented in Fox News and CNN, which are often seen as representing the two main political sides in the United States. As a first step, I applied topic modeling to all documents related to abortion from the year 2022. The results clearly show that the two outlets cover the topic in different ways. For Fox News, the main themes are shown in Table 4.9.

Fox News has three main topics. The largest cluster focuses on family-related words and child care, indicating that abortion is discussed alongside other family concerns. This topic also criticizes the government at the time; during this period, there was a shortage of baby formula in the U.S. The topic includes mentions of Democrats and the administration, showing some focus on political leadership and responsibility criticism.

The second topic, named "Legal and Judicial Framing," presents abortion in terms of laws, court decisions, and official processes. This section is mostly neutral in tone and does not include strong emotional language. However, the third topic, "Protest and Public Response", includes words like "violence," "angry," and "protester." These words suggest a more negative tone when describing people reacting to the law changes. It is very likely about protests after the Dobbs leak (words such as "law," "opinion," and "leak" support this).

The topics identified from CNN news transcripts are shown in Table 4.10. The largest cluster, called "State-Level Electoral Campaign Framing," places abortion in a political campaign context, highlighting the topic's importance and influence on elections, especially in swing states like Pennsylvania and Georgia. Words such as "candidate," "poll," "election," and "Re-

28

publan" confirm this. Additionally, this cluster partly discusses the effects of the Roe v. Wade overturn on states with stricter abortion access, such as Georgia.

The second-largest cluster frames abortion in legal and judicial terms, reflecting CNN's coverage of the Dobbs decision.

The last cluster frames the issue under reproductive rights and access to abortion as a medical and health concern, suggesting that CNN views the topic as a rights and health issue.

One important point is what FOX News does not include. Unlike CNN, there is very little focus on abortion as a healthcare or rights issue. Based on framing theory, frames are not just about what is included, but also about what is left out or ignored [15]. In this case, FOX's choice not to highlight topics like reproductive rights or healthcare access suggests a different way of presenting the issue—one that does not frame it as an invasion of rights or a matter of medical access. This is especially important since, as shown in Figure 4.1, FOX covered abortion far less than CNN after the leak of the Dobbs draft. The low volume of coverage, along with the choice of focus, points to a clear bias in how FOX News chose to report the issue.

Table 4.9: Topics Identified from Abortion-Related Transcripts from FOX News

Cluster(s)	Topic Label	Keywords	Doc Count
1	Policy and Political Commentary	formula, baby formula, baby, abortion, end, need, american, begin, house, percent, big, happen, democrats, administration, actually, let, border, white, woman, hungary, black, kid, smith, bongino, child, allow, twitter, long, mask, business	82
2	Legal and Judicial Framing	abortion, decision, court, leak, law, opinion, justice, draft, woman, democrats, overturn, levin, vote, party, case, draft opinion, life, ruling	21
3	Protest and Public Response	justice, abortion, home, protest, outside, law, protester, violence, house, pro, justice home, opinion, angry, sexton, leak, white, activist, vaughn, passive	13

Although the topic of Legal and Judicial Framing includes many shared keywords between CNN and FOX, there are notable differences in emphasis. While FOX focuses more on partisan

language—mentioning figures like "Levin"[1] and emphasizing the role of "Democrats"—CNN's framing includes broader institutional concerns. It highlights issues of transparency and legitimacy, with references to legal processes such as "committee", "hearings", "trials", "subpoenas", and "investigations". This suggests that CNN frames the Dobbs ruling as part of wider political and governmental accountability processes.

"committee" and "hearings" suggest formal legislative or judicial scrutiny, where elected representatives examine issues publicly, allowing for transparency and public participation. "trials" imply judicial processes where evidence is evaluated under rule of law, reinforcing the legitimacy of outcomes. "Subpoenas" and "investigations" denote efforts to uncover facts and hold individuals or institutions responsible, emphasizing checks and balances within government. So, together, it can provide support that CNN frames the Dobbs ruling not just as a standalone legal event but as part of a process of political and institutional accountability.

Unlike FOX, which referenced the shortage of baby formula as a critique of current policy and governance, CNN does not highlight this issue. This omission may reflect CNN's partisanship, as it tends not to emphasize shortcomings associated with Democratic leadership.

CNN also draws attention to the electoral impact of abortion policy, especially in key swing states. Which is absent from FOX News topics.

Most distinctively, CNN emphasizes the health-related framing and importance of women's bodily autonomy. This health- and rights-focused framing presents abortion not just as a legal or political issue, but also as a matter of personal well-being and medical access—an aspect that receives no attention in FOX's coverage.

Table 4.10: Topics Identified from Abortion-Related Transcripts from CNN

Cluster(s)	Topic Label	Keywords	Doc Count
1	State-Level Electoral Campaign Framing	abortion, georgia, herschel, election, warnock, voter, republican, vote, race, candidate, kemp, campaign, senate, pennsylvania, democrats, oz, county, poll, florida, desantis, governor, kansas, democratic	219
2	Legal and Judicial Framing	abortion, justice, court, woman, decision, opinion, law, draft, overturn, case, vote, leak, legal, committee, testify, trial, house, privilege, contempt, executive privilege, subpoena, criminal, attorney, judge, investigation	190

---

[1]Mark Levin is a conservative political commentator, radio host, and author known for his influential commentary on legal and political issues.

Cluster(s)	Topic Label	Keywords	Doc Count
3	Abortion Access and Reproductive Rights	abortion, woman, law, ban, week, decision, care, life, need, access, doctor, health, case, child, court	105

In the next stage of the analysis, I explore how abortion was framed in news coverage before and after the leak of the Dobbs draft decision. The goal is to assess how this significant legal event may have influenced topic emphasis and discourse structure on opposite sides of the political spectrum.

As shown in Table 4.11, FOX News coverage prior to the leak often presented abortion alongside broader political and social issues. In particular, one topic cluster framed abortion within general political discourse, including references to political figures, legislative action, and policy concerns. Another cluster was more focused on family and parenting, with co-occurring terms such as "child," "baby," "school," and "parent." These associations may suggest that the topic was integrated into broader discussions around family and societal values.

Following the leak, the emergence of new themes is evident. One cluster focuses on privacy-related language, linking abortion to terms such as "policy" and "privacy." Another cluster addresses economic issues, such as the baby formula shortage, which appeared in the same context as abortion along with other references to household or social welfare and economic terms like "inflation." The occurrence of terms like "Democrat" and "Biden" again highlights a critical view of Democratic rule. FOX News coverage also included clusters connected to the legal and judicial response to the draft, as well as protest-related discourse that referenced demonstrations and public reactions. The "Political and Judicial Framing" uses an emotionally neutral tone, covering the news around law changes. Naturally, protests by pro-choice groups are covered under the "Protest and Public Response" topic.

In CNN's coverage (Table 4.12), legal and political dimensions of abortion were already present before the leak. Clusters contained references to the Supreme Court, judicial figures ("Breyer[2]"), and legal processes, showing coverage of abortion-related news and the threats surrounding the "end" of reproductive rights. State-level policies also appeared in early clusters, particularly focusing on specific legislation and actions in states like Oklahoma, Florida, and Texas.

After the leak, CNN continued to focus on legal framing, with an expanded emphasis on decision-making processes, judicial opinions, and legal precedent. In addition, clusters emerged highlighting the relationship between abortion and broader electoral dynamics—especially at the state level—and other policy areas such as economic trends and gun violence. Compared

---

[2]Justice Stephen Breyer is a former Supreme Court justice known for his liberal views and support for abortion rights.

to FOX, CNN's topic structure included more explicit references to healthcare, access, and institutional accountability.

In summary, both outlets demonstrated a shift in topical emphasis following the draft leak, but with different areas of focus. FOX introduced themes related to privacy, economic context, and protest, while CNN expanded its framing into legal investigation, electoral relevance, and access-related concerns. These shifts suggest the Dobbs draft served as a turning point in how and to what extent abortion was discussed in both news environments.

Table 4.11: Comparison of Key Topics in Abortion-Related Transcripts from FOX News Before and After the Supreme Court Draft Opinion Leak

Cluster(s)	Topic Label	Keywords	Doc Count
Before Leak			
0	Political Framing	people, biden, jackson, abortion, end, let, need, american, senator, percent, oil, sanction, begin, justice, life, course, believe, energy, big, law	25
1	Family Oriented Framing	people, kid, baby, child, smith, mask, teacher, happen, maybe, woman, tiktok, parent, fake, need, life, school, law, long, bongino, biden	19
After Leak			
3	Privacy Policy	abortion, justice, democrats, leak, policy, privacy, angle, law, radio, home, listen, report, privacy policy, menu home, policy term, policy term use, matter play, facebook twitter email, menu, episode authorize reduce	44
1	Economic and Political Commentary	biden, people, formula, baby formula, baby, abortion, shortage, house, need, border, actually, happen, american, white, big, problem, inflation, end, democrats, formula shortage	38

(continued on next page)

Cluster(s)	Topic Label	Keywords	Doc Count
2	Political and Judicial Framing	abortion, decision, people, court, leak, justice, opinion, woman, law, draft, democrats, biden, levin, overturn, vote, case, party, democracy, constitution, draft opinion	23
0	Protest and Public Response	justice, people, abortion, home, protest, rivera, baby, violence, outside, montgomery, protester, house, law, formula, yellen, baby formula, pro, host, believe, biden	13

Table 4.12: Comparison of Key Topics in Abortion-Related Transcripts from CNN Before and After the Supreme Court Draft Opinion Leak

Cluster(s)	Topic Label	Keywords	Doc Count
		Before Leak	
4, 2	Political and Judicial Framing	court, justice, breyer, people, abortion, law, biden, case, woman, end, judge, jackson, election, senator, point, justice breyer, vote, term, let, house, election, foreman, overturn, pence, thomas, vice, republican, mike, cheney	80
0	State-Level Abortion Policy	abortion, woman, oklahoma, law, ban, week, governor, texas, people, republican, romero, end, florida, pass, case, abortion week, make, sign, life	28
		After Leak	

Cluster(s)	Topic Label	Keywords	Doc Count
44, 46	State-Level Campaign Framing	pennsylvania, oz, abortion, voter, vote, people, race, republican, election, fracke, candidate, democrats, mastriano, governor, republicans, debate, support fracke, senate, end, campaign, georgia, kemp, warnock, herschel, abrams, woman, brian kemp	123
20	Legal and Judicial Framing	abortion, justice, court, decision, people, woman, opinion, law, draft, overturn, draft opinion, zakaria, end, vote, today, leak, case, biskupic, precedent, legal	92
21	Transparency and Investigation	griffin, committee, people, harnwell, election, house, parscale, griffin camera, happen, th, abortion, white, precinct, testify, white house, subpoena, diodati, biden, trial, end	68
37	Economic Concerns	inflation, people, abortion, percent, rate, economy, democrats, biden, end, recession, price, republicans, week, bank, solomon, house, election, fed, high, point	63

The topics identified across the full media dataset before and after the leak in Table 4.13 show that abortion consistently functions as an important point in political debates. However, the overturning of Roe v. Wade shifted the debate toward more specific and directed discussions. Abortion was embedded in more general topics with associations to election, law, parenting, and ideological debates, but the Dobbs draft leak in May 2022 brought the issue into a wider range of political conversations, increasing its importance in different topics and intensifying debates over it in specific ways as well.

Abortion appears in the "State-Level Electoral Framing" topic, probably indicating its significance in swing states and its influence on election outcomes, with words like "Georgia," "vote," "candidate," and "party." It is also associated with racial inequality effects in politics and potential problems in restricting access for marginalized groups. This framing appears to

intensify partisan divides, as evidenced by the frequent mention of politicians and parties within the clustered topics. The results also highlight how different state-level legislations became central to the framing, with references to how abortion policy influences election campaigns and political strategies.

Additionally, new concerns emerged, such as the health risks associated with abortion bans and broader debates over bodily autonomy and individual rights. Words like "care," "doctor," and "health" emphasize medical concerns, while references to states like "oklahoma" and "texas" highlight focus on states with restrictive abortion laws, especially those more affected by the overturn.

Table 4.13: Comparison of Key Topics in Abortion-Related Transcripts from News Outlets Before and After the Supreme Court Draft Opinion Leak

Cluster(s)	Topic Label	Keywords	Doc Count
Before Leak			
0, 1, 3	General Debate on Abortion	people, end, election, begin, court, woman, ve, happen, abortion, law, hospital, ban, school, foreman, need, biden, case, today, justice, government, american, let, course, percent, week, question	217
5	State-Level Abortion Legislation (FL, OK)	abortion, romero, week, florida, governor, desantis, oklahoma, parker, republican, fight, law, ban, woman, house, make, begin, happen, people, end, need	32
2, 4	Scattered Debates	woman, baby, bongino, smith, kid, li, fake, chasten, montgomery, sex, censorship, crime, mask, leave, gingras, ostafieva, olesya, judge, sanction, war, child	28
After Leak			

Cluster(s)	Topic Label	Keywords	Doc Count
49, 36	State-Level Electoral Framing (GA, FL, AZ)	georgia, vote, kemp, election, voter, warnock, herschel, republican, abrams, race, governor, candidate, democrats, republicans, brian kemp, win, senate, democrats, arizona, desantis, florida, biden, house, democratic, party, kansas	110
33, 28	Legal and Governmental Framing	abortion, justice, court, decision, people, woman, opinion, law, draft, overturn, zakaria, biskupic, case, draft opinion, end, leak, today, precedent, vote, legal, committee, house, trial, testify, subpoena, contempt, privilege, executive privilege, white house, waive, biden, case	95
34, 25	Policy, Health, and Access Debates	abortion, people, oil, woman, end, ban, gioko, eu, world, chatterley, quest, course, decision, law, war, sanction, energy, european, care, patient, access, oklahoma, pill, governor, prosecute, provider, texas	89
52	Narratives on Individual Cases	abortion, woman, indiana, story, care, old, case, law, bernard, doctor, rape, report, health, dr, girl, stelter, life	40

Based on the combined topic modeling and the dynamic distribution plot of topic shares across CNN and FOX News (Figure 4.7 and Table 4.14), it is evident that abortion was framed through multiple lenses in both media outlets. The dominant topic, Electoral Politics (Cluster 13), suggests that abortion was closely tied to partisan dynamics, election strategies, and political figures. As the graph shows, this frame became increasingly prominent around the time of the draft opinion leak and continued after the final Supreme Court ruling. Both outlets appeared to emphasize this connection, possibly reflecting broader debates about how abortion

could impact voter behavior and midterm campaigns.

The topic labeled Legal and Judicial Framing (Cluster 7) also shows a significant increase in coverage post-leak, especially in CNN. The rise of this frame after May 2 indicates that the legal implications of the Dobbs draft, including the court's role and party stances, became central in the narrative. Meanwhile, FOX had a relatively larger share in framing abortion through Economic (Cluster 27) and Infant Care (Cluster 0) lenses—especially in the period leading up to the leak—where topics like baby formula shortages and inflation were tied to broader critiques of Democratic governance. On the other hand, CNN seemed to place more emphasis on "Abortion Access and Policy Impact" (Cluster 41), reinforcing a rights- and policy-oriented narrative. These trends highlight not only the shared focus on abortion as a political issue but also the different framing priorities each outlet maintained during a time of major judicial and political change.

I also measured sentiments in documents within each cluster and calculated the percentage of positive labels relative to the total. While the absolute values have limited meaning—since positivity is largely driven by the dominance of CNN documents in the dataset—they can be used for comparative purposes. The results show that most documents in the "Legal and Judicial Framing" cluster have the lowest proportion of positive sentiment, indicating a generally negative attitude toward the overturning of Roe v. Wade. Other clusters exhibit relatively balanced sentiment distributions, though the cluster related to infant care carries a more positive tone.

Table 4.14: Topics Identified from News Outlets with Positive Sentiment Label Percentage

Cluster	Topic Label	Keywords	Doc Count	Pos Sent (%)
13	Electoral Politics	abortion, people, election, republican, voter, vote, democrats, republicans, race, end, woman, candidate, biden, house, georgia, senate, begin, win, happen, party	643	62.52
7	Legal and Judicial Framing	abortion, people, justice, court, decision, woman, law, opinion, case, house, end, overturn, vote, today, committee, happen, democrats, biden, draft, begin	271	49.45
27	Economic Framing	abortion, people, woman, end, week, inflation, law, need, ban, democrats, percent, biden, happen, begin, vote, house, price, economy, today, decision	216	62.04

(continued on next page)

Cluster	Topic Label	Representative Keywords	Doc Count	Pos Sent(%)
41	Abortion Access and Policy Impact	abortion, woman, law, people, ban, decision, happen, life, week, overturn, today, court, care, need, end, justice, hear, access, ban abortion, begin	66	63.64
0	Infant Care	formula, baby, baby formula, biden, abortion, need, people, parent, shortage, week, house, help, family, child, problem, woman, end, election, united, today	24	79.17
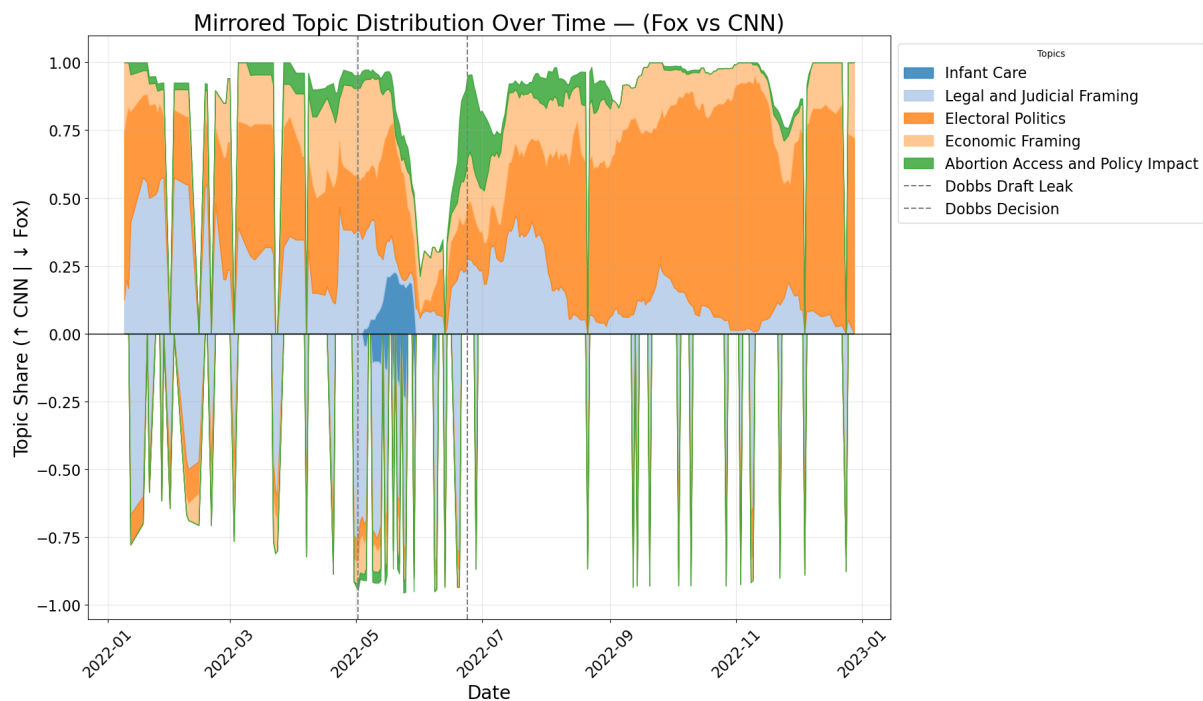
Figure 4.7: Topic distribution between FOX News and CNN in 2022. The upper panel shows the normalized to relative frequencies distribution of each topic across time in CNN, while the lower panel illustrates the distribution of topics in FOX (positive values indicate coverage on CNN). Vertical lines mark the draft opinion leak (May 2) and the final Supreme Court decision (June 24).

4.3.2 Reddit Frames

Reddit is a social forum that can be used to study public opinion. To answer the question of how public debate over the issue of abortion is conducted, I used Reddit posts and comments from subreddits related to politics, women's issues, and abortion (a list is provided in Appendix A.2).

This helps to examine, first, how the framing of abortion differs between mainstream media and public discussions, and second, to see if and how the change in law influenced that framing.

The topic modeling results for Reddit are shown in Table 4.15. Some major themes are shared across both domains—news outlets and Reddit—such as political and ideological debates, and concerns around bodily autonomy, health, and care. However, the way these themes are expressed differs. While the media often presents the issue in a more neutral or news-oriented tone, public discussions tend to use more emotionally charged language. For example, regarding bodily autonomy, Reddit users often refer to concepts like "slavery","enslavement," and "forced" in the topic "Body Autonomy," whereas in CNN coverage, the same issue is reflected through terms like "health," "right," and "care" (see Table 4.10). The health concerns on Reddit are also more personal and focused on lived experiences, such as threats to the life of the pregnant person or dangers of pregnancy, with terms like "ectopic," ectopic "pregnancy," and treatments associated with "life," "treatment," and "save" in the "Health and Medical Framing" cluster.

Political framing is present in both Reddit and media discourse. While both refer to parties and politicians, Reddit discussions include more frequent and direct references to political affiliations, suggesting a more polarized tone. Not only is the repetition of political affiliation evident, but words like "murder," "freedom," "choice," and "pro-life" also appear in the "Political and Ideological Debates" cluster, suggesting a more adversarial tone in discussions.

Reddit also offers unique nuances, such as opportunities for community-building and emotional or practical support. This is visible in Clusters 5, 13, 19, 53, and 651, which involve sharing experiences, offering information and guidance, and providing resources and emotional support. The appearance of words like "feeling," "feel," "pain," "help," and "love" is a sign of the emotional tone present on Reddit. On the other hand, the platform also hosts more polarizing and emotionally intense debates. For example, some pro-choice users refer to the fetus as a "parasite" or "zef," or use analogies like slavery (as seen in Clusters 231, 127, and 2010). Meanwhile, pro-life users emphasize terms like "kill," "murder," "child," and "life" (Clusters 127, 2010, 1898, and 231).

The sentiment scores associated with each Reddit topic cluster reveal important nuances in the emotional tone of public discussions on abortion. As shown in Table 4.15, clusters such as "Personal and Emotional Experiences" exhibit relatively higher proportion of positive sentiment (46.38%), reflecting supportive and empathetic exchanges within the community. In contrast, clusters focused on ideological and political debates, such as "Political and Ideological Debates" and "Right to Life and Consent Arguments," show more mixed or lower positive sentiment percentages (29.41% and 35.80% respectively), indicating the contentious and polarized nature of these discussions. Interestingly, the "Legal Enforcement" cluster exhibits an overwhelmingly high positive sentiment (99.95%), likely due to the predominance of advocacy and calls for rule adherence within those texts in topic.

Table 4.15: Topics Identified from Reddit Discussion on Abortion with Percentage of Positive Sentiment Scores Label Percentage)

Cluster(s)	Topic Label	Keywords	Doc Count	Pos Sent (%)
13, 5, 19	Informational Resources and Community Support	birthcontrol, resource look right, birthcontrol comment experience, rule resource look, assistance information abortion, funding assistance information, link work subreddit, subreddit source information, intend substitute medical, immediately action, mistake pregnancy risk, work subreddit, local number, number immediately, advice think experience, experience medical local, substitute medical treatment, add rule, perform birthcontrol, question concern	41,093	17.93
651, 53	Personal and Emotional Experiences	feel, abortion, regret, decision, want, know, feeling, love, time, pregnancy, cramp, pain, pill, day, week, period, pregnant, pro, help, start, life, right, baby, woman, hard	23,208	46.38
1898	Right to Life and Consent Arguments	abortion, life, woman, right, human, child, pregnancy, person, body, people, consent, kill, fetus, force, sex, want, argument, think, pro, rape	6,111	35.80
2224, 1639	Political and Ideological Debates	democrats, vote, republicans, party, republican, trump, voter, election, biden, gop, libertarian, liberty, government, pro-life, think, believe, choice, freedom, issue, abortion, state, murder, people, democratic	4,227	29.41

(continued on next page)

40

Cluster(s)	Topic Label	Keywords	Doc Count	Pos Sent (%)
127, 2010	Biological and Philosophical Analogies	parasite, parasitic, fetus, host, immune system, zef, pregnancy, cell, donation, kill zef, organism, relationship, life, human, fetus parasite, specie, parasitism, placenta, zefs, nutrient, definition, right, pregnant person, abortion kill	2,897	28.62
375	Health and Medical Framing	ectopic, ectopic pregnancy, pregnancy, abortion, tube, treatment, life, treatment ectopic, treatment ectopic pregnancy, fallopian, fallopian tube, mother, miscarriage, woman, pregnancy ectopic, doctor, save, medical, treat ectopic	2,822	31.18
51	Legal Enforcement	like remind rule, people action perform, unwelcome ban rule, comment question long, talk people action, long pro choicer, play safe talk, demonstrate open minded, discretion want avoid, advocacy trolling, simply advocacy, question long pro, lot discretion want, avoid ban play	1,931	99.95
231	Body Autonomy	slavery, slave, abortion, force, right, people, woman, human, pregnancy, owner, fetus, argument, body, enslave, person, compare, ban, want, child	1,500	23.33

Based on the Reddit topic modeling results before and after the Supreme Court draft leak, shown in Table 4.16, there are noticeable shifts in how users framed the abortion debate. Before the leak, discussions were primarily centered on broad ethical considerations. Cluster 1 focused on moral arguments about abortion, bodily rights, and the value of life, often expressed through ethical reasoning and personal experiences. Words such as right," question," life," human," and "kill" are typical in ethical discussions, reflecting the spectrum of beliefs around abortion. Another cluster addressed judicial and legal discourse, referencing courts, trials, and legal appeals. This indicates that even before the leak, abortion was framed not only as a political issue but also as a legal one, with emphasis on individual cases and legislative roles.

After the leak, however, the framing became more layered and diversified. A significant portion of the discussion shifted toward practical concerns, as seen in the "Informational

Resources and Community Support" cluster, where users shared guidance, links, and emotional support to help others navigate abortion access. Simultaneously, a legal-political frame emerged through the "Constitutional Framing" cluster, where users debated abortion as a matter of state control versus individual autonomy. Words like "right," "libertarian," and "government" support the idea that control over a woman's body became a central discussion point regarding state versus individual authority.

Finally, the second largest cluster after the leak combined ethical concerns with highly emotional and embodied experiences—covering not only concepts of consent and personhood but also feelings, pain, and the physical realities of pregnancy and abortion. Terms such as "baby," "child," "consent," "kill," "sex," and "feel" reflect ongoing moral and ethical debates between pro-life and pro-choice groups. Overall, the leak appeared to trigger more diverse, emotionally charged, and ideologically framed discussions on the issue.

Table 4.16: Comparison of Key Topics in Abortion-Related Posts and Comments from Reddit Before and After the Supreme Court Draft Opinion Leak

Cluster(s)	Topic Label	Keywords	Doc Count
Before Leak			
1	Ethical Debates	abortion, pregnancy, woman, right, question, birthcontrol, people, life, action, human, want, think, kill, body, like, child, experience, sex, risk, fetus	103853
0	Court and Judicial Framing	abortion, vs, woman, man, black, trial, right, case, try, court, white, rape, supreme, supreme court, people, party, legal, appeal, fetus, birth	11
After Leak			
19	Informational Resources and Community Support	abortion, trimester abortion, information abortion, trimester, link, help, abortion question concern, perform abortion question, action perform abortion, women help, funding assistance, receive harassment, hard supportive, harassment troll, select trust, seek abortion usa, usa need clinic	33726

(continued on next page)

Cluster(s)	Topic Label	Keywords	Doc Count
6, 15	Constitutional Framing	abortion, right, libertarian, state, people, life, fetus, government, woman, law, roe, think, like, human, want, court, child, person, argument, constitution	7367
4, 17, 116, 1478	Ethical and Emotional Framing	abortion, pregnancy, woman, right, life, people, fetus, want, child, think, person, human, baby, like, consent, kill, pregnant, sex, body, pro, state, feel, know, week, time	10046

4.4 Discussion

This thesis aimed to answer two main questions: first, how media outlets with different political affiliations and public users frame the abortion debate, and how this framing changed following the *Dobbs v. Jackson* decision; and second, what attitudes and emotional tones are expressed by media and individuals with different political affiliations, and how these differ or shift after the policy change.

To address these questions, I analyzed data from CNN and FOX News—two ideologically distinct media outlets—and public discussions on Reddit. The sentiment analysis revealed a significant difference between the news sources. CNN maintained a more positive tone overall, while FOX News was more negative. This pattern aligns with their respective political leanings, with CNN generally supporting abortion rights and FOX taking a more critical stance. In both outlets, fear and anger were the two most frequently expressed emotions.

Framing analysis further highlighted substantial differences. CNN devoted notable attention to topics such as "Abortion Access and Reproductive Rights," a framing absent in FOX coverage. Instead, FOX focused on themes such as "Protest and Public Response," describing events using emotionally charged language like "violence" and "angry," particularly in response to demonstrations after the draft leak. FOX also framed abortion alongside broader criticisms of Democratic governance, such as economic concerns and the baby formula shortage, whereas CNN emphasized electoral implications, especially in swing states like Georgia and Pennsylvania, and discussed abortion through legal and rights-based lenses.

Overall, CNN framed abortion more positively and linked it to reproductive health and political accountability, while FOX's framing focused on protest, policy failures, and critiques of the Democratic administration. These differences illustrate how editorial choices selectively amplify or omit frames, reflecting the ideological priorities of each outlet.

Reddit discussions exhibited more emotionally diverse content. While fear and anger were again dominant, a wider range of emotions such as sadness and surprise also appeared. Reddit's public discourse was more polarized and morally charged, using terms like "slavery," "murder," and "parasite" in debates over bodily autonomy and fetal rights. These emotionally intense framings confirm earlier findings that social media platforms enable more polarized and affectively charged discourse than traditional media. For example, the topic "Body Autonomy" featured moral comparisons to slavery, while others centered around fetal personhood or medical emergencies such as ectopic pregnancies. These findings are in line with the work of Pan et al. (2023) [6] and Moyer (2022) [3], showing that social media facilitates the expression of moral and emotional arguments that might be absent or softened in institutional news reporting.

Although sentiment on Reddit remained largely negative both before and after the Dobbs leak—similar to the tone observed on FOX News—framing evolved after the leak. Topic modeling revealed an increased emphasis on practical support, access to care, and community resource-sharing. This shift toward informational and supportive posts reflects how public discourse responded to legal uncertainty by adapting in tone and purpose. Emotional language remained high, but the emergence of support-focused frames shows that online communities became not only arenas of debate but also sources of solidarity and information during the post-leak period.

To assess whether these observed differences were statistically meaningful, I conducted Chi-Square tests and calculated Jensen–Shannon Divergence (JSD). Before the leak, the sentiment distribution between CNN and FOX was already significantly different ($\chi^2 = 11.71$, $p = 0.00062$). After the leak, the difference became much more pronounced ($\chi^2 = 152.65$, $p < 0.00001$). This result was further supported by the JSD scores: the divergence between CNN and FOX increased from 0.2058 before the leak to 0.4027 after it. This shows not only that the differences were statistically significant but also that the degree of polarization between the two outlets intensified, supporting earlier research by Siekiera and Szews (2024) [7].

On Reddit, sentiment differences between ideological groups (e.g., pro-choice vs. pro-life communities, or Democrat-leaning vs. Republican-leaning subreddits) remained statistically stable in most cases. However, sentiment divergence became more pronounced across subreddit-level communities after the leak, suggesting increased polarization at the group level, even if overall emotional tone remained steady.

While the emotion analysis did not reveal substantial shifts following the Dobbs leak, the combined findings from sentiment and topic modeling indicate that legal changes such as the overturning of Roe v. Wade can serve as inflection points in public discourse. These changes do not necessarily alter emotional tone but can significantly influence sentiment and framing patterns across both institutional and public domains.

In summary, CNN framed abortion more positively and emphasized access, rights, and political accountability. FOX, in contrast, highlighted protest and moral concerns, using more negative sentiment. Reddit discussions were more emotionally intense and morally diverse,

with framing patterns shaped by ideological alignment and community context. After the Dobbs leak, polarization increased across both media and Reddit, suggesting that legal policy shifts can intensify differences in public and institutional discourse around reproductive rights.

Chapter 5

Conclusion

This study examined how abortion was discussed and framed in different media contexts, specifically, in liberal and conservative news outlets (CNN and Fox News) and on social media platform (Reddit), before and after the leak of the *Dobbs v. Jackson Women's Health Organization* draft opinion. Using computational methods such as sentiment analysis, emotion recognition, and topic modeling, the research aimed to understand how media and public discourse responded to the legal shift following the potential reversal of *Roe v. Wade*.

The findings show clear differences in how CNN and Fox News covered abortion. CNN, generally considered left-leaning, framed the issue around reproductive rights, healthcare access, and the political impact of abortion laws, especially in swing states. Fox News, on the other hand, focused more on political commentary, public protests, and broader social issues, while not discussing abortion as a healthcare or rights-based issue. Sentiment analysis confirmed that CNN generally had a more positive tone, while Fox News maintained a mostly negative stance. These differences became even more noticeable after the Dobbs leak, suggesting that the media became more polarized around this issue.

Reddit, as a platform for public discussion, offered a wider range of emotional responses and framing. While overall sentiment on Reddit, in abortion related posts or comments, remained mostly negative both before and after the leak, the emotional content varied across communities. Anger and fear were common emotions, especially in pro-choice and Democrat-leaning subreddits, while pro-life and Republican-leaning subreddits showed slightly more positive sentiment. Interestingly, user-level political affiliation did not have a strong effect on sentiment, but community-level context (i.e., subreddits) did.

The topic modeling results further showed how framing differed across sources. News outlets focused on legal, political, and policy-related themes, with CNN including more content about reproductive health and elections. Reddit topics were more personal and emotionally driven, often reflecting lived experiences, activism, and ideological debates.

Overall, this study highlights the different ways that abortion was discussed in the media and among the public at a time of major legal change. Traditional media outlets mostly keep their neutral tone, while Reddit provided a more complex and participatory space for people to

46

express their views and emotions. These findings suggest that both media types play important roles in shaping how controversial issues like abortion are framed and understood.

This research contributes to a better understanding of media influence and public opinion during times of political and legal change. It also demonstrates the usefulness of computational methods in analyzing large-scale textual data. Future work could build on this study by including more media sources, longer time periods, other social platforms, or across countries to deepen our understanding of how such debates evolve.

Bibliography

[1] S. E. Davis Kempton, "My body, my voice: Analyzing news sources in the Roe v. Wade Reversal," *Newspaper Research Journal*, vol. 45, no. 3, pp. 299–310, 2024.

[2] A. Beck, B. Amiri, J. Dalven, J. Arons, L. Chaiten, and M. Roberts, "Trump on abortion," 06 2024. Accessed: 2025-05-22.

[3] L. P. Moyer, ""She Blinded Me with Science": The use of science frames in abortion litigation before the Supreme Court," *Justice System Journal*, vol. 43, no. 2, pp. 153–173, 2022.

[4] M. Ali and N. Hassan, "A survey of computational framing analysis approaches," pp. 9335–9348.

[5] D. Valdez and P. Goodson, "Neutral or framed? a sentiment analysis of 2019 abortion laws," vol. 19, no. 3, pp. 936–945. Publisher: Springer.

[6] J. Pan, Z. Wang, W. Qi, H. Lyu, and J. Luo, "Understanding divergent framing of the supreme court controversies: Social media vs. news outlets," pp. 5880–5887, IEEE.

[7] R. Siekiera and P. Szews, "News frames and differences in their application according to the author's beliefs. Polish conservative vs. liberal press on the protests against tightening the abortion law," vol. 32, no. 2, pp. 233–249. Publisher: Taylor & Francis.

[8] O. I. Ujah, P. Olaore, O. C. Nnorom, C. E. Ogbu, and R. S. Kirby, "Examining ethno-racial attitudes of the public in Twitter discourses related to the United States Supreme Court Dobbs vs. Jackson Women's Health Organization ruling: A machine learning approach," vol. 4, p. 1149441. Publisher: Frontiers Media SA.

[9] S. Ahmed, J. Cho, and K. Jaidka, "Framing social conflicts in news coverage and social media: A multicountry comparative study," vol. 81, no. 4, pp. 346–371. Publisher: SAGE Publications Sage UK: London, England.

[10] Z. Pan and G. M. Kosicki, "Framing as a strategic action in public deliberation," in *Framing public life*, pp. 51–82, Routledge.

[11] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a feeling: Accuracy and application of sentiment analysis," *International Journal of Research in Marketing*, vol. 40, no. 1, pp. 75–87, 2023.

[12] S. D. Reese, O. H. Gandy Jr, and A. E. Grant, *Framing public life: Perspectives on media and our understanding of the social world*. Routledge.

[13] J. Bryant and D. Miron, "Theory and research in mass communication," vol. 54, no. 4, pp. 662–704. Publisher: Blackwell Publishing Ltd Oxford, UK.

[14] R. M. Entman, "Framing: Toward clarification of a fractured paradigm," vol. 43, no. 4, pp. 51–58. Publisher: Blackwell Publishing Ltd Oxford, UK.

[15] T. Gitlin, *The whole world is watching: Mass media in the making and unmaking of the new left*. Univ of California Press.

[16] J. W. Tankard Jr, "The empirical approach to the study of media framing," in *Framing public life*, pp. 111–121, Routledge.

[17] K. Woodruff, "Coverage of abortion in select U.S. newspapers," *Women's Health Issues*, vol. 29, pp. 80–86, Jan–Feb 2019.

[18] G. Tuchman, "The news net," pp. 253–276. Publisher: JSTOR.

[19] R. M. Entman, "Cascading activation: Contesting the white house's frame after 9/11," vol. 20, no. 4, pp. 415–432. Publisher: Taylor & Francis.

[20] M. S. Güran and H. Özarslan, "Framing theory in the age of social media," no. 48, pp. 446–457. Publisher: Selcuk University.

[21] J. V. Pavlik, "News framing and new media: Digital tools to re-engage an alienated citizenry," pp. 311–321. Publisher: NJ: Erlbaum Mahwah.

[22] E. A. Mabry, "Textual framing as a communication climate factor in online groups," pp. 323–336.

[23] W. A. Gamson and A. Modigliani, "Media discourse and public opinion on nuclear power: A constructionist approach," *American journal of sociology*, vol. 95, no. 1, pp. 1–37, 1989.

[24] Z. Pan and G. M. Kosicki, "Framing analysis: An approach to news discourse," *Political communication*, vol. 10, no. 1, pp. 55–75, 1993.

[25] I. E. Kalabikhina, E. P. Banin, I. A. Abduselimova, G. A. Klimenko, and A. V. Kolotusha, "The measurement of demographic temperature using the sentiment analysis of data from the social network VKontakte," vol. 9, no. 9, p. 987. Publisher: MDPI.

[26] A. Roberti, ""women deserve better:" the use of the pro-woman frame in anti-abortion policies in US states," vol. 42, no. 3, pp. 207–224. Publisher: Taylor & Francis.

[27] J. P. Vossen, G. L. de Pooter, and P. Meier, "Conceptualizing morality policy: a dyadic morality frame analysis of a gendered legislative debate on abortion," vol. 55, no. 1, pp. 185–207. Publisher: Springer.

[28] G. Pinto, "Analyzing the overturning of Roe vs Wade on Twitter using natural language processing techniques," Master's thesis, Chapman University, 2023.

[29] T. Stanier and H. Shin, "Polarization and morality: Lexical analysis of abortion discourse on Reddit," *arXiv preprint arXiv:2407.00455*, 2024.

[30] A. Rao, F. Nanni, and H. Wachsmuth, "Discovering hostile framing patterns in online discussions," *arXiv preprint arXiv:2311.16831*, 2024.

[31] G. Sood, "CNN transcripts 2000–2025." Section: 2017-09-15 19:19:11.037.

[32] G. Sood, "Fox News transcripts (2003–2025)." Section: 2022-06-06 13:04:15.329.

[33] AllSides, "CNN editorial review," 2024. Accessed: 2025-05-13.

[34] AllSides, "Fox News editorial review," 2024. Accessed: 2025-05-13.

[35] J. Baumgartner, "Pushshift Reddit dataset," 2023. Accessed: 2025-05-13.

[36] J. Gerstein and A. Ward, "Supreme court has voted to overturn abortion rights, draft opinion shows," 2022. Accessed: 2025-05-13.

[37] Supreme Court of the United States, "Dobbs v. Jackson Women's Health Organization," 2022. Decided June 24, 2022. Accessed May 13, 2025.

[38] A. Rao, R.-C. Chang, Q. Zhong, K. Lerman, and M. Wojcieszak, "Polarized online discourse on abortion: Frames and hostile expressions among liberals and conservatives," *arXiv preprint arXiv:2311.16831*, 2023.

[39] Y.-R. Lin, B. Keegan, D. Margolin, and D. Lazer, "Rising tides or rising stars?: Dynamics of shared attention on Twitter during media events," *PloS one*, vol. 9, no. 5, p. e94093, 2014.

[40] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using Siamese BERT-Networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 2019.

[41] L. McInnes, J. Healy, N. Saul, and L. Großberger, "UMAP: Uniform manifold approximation and projection," *Journal of Open Source Software*, vol. 3, no. 29, p. 861, 2018.

[42] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on Hierarchical Density Estimates," in *Advances in Knowledge Discovery and Data Mining (PAKDD 2013)* (J. Pei, V. S. Tseng, L. Cao, H. Motoda, and G. Xu, eds.), vol. 7819 of *Lecture Notes in Computer Science*, pp. 160–172, Springer, Berlin, Heidelberg, 2013.

[43] Z. Drus and H. Khalid, "Sentiment analysis in social media and its application: Systematic literature review," *Procedia Computer Science*, vol. 161, pp. 707–714, 2019.

[44] M. A. Mersha, J. Kalita, *et al.*, "Semantic-Driven topic modeling using transformer-based embeddings and clustering algorithms," *Procedia Computer Science*, vol. 244, pp. 121–132, 2024.

51

Appendix A

A.1 List of Removed Terms Due to Redundancy or Low Informational Value

Table A.1: Terms Removed After Embedding for Topic Modeling

General Words	Vague/Broad Terms	Media Names	Redundant Terms
say	thing	timpf	supreme court
know	issue	bartiromo	court justice
think	time	steve	roe wade
want	year	steve bannon	right
come	day	orban	people
look	state	tyrus	biden
like	country	fox	trump
yes	new	cnn	joe
thank		pay programming	psaki
mean		weather	roe
work		copy	wade
try		picture	roe v wade
good		avatar	dark mode
		san	pelosi
		francisco	hearing
		nancy	crist
		nancy pelosi	charlie
		pelosi husband	charlie crist
		paul pelosi	bannon
		san francisco	mckend
			raphael
			amanpour
			donald
			walter
			kathy
			zeldin
			january

A.2 Subreddits Used for Data Collection

Table A.2: List of Subreddits Used in the Study

Subreddit	Subreddit
AARP_Politics	Abortiondebate
Democrat	DemocraticSocialism
DemocraticUnderground	DemocratsforDiversity
LeftistGamersUnion	LeftvsRightDebate
Leftyguitarists	Liberal
LibertarianLeft	LibertarianPartyUSA
LibertarianUncensored	Libertarian
PoliticalCompass	PoliticalDebate
PoliticalDiscussion	PoliticalOpinions
PoliticalRaceplay	Political_Revolution
PoliticsDownUnder	PoliticsPeopleTwitter
abortion	birthcontrol
conservatives	conservativeterrorism
democrats	leftist
politicly	politics2
politicsdebate	politicus
politicyl	prochoice
prolife	republicans
rightistvexillology	women