The University of Chicago

A Case for Regulating Political Microtargeting:
How the Messaging of Meta Ads in the 2024 U.S. Presidential Election Varied by Age, Gender, and State

By Kate Davis

A thesis submitted for partial fulfillment of the requirements for the degree of BACHELOR OF ARTS in PUBLIC POLICY STUDIES at THE UNIVERSITY OF CHICAGO

BA Faculty Advisor: Maria Angelica Bautista
Public Policy Studies Preceptor: Brian Curran
Second Reader: Jeffrey Tharsen

April 14th, 2025

ABSTRACT

This thesis investigates political microtargeting during the 2024 U.S. presidential election by analyzing over 120,000 unique Meta (formerly Facebook) ads from the official Trump and Harris campaigns, as well as from outside groups. Using sentiment analysis powered by OpenAI's GPT-4o, I assess how messaging in ad captions—categorized by central theme, emotional tone, and candidate alignment—varied by the age range, gender, and state of targeted Meta users. My findings reveal robust, statistically significant variation in political ad content across demographic traits, demonstrating how tone, theme, and candidate alignment shifted markedly depending on the age, gender, and location of intended audiences. These messaging differences shed a light on how official campaigns and outside groups leverage Meta's microtargeting technology to deliver distinct narratives to specific population segments, potentially shaping tailored perceptions of political candidates and issues. Through a comprehensive literature review demystifying political microtargeting, a novel case study of political microtargeting at-scale, and feasibility-minded policy recommendations, this thesis argues that political microtargeting erodes democratic legitimacy and demands urgent regulatory scrutiny. To promote transparency, replicability, and further applications of this work, all results from this thesis will be hosted in the University of Chicago Library's Institutional Repository under a DOI.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the guidance and support of numerous individuals. First and foremost, I thank and acknowledge Dr. Jeffrey Tharsen, Associate Director of Technology at the University of Chicago Forum for Digital Culture, for going above and beyond typical expectations of a second B.A. thesis reader. Dr. Tharsen generously took time out of his extremely packed schedule to regularly meet with me in-person, offering not only critical technical expertise for this thesis' methodology but also a collaborative space to brainstorm, troubleshoot countless challenges, refine ideas, copyedit, and receive (very needed) personal encouragement—from the early brainstorming stages of this project to the final days before submission. I have endless gratitude for his unwavering mentorship, for instilling in me a newfound confidence in my abilities as a data scientist, and for continually grounding me in the big-picture purpose of this thesis: protecting against democratic erosion.

I would also like to acknowledge and thank my preceptor, UChicago PHD candidate Brian Curran, for his unwavering patience and trust when I submitted assignments past suggested due dates, as well as for his timely responsiveness, supportive weekly check-ins, careful reading of my work, detail-oriented feedback, and early-stage guidance on structuring a thesis topic that provided a strong foundation for this thesis. Brian went out of his way to accommodate my and my peers' schedules, always making himself available when help was needed. Similarly, I am deeply grateful to my BA Faculty Advisor, UChicago Harris Senior Research Associate Maria Angelica Bautista, for her compassionate encouragement and for providing a comprehensive lecture series in the fall that clearly outlined every stage of the thesis process.

I would also like to thank and acknowledge UChicago Harris Sydney A. Stein Jr. Professor Anthony Fowler for generously taking the time to meet with me on two occasions to discuss statistical approaches for data analysis and for providing direct Stata code edits that were instrumental during my data cleaning process; UChicago undergraduate students JJ Davidoff, Ilyas Qureshi, and Yasmin Lee for spending several hours both troubleshooting coding challenges and creating code templates for me; David Axelrod, former Senior Advisor to U.S.

President Barack Obama, for taking the time to participate in an informational interview with me that provided helpful context on the rise of political microtargeting; and both former U.S. Senator Jon Tester (D-MT) as well as Erin Simpson, Program Officer with the Technology and Society program at the Ford Foundation, for meeting with me to discuss my policy recommendation ideas and their political feasibility.

Last, I explicitly acknowledge my use of ChatGPT-4o throughout this project. Stata, R, and python code generated by the model were used in every stage of my data cleaning, sentiment analysis, and statistical analysis. I also used GPT-4o as a copy-editing tool—for word-choice suggestions, grammar corrections, and sentence-level adjustments aimed at smoothing awkward phrasing. That said, all descriptions and analyses reflect my original ideas and interpretation; GPT-4o was used to refine my personal writing, never to generate text.

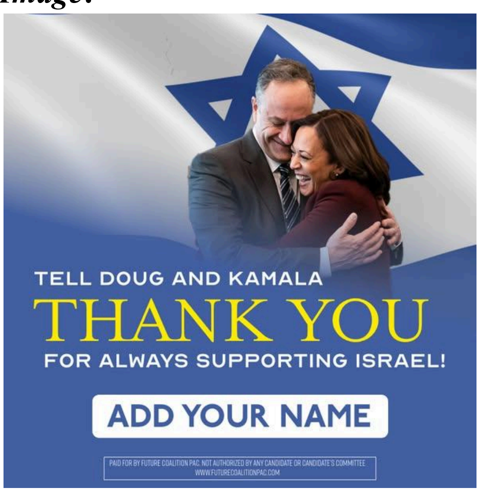
TABLE OF CONTENTS

INTRODUCTION

Kamala Harris: the Pro-Israel and Pro-Palestine Candidate?

Search the "Future Coalition PAC" on Meta's Ad Library—the company's "comprehensive ads transparency surface" (Meta Platforms, Inc. 2021)—and you'll find the two following Facebook and Instagram ads.

<table>
<tr><th>Ad #1</th><th>Ad #2</th></tr>
<tr><td>Ad Library ID: 553539253691695</td><td>Ad Library ID: 8359938514100359</td></tr>
<tr><td>Caption: "Join us in thanking Doug & Kamala for always standing with Israel. They stand for what is right. It's time we put a real pro-Israel president in the White House. Sign your name today to tell them THANK YOU!"</td><td>Caption: "Kamala Harris is secretly campaigning on her pro-Palestine beliefs and trying to get away with it! 👀"</td></tr>
<tr><td>Image:



</td><td>Image:

**Graphic cropped to fit page*</td></tr>
</table>

One thing should stand out off the bat: these ads directly contradict each other. What isn't obvious, though, is that these ads were created, purchased, and managed by the same organization. Yes, you read that right. Ad #1, celebrating Kamala Harris as a "pro-Israel" candidate, ran from September 23rd to September 30th, 2024, and appeared on users' screens between 125,000 to 150,000 times. It was targeted 100% to voters in Michigan, almost entirely to individuals between the ages of 18-34. Ad #2, attacking Harris for having “pro-Palestine beliefs,” ran from September 18th to September 30th, 2024, and appeared on users’ screens over 1 million times. It was targeted 99% to individuals in Pennsylvania, across a range of voters aged 18 to 65+.

Why would it be in the interest of an ad buyer to paint opposing depictions of the same candidate? The GOP-aligned Future Coalition PAC—funded by $3 million from Elon Musk’s Building America’s Future organization—claims in Federal Election Commission (FEC) filings that its digital ads are in support of Kamala Harris (Perkins 2024). Yet, in Pennsylvania, the swing state with the highest proportion of Jewish American voters (Saxe et al. 2021), the PAC’s ads condemn Harris for “pandering to Palestine.” Meanwhile, in Michigan, the swing state with more Arab American voters than any other (Arab American Institute 2024), the group’s ads identify Harris as a pro-Israel hardliner. The Future Coalition PAC’s tactic isn’t subtle: undermine support for Harris by sending out contradictory messages that match presumed foreign policy stances of targeted audiences, a move made possible by strategic use of Meta’s highly granular personal data on users’ political leanings.

Though perfectly legal, a simple case can be made for why the Future Coalition PAC’s clever operation is unethical. The ads are, by nature, deceptive; they cast Harris as both “pro-Israel” and “pro-Palestine,” labels conventionally considered mutually exclusive in the context of the ongoing war in Gaza. Instead of presenting a consistent portrayal of Harris’ views, the PAC tailors messaging to strategically exploit perceived divisions within critical voting blocs. As Hannah Arendt famously warns: “freedom of opinion is a farce unless factual information is guaranteed and the facts themselves are not in dispute” (Arendt 1968, 238). When voters are presented with inconsistent information that obscures fact—in this case, Harris' true opinions on Israel and Gaza—their opinions are at risk of being formed not freely, but by manipulation.

Thesis Motivation and Aims

Genuine democratic deliberation necessitates shared understandings of reality—and today, these understandings are arguably more under threat than ever before. Our ability to partake in informed, collective decision making is made increasingly difficult in an age where campaigns, PACs, and other outside groups can utilize Big Data and technology to microtarget individuals with varying, misleading, and sometimes outright contradictory information. The central role social media plays in modern life intensifies these challenges: Americans spend an average of 2 hours and 25 minutes a day on platforms (Bajarin 2023), with 54% reporting to get news from them too (Pew Research Center 2024). Young people may be especially susceptible to microtargeted content, considering the average U.S. teen spends 4.8 hours a day on social media (Rothwell 2023). Given the undeniable power platforms like Facebook and Instagram have in shaping political opinions, it is critical that scholars investigate and raise awareness about microtargeting in practice.

This thesis argues that political microtargeting poses a serious threat to democratic integrity by fueling the dispersal of political information selectively tailored to voters' identities. In addition to defending this normative framework, I seek to answer the following question: *How did the targeted messaging of Meta ads in the 2024 U.S. presidential election, from the official Harris and Trump campaigns as well as outside groups, differ by users' age, gender, and state?* I use sentiment analysis powered by OpenAI's GPT-4o to help interpret the independent variable: ad "messaging," as represented by ads' captions. Sentiment analysis is a natural language processing technique used to assign categorical labels to text. Typically, this process involves providing a large language model with access to the text of interest and thereafter specific instructions in the form of a prompt that guide the model on how to classify text inputs. Labels can take infinite forms; for instance, one might use labels to evaluate the emotional tone of text—with terms like “positive,” “negative,” and “neutral”—or to summarize a piece of text’s intended purpose—identifying whether text seeks to be “anti” or “in-favor of” a something. In the case of this study, I use sentiment analysis to categorize ads’ emotional tone, central theme, and candidate alignment (pro-Trump, anti-Harris, etc). I then apply statistical analysis to explore how the dependent variables—impressions by age, gender, and state—affect the content of ad messages.

I distinctly contribute to the literature in three key ways. First, I provide comprehensive analysis unpacking the origins of Big Data, the evolution of corporate—and eventually political—microtargeting, and the prevailing philosophical and ethical frameworks used to evaluate their impact on democratic societies. Second, I offer a novel case study that characterizes political microtargeting in practice—not by highlighting isolated examples, but by using sentiment and statistical analysis to identify messaging trends by age, gender, and state across over 100,000 unique political ads from Meta’s Ad Library. Third and last, I leverage my findings to craft public policy recommendations grounded in political feasibility.

Following this introduction, this thesis is divided into 6 additional sections: Literature Review, Data & Methods, Findings & Analysis, Policy Recommendations, Bibliography, and Appendix. The Literature Review traces the historical and conceptual development of Big Data, mapping how its commercial applications have evolved into strategic political tools. This section also examines how scholars across disciplines have evaluated the implications of microtargeting for personal autonomy and democratic legitimacy. In Data & Methods, I outline the basic mechanics of the Meta Ad Library, detail my ad search process, walk through a multi-stage data cleaning procedure, and explain my sentiment analysis approach labeling the relevancy, theme, tone, and candidate alignment of ads. Next, in Findings & Analysis, I summarize ad counts and estimated impressions across demographics, present visual analyses of population-specific messaging with colored-coded bar graphs, break down statistically significant results indicating variation across demographics, highlight the most striking differences in messaging, and reflect on the study's limitations, lessons learned, and further applications. In Policy Recommendations, I dissect Section 230 of the 1996 Communications Decency Act, propose reforms aimed at mitigating political microtargeting's democratic risks, and conclude with a note to the reader. My thesis ends with a full Bibliography and Appendix, which respectively list all cited sources and include complete demographic microtargeting results—contingency tables and summary statistics—provided for transparency purposes.

LITERATURE REVIEW

"Big Data is the new oil." (Virginia Rometty, 2013)

If you're reading this, it's likely that you've encountered "Big Data" often enough that the term isn't one you stop to ponder. A buzzword whose popularity—as measured by Google search records—skyrocketed between 2011 and 2014 (Google Trends 2024), Big Data has become too big for a "univocal" definition among leading sociologists and psychologists (Favaretto 2020, 1). One explanation for the term's ambiguity lies in its disputed origins. "Big Data" is referenced patchily and without explanation in published academic articles throughout the 1980s and 90s (Diebold, 2012 3). On the non-academic, non-published side of things, Big Data is alluded to as a phenomenon in a 1999 slide deck made by John Mashey, a Chief Scientist at Silicon Graphics (Mashey 1999). Mashey, considered by some to be the father of Big Data, downplays this title, explaining in a New York Times interview that Big Data can be used as "one label for a range of issues," and is "the simplest, shortest phrase to convey that the boundaries of computing keep advancing" (Lohr 2013). Lacking a concrete conceptual history, Big Data's broad and ever-evolving applicability may be what allows it to be as ubiquitous as it is, so much so that we've stopped questioning what it means in the first place.

Generic definitions of the term are pretty simple. Merriam-Webster Dictionary defines Big Data as "an accumulation of data that is too large and complex for processing by traditional database management tools" (Merriam-Webster.com Dictionary 2024). Similarly, the Oxford English Dictionary defines the term as "data of a very large size, typically to the extent that its manipulation and management present significant logistical challenges; (also) the branch of computing involving such data" (Oxford English Dictionary 2024). These definitions share a common thread: the idea that Big Data's size and complexity create methodological hurdles. Big Data certainly is different from structured data. Instead of relying on relational data stores—tables with rows and columns that represent relationships between different data points—Big Data does not operate in isolation and requires non-relational data stores, also known as "NoSQL databases" (Tamane 2016, 1). Big Data also requires distinct data analytics strategies (Tsai 2015, 1), such as machine learning and other subsets of artificial intelligence (AI), to draw meaningful insights. However, it's apparent that these definitions, which reduce Big Data to a field posing technical challenges, fail to pinpoint why Big Data is the buzzword that it is.

Some, like De Mauro et al., have attempted to address Big Data's conceptual vagueness by proposing a new formal definition of the term. Analyzing "a conspicuous corpus of industry and academia articles" to settle on an official meaning, De Mauro et al. assert that "Big Data is the Information asset characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value" (De Mauro 2016, 1). In other words, the authors paint Big Data—distinct from other forms of data in its vast quantity (volume), diversity (variety), and speed demands (velocity)—as a resource with unique processing requirements, whose value is not intrinsic but rather vested in a transformational potential to generate insights and change outcomes. Notably, the "3V" framework—identifying volume, velocity, and variety as the key features that set Big Data apart—is widely cited in the literature (Emmanuel et al. 2016, 1). Understanding the 3Vs in Big Data, furthermore, is considered necessary for generating the "innovation success" and "improved insights" that Big Data has to offer (Johnson et al. 2017, 640).

Despite common recognition of the 3V framework, De Mauro et al were the first to pin down Big Data's formal definition beyond its characteristics or methodological requirements; they define Big Data as a resource, almost like a raw commodity, that, when harvested and manufactured in specific ways, creates a new, more valuable product. This sentiment echoes IMB Chief Executive Officer Virginia Rometty's famous 2013 quote: "Big Data is the new oil" (Rometty 2013). Big Data "promises a plethora of new uses," from pandemic identification to the creation of entirely new business sectors, just as oil "generated useful plastics, petro-chemicals," and many other products (Hirsch 2013, 374). Among the many applications of Big Data, one that warrants more attention is the use of Big Data as a means of increasing political control—the focus of this thesis.

The use of data as a strategic political tool is far from novel. Forward-thinkers have recognized the role large quantities of data play in protecting and enhancing political hegemony as far back as the start of the third century BC. Take Ptolemy I Soter, Alexander the Great's successor in Egypt. Considered the most probable founder of the Library of Alexandria (Heller-Roazen, 2002 135), Ptolemy I took on a monumental effort in amassing data for the sake of legitimizing his new position in the eyes of the Greeks. Invoking Alexander's name and drawing upon intellectual practices of the Aristotelian school, establishing the Library safeguarded the Ptolemaic Kingdom with a public, indisputable cultural link to Greece, despite its rule over a predominantly non-Greek population (Erskine 1995, 41-42). At the same time, by establishing a "monopoly of Greek culture" through the claiming, organizing, cataloging, and editing of Greek books and texts, Ptolemy I and those that came after him collected data in an "imperialist" sense, too (Erskine 1995, 45). Eventually encompassing, according to one account, 400,000 "composite rolls" and 90,000 "single rolls", (Heller-Roazen 2002, 140), the Library's position as a centralized repository and arbiter of knowledge provided the Ptolemies consolidated power over information—and thus the ability to curate and disseminate specific understandings of history, philosophy, science, and literature across the Hellenistic world.

Of course, it goes without saying that the Library of Alexandria is not an example of Big Data in-line with De Mauro's "3V" framework. What this ancient example does highlight, though, is that amassing and controlling large quantities of data provides opportunities to shape political futures. With great amounts of data, Big (in its volume, variety, or velocity) or not, can come the ability to curate tailored messages, amplify specific narratives over others, and consolidate power over populations—three forms of data-generated value, a key component of De Mauro et al.'s definition. Just as Ptolemaic kings of the 3rd century BC leveraged the Library of Alexandria's data to enhance political dominance, modern political leaders leverage social media platforms' Big Data for similar objectives.

A Big Data Product: Microtargeting in Marketing

Before diving into the political power applications of Big Data, it's necessary to take a step back and examine these applications' origins: the marketing industry. Targeting consumers in corporate marketing is far from a new phenomenon; traditionally, advertising allocation is approached at the macro-level, with marketers establishing advertising budgets and strategically dispersing media across selected channels (Danaher 2023, 564). For example, after setting a budget, a business selling beach gear might disseminate media in targeted communities likely to frequent the beach, like those near the coastline. With the advent of behavioral Big Data, however, marketers can now disseminate media using far more precise metrics.

From online shopping to wearable fitness trackers, datafication across all facets of society is rendering human behavior measurable (Leonardi 2020, 1601). Through the analysis of vast, diverse, and rapidly accumulated data, marketers can pinpoint patterns and trends in human behavior that were previously invisible to them (Erevelles et al. 2016, 897). For instance, Big Data might allow a beach gear company to utilize location records from social media activity to pinpoint individuals who visit the beach every weekend. Targeting definitive beachgoers instead of suspected ones, the beach gear business is more likely to find consumers who, upon seeing their ads, will give-in and buy a new umbrella or towel set. By making decisions "on the basis of evidence,"—that is, data-backed trends or patterns—"rather than intuition," advertisers better predict consumers' wants and needs (McAfee 2012, 5). It's no surprise that, as Everelles et al. aptly put it, "consumer analytics is at the epicenter" of the "Big Data revolution" (Erevelles 2016, 897); traditional targeting methods in marketing—like the coastline strategy—become obsolete as our individual actions and sentiments are converted into actionable data points.

The practice of strategically using data on individual human behavior to deliver customized messages in pursuit of predicted outcomes is known as microtargeting (Agan 2007, 2). As a member of today's digitalized world, it's more likely than not that you experience microtargeting each and every day. From online ads and email promotions to streaming recommendations and social media feeds, marketers channel hyper-customized media our way whether we realize it or not. In the face of the 2000s "digital disruption,"—the rapid expansion of Internet access that threw "communication and the physical world of processes and goods into disorder" (Arther, 2013 9), marketers who failed to hop on the microtargeting train risked being left behind.

Corporate Microtargeting and Eroding Autonomy

Philosopher Shoshana Zuboff famously describes this new microtargeting imperative as part of a phenomenon she calls surveillance capitalism. A new economic order dependent on the "unilateral dispossession of human experience for the sake of profit," (Zuboff 2019, 7) surveillance capitalism first originated at Google between 2001 and 2004 (Zuboff 2019, 9). Responding to increasing pressure from investors to generate revenue from its search engine, Google made the monumental move of using data logs on user queries—originally viewed as "waste"—to transform their ad targeting strategy (Zuboff 2019, 10). These user queries, suddenly "an asset class of vital raw materials" created "at zero marginal cost" (Zuboff 2019, 12) in a marketing-focused production line, set a "historic turning point" (Zuboff 2019, 11) not just for Google but for all companies. Within just a few years, an economic model surrounding cheaply sourced—and often freely sourced—personal data, or "behavioral surplus," became the basis for generating profits across all sectors of the economy (Zuboff 2019, 11).

Zuboff raises serious ethical concerns about microtargeting within the framework of surveillance capitalism. Transforming individual people and their actions into quantifiable "data objects," Zuboff asserts that microtargeting practices dehumanize us (Zuboff 2019, 5). Beyond the moral discomfort that comes with datafying lived human experience, Zuboff also contends that microtargeting undermines personal autonomy by generating new forms of domination. Sparking what Zuboff calls the "prediction imperative"—a desire to accurately forecast human behavior so intense that it necessitates modifying behavior to guarantee certainty—surveillance capitalists analyze vast swaths of data with the goal of shaping our future actions (Zuboff 2019, 17). Denying what political theorist Hannah Arendt refers to as the "right to the future tense"—the ability to engage in actions in pursuit of the futures we aspire for—Zuboff argues that surveillance capitalism undermines our capacity for autonomous decision making (Zuboff 2019, 37). Unlike past forms of coercion, which posed various limits on our actions, microtargeting practices attempt to preempt our actions by steering us toward selected outcomes. Setting the script in advance, Zuboff warns that surveillance capitalism threatens our ability to engage in action at all—a practice that jeopardizes our "right to have rights" in the first place (Zuboff 2019, 38).

The Rise of Political Microtargeting

Despite the fact that Zuboff's surveillance capitalism theory examines microtargeting in a corporate context, one can also use it to draw meaningful comparisons in the context of political campaigning. Just as targeted messaging strategies existed in the marketing industry before the age of Big Data, they also existed in the world of elections. From direct mail initiatives (Rosenbaum 1997, 209) to the "explosive growth of opinion polling" from 1930 to 1950 (Fulda

2011, 1), political parties, candidates, and campaigns have an extensive history of using targeted messaging to influence turnout among specific parts of the electorate (Bodó et al. 2017, 3). Two notable 1960s examples of such targeting include President Richard Nixon's "Southern Strategy," the tailoring of messaging toward white southern voters opposed to racial desegregation (Brown 2004, 191), as well as President John F. Kennedy's careful usage of private polling to pinpoint popular issues and change voter perceptions of his personal attributes (Jacobs et al. 1994, 527). While certainly innovative for their time, these targeting strategies are not remotely in the same playing field as the game-changing techniques of the 2000s.

Just as Google shifted the paradigm for corporate marketing, President Barack Obama's first presidential campaign transformed electoral strategy with the launch of political microtargeting. The 2008 U.S. presidential race was "the first major election that perfected" political microtargeting models (Bodó et al. 2017, 3). In *Obama and the Power of Social Media and Technology,* Hughes et al, characterize Obama's remarkable journey from a little-known freshman senator in 2007, to beating out frontrunner U.S. Sen. Hillary Clinton for the Democratic nomination and ultimately becoming the first Black U.S. president in just 2 years (Hughes et al. 2010, 1).

One critical factor in Obama's success was his Chicago-based campaign team's top-priority focus on Big Data and digital outreach. While data-based targeting strategies had been used in political campaigning before Obama (Blaemire 2018, 5-8), they almost always focused exclusively on fundraising (Dutta et al. 2008). Uniquely incorporating voter engagement as a focus of digital outreach—not fundraising alone—the Obama campaign achieved unprecedented success with microtargeting. For instance, Obama's website, My.BarackObama.com (MyBo) encouraged users to create profiles to sign up for events and build relationships with other supporters—online activity that could be tracked to pinpoint and mobilize the most active volunteers (Hughes et al. 2010, 2). Using a new tool called Neighbor-to-Neighbor, the Obama campaign also streamlined the process of matching volunteers to undecided voters based on shared identity traits recorded in internal databases, such as "age, geography, profession, language," and "military service" (Hughes et al. 2010, 16).

Collecting and capitalizing on Big Data in this manner, the Obama campaign was able to pioneer a groundbreaking segmentation approach in their circulation of emails, text messages, and social media posts. First, the campaign sent out 1 billion emails to a 13-million-member email list—emails that were not generic but rather had 8,000 to 10,000 unique messages personalized to voters' residences, issue preferences, and donation histories (Hughes et al. 2010, 2). 3 million mobile and SMS subscribers also received similarly personalized text messages (Hughes et al. 2010, 2). Moreover, maintaining a hugely influential presence on social media, the campaign also curated content to specific online communities, disseminating tailored messaging to sites such as AsianAve.com, MiGente.com, and BlackPlanet.com, as well as a disabled American social network, Disaboom (Hughes et al. 2010, 6).

Winning the election with a margin of nearly 200 electoral votes, an 8.5 million popular vote lead, and support from 66% of voters under 30 (Pew Research Center 2008), President Obama's sweeping victory in 2008 made one thing clear: campaigns couldn't afford to sit back and continue doing things the old way. Of course, candidates win elections based on far more than effective microtargeting efforts—just as marketers' profits hinge on a multitude of factors outside of Big Data. That said, in the same way that Zuboff describes a permanently altered marketing landscape post Google's Big Data ad strategy, the Obama campaign's impressive use of microtargeting transformed the strategic landscape for campaigning. This transformation is evidenced by the extensive use of microtargeting in post-2008 campaigns—both in the U.S. and abroad. Notable examples include races fueled by "factional politics" during the U.S. 2016 Republican Presidential Primary (Raynauld 2016, 16), the campaigns of India's two leading parties since 2014 (Kröger et al. 2024, 24), and U.K. Brexit-centered campaigns in 2016 (Ryabtsev 2020, 75-76). The Big Data shift in politics is also represented by the emergence of companies focused specifically on selling microtargeting services to politicians worldwide. The tactics of one of these companies, Cambridge Analytica, underscore how both surveillance capitalism and political microtargeting create what Zuboff calls prediction imperatives.

A Story of Political Microtargeting: Cambridge Analytica

In the spring of 2018, Cambridge Analytica made headlines worldwide after it was revealed that the company used the personal data of 200,000 Facebook users to create "detailed psychological profiles of up to 87 million" potential voters (Heawood 2018, 429). To achieve this feat, an academic associated with the company got an initial set of Facebook users to voluntarily complete a personality test through Amazon Mechanical Turk (Fruchter et al. 2018). What was unknown to these users, however, was that the test utilized a loophole in Facebook's API that granted Cambridge Analytica access to not just their personal data, but also that of their Facebook "friends" (Fruchter et al. 2018). Armed with an unprecedented Big Data set, the company then used a profiling tool called OCEAN that characterized millions of Facebook users based on their "Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism" (Heawood 2018, 429). After profiling users for these formative personality traits—traits known as the Big Five (McAdams 1992)—Cambridge Analytica could then create ads tailored to the datafied psyches of specific voters.

Appealing to deeply personal and often unspoken information—like individuals' worst fears and greatest hopes—such ads go beyond resonating well with voters' existing preferences. Just like corporate microtargeting, political microtargeting attempts to steer individuals toward predetermined thoughts and actions. However, instead of trying to preempt independent purchasing behavior like Zuboff's *corporate* prediction imperative, the *political* predictive imperative tries to undermine genuine democratic engagement through voter manipulation.

Working on political campaigns in 68 countries—including the campaigns of 2016 U.S. presidential candidates Donald Trump and Ted Cruz (Detrow 2018)—Cambridge Analytica conducted voter manipulation on an "industrial scale" (Cadwalladr 2020). In May of 2018, within a month of news breaking surrounding the OCEAN operation, the company filed for bankruptcy and announced the ceasing of its operations (Watkins 2018). It's unclear whether this decision stemmed from genuine financial hardship or a desire on behalf of company executives to conceal additional records of what transpired. Shortly after the filing, the Federal Trade Commision sued the company for "falsely claiming they did not collect any personally identifiable information" on the Facebook users who took the personality test (Federal Trade Commision 2019).

While Cambridge Analytica conducted political microtargeting on a historic scale and broke U.S. law by lying about its data practices, the company's central objective—changing voter behavior through individualized ads—is no different than that of most political campaigns today. Since the Cambridge Analytica scandal, elections across the globe continue to be influenced by hyper-customized messaging that provides voters with individually tailored and potentially misleading narratives. Lowering the bar for candidates to create compelling policy agendas that appeal to broad coalitions, political microtargeting allows campaigns to pick and choose what stances to highlight to specific voters. This selectiveness not only makes it harder for voters to understand "the full picture" of a particular candidate; it further limits real deliberation by fragmenting voters and increasing political polarization, reducing productive discourse and driving unproductive echo chambers.

Critical Analysis of Political Microtargeting

This thesis is not alone in calling attention to the ethical concerns and democratic risks posed by political microtargeting. In *Algorithmic Elections*, Sarah Bender rings alarm bells that echo this thesis' characterization of the political prediction imperative. According to Bender, AI isn't just automating voting roll maintenance, signature verifications, or redistricting efforts; it's transforming political advertising by providing ways to produce digital content that is most likely to influence voter behavior in desired ways (Bender 2022, 513).

Other authors raise new concerns about microtargeting's ethical implications. In *Social media ethics: A Rawlsian approach to hypertargeting and psychometrics in political and commercial campaigns*, Morten Bay argues that "hypertargeting" in political campaigns "creates inequities in the access to information essential for voting (Bay 2019, 1). Using a Rawlsian ethical framework rooted in the idea of a "veil of ignorance"—the notion that if placed in the hypothetical "original position," that is, an impartial standpoint where personal circumstances are completely unknown, individuals would prioritize the well-being of society's most disadvantaged (Rawls 2005)—Bay argues that microtargeting denies some a "primary good" that would be prioritized under the "veil": information access (Bay 2019, 4). Bay also contends that the "uneven information access" caused by political microtargeting "can be expressed in power relations," where some are able to more effectively engage in democratic discourse over others (Bay 2019, 10).

Similarly, in "Microtargeting people as a mere means," Jongepier and Wieland critique the "trickery" that microtargeting encourages: presenting disingenuous or contradictory messages to groups that inhibit voters from knowing candidates' actual beliefs (Jongepier and Wieland 2022, 169-170). Moreover, the authors also suggest that microtargeting is morally impermissible in the way that it disregards voters' "capacity as democratic agents" for the sake of "getting their vote." (Jongepier and Wieland 2022, 170-171). According to them, if a candidate genuinely cared about "people's consent to giving their vote," they would avoid microtargeting tactics that spread misleading information about their platform or outright lies (Jongepier and Wieland 2022, 170-171). Ultimately, while abstract arguments surrounding ethical concerns and democratic risks are certainly helpful for advocacy on the harms of political microtargeting, they are less actionable than arguments grounded on concrete evidence.

What's Missing from the Political Microtargeting Conversation

Despite some discussion of the problems political microtargeting poses to democratic societies, literature on this topic hardly contextualizes them with in-depth analysis of particular case studies. Just two academic papers were found that attempted something similar to this task. The first, Islam Mohamed Hegazy's "The effect of political neuromarketing 2.0 on election outcomes: The case of Trump's presidential campaign 2016," sets out to answer how "political neuromarketing 2.0 tools and techniques" contributed to Trump's 2016 election win (Hegazy 2021, 236). Hegazy uses other authors to trace the Trump campaign's strategic usage of Big Five psychographic profiling (Hegazy 2021, 244), Facebook Lookalike Audience ads—ads that clone audiences with shared attributes to those that have already been microtargeted (Hegazy 2021, 246),— and "Dark posts"—nonpublic paid posts that are strictly shown to some audiences over others (Hegazy 2021, 246). While Hegazy certainly provides a helpful overview of the Trump campaign's microtargeting tactics, his paper lacks original analysis of the ads themselves. Failing to characterize specific instances of voter manipulation, Hegazy's research leaves questions related to the ads' effectiveness, along with their ethical implications, unanswered.

The only other paper found to take a case study approach, "Quantifying the potential persuasive returns to political microtargeting" by Ben M Tappin et al., attempts to answer questions related to the effectiveness of microtargeted ads compared to simpler messaging strategies through artificial survey experiments. Tappin at al., undertake two studies. In the first study, they focus on persuasion in the context of selected policy issues: the U.S. Citizenship Act, a bill to establish paths to citizenship for certain undocumented individuals, and universal basic income (UBI). Showing ads to three groups—a microtargeted group with ads tailored to demographic characteristics including age, gender, and partisanship, a group shown the same top-performing ad, and a group shown random ads—the researchers find that for ads on the U.S. Citizenship Act, microtargeted ads are 70% more persuasive than the top-performing ad, and over 200% more persuasive than ads shown at random (Tappin et al. 2023, 4). For UBI, the group finds smaller and non-statistically significant effects on persuasion power for microtargeting: microtargeted UBI ads are 40% more effective than the single top-performing ad and only 25% more effective than the random ads (Tappin et al. 2023, 4). From these policy-specific experiments, the researchers conclude that microtargeting's persuasive advantage depends on the context in which it is used.

In their second study, Tappin et al. focus on ads' ability to invoke broader ideological shifts rather than ones on isolated policies. Using survey data that recorded participants' attitudes toward a variety of issues, such as immigration, climate change, and minimum wage, the researchers use machine learning to predict what type of messaging would be most salient for individuals based on demographic and psychological traits, including age, gender, political affiliation, and cognitive reflection. Randomizing participants into three groups, participants are shown ads using the "naive strategy," a random ad, the "single-best-message strategy," the ad with the highest average persuasiveness across the population, and the "microtargeting strategy," ads predicted to be the most persuasive by the machine learning model (Tappin et al. 2023, 2). While the microtargeting strategy is found to have a statistically significant persuasive impact approximately double that of the naive strategy, it is a non-statistically significant 3 percentage points less persuasive than the single-best-message strategy. The researchers conclude that despite offering a clear advantage over the naive strategy, microtargeting may be just as effective, or even less effective, than broad, widely appealing messaging in some cases.

To summarize their conclusions, one of the paper's authors, Professor at the MIT Sloan School of Management David G. Rand, states that he and his colleagues did not "find much evidence that microtargeting works" and that just as much "persuasive advantage from targeting based on just one attribute" was found when "targeting on more attributes" (MIT News Office 2023). This summary is inaccurate: while the study acknowledges that targeting using one attribute alone can be effective in some scenarios (i.e., the single-best-message strategy), it also evidences that microtargeting can be far more effective in others. Why then, does Rand make this claim? Notably, Rand and at least one additional author of the paper, Adam J. Berinsky, have conducted work funded by Meta and Google (MIT News Office 2023), companies likely to oppose regulations on political microtargeting because they jeopardize profits.

Tappin et al's attempt to evaluate the effectiveness of political microtargeting, the first of its kind, has several shortcomings which undermine its conclusions and render the paper insufficient as a basis for regulatory decisions. Firstly, both studies are undertaken in a static context that does not account for the dynamic, constant flow of competing messaging by campaigns. Rather than speaking to the effectiveness of repeated microtargeting efforts over a period of time, the results reflect immediate responses that fail to capture how campaigns work to shift public opinion over the course of an election cycle. Similarly, the results do not translate the real-world impacts of political ads well because they rely on the self-reported attitudes of participants, not actual voting behavior or engagement. Unlike recorded votes or organizing activities, self-reported attitudes are susceptible to social desirability bias—the tendency of individuals to give responses they think will be viewed favorably by others (Grimm 2010)—and participants' lack of follow-through. Additionally, the study generalizes its findings across campaign contexts, without considering how microtargeting's effectiveness might be dependent on these contexts' nuances. For instance, it is likely that microtargeting is more effective in competitive swing districts, where voter preferences are wide-ranging, than in safer districts, where voter preferences are more uniform. In swing districts, granular messaging can capitalize on differences between specific voter subgroups to sway voters. In safe districts, the homogeneous nature of the electorate means that less tailored, more broad messaging can effectively address collective preferences.

Overall, literature on political microtargeting, consisting of sufficient theoretical analysis but lacking in reliable case-studies, fails to characterize the issue as one that is relatable and pressing. Firstly, the conversation does indeed contextualize political microtargeting as a practice fraught with problematic ethical implications for democracy and personal autonomy—but it does so relying on complex social theory and abstract argumentation rather than tangible examples capable of resonating with the broader public. Of the only two attempts in the literature that examine actual case studies of political microtargeting, one simply catalogs overarching microtargeting strategies while the other problematically suggests that microtargeting might not be as influential as some expect it to be. In addition to research that more accurately evaluates the real-world effectiveness of political microtargeting, research unpacking political microtargeting in practice is urgently needed. Until advocates start raising awareness toward real-life microtargeting examples—examples that wake the public up to political microtargeting's potential societal harms—policy makers will not face the political pressure necessary for meaningful reform.

DATA & METHODS

The Facebook Ad Library

This thesis explores microtargeting during the 2024 U.S. Presidential election by analyzing campaign ads stored in Meta's publicly available Ad Library, a website that allows users to access information on active and inactive ads posted to Facebook and Instagram. Facebook (now Meta) launched the Ad Library in March 2019, a year after the Cambridge Analytica scandal sparked global outcry over foreign election interference via online advertising. The Ad Library's current "about" page echoes this context, with the company stating that "transparency is a priority for us to help prevent interference in elections" (Facebook 2025). In a 2019 press release announcing the Library's creation, Facebook also asserted that "we know we can't protect elections alone," presenting the resource as a collaborative tool for enhancing transparency and accountability in the democratic process (Facebook 2019). Using the Library with the same aim, this thesis examines microtargeting tactics' broader implications for U.S. election integrity and democratic discourse.

Though the Ad Library was supposedly created to increase transparency in online advertising, it's worth noting that the site is rather limited in the U.S.-based information it discloses. Library users are provided estimates—not precise figures—of impressions for every ad (the total number of times an ad is shown on users' screens) as well as three ad targeting metrics: age ranges (18-24, 25-34, 34-44, 45-54, 55-64, and 65+), gender (restricted to male, female, or unknown), and region (country or U.S. state). It is largely unknown what ad targeting metrics are *not* disclosed by the Library. According to a publicly available guide in Meta's "Business Help Center," the company allows ad buyers to use extremely precise geographic data—singling out individuals within as tight as a 1-mile radius of a particular address.

Beyond user age, gender, and location, Meta touts other forms of "detailed targeting" on their help center page, including user "interests," "behaviors," and "connection" (Meta Platforms, Inc. 2025a). Additionally, once ads start running, Meta shares that its "system will learn who is engaging" and "narrow" audiences to "reach more of the right people" (Meta Platforms, Inc. 2025b). This automated "learning" phase, described in vague terms only, underscores a major oversight gap in microtargeting at large. The public can only speculate how Meta—and companies doing similar work—reinterpret, reclassify, and optimize personal data, leaving questions unanswered about whether these processes favor certain demographics, reinforce stereotypes, or discriminate against vulnerable groups.

In addition to placing limits on which targeting metrics are visible to users, the Library website is also restrictive in its navigational functionality. Users can only search ads by keyword, advertiser account, and category ("all ads," "issues, elections, or politics," "housing," "employment," and "financial products and services,"). Further filters on searches that result from the previous three criteria are limited to delivery by region, language, platform (e.g., Instagram, Facebook, or Threads), media type (e.g., images with text or videos), active status (active, inactive, or both), impressions by date, ad buyer spending ranges, the presence of specific disclaimers (labels that say who paid for an ad), and estimated audience ranges. Importantly, users cannot search or filter ads based on demographic distribution or other Meta targeting parameters (user "interests," "behaviors," and "connection").

Once a specific search is completed, Library users have the option of downloading a spreadsheet (CSV) containing information on all ads meeting the specified criteria. Each row in the table captures a specific ad, with columns in the table denoting the variables associated with that ad (see Appendix Table I for variable definitions). Crucially, the Library doesn't permit a downloading free-for-all: the site blocks users from downloading more than 3 spreadsheets (CSVs) a day, making large-scale or multi-query analysis slow and difficult.

Technically, users have two means of Library navigation: directly through the Library's website—the approach used in this study—or through the Meta Ad Library API (application programming interface). Though the API offers slightly more customizable search options, it requires programming expertise and adherence to Meta's platform policies about storing, redistributing, and publishing accessed data. Crucially, the API also does not provide any relevant ad data beyond what is provided by the website-downloaded CSVs (Meta Platforms, Inc. 2025c), leaving no compelling case for its usage for a project of this nature.

Immediate Data Limitations in CSV Exports

Upon downloading and examining a CSV from the Library, several challenges became apparent. First, I discovered that the `ad_creative_bodies` variable—the primary variable that encapsulates ad content—only includes text from the ads' captions, i.e., the short bodies of text that appear above or alongside a graphic, image, or video in a post. Any text embedded in graphics or images, as well as video transcript text, is excluded from this variable. Technically speaking, one could manually screenshot or screen record the videos and graphics included within specific ads to retrieve this information. However, given the vast amount of ads being examined, an automated approach using OCR (optical character recognition) technology would be necessary to make this strategy practically feasible. Moreover, Meta's Terms of Service explicitly prohibits automated data scraping without prior permission (Meta Platforms, Inc. 2025d), making such an approach—if even possible in the face of Meta safeguards—legally and ethically problematic.

Though the `ad_creative_bodies` variable tells an incomplete story about an ad's messaging approach, it still offers valuable insight into the tone, framing, and core messaging. Captions tend to summarize or directly repeat the content within graphics, images, and videos. Usually serving to emphasize an ad's central point, they are a useful—if imperfect—proxy for analyzing patterns in online advertising. While they do not, for instance, capture visual imagery designed to inflict a specific emotional response, they nonetheless reflect the verbal choices campaigns and outside groups make when reaching out to targeted audiences.

Another limitation that was quickly apparent from the original CSVs was that some ads were inexplicably missing demographic data. These ads always had blank inputs in both the `demographic_distribution` *and* `delivery_by_region` columns, with inputs for all other variables provided. It is unknown why this data is missing for some ads over others. I provide exact figures and further analysis of missing demographic data by group in the Findings & Analysis section (see Table III).

Conducting The Search

The data spans 3 ad cohorts that resulted from 4 specific searches. First, in order to investigate political microtargeting at the official campaign level, I collected ads from two specific advertisers: the blue-check verified Meta profiles of Kamala Harris and Donald Trump. For each, I selected "United States" as the region, "Issues, elections, or politics" as the category, and then input their verified profiles into the search bar. I refer to these first two groups of ads as part of the "Official" ad cohort, and their original respective file names are `Official_Trump.csv` and `Official_Harris.csv`. Second, to examine political microtargeting for the 2024 presidential campaign across *all* political advertisers—PACs and Super PACs, political parties, advocacy and issue-based nonprofits, unions and trade associations, political media outlets, state and local campaign groups, and so on—I collected ads using the keywords "Donald Trump" and "Kamala Harris," a search option that returns ads containing either or both words. I refer to these second two ad groups as part of the "General" ad cohort, and their original respective file names are `General_Trump.csv` and `General_Harris.csv`.

Other than varying advertiser, as in the case of the first set of searches, and keyword, as in the case of the second set of searches, I followed an identical search procedure. For all initial searches, I selected "United States" as the region and "Issues, elections, or politics" as the category. Next, for filters, I only removed the requirement that ads be "active" and selected a specific range for impressions by date: Sunday, September 1st, 2024 to Tuesday, November 5th,

2024. This period marks the busiest stretch of activity for Meta advertising by presidential campaigns and outside organizations, with almost half of all Meta ad spending occurring during this time (Institute for Democracy, Journalism & Citizenship 2024, 12). Additionally, considering former President Joe Biden stepped out of the race on July 21st, 2024 (Miller et al.2024), this window was ideal for capturing a period of relatively stable campaign messaging focused on the two ultimate candidates.

A Painstaking Data Cleaning Process

Inconveniently, CSVs downloaded directly from the Library clump together this study's dependent variables—impressions by age, gender, and state—into just two variable columns: first, `demographic_distribution`, which uses commas to separate the percent of impressions shown to 18 cohorts of different gender and age range combinations (e.g., females aged 18-24, males aged 18-24, individuals of unknown gender aged 18-24, females aged 25-34, etc.) and second, `delivery_by_region`, which uses commas to separate the percent of impressions shown to voters in various states. With the aim of using Stata to conduct statistical analysis comparing numerical demographic data with varying messaging, I sought to separate these two variable columns into 72 new ones: 18 columns presenting impression percents for each gender and age range combination, 50 for each state, one for swing states, one for non-swing states, and one for impressions coming from "unknown" regions.

Though a somewhat simple data reorganization task, this hurdle posed quite a challenge for me, a self-identifying "beginner" in data science and programming languages. After dividing the task into two sections—one section for the age/gender columns and one for the state columns—I outlined steps for each. Next, after choosing to start with the `Official_Trump.csv`, I drafted several versions of unsuccessful Stata code on my own, eventually resorting to prompting GPT-4o with detailed instructions and a modified sample table representing my desired outcome. GPT-4o was a poor helper. For each section, I ran different versions of AI-generated code dozens of times in Stata, reporting Stata error messages back into GPT-4o and asking the system to generate new code with these errors in-mind.

On its own, this iterative process proved ineffective, leading me to reach out to a UChicago professor of Quantitative Methods for Public Policy for direct edits. While this professor was able to quickly generate successful code for the age/gender section of the task on his first attempt, he was only able to make partial progress for the state section: the code he provided ran without errors but frequently filled the state columns with numbers well above 1—a sign the underlying data in the table was malformed (i.e., inconsistently formatted across different ads). Eventually, after 108 additional rounds of GPT-4o editing—a number I found by tracking my GPT chat history and counting the number of times I prompted the system—GPT-4o finally produced successful code for the state columns. Finally, with a template for both sections of the data reorganization task, I adjusted and ran the code for all the CSVs, producing four new datasets: `official_trump_ads_clean.csv`, `official_harris_ads_clean.csv`, `general_trump_ads_clean.csv`, and `general_harris_ads_clean.csv`.

Thankfully, the other primary cleanup task—removing Official ads from the General cohorts, as well as duplicates within the General cohorts themselves—proved easier to implement in Stata. Starting with the General Trump ads, I successfully merged `general_trump_ads_clean.csv` with `official_trump_ads_clean.csv` on the first attempt. Next, I created code that flagged overlapping instances of the `ad_archive_id` variable (i.e., Official Trump ads included in the General Trump cohort). At this point, I noticed something peculiar: the number of duplicate ads was very close—but not exactly equivalent—to the number of ads recorded in the Official Trump dataset. When conducting the General search, I had expected that using the keywords "Donald" and "Trump" would capture all ads from the Official cohort. However, 201 Official Trump ads were absent from the General dataset, an indication that there may be slight metadata inconsistencies in how the Ad Library indexes ads across keyword-based and account-based searches (see Appendix Table II for exact figures on ad counts for CSV mergers during data cleanup). With this finding in mind, I deleted both duplicate and unique Official Trump ads from the merged dataset and downloaded a new CSV containing only the remaining ads: `general_trump_ads_clean_Tdeduped.csv`.

Next, I repeated this process to identify Official Harris ads potentially present in the General Trump search, merging `general_trump_ads_clean_Tdeduped.csv` and `official_harris_ads_clean.csv`. I found that a majority of the ads in the Official Harris dataset were unique, which makes sense considering the keywords "Donald" and "Trump" would not be included in all Official Harris ads. After deleting duplicate and unique Official Harris ads, I downloaded a new CSV for the General Trump cohort: `general_trump_ads_clean_THdeduped.csv`.

I then repeated this entire process for the General Harris dataset, first merging `general_harris_ads_clean.csv` with `official_harris_ads_clean.csv`. Unlike before, this merged dataset had no unique Official Harris ads, indicating that the General Harris search captured all Official Harris ads. After deleting duplicates, I created a new dataset (`general_harris_ads_clean_Hdeduped.csv`) and merged it with `official_trump_ads_clean.csv`. The General Harris search with keywords "Kamala" and "Harris" only captured 703 Official Trump ads, indicating that these words were not commonly used across the Official Trump cohort. After deleting duplicates and unique Official Trump ads, I downloaded a new CSV for the General Harris cohort: `general_harris_ads_clean_THdeduped.csv`.

Finally, after removing duplicates of the Official ads from the General cohorts, I addressed the possibility that the same General ads appeared in both searches. To resolve this, I merged `general_harris_ads_clean_THdeduped.csv` with `general_trump_ads_clean_THdeduped.csv`, identified duplicates using the `ad_archive_id`, and retained just one copy of each ad to ensure all General ads were uniquely represented in a final new CSV: `general_combined_unique.csv`.

ChatGPT-Powered Sentiment Analysis

The final portion of the methodology involved conducting GPT-4o sentiment analysis through R code connected to the OpenAI API. I obtained a template for this code from this thesis' secondary reader, the Associate Director of Technology at the University of Chicago Forum for Digital Culture, and modified the code with task-specific prompts for label assignments. This work required establishing a personal OpenAI account with automatic billing set up, as all API calls incur a cost based on the number of tokens—units of text—involved in both the prompt and the model's response.

The first sentiment analysis task was aimed at completing a final cleanup step for the General ad cohorts: removing irrelevant ads—which I call "Donald Ducks." Intuitively, using the keywords "Donald," "Trump," "Kamala," and "Harris" in the Library's search function could result in the inclusion of political ads that use these words for irrelevant purposes (i.e., purposes other than communicating messages about candidates or issues in the 2024 U.S. presidential election). For instance, private companies not affiliated with campaigns may use nominee names to sell unofficial merch, and some ads may reference unrelated individuals, entities, or characters (e.g., Donald Duck). Though identifying Donald Ducks on an individual basis is quick and easy, manual removal across both General CSVs would require reviewing over 50,000 ads—a practically infeasible task for a single researcher.

To conduct this task with sentiment analysis, I first created two "test" CSVs: one with examples of relevant ads and one with examples of Donald Ducks. I selected these examples to test how accurately the prompt would classify ads across a wide spectrum of messaging approaches, with special attention placed on trickier scenarios (e.g., ads run by politically affiliated organizations promoting nonpartisan reporting that implicitly favored a candidate, ads referencing presidential nominees to frame issues or candidates further down the ballot, and ads for unofficial merch that indirectly praised a candidate solely for driving up sales). Then, using the obtained template, I modified the code with help from GPT-4o so that it imported and merged the example CSVs, instructed GPT-4o via the API to generate a binary relevance label based on each ad's caption (with "1" indicating relevant and "0" indicating irrelevant), inserted this output into a new `ad_relevancy` column, and exported the CSVs in their original formats. Testing several versions, I settled on the prompt below after it correctly labeled all ads in the example CSVs.

<u>Sentiment Analysis Prompt #1: Identifying Donald Ducks</u>

Return 1 if the ad expresses sentiment about Kamala Harris or Donald Trump as 2024 candidates, former leaders, or political figures, including their policies or election issues. Include other races that reference them politically. Return 0 if their names are used for unrelated people, selling unofficial merch or goods, or non-political content. Reply only with 0 or 1.

Several challenges became apparent during the first sentiment analysis test. First, with output time and API costs directly correlated with the number of tokens involved in a task, I had to craft prompts that were as detailed as possible—so as to ensure accurate labels—but also as short as possible—so as to minimize time and money spent. This task was further complicated by the need for GPT-4o to output responses in a perfectly consistent format. If the model deviated even slightly—adding punctuation or reordering the response—R would fail to successfully parse and assign responses to their appropriate row and column in the dataset. GPT-4o proved to be a useful tool in finding a balance between specificity and brevity, helping flag unnecessary or repetitive words and phrases. Additionally, as token count includes not just the prompt length but also the input text GPT-4o reads, I needed a way to prevent the model from processing the same ad captions repeated several times within a single entry. To address this issue, I implemented a truncation strategy that capped the caption text GPT-4o read at 500-characters, ensuring complete captions were reviewed while reducing wasteful token usage.

With a finalized prompt and tested procedure that addressed the aforementioned challenges, I ran the code on `general_combined_unique.csv`. After approximately 10-hours of running R uninterrupted, I successfully generated a new version of the file with the relevancy column included. I then removed all Donald Ducks—ads with a relevancy label of "0"—resulting in a final, cleaned CSV of unique and relevant General ads: `general_combined_unique_relevant.csv`.

Next, I repeated this process to develop a prompt and functioning code for the second sentiment analysis task: labeling ads based on policy theme and emotional tone. This time, instead of creating a single column, I generated two—one for the ad's primary policy theme (among a choice of 15) and one for the ad's primary tone (among a choice of 5). I also modified the code to merge `general_combined_unique_relevant.csv` with `official_trump_ads_clean.csv` as well as `official_harris_ads_clean.csv`, conducting sentiment analysis for the first time on the Official ads cohorts. With nearly twice the ads to process and a more complex labeling task, this round required approximately 36 hours of running R uninterrupted before new files for the 3 cohorts—`general_combined_unique_relevant_themetone.csv`, `official_trump_ads_clean_themetone.csv`, and `official_harris_ads_clean_themetone.csv`—could be downloaded. After testing three rounds of prompts on sample datasets that included all pre-specified themes and tones, I settled on the finalized prompt below.

> **<u>Sentiment Analysis Prompt #2: Ad Policy Theme and Emotional Tone</u>**
>
> For the given ad text, return the theme and emotional tone using the categories and loose definitions provided. Choose only one of each, selecting the dominant theme/tone if multiple are present.
>
> Themes (pick one):
> Economy – jobs, taxes, inflation, spending, prices
> Healthcare – insurance, Medicare, Medicaid, pharma
> Education – schools, childcare, curriculum
> Immigration – border, migrants, asylum, immigrants, deportation, visas
> Climate – environment, energy, clean tech, gas, climate change
> Crime – policing, safety, illegal drugs
> Guns – gun rights, control, 2nd Amendment
> Military – armed forces, vets, defense
> Abortion – rights, bans, dobbs decision
> LGBTQ+ – rights, gender, trans issues
> Foreign Policy – war, trade, global security, Israel, Palestine, China, Russia
> Democracy – voting, fraud, Jan 6, misinformation
> Welfare – food stamps, disability, social security, benefits

Candidate – candidate traits, leadership, trust
Other – if none apply

Tones (pick one):
Fearful – tries to scare
Angry – expresses outrage, blame
Hopeful – optimistic about change, pride, progress
Patriotic – appeals to love of country, tradition, unity
Neutral – dry, factual, unemotional

Format: Theme: ___; Tone: ___

Last, I adapted this process to complete the third and final sentiment analysis task: labeling ads based on candidate alignment. I adjusted the code to create two additional variable columns: `pro_Trump` and `pro_Harris`, with outputs of "1" indicating support for the candidate, "0" indicating opposition, and "3" marking ads that were neutral or did not mention the candidate. Similar to the theme and tone assignment, this labeling round took approximately 36 hours. After three rounds of prompt testing on curated sample datasets containing all pre-specified candidate alignment combinations, I ran the analysis with the finalized prompt below.

Sentiment Analysis Prompt #3: Ad Candidate Alignment

For the given ad text, assign a score for each of the following candidates:
Donald Trump
Kamala Harris

Use this scale:
1 if the ad expresses support for the candidate (positive tone or promotional language)
0 if the ad expresses opposition to the candidate (critical or negative tone)
3 if the ad is neutral toward the candidate or does not mention them

Return your answer in the following format:
pro_Trump: [0, 1, or 3]; pro_Harris: [0, 1, or 3]

FINDINGS & ANALYSIS

Ad Counts, Estimated Impressions, and Missing Demographic Data

Table I: Ad Count Across Datasets

Dataset Name	Number of Ads	% of Total Ads
`official_trump_ads_clean.csv`	17,235	14.28%
`official_harris_ads_clean.csv`	51,988	43.07%
`general_combined_unique_relevant.csv`	51,489	42.65%
Total Ads	120,712	

Table II: Impression Count Across Datasets

Dataset Name	Estimated Impressions	% of Total Estimated Impressions
`official_trump_ads_clean.csv`	331,016,428	6.76%
`official_harris_ads_clean.csv`	2,292,432,718	46.78%
`general_combined_unique_relevant.csv`	2,276,761,758	46.46%
Total Estimated Impressions	4,900,210,904	

Table III: Missing Demographic Data Across Datasets

Dataset Name	Number of Ads Missing Demographic Data	% of Ads Missing Demographic Data

official_trump_ads_clean.csv	4,023	23.34%
official_harris_ads_clean.csv	5,408	10.40%
general_combined_unique_relevant.csv	6,274	12.18%
Total Ads Missing Demographic Data	15,705	
% of Total Ads Missing Demographic Data	13.01%	

Table I demonstrates that after cleaning and sentiment analysis, a total of 120,712 unique Meta ads were examined for statistical analysis: 17,235 from the Official Trump cohort, 51,988 from the Official Harris cohort, and 51,489 from the combined General cohorts. While the General and Official Harris ads each comprise roughly 43% of the dataset (42.65% and 43.07%, respectively), Official Trump ads account for just 14.28%—less than one-third the size of the other groups. As shown in Table II, ad impressions—the number of times an ad appeared on Meta users' screens, not the number of unique individuals reached—roughly mirror the distribution of ads, though with even greater disparity between the groups: Official Trump ads accounted for approximately 331,016,428 impressions (6.76% of total impressions), compared to 2,292,432,718 impressions for Official Harris ads (46.78%) and 2,276,761,758 impressions for General ads (46.46%). It is important to note that these values are estimates rather than exact figures, as Meta reports only impression ranges for each ad. Total impressions were approximated by summing the midpoint of each ad's range.

These results indicate a substantial decrease in both the volume and reach of Official Trump ads compared to those from the Harris campaign and outside groups, aligning with 2024 reporting that the Trump campaign invested significantly less in online advertising (Goldmacher and Nehamas 2024). This consistency reinforces confidence that the search methods effectively captured relevant ads across the cohorts. Moreover, the relatively limited presence of Official Trump ads on Meta platforms suggests the campaign's strategic shift away from apps like Facebook and Instagram in favor of other voter outreach channels.

Strategy evaluations like that one, however, must contend with data quality concerns that arose. As demonstrated in Table III, 13.01% of all ads—23.34% of Official Trump ads, 10.40% of Official Harris ads, and 12.18% of General ads—were missing demographic data. This missing data should not be interpreted as evidence that ads were not targeted by advertisers; instead, the gaps are more likely an indication of data logging failures leading to incomplete records, an issue flagged by other researchers (Rosenberg 2019). Notably, the Official Trump ads are missing demographic information at more than twice the rate as the other datasets. This discrepancy disproportionately limits our ability to fully understand the Trump campaign's demographic targeting strategy, potentially obscuring important patterns in how certain audience groups were or were not reached.

Though missing demographic data inhibits us from having a complete view into how ad messaging varied by demographic traits in the 2024 election, the large sample size and high proportion of ads with targeting criteria (over 75% for each dataset) still provide a sufficient foundation for meaningful and generalizable analysis. That said, the presence of incomplete ad records—paired with the fact that existing data inputs were often malformed, as evidenced by the challenges of the cleanup process—should raise eyebrows. Given Meta's vast resources and engineering talent, one would expect such errors to be easily addressable. One possible explanation for Meta's failure to standardize demographic ad data is that such inconsistencies discourage independent research, making studies like this one unnecessarily burdensome and time-consuming.

So, Who Was Targeted Overall?

Table IV: Estimated Impressions by State Across Datasets

Dataset Name	State Type	Estimated Impressions	% Share of Impressions
official_trump_ads_clean.csv	Swing State	246,192,238	74.37%
	Non-Swing State	82,814,698	25.02%
official_harris_ads_clean.csv	Swing State	3,816,430,363	90.91%
	Non-Swing State	381,516,264	9.09%
general_combined_unique_relevant.csv	Swing State	1,574,045,105	69.23%
	Non-Swing State	699,582,596	30.77%

Table V: Estimated Impressions by Age Group Across Datasets

Dataset Name	Age Group (Years Old)	Estimated Impressions	% Share of Impressions
official_trump_ads_clean.csv	18-24	19,153,217	5.82%
	25-34	46,561,175	14.15%
	35-44	58,132,329	17.67%
	45-54	60,759,041	18.47%
	55-64	72,652,209	22.08%
	65+	71,748,968	21.81%
official_harris_ads_clean.csv	18-24	397,008,640	17.34%
	25-34	602,565,205	26.32%
	35-44	465,236,377	20.32%
	45-54	309,545,045	13.52%

	55-64	252,549,623	11.03%
	65+	262,826,531	11.48%
general_combined_unique_relevant.csv	18-24	253,523,609	11.15%
	25-34	471,534,035	20.74%
	35-44	441,350,983	19.41%
	45-54	342,973,219	15.08%
	55-64	345,345,423	15.19%
	65+	418,900,476	18.42%

Table VI: Estimated Impressions by Gender Across Datasets

Dataset Name	Gender	Estimated Impressions	% Share of Impressions
official_trump_ads_clean.csv	Female	161,481,348	49.08%
	Male	162,780,940	49.48%
	Unknown	4,744,651	1.44%
official_harris_ads_clean.csv	Female	1,339,733,454	58.51%
	Male	917,704,114	40.08%
	Unknown	32,293,854	1.41%
general_combined_unique_relevant.csv	Female	1,261,789,957	55.50%
	Male	960,897,390	42.26%
	Unknown	50,940,399	2.24%

As shown in Table IV, ads directed toward Meta users in swing states make up a majority of total impressions across all three cohorts, reflecting a consistent outreach strategy prioritizing politically competitive areas over less contested ones. The Harris campaign appears to have prioritized swing state social media outreach the most, with 90.01% of its total estimated impressions aimed at users in Arizona, Georgia, Michigan, Pennsylvania, Wisconsin, North

Carolina, and Nevada. While the Trump campaign allocated a relatively greater share of impressions to non-swing states—25.02% of impressions as opposed to Harris' 9.09%—it still delivered nearly five times fewer total impressions to battleground states due to substantial differences in overall ad volume. Ads from the General cohort were the least swing-focused, with 30.77% of impressions targeted toward non-swing states.

Despite sharing an emphasis on swing state outreach, the three ad cohorts diverged more notably in their targeting of specific age groups. As shown in Table V, Favoring older voters, the Trump campaign allocated almost two-thirds of its impressions toward users 45 and up, while the Harris campaign took the inverse approach, allocating almost two-thirds of its impressions toward users aged 18-44. General advertisers appear to have taken a more evenly distributed age strategy, though they too appear to have prioritized younger voters more than Trump's campaign, with over 51% of impressions going toward users aged 18-44.

Table VI demonstrates gender-based differences across the datasets too. While the Trump campaign split impressions almost evenly between male and female users, the Harris campaign's impressions skewed female, with 58.51% sent to women. Similarly, General ads targeted women 55.50% of the time, with men receiving just 42.26% of the total impression share. These disparities suggest that the Harris campaign and many outside groups viewed female voters as a pivotal voting block in the 2024 presidential race.

Variations in Ad Messaging by Swing vs. Non-Swing State

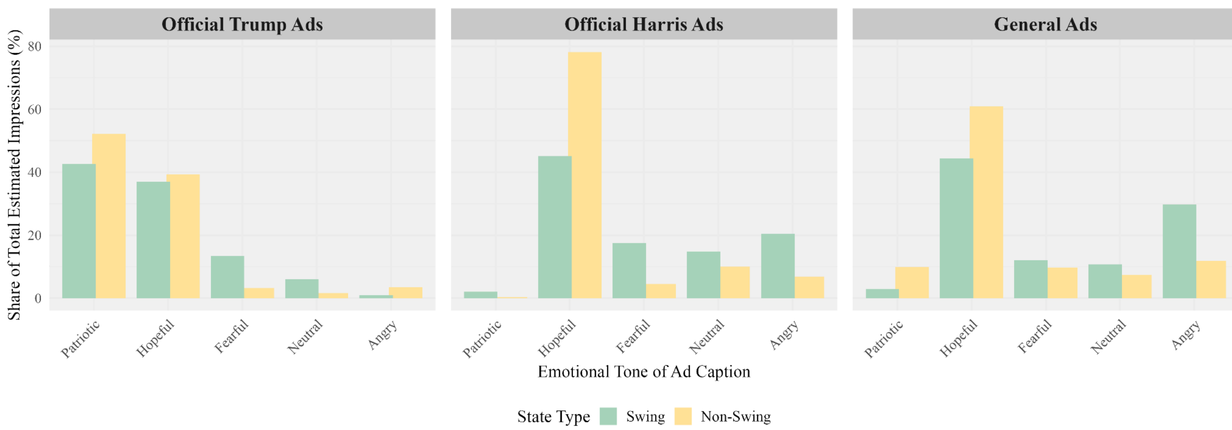
To visually demonstrate message variation across the three dependent variables—impressions by state type, age group, and gender group—I created a series of color-coded bar graphs depicting distributions of ad tone, theme, and candidate alignment. Each graph is divided into three panels: one for Official Trump ads, one for Official Harris ads, and one for General ads. The x-axes of the graphs display specific label categories (e.g., “fearful,” “economy,” or “pro-Trump”), and the y-axes indicate the percentage of total estimated impressions tied to each label. In other words, each bar reflects how much of an ad cohort’s impressions were associated with a particular tone, theme, or alignment label. Contingency tables that display exact proportions for all labels—the tables used to generate the graphs—are provided in the Appendix.

To assess the statistical significance of observed differences in messaging, I also conducted a Pearson’s Chi-Squared test to accompany the graphs. Each chi-squared test evaluated a null hypothesis that the label distributions (tone, theme, or alignment) were independent of the dependent variables (state type, age group, and gender). In that sense, the tests all asked a version of the following question: *is the proportion of ads with these messaging labels consistent across different demographic groups?* The tests provided two values for interpretation, an X-squared value and a p-value. The X-squared value quantifies how much the observed distribution deviates from what would be expected under independence, with the higher the X-squared value, the greater the divergence. The p-value indicates how likely it is that such a divergence could have occurred by chance, with p-values below 0.001 allowing us to reject the null hypothesis and point toward statistically significant evidence that messaging varied depending on the demographic variable in question.

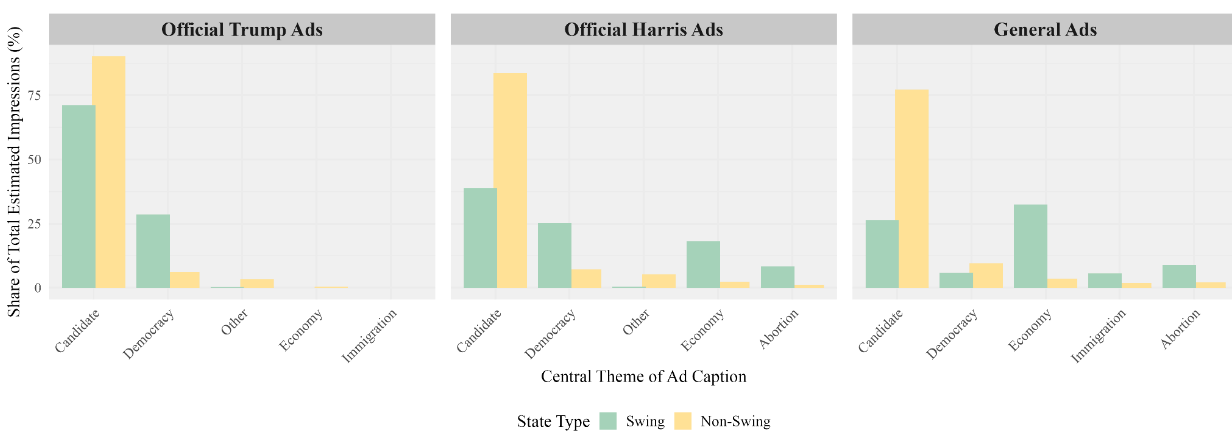
The first set of graphs (Graphs I-III) focus on state type. For each label, there are two bars: one in green, representing impressions in swing states, and one in yellow, representing impressions in non-swing states. The height of each bar communicates the share of a cohort’s impressions tied to that label. Though this design allows for easy side-by-side comparisons of ad messaging between state types, the graphs may initially feel dense or visually overwhelming.

Rather than trying to immediately grasp the meaning of every bar, I encourage you to start with a simpler task: look at whether the pairs of green and yellow bars are the same height. If messaging didn't vary significantly by state type, we'd expect each label's two bars to be nearly identical. As graph after graph reveal, however, the green and yellow bars consistently diverge, illustrating how the official campaigns and outside groups tailored content depending on the state users lived in.

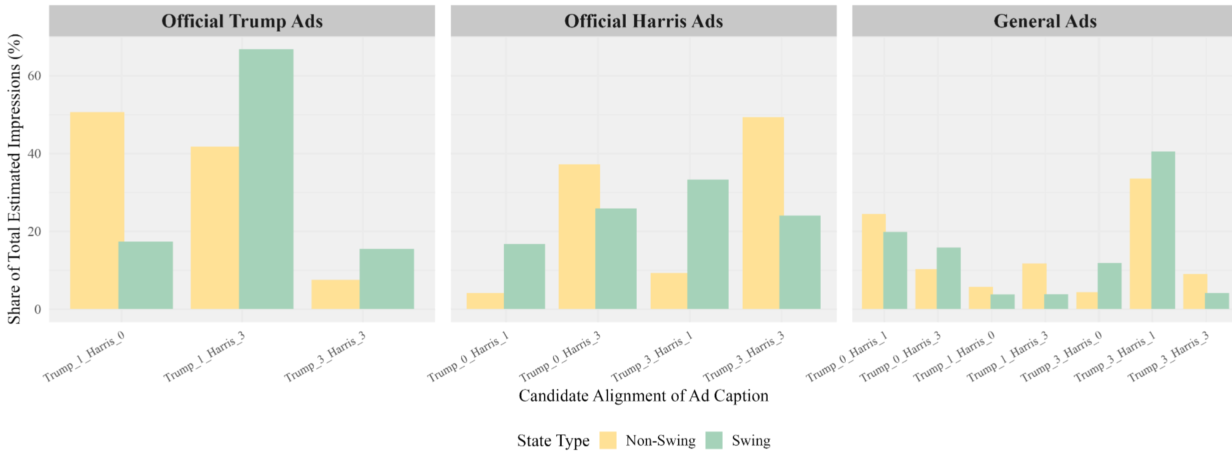
Graph I: Emotional Tones of Ad Captions Targeted to Swing vs. Non-Swing States



Graph II: Top 5 Central Themes of Ad Captions Targeted to Swing vs. Non-Swing States



Graph III: Candidate Alignments of Ad Captions Targeted to Swing vs. Non-Swing States



Graph I: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	12987661	155270145	150803918
P-value	2.2e-16	2.2e-16	2.2e-16

Graph II: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	24562332	358232292	629896368
P-value	2.2e-16	2.2e-16	2.2e-16

Graph III: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	37973259	195871370	132557596
P-value	2.2e-16	2.2e-16	2.2e-16

As demonstrated in Graph I and its corresponding chi-squared table, the emotional tone of ads in swing versus non-swing states varies in a statistically significant manner, with p-values so small they are reported in R's default lower threshold: 2.2e-16. Among Official Trump ads, the most stark difference occurs in the share of fearful messaging: while just 3.2% of impressions in non-swing states have captions with fearful labels, this share is 13.4% for impressions in swing states—a fourfold increase. Similarly, the Harris campaign relies on fearful messaging in only 4.6% of non-swing state impressions, almost four times less than in swing states (17.5%). Anger-based appeals also surge in swing states among both Official Harris and General impressions, with “angry” labels accounting for 20.5% and 29.8% of swing state impressions, respectively, compared to only 6.9% and 11.9% in non-swing states. These trends shine a light on a consistent shift toward more emotionally charged messaging in competitive battleground states across all three ad cohorts (see Appendix Table III for all tone by state type proportions).

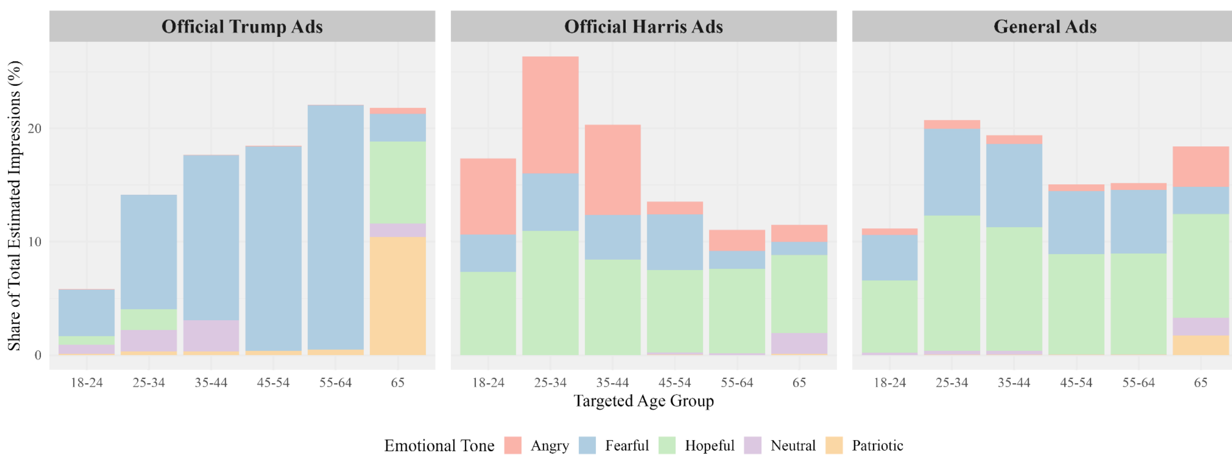
Graph II—supported by chi-squared tests also with p-values of 2.2e-16—demonstrates how swing state messaging is consistently more policy-focused than non-swing state messaging. Some of the most striking differences include the Official Trump ads linking 29.5% of swing state impressions to the theme of democracy, compared to just 6.06% in non-swing states; the Official Harris ads respectively devoting 25.3% and 18.17% of swing state impressions to democracy and economy themes, compared to just 7.2% and 2.3% in non-swing states; and the General ads devoting a whopping 32.5% of swing state impressions to economy-appeals, compared to only 3.5% in non-swing states. Across all cohorts, non-swing state impressions were far more focused on candidate-related messaging, reflecting a widespread microtargeting approach: campaigns and outside groups relied on name recognition and personal branding in less competitive states, likely to drive fundraising and maintain visibility, and reserved policy-specific appeals for persuadable voters in battlegrounds (see Appendix Table IV for all theme by state type proportions).

Lasty, Graph III and its corresponding chi-squared table demonstrate statistically significant variation in candidate alignment between swing and non-swing state impressions (all p = 2.2e-16). Among Official Trump impressions, messaging that combines pro-Trump and explicitly anti-Harris content drops from 50.6% in non-swing states to just 17.4% in swing states, while messaging that is solely pro-Trump rises from 41.8% in non-swing states to 66.8% in swing states. These differences highlight a strategic choice on behalf of the Trump campaign to avoid direct attacks on Harris in battlegrounds, perhaps to appear less negative and deliver a more straightforward, positive pro-Trump message. Interestingly, Official Harris ads take a different approach: messaging that is both explicitly pro-Harris and anti-Trump climbs from 4.0% in non-swing states to 16.70% in swing states, while messsages that are solely pro-Harris also rise from 9.3% to 33.3%. The Official Harris cohort devotes a majority of its non-swing state impressions (49.4%) toward messaging that doesn't mention either candidate directly and instead focuses on the election more broadly. These proportions suggest that the Harris campaign prioritized making clear contrasts and candidate-centered appeals in battleground states, while relying on broader, civic-focused appeals in less competitive regions. General ads followed a similar pattern, showing a consistent shift toward polarized messaging in swing states: messages that are both pro-Harris and Anti-Trump increase from 24.5% in non-swing states to 33.5% in swing states, while messages that are both pro-Trump and Anti-Harris rise from 9.1% to 15.9% (see Appendix Table V for all alignment by state proportions).

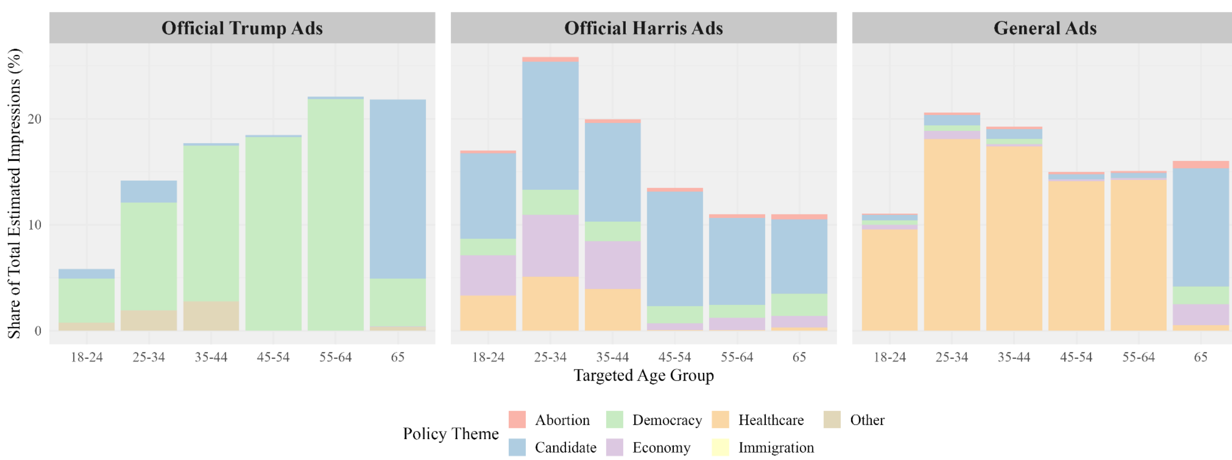
Variations in Ad Messaging by Age Group

The next set of graphs (Graphs IV-VI) explore message variation by age group. Unlike the previous side-by-side bar layout, these graphs use stacked bars. The height of each bar corresponds to the total share of impressions targeted to a given age group, and the colored segments within each bar correspond to the proportion of impressions in that group associated with specific tone, theme, or alignment labels. In other words, the taller the bar, the more ad impressions were targeted to that age group, and the larger the color segment within each bar, the greater the portion of those impressions tied to a specific label. If messaging was not age group-specific, we'd expect the color proportions within each bar to be consistent. It's clear, however, that major shifts in color composition occur, pointing toward meaningful variation in how messages were tailored according to users' ages.

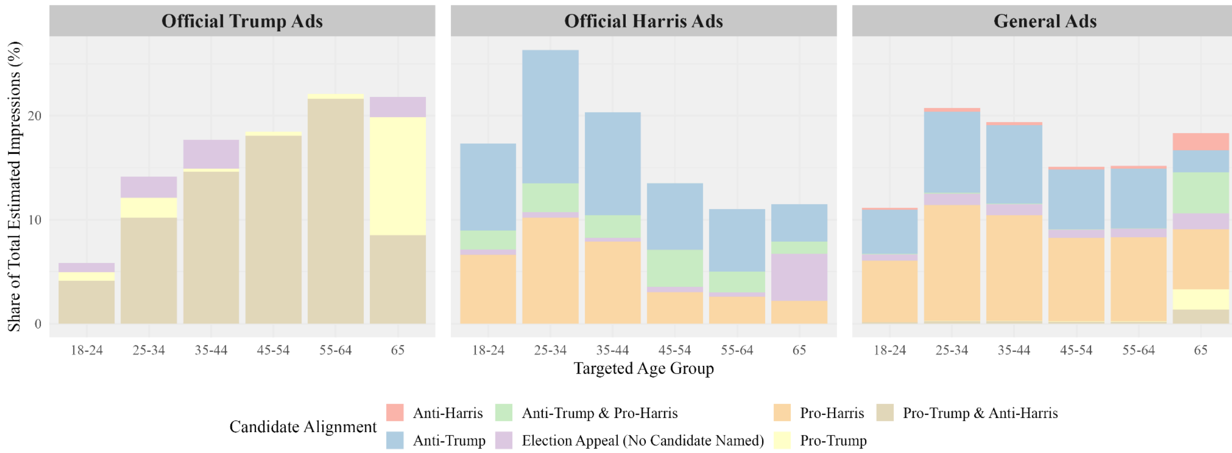
<u>Graph IV: Emotional Tones of Ad Captions Targeted to Different Age Groups</u>



<u>Graph V: Central Themes of Ad Captions Targeted to Different Age Groups</u>



Graph VI: Candidate Alignment of Ad Captions Targeted to Different Age Groups



Graph IV: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	3.9341e+11	1.7179e+10	9.0427e+10
P-value	2.2e-16	2.2e-16	2.2e-16

Graph V: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	3.7374e+11	3836914848	4.3408e+11
P-value	2.2e-16	2.2e-16	2.2e-16

Graph VI: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	3.215e+11	3096072535	1.8269e+10
P-value	2.2e-16	2.2e-16	2.2e-16

Across all three cohorts, Graph IV and its corresponding chi-squared tests (all p = 2.2e-16) highlight statistically significant differences in emotional tone between age groups.

Among the Official Trump ads, fearful tones dominate impressions toward non-elderly users, peaking at 97% for 45-64 year-olds. In contrast, users aged 65+ only receive fearful impressions 11.2% of the time, receiving patriotic impressions 47.8% of the time instead. Similar patterns emerge in the Official Harris cohort: anger-based messaging is prominent for users 44 and under (about 40%) but drops sharply among older users, who see a disproportionate increase in hopeful messaging instead. The General cohort somewhat abides by this trend as well, with users aged 18-64 receiving fearful messaging about 37% of the time, compared to 12.9% for elderly users. Interestingly, however, General impressions for 65+ year-olds are 19.5% anger-based, dropping to around 4% for middle and younger-aged users. Overall, these differences communicate a consistent trend of fear and urgency being tapped into among younger and middle-aged audiences, while senior audiences receive messaging more focused on tradition and optimism (See Appendix Table VI for all tone by age group proportions).

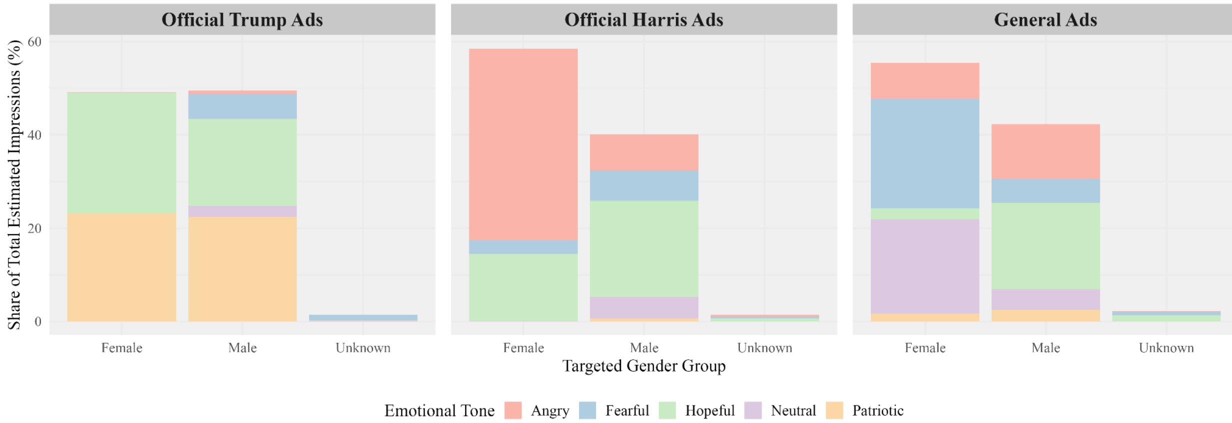
Graph V, also accompanied with evidence of statistically significant variation by age group across all cohorts (all $P = 2.2e-16$), relays stark patterns in how themes are strategically targeted by age. Official Trump impressions for users aged 18-64 are largely focused on democracy (about 71%) with about 14% of impressions emphasizing Trump himself. Middle-aged users receive an even greater proportion of democracy-focused impressions, peaking at 99% of impressions for those 55-64. Among voters aged 65 and up, the inverse occurs: candidate messaging jumps to 77.6%, and democracy messaging declines to 20.6%. The Harris campaign takes a somewhat similar approach: younger audiences receive a more diversified message mix than the other cohorts—focused primarily on candidates (about 45%), economy (about 22%), and healthcare (about 20%) themes. Candidate-appeals surge among middle-aged and elderly audiences, reaching 79.9% for 44-54 year-olds and 74.3% for 55-64 year-olds. Messaging on candidates, however, dips at 60.9% for voters aged 65+, who receive a renewed emphasis on democracy (18.2%). The General cohort further illustrates generational divides: users 45 and under see messaging dominated by the theme of healthcare (over 85% of impressions), with minimal messaging related to candidates or democracy. While healthcare is also the dominant theme for those aged 45-65 (over 93%), voters 65+ see a rise in candidate-based appeals (60.4%), as well as democracy (9.0%) and economy-centered (10.8%) messaging. Overall, these differences suggest a broader microtargeting trend of campaigns and outside groups sending younger voters messaging that is disproportionately policy-focused, while sending older voters messaging that is disproportionately candidate-focused (see Appendix Table VII for further theme by age group proportions).

Last among those depicting message variation by age group, Graph VI illustrates how candidate alignment changes dramatically depending on audience age, with statistically significant variation across all cohorts (all P = 2.2e-16). Among Official Trump ads, impressions to users under 65 are predominantly pro-Trump and anti-Harris, peaking at about 97% for those aged 45-64. Seniors, however, receive a sharp decline in anti-Harris messaging (only 39.1% pro-Trump and anti-Harris), instead being largely targeted with content that is solely pro-Trump (52.0%). Unlike the Trump campaign, which tends to maintain pro-Trump messaging across all age groups, the Harris campaign makes a different move: it sends non-senior voters explicitly anti-Trump (and only anti-Trump) messaging about 50% of the time, focusing less on solely pro-Harris content. Anti-Trump-only impressions among seniors voters drop to 31.3%, with impressions that are neutral toward both candidates—general appeals about the election—surging to almost 40% (as opposed to all other age groups, who receive this alignment combination 2.0–3.8% of the time). Finally, the General cohort reinforces shifts in alignment based on age, but in a different manner. First, senior voters are sent messaging that is both anti-Trump and pro-Harris 21.6% of the time, while all other age groups receive this specific alignment combination less than 1% of the time. Conversely, younger and middle aged voters are largely targeted with messaging that is solely pro-Harris—about 53% for users aged 18-64—compared to 31.3% for those 65 and up. Additionally, anti-Trump-only impressions make up about 38% of impressions across younger and middle-aged groups, whereas they drop to just 11.5% among those 65 and older. Taken together, these patterns point toward two overarching trends: among the official campaigns, younger and middle-aged voters are targeted more often with direct attacks on opponents, while seniors more often receive less combative messaging, including neutral appeals and positive content focusing exclusively on a favored candidate. Among outside groups, however, this dynamic is flipped, with younger and middle-aged voters being sent more singularly positive appeals than older voters.

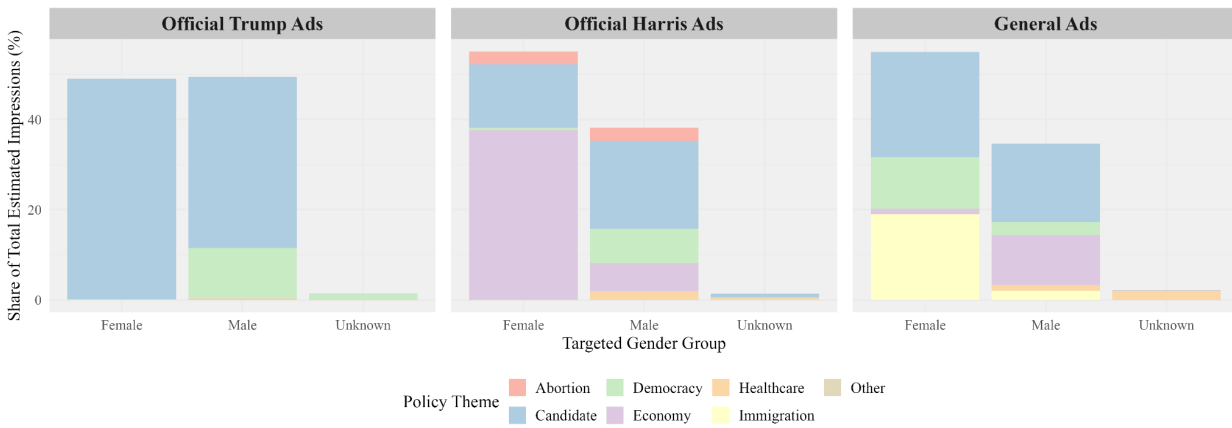
Variations in Ad Messaging by Gender

The final set of graphs (Graphs VII-IX) look at message variation by gender, using a stacked bar layout just like the set focused on age group. The taller the bar within a panel, the more ad impressions were targeted to that gender, and the larger the color segment within a bar, the greater proportion of those impressions were associated with a specific label. Again, color proportions within each bar vary significantly, highlighting careful targeting choices based on users' gender.

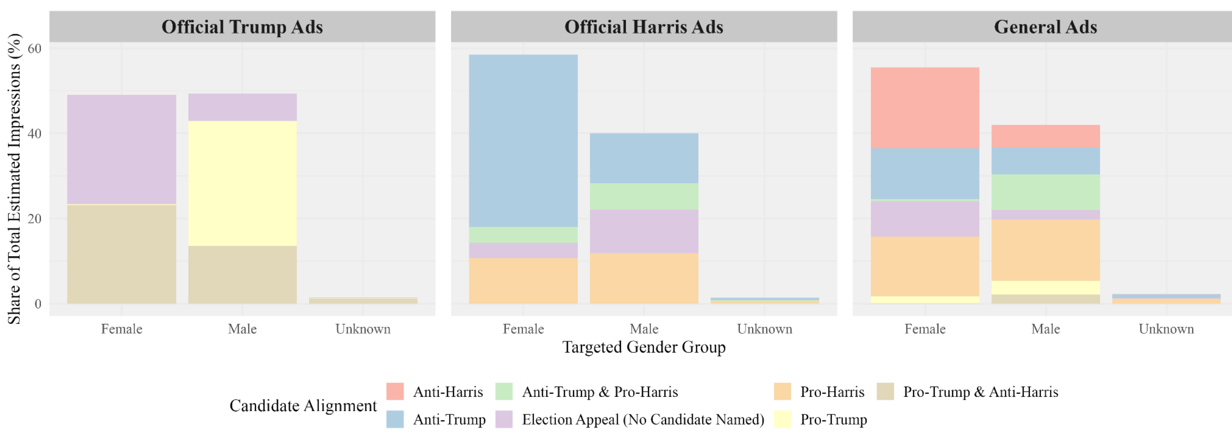
Graph VII: Emotional Tones of Ad Captions Targeted to Different Genders



Graph VIII: Emotional Tones of Ad Captions Targeted to Different Genders



Graph IX: Emotional Tones of Ad Captions Targeted to Different Genders



Graph VII: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	2.442e+11	1.2603e+11	3.0728e+12
P-value	2.2e-16	2.2e-16	2.2e-16

Graph VIII: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	2.1435e+11	3.2369e+11	5.8445e+12
P-value	2.2e-16	2.2e-16	2.2e-16

Graph IV: Chi-Squared Test Results

	Official Trump	Official Harris	General
X-Squared	3.9411e+10	6.5829e+10	1.8269e+10
P-value	2.2e-16	2.2e-16	2.2e-16

Graph VII and its corresponding chi-squared table indicate statistically significant (all P = 2.2e-16) gender-based differences in messaging tone. Among Official Trump impressions, users with unknown genders are sent fearful messaging 82.7% of the time, in contrast to men (10.6%) and women (0.2%), who instead receive higher rates of hope-based messages (52.6% and 37.8%, respectively) and patriotic messages (45.2% and 47.2%, respectively). Impressions in the Official Harris cohort follow a different pattern: female viewers receive content that is 70.2% anger-based (compared, respectively, to 19.3% and 25.2% for men and individuals with unknown genders), while men and individuals with unknown genders receive more hopeful messaging (51.7% and 49.3%, respectively, compared to females' 24.7%). Among impressions in the General cohort, females also receive more fearful messaging (42.3% compared to males'

12.0%), with hopeful messaging dominating impressions for males and individuals of unknown genders (44.0% and 58.4%, respectively, compared to females' 4.2%). Overall, it's apparent that while Official Trump impressions tend to target women with more positive emotional tones, the Official Harris and General cohorts disproportionately target women with negative ones (see Appendix Table IX for all tone by gender proportions).

Graph VIII and its associated chi-squared tests also reveal statistically significant gender-based messaging variations (all P = 2.2e-16)—this time, by messages' central themes. Among Official Trump impressions, women are almost entirely targeted with candidate-focused messaging (99.7%), while men and unknown-gender users are sent more democracy-focused content (22.1% and 88.7%, respectively). In the Official Harris cohort, on the other hand, women are largely sent economy-related messaging (64.4%), far more than men (15.7%) and individuals with unknown genders (3.7%), who mostly receive candidate-focused messages (48.7% and 56.1%, respectively). In the General cohort, women are disproportionately sent messaging about democracy (20.6%, compared to men with 6.7% and unknown-gender users with 1.5%), as well as immigration (34.2% compared to men with 4.8% and unknown-gender users with 0.2%). Across the board, these differences highlight two opposing microtargeting tactics: while the Trump campaign targets women with almost exclusively candidate-centered content, sending men and unknown-gender users more policy-related content, the Harris campaign and outside groups do the opposite (see Appendix Table X for further theme by gender proportions).

Last, Graph IX and its corresponding chi-squared table communicate statistically significant variations in candidate alignment by gender across all cohorts. In the Official Trump cohort, Women and unknown-gender users are disproportionately sent messages that are both pro-Trump and anti-Harris (47.2% and 82.9%, respectively, compared to 27.5% for men), while men mostly receive content that is solely pro-Trump (59.3%, compared to 0.6% for women and 7.6% for unknown-gender users). The opposite pattern emerges among Official Harris impressions: women overwhelmingly are targeted with anti-Trump-only messaging (69.2%), far less than men (29.4%) and unknown-gender users (39.7%). Finally, the General cohort follows its own unique pattern, targeting women with mostly pro-Harris-only content (25.3%), anti-Trump and pro-Harris content (34.2%), and anti-Trump-only content (21.4%). Men, in contrast, most commonly are targeted with solely pro-Harris content (34.0%), with other alignment combinations more evenly distributed. Users of unknown gender in the General cohort are most comonly targeted with anti-Trump, pro-Harris messaging (38.4%), as well as solely pro-Harris messaging (53.55%). Overall, these differences suggest that across the cohorts, women are more often sent mixed-alignment or opponent-focused messages, while men more often receive messaging focused only on promoting a favored candidate.

Reflecting on Results: Why Should We Care?

After all that number crunching, an important reminder: the patterns uncovered in this study aren't just statistically significant—they're normatively and socially significant. The demographic-based variations in messaging I've described—which, for that matter, only capture a small part of the massive data story conveyed by the Appendix's contingency tables—affirm widespread critiques of political microtargeting. By comprehensively examining a particular domain of surveillance capitalism—not just in theory, but in practice—these findings underscore concerns across the literature surrounding the potential for political entities to segment, manipulate, and mislead voters via personalized emotional and informational microtargeting tactics. As these strategies threaten both our individual autonomy and the health of our democracy, carefully identifying them and their potential impacts is paramount.

First, the results underscore the frequent, audience-specific use of emotionally charged messaging in political microtargeting. Across the all three ad cohorts, competitive battleground states were consistently sent emotionally charged messaging—most notably, fear. This finding corroborates concerns raised by Zuboff, who contends that microtargeting is designed not just to predict behavior, but to shape it by praying on human nature (Zuboff 2019, 17). Acting in-line with what Zuboff coins the "prediction imperative," campaigns and outside groups during the 2024 presidential election likely sought to exploit emotional triggers among swing-voters. By flooding the social media channels of those more likely to be politically undecided with anxiety-inducing content, political organizations could prime these individuals for future behavioral shifts (e.g., voting for one candidate over another).

Second, the results also highlight the prominent role psychological profiling plays in political microtargeting. A primary example of such profiling is evidenced by the relative decrease in combative sentiments toward seniors across all cohorts—who instead were met with a steady stream of optimistic messaging. Likely operating under the assumption that older voters are more resistant to conflict-ridden content—and alternatively, more responsive to messages conveying reassurance, stability, and tradition—campaigns and outside groups largely avoided sending negative messaging to those 65 and older. This approach echoes the psychographic logic made famous by Cambridge Analytica, which involves segmenting voters based on presumed emotional dispositions and personality traits for the purpose of maximizing the persuasive impact of ads (Headwood 2018, 429).

Third, the results shed a light on microtargeting's formation of asymmetrical information landscapes online. While emotionally charged messaging and psychological profiling may be effective voter mobilization strategies, they also force fragmented understandings of candidates and issues onto the electorate. For instance, in the fall of 2024, microtargeting toward seniors likely caused their online political environments to be artificially sanitized, disproportionality shielding those 65 and up from the contentious debates and hot-button issues constantly confronting younger and middle-aged voters. Similarly, vast differences in the targeting of substantive policy-based appeals versus vague, character-centered messaging likely left some voters (particularly swing state voters and younger voters) more equipped to engage in meaningful political evaluations than others. These examples underscore the informational "inequity" scholars like Bay associate with political microtargeting (Bay 2019, 11). When demographic groups are targeted with vastly different points of emphasis, the result is not simply unequal access to political information; the result is unequal access to democratic deliberation. By simultaneously privileging and denying election information to different audiences, microtargeting erodes the shared informational foundation necessary for collective reasoning and informed participation in the democratic process.

All in all, rather than providing voters consistent framing and information on candidates' issue-stances, trusting them in their full capacity as "democratic agents" to make sensible voting decisions (Jongepier and Wieland 2022, 170), microtargeting reinforces problematic assumptions about our emotional susceptibility, persuasion potential, and political utility. As this study demonstrates, these assumptions are imposed unevenly across age, gender, and geography, potentially affecting our political understandings and engagement levels—often without us realizing it. Our capacity as free and equal individuals to exercise political power—through voting and discourse—can be eroded, as Zuboff reminds us, not only through overt coercion but through granular, algorithm-driven behavioral nudges deployed at scale (Zuboff 2019, 37). Given the gravity of this threat, microtargeting demands both regulatory and individual-level intervention.

Limitations, Lessons Learned, and Further Applications

While results across all cohorts showed statistically significant variation in messaging by demographic group, this study's findings are not without limitations. Several factors constrain the comprehensiveness of analysis, including missing demographic data for 13.01% of all examined ads, Ad Library search options that were not perfectly reliable (e.g., the 201 missing Official Trump ads from the General cohort before CSV cleaning), and the exclusive provision of impression ranges (rather than precise futures) that necessitated the use of midpoint estimates to approximate ad reach. Furthermore, analysis may also have been impacted by the inability to definitively verify the accuracy of GPT-4o labeling; scoring the model's labeling success rate would require reading through the 120,000+ ads, a logistically impractical task that would undermine the point of automated labeling in the first place.

Appendix Tables V, VIII, XI—which display the proportional breakdown of impressions by candidate alignment across state type, age group, and gender, respectively—point toward evidence of extremely infrequent, but still present, labeling mistakes. For instance, in a few instances, nonsensical candidate alignment combinations occur—such as Official Trump impressions that were labeled as anti-Trump. While these sorts of nonsensical labels appear to be randomly distributed—not biased toward a particular cohort or demographic type—I still sought to conduct a follow-up reliability check for all labels. This check involved randomly selecting sets of 5 ads from the cohorts, labeling them myself, and cross-checking my labels with those generated by GPT-4o. After conducting two trials for each label type (tone, theme, and alignment), totaling in 6 trials and 30 ads, I found that the model's labels matched my own with complete accuracy, reinforcing confidence that labeling errors were random anomalies that do not undermine the validity of the study's broader conclusions.

Despite limitations, this thesis' 7-month research, testing, and writing process offered many valuable lessons. At the top of the list, this work is a testament to the power of GPT-4o. Beyond successfully automating a 4-round labeling process for 120,000+ ads, GPT-4o empowered me, an undergraduate Public Policy major with entry-level R, Stata, and Python skills, to embark on dozens of complex coding endeavors with a full-time, 24/hour, free, personal tutor guiding me along the way. GPT-4o by no means was a perfect helper: it constantly repeated the same errors, forgot outputs it had just given me, and sometimes solved problems by creating new ones. That said, this completed thesis stands as evidence that GPT-4o is an incredibly transformative tool—one with the ability to serve as an equalizer in research. Before the existence of the model, the programming and data science techniques deployed in this thesis would have been out of the question for a student with my academic and professional experiences. Yet, with a GPT-4o account, determination, lots of coffee, and human input to fill in the gaps, I was able to complete a technically rigorous, large-scale exploration of microtargeting in practice.

Conducting sentiment analysis with GPT-4o also taught me how to approach labeling tasks more effectively. In hindsight, the decision to have the model select a single dominant tone and theme oversimplified ads that contained multiple emotional and/or policy elements. In future iterations of this study, using a multi-label classification system with confidence scoring may better capture more nuanced, multi-faceted messaging. The sentiment analysis component of this thesis also underscored the importance of minimizing token inputs and outputs to control costs when using OpenAI's API. Strategies such as using shorthand notation, cutting filler words in instructions, limiting output verbosity, and narrowing the scope of the input text the model had to process proved essential for scalability.

This thesis also offers practical and political insights that extend beyond its immediate focus. Most notably, this work's methodology and underlying dataset could be used to investigate the effectiveness of specific campaign microtargeting tactics. For instance, one might compare messaging strategies across demographic groups with corresponding voter turnout data, potentially helping candidates and political organizers pinpoint the most successful (or least successful) messages for mobilizing particular segments of the electorate. Additionally, the results also inspire new questions about the campaigns' strategic choices and potential oversights. For instance, after examining Appendix Tables IV, VII, and X, one might ask why abortion-related ads appear largely underrepresented across the cohorts, despite abortion being a central component of Democrats' messaging in 2024. Lastly, the methodology developed here offers a promising tool for identifying dark money actors and ads that directly contradict one another. In future iterations of this study, sentiment analysis could be used to flag conflicting messages distributed by the same advertiser across different demographic groups—providing a scalable way to detect deceptive microtargeting tactics and enhance transparency in digital political advertising.

POLICY RECOMMENDATIONS

Where Section 230 Leaves Us

The pervasive expansion of microtargeting, across both commercial and political contexts, is not simply a product of technological advancement. Microtargeting tactics continue to grow in sophistication and reach because regulatory environments have failed to evolve alongside them. In the absence of legal constraints, companies, organizations, and political actors are incentivized to constantly find new and bigger ways to capitalize on Big Data, whether for profit, political power, or both. The following section unpacks the legacy of one of the most influential policies that enabled an internet ecosystem driven by the prediction imperative: Section 230 of the Communications Decency Act (CDA).

Enacted in 1996, Section 230 provides online platforms like Meta broad immunity from civil liability for content created by third-parties. Specifically, it declares that “no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider” (Brannon and Holmes 2024). Courts have interpreted this provision to extend well beyond simply hosting content on online servers. For instance, in Z*eran v. America Online*, the court granted sweeping immunity to platforms in their capacity not just as publishers but as distributors, regardless of receiving notification about harmful content (Lukmire 2010, 385). Furthermore, in *Force v. Facebook,* the court reinforced platforms’ role as content curators, affirming the use of algorithms to recommend material (Robertson 2020). As summarized by Public Knowledge Legal Director John Bergmayer, the provision means that platforms are shielded from being held legally accountable for user-generated content in practically all scenarios—even when they explicitly “choose to highlight and promote” specific material or “use it in online advertisements” (Bergmayer 2019).

Before diving into the present-day consequences of Section 230’s (essentially boundless) immunity for online platforms, it’s critical to emphasize the provision’s important, early contributions to the “open” internet. As Jeff Kosseff, professor of cybersecurity law at the United States Naval Academy, argues, without the original enactment of Section 230, access to the internet would likely be far less equal today (Kosseff 2010, 125). Had the internet emerged during a period of increased legal risk for platforms, it's reasonable to assume online stakeholders wouldn't have embraced the decentralized innovation structure that transformed the internet into a collaborative social hub and economic driver. Rather than empowering users early on to freely create, post, and communicate online, platforms would have faced legal incentives to place rigid limits on user activity in order to avoid liability (Kosseff 2010, 152). Put simply, in the early internet era—an era characterized by small-scale websites—Section 230 was a vital tool that promoted free expression and innovation.

Today's internet landscape, however, looks radically and indisputably different. Dominated by tech giants like Meta, the internet's early community-driven web pages have been replaced by platforms like Instagram and X (formerly Twitter)—sites used by hundreds of millions globally that function as essential public infrastructure. Instead of safeguarding vulnerable innovators, Section 230 safeguards companies, as Haucap and Heimeschoff note, like "Google…who currently dominate their relative markets" (Haucap and Heimeschoff 2014, 50). Creating, as they describe, "highly concentrated market structures" through the facilitation of "direct and indirect network effects" (Haucap and Heimeschoff 2014, 52 & 50), these platforms dominate the internet ecosystem, creating insurmountable barriers for smaller players to enter and survive in the market. As such, Section 230, though it once served a crucial purpose, does more harm than good today; though it still helpfully protects a shrinking group of smaller tech players from liability, it does so at the cost of emboldening mega-corporations with free-reign, zero accountability, and no mechanisms for official oversight.

These harms are especially evident in the context of political microtargeting. Providing legal standing for platforms to publish, maintain, manage, and curate content at a massive scale, Section 230 sets in stone the role of Meta and other platforms as invisible gatekeepers of political information. As Citron and Wittes note, “courts have built a mighty fortress protecting platforms from any accountability for unlawful activity on their systems—even when they actively encourage such activity or deliberately refuse to address it” (Citron & Wittes 2018, 458). Completely safe in this “mighty fortress,” platforms are free to collaborate with ad buyers however they please, leveraging Big Data to segment the electorate and deliver hyper-tailored messaging with no restraints—and to extents that still remain extraordinary hard to trace (as evidenced by this thesis’ complex and resource-heavy methodology).

Where Do We Go From Here?

Efforts to repeal or reform Section 230, though gaining political momentum, have consistently failed. It’s not hard to guess one reason why: last year alone, Meta, Alphabet (the parent company of Google and YouTube), Microsoft, ByteDance (the parent company of TikTok), X, and Snap (the parent company of Snapchat) spent a combined “staggering $61.5 million on lobbying” (Minkin 2025). Partisan divides, however, have also played a major role in hindering progress. While past Democratic proposals have centered around holding platforms accountable for harmful content, Republican proposals have focused on punishing platforms for restricting certain kinds of content—goals often in direct opposition to one another (Feiner 2025).

New legislative developments, however, are underway: as of reporting in late March, 2025, U.S. Senators Dick Durbin (D-IL) and Lindsey Graham (R-SC) are expected to introduce a bipartisan bill aimed at sunsetting Section 230 over a period of two years (Nazzaro 2025). While the bill’s exact contents are not currently public, it appears as though the bill will call for the repealing of Section 230 in its entirety. In my opinion, such a move—unless part of a larger negotiation strategy—is not strategically sound. A quick search on Google will land you in a sea of complaints from various stakeholders—civil liberties groups, trade associations, independent commentators, and so on—outlining how eliminating Section 230 protections would chill free speech and push out already-struggling small and medium-sized platforms. Unlike companies like Meta, which can respond to legal challenges with armies of lawyers and billions of dollars, many contend that startups and smaller firms are ill-equipped to handle new liability burdens. Given the range of powerful and ideologically diverse stakeholders invested in preserving parts of Section 230, a sunset bill would likely face intense political pushback from all sides—rendering its passage through Congress politically implausible.

Instead of attempts to repeal Section 230 in one go, I would encourage incremental legislation that is context-specific. In regard, specifically, to the issue of political microtargeting, policy proposals should do the following:

1. Remove Section 230 protections for paid advertising content;
2. Establish a federal political ad registry.

Designed to avoid the pitfalls of a complete Section 230 repeal, these reforms have a more defined scope aimed at curbing the most urgent risks posed by political microtargeting. The first, an idea embraced by advocacy organizations like Public Knowledge (Macpherson 2025), would strip platforms of immunity when profiting from paid advertisements—political or non-political. Such a reform would open the door for companies like Meta to be legally held accountable for misleading or manipulative ads they choose to host and amplify algorithmically. For instance, if a court were to deem the Elon Musk-funded Future Coalition's ads on Harris' stance toward Israel and Palestine materially deceptive or coercive, this reform would create a legal basis to argue that Meta played a role in distributing that content—and therefore should be liable for it. As such, Meta and other tech giants would face new pressure to think twice before accepting or promoting contradictory or misleading political ads that could get them in legal trouble. At the same time, however, this adjustment to Section 230 still maintains the crucial protections that foster innovation and fairness for smaller firms and individual users, focusing on holding tech giants accountable without dragging unrelated actors into the political crosshairs.

Second, by establishing a federal political ad registry, regulators could finally impose transparency standards onto the dark-money-ridden world of digital campaigning. Modeled in a similar manner to existing disclosure rules for print and broadcast political ads, the registry would provide scholars like me the Meta Ad Library of their dreams: a centralized database with consistent, clear, and accurate information. In addition to disclosing specifics on demographic targeting criteria, sharing accurate impression counts (not ranges), and being easy to navigate with robust search options, the database would require advertisers to provide verifiable information on who paid for ads. Crucially, non-registered entities—such as loosely organized political nonprofits or nondescript PACs—would be obligated to list a publicly visible legal name, funding source, and valid contact information. Closing current loopholes that enable bad actors to obscure their identities and intent, this reform would empower watchdogs, journalists, and voters alike to identify instances of manipulation in real time. Moreover, because this reform does not restrict content or speech—solely requiring disclosure—it is far more likely to be politically salient and withstand First-Amendment scrutiny.

Finally, I leave the reader with a tangible action item (beyond advocating for the aforementioned policies). Meaningful political microtargeting reform is going to take time—probably far longer than most of us would hope for. That said, each of us—as empowered and informed democratic agents—have an immediate role to play. By remaining vigilant toward microtargeting efforts aimed at entrenching us deeper into echo chambers, manipulating our psyches, and eroding our ability to partake in collective reasoning, we begin to undermine the power of the prediction imperative. As Zuboff warns, surveillance capitalism operates not by directly restraining our actions, but by setting scripts in advance that steer us toward predetermined outcomes. Proactively working to expand our understanding of these scripting mechanisms is the first step toward reclaiming what Arendt refers to as our “right to the future tense”: the capacity to imagine, deliberate, and act toward futures of our own choosing. While it won’t dismantle the system on its own, awareness is the only way to restore the possibility of action that the system seeks to erase.

WORKS CITED

Agan, Tom. 2007. "Silent marketing: Micro-targeting." *Penn, Schoen and Berland Associates*. LINK

Arab American Institute. 2024. *National Arab American Demographics*. Accessed March 30, 2025. LINK

Arendt, Hannah. 1968. "Truth and Politics." In *Between Past and Future: Eight Exercises in Political Thought*, 227—264. New York: Viking Press. LINK

Arthur, Lisa. 2013. *Big Data Marketing: Engage Your Customers More Effectively and Drive Value*. John Wiley & Sons. LINK

Bay, Morten. 2018. "Social Media Ethics: A Rawlsian Approach to Hypertargeting and Psychometrics in Political and Commercial Campaigns." *ACM Transactions on Social Computing* 1 (4): 1-14. LINK

Bender, Sarah M. L. 2022. "ALGORITHMIC ELECTIONS." *Michigan Law Review* 121 (3): 489-522. LINK

Bergmayer, John. 2019. "What Section 230 Is and Does — Yet Another Explanation of One of the Internet's Most Important Laws." *Public Knowledge*, May 14, 2019. LINK

Bodó, Balázs, Natali Helberger, and Claes H. de Vreese. 2017. "Political micro-targeting: a Manchurian candidate or just a dark horse?." *Internet Policy Review* 6 (4): 1-13. LINK

Brannon, Valerie C., and Eric N. Holmes. 2024. *Section 230: An Overview*. Congressional Research Service, January 4, 2024. LINK

Brown, Frank. 2004. "Nixon's 'Southern Strategy' and Forces against Brown." *The Journal of Negro Education* 73 (3): 191—208. LINK

Cadwalladr, Carole. 2020. "Fresh Cambridge Analytica leak 'shows global manipulation is out of control'." *The Guardian*, January 4, 2020. LINK

Citron, Danielle K., and Benjamin Wittes. 2018. "The Problem Isn't Just Backpage: Revising Section 230 Immunity." *Georgetown Law Technology Review* 2, (2): 453. LINK

Danaher, Peter J. 2023. "Optimal microtargeting of advertising." *Journal of Marketing Research* 60 (3): 564-584. LINK

De Mauro, Andrea, Marco Greco, and Michele Grimaldi. 2016. "A formal definition of Big Data based on its essential features." *Library review* 65 (3): 122-135. LINK

Detrow, Scott. 2018. "What Did Cambridge Analytica Do During the 2016 Election?" *NPR*, March 20, 2018. LINK

Diebold, Francis X. "*On the Origin (s) and Development of the Term "Big Data.""* Working Paper No. 12-037. Penn Institute for Economic Research, University of Pennsylvania. LINK

Dutta, Soumitra, and Matthew Fraser. 2008. "Barack Obama and the Facebook Election." *U.S. News & World Report*, November 19, 2008. LINK

Erskine, Andrew. 1995. "Culture and power in ptolemaic Egypt: The Museum and Library of Alexandria." *Greece & Rome* 42 (1): 38-48. LINK

Erevelles, Sunil, Nobuyuki Fukawa, and Linda Swayne. 2016. "Big Data consumer analytics and the transformation of marketing." *Journal of Business Research* 69 (2): 897-904. LINK

Facebook. 2019. *"A Better Way to Learn About Ads."* Facebook Newsroom, March 28, 2019. LINK

Facebook. 2025. *"About the Ad Library."* Accessed March 9, 2025. LINK

Favaretto, Maddalena, Eva De Clercq, Christophe Olivier Schneble, and Bernice Simone Elger. 2020. "What is Your Definition of Big Data? Researchers' understanding of the phenomenon of the decade." *PLOS ONE* 15 (2): e0228987. LINK

Federal Trade Commission. 2019. "FTC Sues Cambridge Analytica, Settles with Former CEO and App Developer." *Federal Trade Commission,* July 24, 2019. LINK

Feiner, Lauren. 2025. "Lawmakers Are Trying to Repeal Section 230 Again." *The Verge*, March 21, 2025. LINK

Fruchter, Nathaniel, Michael Specter, and Ben Yuan. 2018. "Facebook/Cambridge Analytica: Privacy lessons and a way forward." *MIT Internet Policy Research Initiative*, March 20, 2018. LINK

Fulda, Bernhard. 2011. "The Market Place of Political Opinions: Public Opinion Polling and its Publics in Transnational Perspective, 1930—1950." *Comparativ* 21 (4): 13-28. LINK

Goldmacher, Shane, and Nicholas Nehamas. 2024. "Kamala Harris Has Raised $1 Billion Since Entering 2024 Presidential Race." *New York Times*, October 9, 2024. LINK

Google Trends. 2024. "Big Data." Accessed November 28, 2024. LINK

Grimm, Pamela. 2010. "Social desirability bias." *Wiley International Encyclopedia of Marketing*. LINK

Haucap, Justus, and Ulrich Heimeshoff. 2014. "Google, Facebook, Amazon, eBay: Is the Internet Driving Competition or Market Monopolization?" *International Economics and Economic Policy* 11: 49–61. LINK

Heawood, Jonathan. 2018. "Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal." *Information polity* 23 (4): 429-434. LINK

Hegazy, Islam Mohamed. 2021. "The effect of political neuromarketing 2.0 on election outcomes: The case of Trump's presidential campaign 2016." *Review of Economics and Political Science* 6 (3): 235-251. LINK

Heller-Roazen, Daniel. 2002. "Tradition's destruction: On the library of Alexandria." *October* 100: 133-153. LINK

Hirsch, Dennis D. 2013. "The glass house effect: Big Data, the new oil, and the power of analogy." *Maine Law Review 66: 373*. LINK

Hughes, Sam Graham-Felsen, Kate Allbright-Hannah, Scott Goodstein, Steve Grove, Randi Zuckerberg, Chloe Sladden, and Brittany Bohnet. 2010. "Obama and the power of social media and technology." *The European Business Review* 16: 21. LINK

Institute for Democracy, Journalism & Citizenship. 2024. *Final 2024 Report: The Breadth and Scope of Online Ads.* Syracuse University, December 10, 2024. LINK

Jacobs, Lawrence R., and Robert Y. Shapiro. 1994. "Issues, candidate image, and priming: The use of private polls in Kennedy's 1960 presidential campaign." *American Political Science Review* 88 (3): 527-540. LINK

Johnson, Jeff S., Scott B. Friend, and Hannah S. Lee. 2017. "Big data facilitation, utilization, and monetization: Exploring the 3Vs in a new product development process." *Journal of Product Innovation Management* 34, (5): 640-658. LINK

Jongepier, Fleur, and Jan Willem Wieland. 2022. "Microtargeting people as a mere means." In *The Philosophy of Online Manipulation*: 156-179. Routledge. LINK

Kosseff, Jeff. 2010. "Defending Section 230: The Value of Intermediary Immunity." *Journal of Technology Law & Policy* 15: 123. LINK

Kröger, Jacob Leon, Emilia Errenst, Niklas Nau, and Sanna Ojanperä. 2024. "Political microtargeting: example cases Global South." *This chapter is an excerpt from:" Mitigating the Risks of Political Microtargeting—Guidance for Policymakers, Civil Society, and Development Cooperation*: 18-25. LINK

Leonardi, Paul M., and Jeffrey W. Treem. 2020. "Behavioral visibility: A new paradigm for organization studies in the age of digitization, digitalization, and datafication." *Organization Studies* 41 (12): 1601-1625. LINK

Lohr, Steve. 2013. "The Origins of 'Big Data': An Etymological Detective Story." *New York Times*, February 1 2013, LINK.

Lukmire, David. 2010. "Can the Courts Tame the Communications Decency Act?: The Reverberations of Zeran v. American Online." *NYU Annual Survey of American Law* 66: 371. LINK

Macpherson, Lisa. 2025. "Public Knowledge Proposes Section 230 Reforms That Address Harms While Protecting Free Expression." Public Knowledge, March 10, 2025. LINK

Mashey, John R. 1999. "Big Data and the Next Wave of InfraStress Problems, Solutions, Opportunities." Slide deck presented at *USENIX Annual Technical Conference (USENIX ATC 99)*. Berkeley, CA: USENIX Association. LINK

McAdams, Dan P. 1992. "The five‐factor model in personality: A critical appraisal." *Journal of personality* 60 (2): 329-361. LINK

McAfee, Andrew, Erik Brynjolfsson, Thomas H. Davenport, D. J. Patil, and Dominic Barton. 2012. "Big data: the management revolution." *Harvard Business Review* 90 (10): 60-68. LINK

Merriam-Webster.com Dictionary. 2024. S.v. "Big Data." Accessed November 28th, 2024. LINK

Meta Platforms, Inc. 2021. “Understanding the Ad Library.” *Facebook*, December 3, 2021. LINK

Meta Platforms, Inc. 2025a. “Ad Targeting.” *Meta Business Help Center*. Accessed March 25, 2025. LINK

Meta Platforms, Inc. 2025b. “Choose the Locations Where You Want to Show Your Ads.” Meta Business Help Center. Accessed March 25, 2025. LINK

Meta Platforms, Inc. 2025c. "Meta Ad Library API," accessed March 29, 2025. LINK

Meta Platforms, Inc. 2025d. *Terms of Service*. Accessed March 27, 2025. LINK

Miller, Zeke, Colleen Long, and Darlene Superville. 2024. "Biden Drops Out of 2024 Race and Endorses Harris." *AP News,* July 21, 2024. LINK

Minkin, Amelia. 2025. “Big Tech Cozies Up to New Administration After Spending Record Sums on Lobbying Last Year.” *Issue One*, January 22, 2025. LINK

MIT News Office. “Study: Microtargeting Works, Just Not the Way People Think.” Last modified June 21, 2023. LINK

Nazzaro, Miranda. 2025. “Senators Revive Efforts to Strip Tech Companies of Key Legal Protection.” *The Hill*, April 2, 2025. LINK

Oxford English Dictionary. 2024. S.v. “Big Data, n.” Accessed November 28, 2024. LINK

Perkins, Tom. 2024. “Musk-linked PAC Accused of Targeting Jewish and Arab Americans in Swing States.” *The Guardian*, November 4, 2024. LINK

Pew Research Center. 2008. “Inside Obama’s Sweeping Victory.” *Pew Research Center*. November 5, 2008. LINK

Pew Research Center. 2024. “Social Media and News Fact Sheet.” *Pew Research Center*, September 17, 2024. LINK

Rawls, John. 2005. “A Theory of Justice.” Cambridge, MA: Belknap Press. LINK

Raynauld, Vincent, and André Turcotte. 2018. ““Different strokes for different folks”: Implications of voter micro-targeting and appeal in the age of Donald Trump.” *Political marketing in the 2016 US presidential election*: 11-28. LINK

Robertson, Adi. 2020. “Supreme Court Rejects Lawsuit Against Facebook for Hosting Terrorists.” *The Verge*, May 18, 2020. LINK

Rometty, Virginia. 2013. “A Conversation with Ginni Rometty.” *The Council on Foreign Relations*, March 7, 2013. LINK

Rosenbaum, Martin. 1997. “Campaign Materials and Direct Mail.” In *From Soapbox to Soundbite: Party Political Campaigning in Britain since 1945*: 209-222. London: Palgrave Macmillan UK. LINK

Rosenberg, Matthew. 2019. "Ad Tool Facebook Built to Fight Disinformation Doesn’t Work as Advertised." *The*

New York Times, July 25, 2019. LINK

Rothwell, Jonathan. 2023. "Teens Spend Average of 4.8 Hours on Social Media Per Day." *Gallup*, October 13, 2023. LINK

Saxe, Leonard, Daniel Parmer, Elizabeth Tighe, Raquel Magidin de Kramer, Daniel Kallista, Daniel Nussbaum, Xajavion Seabrum, and Joshua Mandell. 2021. "American Jewish Population Estimates: Summary & Highlights." Waltham, MA: *Cohen Center for Modern Jewish Studies and Steinhardt Social Research Institute, Brandeis University*. LINK

Tamane, Sharvari. 2016. "Non-Relational Databases in Big Data." *In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies (ICTCS '16), Udaipur, India, March 4–5, 2016,* 1–4. LINK

Tappin, Ben M., Chloe Wittenberg, Luke B. Hewitt, Adam J. Berinsky, and David G. Rand. 2023. "Quantifying the potential persuasive returns to political microtargeting." Proceedings of the National Academy of Sciences 120 (25). LINK

Tsai, Chun-Wei, Chin-Feng Lai, Han-Chieh Chao, and Athanasios V. Vasilakos. 2015. "Big data analytics: a survey." *Journal of Big Data*: 1-32. LINK

Watkins, Eli. 2018. "Cambridge Analytica Announces Closure." *CNN*, May 16, 2018. LINK

Zuboff, Shoshana. 2019. "'We Make Them Dance': Surveillance Capitalism, the Rise of Instrumentarian Power, and the Threat to Human Rights." *Human rights in the Age of Platforms*: 3-51. LINK

APPENDIX

Appendix Table I: Variable Definitions for CSV Data from the Meta Ad Library

Variable	Definition
ad_archive_id	The unique identifier assigned to the ad for tracking and referencing.
page_id	The unique identifier assigned to the Facebook or Instagram page that published the ad.
page_name	The public name of the Facebook or Instagram page that ran the ad.
ad_creation_time	The date when the ad was originally created on Meta, potentially distinct from the date the ad began running.
ad_delivery_start_time	The data the ad began running (being shown to users).
ad_delivery_stop_time	The final date the ad was run.
ad_creative_bodies	The caption text of an ad, as it appears to users. This field excludes any text embedded within attached graphics, images, or transcript text from videos. Often, the entirety of this caption is repeated several times within a single entry of the variable, with identical captions separated by commas. These repetitions reflect multiple placements or slightly modified versions of the same ad, not new content.
ad_creative_link_titles	The title associated with any links in the ad.
ad_creative_link_captions	Text from short captions shown near the ad’s linked content, typically summarizing a link.
ad_creative_link_descriptions	Additional descriptive text about the ad’s linked content, typically shown below headlines and captions.
impressions	The estimated number the ad appeared on users’ screens, excluding all measures of engagement.
spend	The estimated cost paid by an ad buyer to run the ad.
currency	The reported currency of the buyer.
demographic_distribution	A breakdown of the ad’s impressions by age and gender, represented as percentages.
delivery_by_region	The geographic distribution of ad impressions by state.
publisher_platforms	The Meta platforms the ad was shown on, typically Facebook and Instagram.
languages	The primary language the ad was delivered to viewers in.

Appendix Table II: Ad Counts for CSV Merges Throughout Data Cleaning

Ad Count Breakdown for the Merger of the General and Official Trump Datasets			
Ad Total in Merged Dataset (Duplicates Included)	Ads Unique to the General Trump Dataset	Ads Unique to the Official Trump Dataset	Duplicates (Ads in Both the General and Official Trump Datasets)
78,725	44,456	201	17,034
Ad Count Breakdown for the Merger of the General Trump (Tdeduped) and Official Harris Datasets			
Total Ads in Merged Dataset (Duplicates Included)	Ads Unique to the General Trump Tdeduped Dataset	Ads Unique to the Official Harris Dataset	Duplicates (Ads in Both the General Trump Tdeduped and Official Harris Datasets)
80,557	28,629	36,161	15,827
Ad Count Breakdown for the Merger of the General and Official Harris Datasets			
Total Ads in Merged Dataset (Duplicates Included)	Ads Unique to the General Harris (Tdeduped) Dataset	Ads Unique to the Official Harris Dataset	Duplicates (Ads in Both the General and Official Harris Datasets)
105,930	53,942	0	51,988
Ad Count Breakdown for the Merger of the General Harris (Hdeduped) and Official Trump Datasets			
Total Ads in Merged Dataset (Duplicates Included)	Ads Unique to the General Harris (Tdeduped) Dataset	Ads Unique to the Official Trump Dataset	Duplicates (Ads in Both the General Harris Hdeduped and Official Trump Datasets)
70,474	53,238	16,532	703
Ad Count Breakdown for the Final Merger of the General Datasets			
Total Ads in Merged Dataset (Duplicates Included)	Ads Unique to the General Harris Dataset	Ads Unique to the General Trump Dataset	Duplicates (Ads in Both the General Harris Trump Datasets)
66,585	37,956	13,346	15,283

Appendix Table III: Proportional Breakdown of Impressions by Emotional Tone and State Type

`official_trump_ads_clean.csv`					
	Angry	Fearful	Hopeful	Neutral	Patriotic
Non-Swing	3.5020827385904900	3.2224276850953500	39.38951240571760	1.6506734119809900	52.23530375861560
Swing	1.0045274084442800	13.401530541362300	37.00612902939560	6.026502591010230	42.56131042978760
`official_harris_ads_clean.csv`					
	Angry	Fearful	Hopeful	Neutral	Patriotic
Non-Swing	6.911501062200480	4.574406525757340	78.19306234052470	10.026839232183100	0.2941908393344160
Swing	20.491862593353600	17.494000554996200	45.12881610146650	14.872196312981400	2.0131244372022600
`general_combined_unique_relevant.csv`					

	Angry	Fearful	Hopeful	Neutral	Patriotic
Non-Swing	11.87884483118000	9.813584124121950	60.907788575752200	7.452224804351050	9.947557664594800
Swing	29.751832971564500	12.08947843493000	44.45902604212380	10.764433041028800	2.935229510352980

Appendix Table IV: Proportional Breakdown of Impressions by Central Theme & State Type

`official_trump_ads_clean.csv`

	Candidate	Democracy	Economy	Other	Immigration
Non-Swing	90.20972099698120	6.056770256630010	0.43931903503334600	3.2941897113554900	0
Swing	71.169819216869	28.52116686643000	2.08454109954473E-04	0.2818910948508560	0.026914367740217900

`official_harris_ads_clean.csv`

	Abortion	Candidate	Climate	Crime	Democracy	Economy	Foreign Policy	Healthcare	Immigration	Military	Other	Welfare
Non-Swing	1.0964968047930400	83.8670519918686	1.159331283836E-04	6.95598770301601E-04	7.194439402462270	2.320927963667180	0.07785565248811970	0.12010521236610100	0.060573376867356500	0.0011602596772564900	5.260577803911380	0
Swing	8.358851783495930	38.79806322257070	0.07475042071434540	0.4339766598296270	25.25700413896960	18.171559516293500	1.291821016662350	5.372956657757300	1.5560261209198700	0.04825665137426000	0.5257658226732580	0.11096798873928100

`general_combined_unique_relevant.csv`

	Abortion	Candidate	Climate	Crime	Democracy	Economy	Education	Foreign Policy	Guns	Healthcare	Immigration	LGBTQ+	Military	Other	Welfare
Non-Swing	2.142919976929930	77.2452511954325	1.501144603156690	0.7830459489502710	9.500219286182290	3.5164213789912500	0.46630903098382900	0.4888302246098320	0.23677320985653200	1.0753616727434600	1.7747254287260500	0.07622655740449590	0.3445329757113560	0.3687415358579890	0.47949697446352500
Swing	8.812189470959460	26.379676531809800	2.7265991095234400	1.7555713400407800	5.820077748982700	32.49004246030950	1.6358458691611600	1.578952928917270	2.5990042970203200	5.1704713597310800	5.596537871253560	0.9710932066578010	0.7384738133054100	0.9443818974020120	2.781082094925660

Appendix Table V: Proportional Breakdown of Impressions by Candidate Alignment & State Type

`official_trump_ads_clean.csv`

	Trump_1_Harris_0	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_3	Trump_0_Harris_3
Non-Swing	50.649771344319700	41.80485708586420	0.008197946135757350	7.537173623680280	0
Swing	17.35037823051430	66.77646745932850	0.027395344726588200	15.483696776741000	0.36206218868960900

official_harris_ads_clean.csv

	Trump_0_Harris_1	Trump_0_Harris_3	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_1	Trump_3_Harris_3
Non-Swing	4.07700500737463	37.24355010305050	0.026110948902083700	1.159331283836E-04	9.260032019987350	49.39318598755710
Swing	16.701913557108600	25.88137634787900	0	0	33.33828815276700	24.078421942245400

general_combined_unique_relevant.csv

	Trump_0_Harris_0	Trump_0_Harris_1	Trump_0_Harris_3	Trump_1_Harris_0	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_1	Trump_3_Harris_3	Trump_1_Harris_1
Non-Swing	0.728397922804863	24.46727319130780	10.291052809881800	5.729070241774170	11.773925509183100	4.396111845388340	33.52491807497540	9.089250404684390	0
Swing	0.17853531985446400	19.802250531546100	15.856360239734600	3.783547346331780	3.80336703208064	11.853578033755400	40.53558074457100	4.184960393845230	0.0018203582808335700

Appendix Table VI: Proportional Breakdown of Impressions by Emotional Tone & Age Group

official_trump_ads_clean.csv

	Angry	Fearful	Hopeful	Neutral	Patriotic
18_24	1.0781652391246900	70.42428397032700	12.941377991956600	13.28719158012190	2.268981218469820
25_34	0.20129784024318400	71.23086815771310	12.831239097678300	13.439418943035400	2.297175961330020
35_44	0.23284902904578300	82.39029410610090	0.006489630319771980	15.545348655146400	1.825018579387190
45_54	0.2759360469850210	97.63045656039990	0.01042831001703370	6.82803563055549E-04	2.0824962790349700
55_64	0.13569774289345000	97.78620336332940	0.008413116307617260	8.30157725727203E-04	2.0688556197437800
65	2.35042306766713	11.219261396619500	33.21805741033660	5.406431560561240	47.80582656481550

official_harris_ads_clean.csv

	Angry	Fearful	Hopeful	Neutral	Patriotic
18_24	38.70040783475120	19.050707117168300	42.2249070815031	0.020419211430946300	0.0035587551464527100
25_34	39.15880029145300	19.289048687855600	41.511671040314500	0.03548282906509550	0.004997151311838330
35_44	39.175549601125700	19.286360178181100	41.50337707610640	0.03110467033105710	0.0036084742557434600
45_54	8.198044518530290	36.286551057681100	53.69289909872430	1.6575697401233800	0.16493558494099700
55_64	16.546871837831000	14.547611853906400	67.35688188369670	1.4276464177750000	0.12098800679082900
65	12.780373977728100	10.322225779703000	60.003071029091900	15.8550100989376	1.0393191145392900

general_combined_unique_relevant.csv

	Angry	Fearful	Hopeful	Neutral	Patriotic
18_24	4.926821977118190	36.2347201318185	56.89954937134330	1.786248053259620	0.152660466460383
25_34	3.8258676487374700	36.79501397834820	57.54015906814580	1.6998953998091000	0.1390639049594550
35_44	3.916255291209300	37.85816095922120	56.33192957272800	1.7505502828250600	0.1431038940164530
45_54	4.036361655985440	36.99685243330610	58.62159030998430	0.19556043888125600	0.14963516184285600
55_64	4.0186719003654300	37.02146621782490	58.62912527994160	0.18128362583351600	0.1494529760345370
65	19.46924631258980	12.935891360199800	49.622123798967000	8.608557971896170	9.364180556347340

Appendix Table VII: Proportional Breakdown of Impressions by Central Theme & Age Group

official_trump_ads_clean.csv					
	Candidate	Democracy	Economy	Other	Immigration
18_24	15.39545766603780	71.31739139526520	1.8247355753486E-07	1.03722701377522E-07	13.28715065250080
25_34	14.425923423062800	72.13480228020830	5.50048602934644E-07	4.9905580672834E-07	13.439273247624500
35_44	1.0190428595611600	83.43606613204660	6.44737693486958E-06	1.26929823386229E-06	15.544883291717100
45_54	1.1302505613101700	98.86958948243720	1.21319421361139E-05	2.71047224982005E-06	1.45113838229211E-04
55_64	0.9722739973657560	99.02747772664290	2.43055336082657E-05	4.1837576898061E-06	2.19786700083923E-04
65	77.5735311046949	20.640632028880000	0.24573075592953200	0.027310660960507400	1.5127954495351000

official_harris_ads_clean.csv												
	Abortion	Candidate	Climate	Crime	Democracy	Economy	Foreign Policy	Healthcare	Immigration	Military	Other	Welfare
18_24	1.5941037436512900	46.537678686843400	6.40008593163857E-05	7.66696115055967E-04	8.876843894015020	21.969404598949700	0.002385203797776930	19.145974088775300	0.002471801102483740	5.29951955764191E-05	1.8701788901825900	7.54005124734801E-05
25_34	1.6281248894466500	45.86209870566870	2.01284568784773E-04	0.0011547262059637600	8.99788570444699	22.2314134513797	0.0031679809895494000	19.377692861830800	0.0038054345185931600	1.07725963137768E-04	1.8941316268188700	2.15608162188624E-04
35_44	1.6151054773681600	45.86279543350540	1.67707733272883E-04	7.97167259907759E-04	8.988689826568380	22.251115578975500	0.0023776474905943400	19.380898917220800	0.0028810851100245400	8.68278954850215E-05	1.8948834040195000	2.00926852943032E-04
45_54	2.5253942452835400	79.90541885559110	0.006543984821583900	0.03177412798497640	12.041459902820900	4.6088481783514000	0.09896422716082300	0.44472444410348200	0.1252690313778040	0.004357794546399200	0.1948496721832520	0.012395535774685900
55_64	3.2050202117707400	74.26493618227650	0.003796188734319500	0.01863601857494770	11.126930872335200	10.670033948995500	0.0649887317107134	0.3437538977443270	0.08867471268948220	0.0035246653330624500	0.19544718860875900	0.014257381226474100

65	4.408432420920660	60.9188939844813	0.03691490539886470	0.15807061474967900	18.235876260999700	9.487678038190280	0.5659567115029780	2.7650740810980800	0.7720379104501140	0.03706346297436550	2.4491905613674400	0.1648110478664830

general_combined_unique_relevant.csv

	Abortion	Candidate	Climate	Crime	Democracy	Economy	Education	Foreign Policy	Guns	Health care	Immigration	LGBTQ+	Military	Other	Welfare
18_24	1.12957528559234	4.627250672776050	2.44406402979329E-04	0.12264602745503300	3.719705712920190	3.8370124235523400	0.020889689019826500	0.010670983697640400	0.029385508140422400	85.74220618552010	0.7004993960221720	7.82088366554807E-05	2.80563604058831E-05	0.02420376249806600	0.03560368120575860
25_34	1.14804461795222	4.658688417688670	4.46586945028767E-04	5.15054518515517E-04	2.4771413383634300	3.8462524915699500	0.02128782327748840	0.0012269995084450100	0.024843281428894200	87.07361557575260	0.71122239063145300	1.38555765180062E-04	8.23086047685229E-05	1.63979142049559E-04	0.036330578851287900
35_44	1.1802705893414700	4.798241409172510	4.23930295860409E-04	0.006132279273201970	2.509247847555470	1.1357489781833200	0.0013716087566391200	0.0011785743075143400	0.02559741509804390	89.57240824853590	0.7316953518187160	1.02475358878843E-04	1.00264520409959E-04	1.39445510000946E-04	0.03734158227203760
45_54	1.2313639730472500	3.109719396305800	3.54207385429249E-04	0.006278183356428350	0.11305638522299400	1.1580324574513900	4.79732684323425E-04	0.0011773831775613800	0.026768936116497200	93.58456372678810	0.7642131894741410	6.53271073789873E-05	8.22896990368383E-05	8.43398746361137E-05	0.0037604723090632600
55_64	1.2342402875380900	3.113134649800990	3.62537869604501E-04	0.006176794078040370	0.03785736652578190	1.1584332288474900	4.14579278695338E-04	0.0011598078591661100	0.01841757105155510	93.66717311774210	0.7587471154374410	5.50575682715378E-05	6.0173116206592E-05	4.55949380067006E-05	0.0037221183486031700
65	3.9054551182855500	60.35668506715380	1.929944273764690	1.0716499004338400	9.097980306513000	10.79014323216400	0.5762846608069030	0.8309386929595060	2.4052947602026700	2.876826924495460	3.7557776595738200	0.3992652446571110	0.670096619014153	0.24272019685580300	1.0909373431197300

Appendix Table VIII: Proportional Breakdown of Impressions by Candidate Alignment & Age Group

official_trump_ads_clean.csv					
	Trump_0_Harris_3	Trump_1_Harris_0	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_3
18_24	1.13435527984201E-05	71.296126167724	14.255840479957700	3.07123146487127E-07	14.448021701642400
25_34	3.06089552658524E-05	72.11208400252380	13.532516245809400	1.67708454294049E-06	14.355367465627000
35_44	3.31309851727652E-05	82.57579189993580	1.8770011874478700	2.95172043524987E-06	15.547170829910700
45_54	3.40775009980354E-05	97.77029040636410	2.227836490345090	2.82908083861132E-06	0.0018361967089357500
55_64	2.77696640000447E-05	97.92664057917740	2.07003085086744	2.74044034742916E-06	0.0032980598508296200
65	0.13806065991855700	39.08117515098930	51.964203614349700	0.011939778460196400	8.804620796282240

official_harris_ads_clean.csv						
	Trump_0_Harris_1	Trump_0_Harris_3	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_1	Trump_3_Harris_3
18_24	10.4431490906691	48.357761057446300	2.51844951935086E-05	0	38.28924452052580	2.9098201468636500
25_34	10.563626111678800	48.69647606288590	2.18312749392237E-05	1.93569465388275E-08	38.77840983445640	1.9614661403470100
35_44	10.549091999736500	48.715401700218300	9.65838271884189E-06	2.783610883525E-08	38.7800608744034	1.9554357394229800
45_54	26.532534425049300	47.359979082501800	2.44321653593482E-04	2.36259481208427E-06	22.455469153952500	3.6517706542479700
55_64	18.02901585726110	54.54802701894760	8.6697278821534E-05	4.67371442442163E-06	23.743776071299800	3.6790896814982900
65	9.982427789826100	31.28119209030780	6.28821540212262E-04	9.72745945065062E-05	19.303671915326900	39.431982108404400

general_combined_unique_relevant.csv									
	Trump_0_Harris_0	Trump_0_Harris_1	Trump_0_Harris_3	Trump_1_Harris_0	Trump_1_Harris_1	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_1	Trump_3_Harris_3
18_24	1.26983619662657E-05	0.5455565937048920	38.20464598808730	1.2277301881656600	3.38019955963747E-09	0.1306216650140150	1.6332734220372600	52.87223587992930	5.385923561319320
25_34	1.44853999972326E-04	0.3602780269873880	37.68460899807520	1.2388245137857900	1.94871215487767E-08	0.12225309111440600	1.6094629849408100	53.63072559886110	5.353701912748270
35_44	1.70215086697922E-04	0.36400468416553200	38.750324764172900	1.2725039990491300	4.02653963167374E-08	0.12592743126772500	1.6568505180144600	52.35941168284910	5.470806665129020
45_54	1.56705031262416E-04	0.36329480843740000	37.8941617510023	1.319553879622750	8.78506463361728E-08	0.13171420653458000	1.734961701075680	53.316015012932200	5.240141847513150
55_64	1.55592183556461E-04	0.3566280857881740	37.89961849754610	1.3208580651988800	2.13638268637349E-07	0.1319060509220950	1.7334256777706900	53.3683689423077	5.189038874644580
65	0.4000670653238790	21.578808322946200	11.503960250437800	7.287177625861440	0.0038739206604000000	10.693607242861500	9.001253505279330	31.305511805341700	8.225740261287700

Appendix Table IX: Proportional Breakdown of Impressions by Emotional Tone & Gender

official_trump_ads_clean.csv					
	Angry	Fearful	Hopeful	Neutral	Patriotic
Female	0.01616759826720860	0.1549516469626050	52.60002785718970	0.04738045901277150	47.18147243856770
Male	1.590718930494430	10.552300137182000	37.805024917024600	4.879159700789620	45.17279631450930
Unknown	0.4161414764087250	82.67154162984000	5.703817239056590	9.35882777835397	1.8496718763406800

official_harris_ads_clean.csv					
	Angry	Fearful	Hopeful	Neutral	Patriotic
Female	70.23919783233880	4.946077935661790	24.696340684666600	0.10695558700744300	0.011427960325380000
Male	19.30981252825300	16.007275701471400	51.70976986384230	11.360997572575700	1.6121443338576300
Unknown	25.2041447180945	25.447658634094400	49.34596612114570	0.002155012575214320	7.55140901020852E-05

general_combined_unique_relevant.csv					
	Angry	Fearful	Hopeful	Neutral	Patriotic
Female	13.993699342392400	42.321573492405500	4.1729492084009200	36.4422619856038	3.069515971197380
Male	27.705762614000800	12.032172805366900	43.99683470908080	10.278444034077700	5.986785837473850
Unknown	3.991388623182490	36.887275708413300	58.44627789584140	0.5758079404217490	0.09924983214109270

Appendix Table X: Proportional Breakdown of Impressions by Central Theme & Gender

official_trump_ads_clean.csv					
	Candidate	Democracy	Economy	Other	Immigration
Female	99.71714675800040	0.2688012426872380	0.0011538867006518000	1.87276394237277E-04	0.012710836217495600
Male	76.71967041031850	22.07771967739750	0.12052059262801200	0.019500323189061900	1.0625889964668900
Unknown	6.921598101561510	83.71957607906270	1.16849538172661E-07	1.83082171573788E-08	9.358825684218030

official_harris_ads_clean.csv												
	Abortion	Candidate	Climate	Crime	Democracy	Economy	Foreign Policy	Healthcare	Immigration	Military	Other	Welfare
Female	4.893846381541390	23.93161058979280	4.32129023527909E-04	0.0020855519826731000	0.9546695979865790	64.36232465758590	0.006164554854439280	0.02733478535327750	0.007764142441638230	2.67552649498397E-04	5.812757711282620	7.4234550567987 5E-04
Male	7.415990171962010	48.73029480526630	0.052797238607701300	0.3939799462253450	18.768160637645400	15.669693111818200	1.2168186477753600	4.8236009891123900	1.3836914555192400	0.0370006175311059	1.433466583098740	0.07450579543814120

Unknown	0.2018660231277080	56.06289467798450	1.73784757688534E-06	1.51381283864241E-05	12.428183966880700	3.741057172828000	5.04160921471232E-05	27.39541835337970	5.77626571202489E-05	1.46959405265935E-06	0.17041338009613300	3.99013840588855E-05

general_combined_unique_relevant.csv															
	Abortion	Candidate	Climate	Crime	Democracy	Economy	Education	Foreign Policy	Guns	Healthcare	Immigration	LGBTQ+	Military	Other	Welfare
Female	0.09458896606817710	42.127134722743500	0.01301028222816130	0.025683851191403900	20.622348877943500	2.1028413457180400	0.5141660039122620	0.13053786320637100	0.07321616750572500	0.10417816319468600	34.17207214312690	0.0036662483709738500	0.0022054005636706200	0.003392136943767860	0.010957827282872200
Male	4.936175200059630	41.04963588115580	2.625107984793390	1.730740091955210	6.733377254726810	26.220091667161500	1.0195244867112800	1.4934105394460700	2.152421644179620	3.189044074661620	4.750775498757270	0.6103023642943180	0.8170009528477050	0.8634946897778960	1.8088976694718300
Unknown	1.2006307902178600	3.4690353170642200	4.26531630332731E-05	0.02939633232261000	1.5175641692757300	2.2707621781350600	9.80083471010298E-04	0.001034515849942900	0.02423021412806860	91.26694917302320	0.19023712649207500	1.76044622715427E-06	8.16954592891092E-07	0.005136095154720430	0.023998774301633900

Appendix Table XI: Proportional Breakdown of Impressions by Candidate Alignment & Gender

official_trump_ads_clean.csv					
	Trump_0_Harris_3	Trump_1_Harris_0	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_3
Female	0.0027046469734908300	47.21783117952540	0.574376130429505	2.11713152105066E-04	52.20487632991950
Male	0.2428453585173120	27.48868244632810	59.31771928379140	0.01769525841762890	12.933057652945600
Unknown	1.84381535523043E-07	82.88221849105500	7.586791749937260	1.16368945495108E-08	9.530989562989330

official_harris_ads_clean.csv						
	Trump_0_Harris_1	Trump_0_Harris_3	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_1	Trump_3_Harris_3
Female	6.394750089700960	69.22268692927540	4.4909198451631E-05	1.82352593688926E-07	18.355051111964100	6.027466777508520
Male	15.38951228338770	29.40807222678220	0.0046246423541775100	2.4715337833847E-05	29.569335721371500	25.628430410766700
Unknown	12.441434950704200	39.73866285294250	6.41229057781613E-07	1.71161251391557E-09	47.16600991062160	0.6538916427909970

general_combined_unique_relevant.csv									
	Trump_0_Harris_0	Trump_0_Harris_1	Trump_0_Harris_3	Trump_1_Harris_0	Trump_1_Harris_1	Trump_1_Harris_3	Trump_3_Harris_0	Trump_3_Harris_1	Trump_3_Harris_3
Female	9.71981346873673E-04	1.034516496786900	21.41676660367090	0.14879980460298700	5.15107741961643E-06	2.94824195975085	34.23162483875340	25.267890975052000	14.95118218895880
Male	0.576546968857552	19.740229271388400	14.882540431090200	5.090528457052990	0.0014485147559364300	7.6322336421553800	12.648801791093400	33.976427897996000	5.4512430256101500
Unknown	8.62534027550834E-05	0.38875225243093200	38.35968869539220	1.2926407516582300	1.19171403658267E-09	0.08251203146074360	1.1436080167400500	53.55799573887200	5.1747162588513600