5-9-2024

Essays on TV Advertising

Donggwan Kim
Washington University in St. Louis, donggwan.kim@wustl.edu

WASHINGTON UNIVERSITY IN ST.LOUIS

Olin Business School

Dissertation Examination Committee:
Raphael Thomadsen, Chair
Tat Y. Chan
Xiang Hui
Zhenling Jiang
Song Yao

Essays on TV Advertising
by
Donggwan Kim

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy in Business Administration

May 2024
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor, Raphael Thomadsen, for his guidance and support throughout my doctoral education. He has not only taught me how to conduct research but also how to do it thoughtfully and with care.

I would also like to thank Zhenling Jiang, who has been both a mentor and a co-author for many years. Her mentorship, encouragement, and generosity with time have greatly helped me prepare for and survive the job market.

Additionally, my gratitude goes to the rest of my dissertation committee, Tat Y. Chan, Xiang Hui, and Song Yao, for their insightful comments and constructive suggestions. I am also grateful to the marketing faculty at Olin, especially to Dennis Zhang, for always giving me invaluable suggestions.

Furthermore, I am thankful to Naser Hamdi and the research team at Equifax for introducing me to interesting topics in consumer finance and helping me develop research ideas.

Last but not least, I would like to thank my parents and my wife for their unconditional love and support. They have been my greatest source of strength.

Donggwan Kim

Washington University in Saint Louis

May 2024

Dedicated to my family.

ABSTRACT OF THE DISSERTATION

Essays on TV Advertising

by

Donggwan Kim

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2024

Professor Raphael Thomadsen, Chair

This dissertation consists of three chapters that attempt to understand the effects of both advertising quantity and content within the consumer finance and political domains.

In the first chapter, I examine the effectiveness of TV advertising in promoting mortgage refinancing. It is well-documented that many mortgage borrowers make sub-optimal financial decisions by failing to refinance when interest rates are low, despite large potential savings. This paper studies the effectiveness of TV advertising on mortgage refinancing decisions. Using six years of mortgage origination and advertising data, I quantify the impact of TV advertising on consumers' decisions to refinance their mortgages, as well as their choice of lenders. I find that TV advertising significantly increases the category-level demand for refinancing (i.e., market expansion effect) and influences the choice of lenders (i.e., business stealing effect). Interestingly, advertising can also benefit competing lenders through the market expansion effect. Beyond the analysis of ad quantity, I investigate the heterogeneous effects of ad content by leveraging zero-shot classification with a pre-trained large language model. I find that ads that highlight low mortgage rates and potential savings are the most effective in increasing both the category and brand-level demand. For policymakers and

mortgage lenders, these findings suggest that advertising can be an effective tool to promote refinancing activities and offer practical guidance for designing advertising content.

In the second chapter, I examine the impact of racial minority representation on advertising effectiveness. I do this by first assembling data on 10 million mortgage refinance loans, along with data on TV ads for mortgage refinance. I construct a measure of minority representation from video ads using computer vision techniques, and extract additional video and transcript features from the advertisements using a variational autoencoder and a text embedding model. I then apply a Double Machine Learning model to estimate how the minority representation in ads affects which lender consumers choose for their refinancing, while controlling for high-dimensional image and text features, as well as a rich set of fixed effects. I find that ads with higher minority representation are more effective in driving consumer choices: as the minority share in ads increases from 15% to 25%, the advertising elasticity increases from 0.037 to 0.042 (a relative increase of 14%). This effect is more pronounced among minority borrowers but is also positive among White borrowers. Across the political spectrum, minority representation has a larger impact among liberal-leaning consumers. In addition to the observational study, I conduct a pre-registered lab experiment ($N = 2,796$) where I manipulate the race of the actors using generative AI technology. The results are consistent with those from the observational study, providing further causal evidence. I discuss potential mechanisms driving these results, as well as the implications of the findings.

In the third chapter, I explore the relationship between the content of political advertising on television and ad effectiveness. Specifically, I investigate how slant – the extremeness of the message – and consistency with the candidate's primary campaign messaging in national ad buys relate to two measures of voter behavior: online word-of-mouth (WOM) and voter preference (captured through daily polls) for the candidates. Using data from the 2016 presidential election, I find that ad messages that are more (1) centrist and (2) consistent with a candidate's primary-election platform associate with increases in online WOM and voter preference for the candidate. I further find that consistency is more important in the early (pre-October) stages of the campaign. These results suggest that while there may be a benefit to candidates moderating their message after winning the primary election, they need to be careful about shedding their messaging from the primary election during the early stages of the general election. Additionally, these results enrich our understanding of the use of extreme messaging in political advertising, a phenomenon that is on the rise, by showing that it may have a cost of decreased candidate-related WOM and voter preference for the candidate.

Chapter 1: The Effects of TV Advertising and Ad Content on Consumer Financial Decisions: Evidence from Mortgage Refinancing

1.1 Introduction

There is ample evidence that consumers often make sub-optimal financial decisions in various contexts (e.g., Madrian and Shea, 2001; Sussman and Alter, 2012; Jiang et al., 2021). How to improve consumer financial decision making has been an important area of research, drawing attention from multiple fields, including economics, finance, and marketing. For example, consumers may benefit from information provision or nudges (e.g., Van Rooij et al., 2011; Beshears et al., 2015; Blaufus and Milde, 2021). In this paper, we examine the effectiveness of TV advertising, a prevalent marketing strategy, on the important financial decision of mortgage refinancing.

There is no doubt that the mortgage market is an important financial sector in the United States. With an outstanding balance of over $12.6 trillion in 2023, it is the largest category of consumer debt by far.[1] A unique feature of the U.S. mortgage market is that a vast majority of loans have fixed rates. Once these fixed rates are determined based on various factors, including prevailing interest rates at loan origination, the borrower's creditworthiness, and the loan amount, they remain unchanged for the entire loan term, typically 15 to 30 years. For mortgage holders, refinancing is the primary way to benefit from lower rates, either due to macroeconomic changes (e.g., a significant decline in prevailing interest rates) or changes in

[1] https://www.newyorkfed.org/microeconomics/hhdc. Accessed on April 14, 2024.

individual financial situations (e.g., an increase in one's credit score). The potential savings from refinancing can be substantial, with average annual savings estimated to be close to $3,000 (Freddie Mac, 2021; Agarwal et al., 2023).

Despite the substantial savings, many mortgage holders do not refinance. The low take-up rate can be attributed to several factors, including a lack of awareness, inattention, procrastination, and mistrust (e.g., Keys et al., 2016; Johnson et al., 2019; Andersen et al., 2020). In this paper, we study the impact of TV advertising on mortgage refinancing decisions. There are reasons to believe that TV advertising could increase refinancing demand by informing consumers about low interest rates or reminding them to take action. Investigating this question is particularly important, given the significant financial impact on mortgage holders and the broader policy implications. For example, if TV advertising is found to be effective, policymakers could leverage it to promote refinancing. Beyond analyzing ad quantity, we also seek to understand what types of advertising content are more effective in influencing refinancing decisions. The findings are practically relevant for designing effective advertising strategies and theoretically relevant for identifying important barriers to refinancing.

More specifically, this paper addresses two main questions. First, we investigate the impact of TV advertising on the category-level demand for refinancing (i.e., market expansion). Second, we examine the effects of TV advertising on borrowers' choice of lenders (i.e., business stealing). For both questions, we start by quantifying the overall impact of TV advertising and then study the heterogeneous effect of different advertising content.

To answer these research questions, we combine loan-level mortgage origination and TV advertising data. We collect the near universe of mortgage loans originated from 2016 to 2021 in the U.S. We merge the mortgage data with detailed TV advertising data, which includes ad expenditures as well as ad creatives (i.e., the video files of TV ads). Leveraging these videos allows us to identify the key content of the advertising messages. To estimate the causal effects of TV advertising, we employ the border strategy introduced by Shapiro (2018), which exploits discontinuities in local advertising at Designated Market Area (DMA) borders. The advertising effects are estimated by comparing demand in neighboring counties located on opposite sides of DMA borders, which are exposed to different levels of advertising.

Leveraging the border identification strategy, we find that TV advertising significantly increases category-level demand for mortgage refinancing (i.e., *market expansion effect*). Specifically, a 10% increase in advertising leads to a 0.82% increase in overall demand for refinancing. A simple back-of-the-envelope calculation indicates that a $1 increase in refinance ad spending leads to $27.3 in total interest savings, suggesting that advertising likely increases consumer welfare.

In addition to category-level demand, we examine how TV advertising influences consumers' choice of lenders. We find that a lender's advertising has a significant impact not only on its demand, but also the demand for other lenders. The positive spillover to competing lenders occurs because of the market expansion effect. However, conditional on borrowers who choose to refinance, we find that a lender's advertising significantly increases its market share (i.e., *business-stealing effect*). The results of significant market expansion and business-stealing effects are likely driven by the fact that ads by private lenders provide both information that can benefit all lenders in the industry (e.g., low interest rates) and information that primarily shifts brand preferences (e.g., brand-building messages).

Given the business stealing effect, one may be concerned about whether advertising leads consumers to refinance with more expensive lenders, which could harm consumer welfare. To address this, we examine the correlation between ad quantity and mortgage rate spreads,[2]

[2] The mortgage rate spread measures the difference between the annual percentage rate of a mortgage loan and the benchmark risk-free mortgage rate, taking into account both the interest rate and origination fees.

controlling for a rich set of borrower characteristics. We find that, although lenders who advertise more tend to have slightly higher rate spreads, the economic magnitude of this difference is very small. Within lenders over time, there is no statistically significant correlation between advertising and mortgage rate spreads.

Beyond analyzing the overall effectiveness based on ad quantity, we further explore the heterogeneous effects of different ad content. To identify the topics within advertising messages, we leverage zero-shot text classification with a pre-trained large language model, which has been shown to be effective in extracting key content from unstructured text data (e.g., Yin et al., 2019; Wei et al., 2021). At a high level, the semantic information of both the input documents (e.g., advertising messages) and candidate topics (e.g., topics about low interest rates or ease of application) is represented using text embeddings. These embeddings are then used to predict the probabilities of each document belonging to the candidate topics. This approach effectively leverages both rich information captured through from text embedding and theory-driven topics from prior research or domain knowledge.

We identify four candidate topics by building on prior literature (Perry et al., 2016) and domain knowledge. For each ad creative, zero-shot classification scores the probabilities that the ad aligns with each of the four candidate topics: low rate/savings, ease of application, homeownership, and brand-building. We then estimate the heterogeneous effects of different ad topics. We find that, conditional on ad expenditures, ads that highlight low mortgage rates and savings are more effective in driving market expansion at the category level than ads that focus on the convenience of online applications or brand-building. In the brand-level analysis, we continue to find that ads emphasizing low rates and savings are the most effective in shifting consumers' lender choices. While brand-focused ads also have a positive impact, it is not statistically significant.

This paper demonstrates the effectiveness of TV advertising in the mortgage refinancing market. The findings are particularly important, considering the well-documented issue of sub-optimal financial decisions regarding refinancing and the substantial amount of foregone savings (e.g., Keys et al., 2016; Johnson et al., 2019). For policymakers, our findings suggest that TV advertising can be an effective tool to promote refinancing, especially when ads highlight interest rates and savings. Interestingly, in line with our insights, we observe that Freddie Mac, a government-sponsored enterprise, launched TV advertising campaigns in 2021 that focused on low interest rate messages. For example, one of their ads says that "homeowners who refinanced their 30-year fixed rate mortgage last year will save on average $2,800 annually" and points out that "many households who could benefit from refinancing have not." However, the scale of Freddie Mac's advertising is significantly smaller compared to government-sponsored advertising in other important markets, such as health insurance (Aizawa and Kim, 2021). Our results suggest that increasing TV advertising efforts can further increase participation in refinancing. Furthermore, our ad content analysis provides practical guidance for advertisers in choosing more effective advertising topics.

Our research is related to three streams of literature. First, we contribute to the body of work on TV advertising effectiveness across various markets, especially those examining its impact on consumer welfare outcomes and policy implications. The border strategy approach has recently been utilized to identify the causal effect of TV advertising in important areas, such as hospital advertising and patient outcomes (T. Kim and Kc, 2020), antidepressant advertising and demand for prescriptions (Shapiro, 2022), political advertising and vote shares (Wang et al., 2018), and e-cigarette advertising and substitution away from traditional cigarettes (Tuchman, 2019). Our paper adds to this literature by investigating the impact of TV advertising on mortgage refinancing decisions, an area where consumers often make costly mistakes of not refinancing. In the same context, D. Kim et al. (2023) study how minority representation in ads impacts the ad effectiveness in the choice of lenders among mortgage holders who choose to refinance. Our paper differs by identifying the market expansion effect (i.e., whether to refinance) beyond the choice of lenders, which is arguably more important from both policy and consumer welfare perspectives.

Second, our research contributes to the growing literature on the effects of advertising content. Compared to the large literature on the impact of advertising quantity, studies focusing on TV ad content are relatively scarce, particularly outside of lab settings. Several papers have manually coded ads into informational or emotional content in automobile ads (Guitart and Stremersch, 2021), action, information, or emotion-focused content in ads for multiple categories (Liaukonyte et al., 2015), and informational or non-informational content in auto insurance ads (Tsai and Honka, 2021). Recent advances in machine learning techniques allow for automatic feature extraction from unstructured data. Yang, Xie, et al. (2021) use machine learning algorithms to measure the energy level in ads, and D. Kim et al. (2023) utilize computer vision techniques to identify the race of actors in ads. Our paper employs a large language model, fine-tuned for natural language inference tasks, to extract topics from unstructured ad texts. This approach allows us to analyze content in a scalable way and can be applied to other contexts where classifying unstructured text data is needed.

Third, our paper is related to the literature that seeks to address the low refinance take-up rate in the mortgage market. Previous studies have identified different sources of friction impacting mortgage refinancing rates, such as home equity constraints (Agarwal et al., 2023), and employment documentation and closing costs (DeFusco and Mondragon, 2020). Beyond these economic factors, psychological factors, such as inattention and inertia, have also been identified as contributing to low refinance participation (e.g., Keys et al., 2016; Andersen et al., 2020). To promote refinancing, Johnson et al. (2019) evaluate the impact of direct mail campaigns on refinancing using field experiments. Closely related to our study, Hu et al.

(2023) find that exposure to business TV channels increases refinancing by leveraging the staggered entry of Fox Business Network across zip codes. Our research adds to this literature by showing that TV advertising can be a cost-effective tool to promote mortgage refinancing.

The rest of the paper is as follows. We describe our data in Section 1.2. We present the category-level demand model and results in Section 1.3. We discuss brand-level analyses in Section 1.4. We then present the heterogeneous effects of ad content in Section 1.5. We discuss the implications of our results and conclude the paper in Section 1.6.

1.2 Data Description

In this section, we describe the two main datasets used in our study: mortgage origination data in Section 1.2.1 and TV advertising data in Section 1.2.2. In addition to these datasets, we also collect 30-year, fixed mortgage rates from the Federal Reserve Economic Data (FRED) database, county-level population and household income data from the U.S. Census Bureau, and county-level home price index data from the Federal Housing Finance Agency. After describing the data, we discuss how we classify TV ads into refinance or home purchase focused ads by using a supervised machine learning model in Section 1.2.3

1.2.1 Mortgage Origination Data

We collect loan-level mortgage origination data from the Home Mortgage Disclosure Act (HMDA) database for the period of 2016 to 2021. HMDA is a federal law in the U.S. that mandates mortgage lenders to disclose detailed information about their lending activities. The data collected under this law cover approximately 90% of all first-lien mortgage originations, with reporting exceptions for small lenders (Bhutta et al., 2017). This dataset provides detailed loan characteristics, including the originating lender, year of origination, property location, loan size, and loan purpose (e.g., home purchase or refinancing). Beginning in 2018, it also includes additional variables, such as interest rate, loan-to-value (LTV) ratio, and debt-to-income (DTI) ratio.

We focus on conventional mortgage loans (i.e., loans that are not part of government programs)[3], which are the most common mortgage type in the U.S. and account for about 80% of total mortgage originations (F. Liu et al., 2022). Following previous studies in household finance (e.g., Bartlett et al., 2022; Bhutta et al., 2022), we apply additional inclusion criteria: these loans must be first-lien mortgages for owner-occupied, site-built, single-family residential homes. Additionally, we exclude jumbo loans that exceed the conventional loan size limits and other "exotic" mortgages, such as reverse mortgages and mortgages with non-amortizing features (e.g., interest-only and balloon payments) or negatively amortizing features.[4] We then include loans originating from counties in the top 101 Designated Market Areas (DMAs), which account for 90% of the mortgage originations in our sample. We provide additional details about the data cleaning process and the number of excluded loans at each step in Appendix A.1.

Using the remaining mortgages in our sample, we present the annual loan origination volume by loan purpose in Figure 1.1, along with the yearly 30-year fixed mortgage rates. The volume of purchase loans remains relatively stable over time, whereas the volume of refinance loans significantly increases starting in 2018, coinciding with the steady decline in mortgage rates. This pattern suggests that refinancing is highly responsive to fluctuations in interest rates.

[3]Government-backed mortgage loans include Federal Housing Association (FHA) loans for low income or low credit score borrowers, Veterans Affairs (VA) loans for veterans or active-duty service members, and United States Department of Agriculture (USDA) loans for rural areas.

[4]Some of these variables, such as reverse, non-amortizing, and negatively amortizing features, are not available in 2016 and 2017. But loans with these features account for a small proportion (e.g., 1.48% in 2021).

Figure 1.1: Annual Mortgage Origination Volume by Loan Purpose

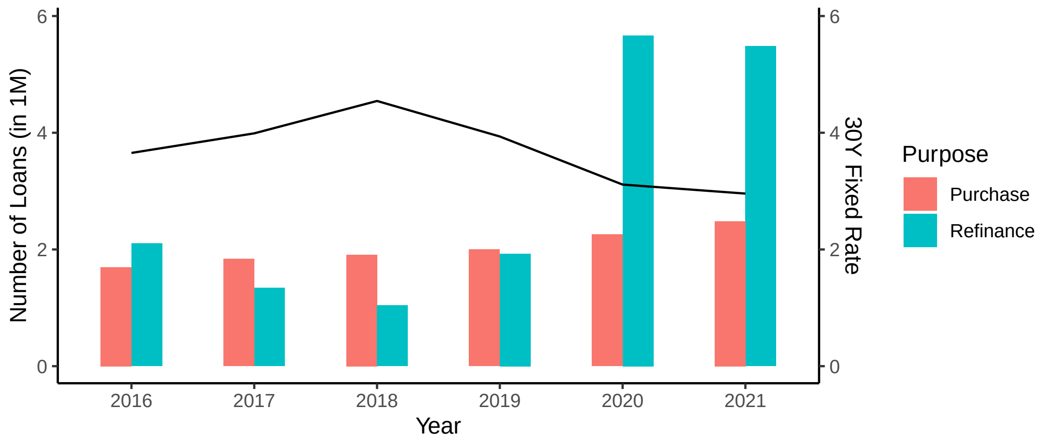


Table 1.1 presents the market share and ad spending of the top 30 lenders by refinancing volume.[5] Column (1) presents the market share in the refinance market. Although not a focus of the paper, we also show the market share in the purchase loan market (i.e., loans for home purchase) in column (2). Comparing the two columns, we see that non-banks (or fintech lenders), such as Rocket Mortgage, Loan Depot, and Amerisave Mortgage, tend to have higher market shares in the refinance market than in the purchase loan market. The mortgage market is relatively competitive, with the top four lenders accounting for about 24% market share in the refinance market and even lower in the purchase loan market.

1.2.2 TV Advertising Data

We obtain TV advertising data from Kantar Media for the same period covered by the mortgage data. This dataset provides detailed information at the ad creative level, including the advertiser, advertised product, ad creative name, year-month, media market (Designated Market Area or DMA), and total ad spending. We also obtain the actual ad creatives (i.e.,

[5]We exclude BB&T and SunTrust, which merged into Truist in December 2019. Because of the merger, there was a time period where they had separate advertising campaigns but reported to HMDA under the new name Truist, which created challenges in matching the advertising data with the loan origination data.

Table 1.1: Market Share and Advertising Spending by Lender per Year

	Lender	(1) Market share in refinance loans	(2) Market share in purchase loans	(3) Total ad spend per 100 capita
1.	Rocket Mortgage	12.30	3.49	$94.52
2.	Wells Fargo Bank	4.42	3.69	$22.83
3.	United Wholesale Mortgage	3.83	3.34	$0.49
4.	JP Morgan Chase Bank	3.12	1.92	$0.10
5.	Loan Depot	2.85	1.40	$7.68
6.	Bank of America	2.19	1.43	–
7.	Nationstar (Mr. Cooper)	1.97	0.07	$1.80
8.	U.S. Bank	1.66	1.11	$6.26
9.	Caliber Home Loans	1.22	1.91	$0.004
10.	Flagstar Bank	1.13	1.03	$1.12
11.	PNC Bank	1.03	0.49	$2.44
12.	Guaranteed Rate	0.97	1.46	$7.67
13.	Fairway Independent Mortgage	0.85	2.25	$13.07
14.	Provident Funding	0.81	0.27	–
15.	Amerisave Mortgage	0.78	0.08	$6.99
16.	Penny Mac	0.75	0.20	$0.17
17.	Freedom Mortgage	0.74	0.34	$0.10
18.	Home Point Financial	0.70	0.61	–
19.	Finance of America	0.67	0.73	$1.05
20.	NewRez Mortgage	0.64	0.22	–
21.	Guild Mortgage	0.61	1.09	$1.01
22.	Fifth Third Bank	0.61	0.42	–
23.	Better Mortgage	0.58	0.24	$0.01
24.	Cardinal Financal	0.57	0.37	$0.40
25.	Broker Solutions	0.55	0.54	$0.24
26.	Citizens Bank	0.53	0.60	$4.94
27.	Regions Bank	0.46	0.43	$13.49
28.	CrossCountry Mortgage	0.44	0.79	$5.54
29.	Huntington National Bank	0.44	0.49	–
30.	Prime Lending Mortgage	0.41	1.13	$1.62
	Total	47.83	32.14	$193.56

Notes. Total ad spending per 100 capita is calculated as the sum of both national and local ad spending per 100 capita within the top 101 DMAs. National and local ads are scaled by the respective population sizes.

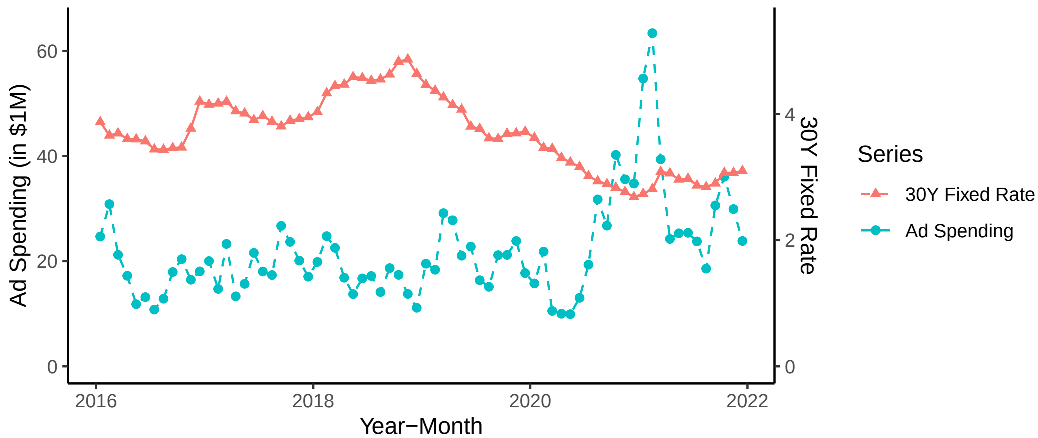
the video files for the TV ads), which allow us to observe the content of each ad and explore how different topics impact ad effectiveness, conditional on spending levels. Note that the advertising data do not distinguish between ads for refinancing vs. purchase loans, and many ads talk about both types. We describe how we classify ads by target borrowers and extract content from ad videos in the subsequent sections.

There are over 13,000 unique ad creatives that aired either nationally or locally within the top 101 DMAs in our dataset, which includes all ads in the broader mortgage industry. We exclude advertisers that specialize in some specific mortgage products that are not relevant to our study, such as reverse mortgage lenders, mortgage foreclosure and modification servicers, and lenders specializing in non-conventional, government-backed loans.[6] After this data cleaning process, we are left with 8,355 unique ads from over 900 lenders, which account for 76.6% of the total TV ad spending in the mortgage industry.

Following prior literature, we scale local ad spending by the population of the corresponding DMA to compute ad spending per capita. Similarly, national ad spending is scaled by the total U.S. population. This adjustment allows us to compare levels of advertising exposure across DMAs with different population sizes. The total ad spending for a DMA is the sum of national and local ad spending per capita within that DMA. Column (3) of Table 1.1 presents the average annual ad spending per 100 capita within the top 101 DMAs for the top 30 lenders. Comparing ad spending across the top lenders, we observe that there are significant variations across lenders. The two largest lenders, Rocket Mortgage and Wells Fargo Bank, also advertise the most. Several major lenders, on the other hand, did not advertise at all during our sample period.

[6]We exclude around 2,350 ads from reverse mortgage lenders (e.g., American Advisors Group), 220 ads from mortgage foreclosure and modification servicers (e.g., American Mortgage Assistance), 1,920 ads from lenders specializing in VA loans (e.g., NewDay USA), and 500 ads for the Home Affordability Refinance Program (HARP).

Figure 1.2: Monthly TV Advertising Spending and Mortgage Rate



We present the total monthly ad spending across all lenders within the top 101 DMAs in Figure 1.2. Overall, we observe a substantial amount of mortgage advertising, with an average expenditure of \$21.8 million per month. This is comparable to ad spending in the health insurance market (Shapiro, 2020; Aizawa and Kim, 2021). While ad spending is quite steady from 2016 to 2019, there is a significant increase during 2020−2021, coinciding with a period of historically low prevailing interest rates. We overlay the monthly 30-year fixed mortgage rate in the figure and find that mortgage ad spending is negatively correlated with the mortgage rate (correlation coefficient $= -0.49$). This correlation shows the classic endogeneity in advertising−lenders tend to increase ad spending when they expect higher demand due to low rates−highlighting the importance of accounting for time trends when modeling the impact of mortgage advertising on refinancing demand.

1.2.3 Classification of Ads by Loan Types

As mentioned in Section 1.2.2, the advertising data do not differentiate between ads that focus on refinance vs. purchase loans. Through manually examining multiple ads, we find that some ads are exclusively about refinance or purchase loans, while others mention both types. To precisely measure the impact of refinance advertising on the demand for refinancing, we need to classify the ads by loan types.

We do so by using a supervised machine learning approach. First, we create labels by identifying ads that exclusively focus on refinance or purchase loans through relevant keywords. Using this labeled data, we train a supervised model that predicts these labels based on the text embeddings of the ads. Lastly, we apply the trained model to all ads to obtain the predicted likelihood of each ad for targeting the two types. We describe each step in more detail below.

To classify ads by loan types, we utilize both ad creative names and ad transcripts. The ad creative names can provide some insights into whether the ads target refinance borrowers (e.g., "refinance and save hundreds") or purchase loan borrowers (e.g., "buying dream house"). In addition, we obtain the transcripts of the video ads using the Transcribe API on Amazon Web Services. The transcriptions are overall highly accurate, and we manually correct a few minor errors (e.g., "buy the house" transcribed as "by the house").

We create training labels using a keyword-based approach. For each ad (i.e., ad creative name and transcript), we identify whether it contains key phrases that indicate refinancing (e.g., "refinance," "lower your current rate") or home purchasing (e.g., "buying a home," "looking for a new home"). The complete list of the keywords/phrases is available in Appendix A.2. With this keyword approach, we find that 24.4% of ads contain only refinance-related phrases, and 29.7% contain only purchase-related phrases. These two types of ads form the training dataset (thus excluding the 13.3% containing both refinance and purchase phrases and the 32.6% containing neither).

We train a random forest model to classify ads as either refinance or purchase using the labels identified through the keyword approach. For input features, we use OpenAI's pre-trained model, "text-embedding-3-large" to represent each ad with embeddings.[7] To prevent underfitting or overfitting, we perform a grid search to fine-tune hyperparameters (including the number of trees, the number of features at each split, and the depth of each tree) and select the set that achieves the highest 5-fold cross-validation F1 score of 0.95.

The model learns a mapping from text embeddings to labels based on a subset of ads. The trained model is then applied to all ads, including those containing both or neither types of phrases. For the model to generalize beyond the training data, the input data should ideally be similar. We examine ads without training labels and find no notable differences in content compared to those with labels, suggesting that the model's predictions should remain effective outside the training data. Although there is no formal test on the unlabeled set, we show a few examples in Appendix A.3, which suggest that the model predictions work well. After this process, we obtain the predicted probabilities of targeting refinance vs. purchase loan borrowers for all ads, which are more appropriate in our context than a binary "hard" classification, because some ads are indeed related to both types. Finally, the refinance (purchase) ad spending is calculated by multiplying the ad spending amount by the probability of the ad targeting refinance (purchase) loan borrowers.

1.3 Advertising Effects on Category-level Demand

In this section, we begin by discussing our identification strategy, which allows us to isolate advertising effects from unobservables, and provide suggestive evidence supporting its validity

[7] We have experimented with other models, specifically "text-embedding-3-small" and "text-embedding-ada-002" from OpenAI, and find that our main findings are robust.

in the context of our study in Section 1.3.1. We then present our empirical model to estimate TV advertising effects on category-level demand and discuss the results in Section 1.3.2.

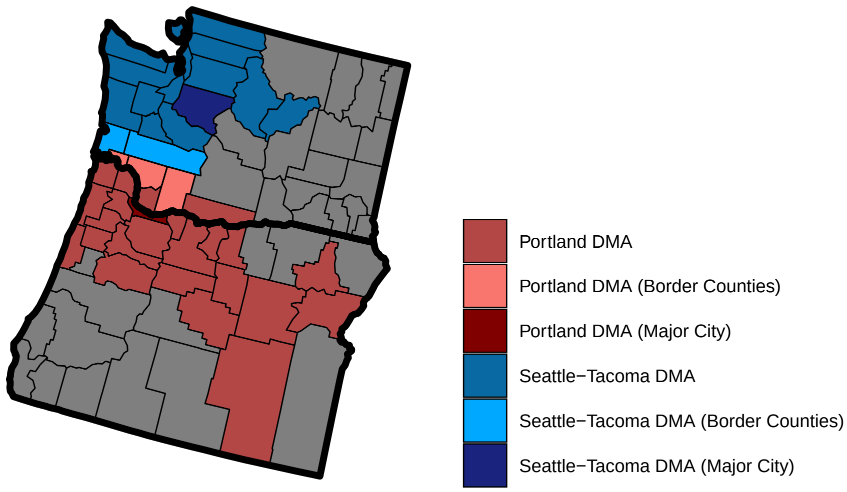
1.3.1 Identification Strategy

Identifying advertising effects is typically challenging because advertising decisions are not random. As shown in Figures 1.1 and 1.2, lenders tend to increase their advertising during periods of lower interest rates, which also coincide with higher expected demand. They may also target certain markets based on demand factors unobserved to researchers, or coordinate TV advertising with other marketing activities.

We employ the widely applied DMA border strategy to account for the endogeneity of advertising (e.g., Shapiro, 2018; Spenkuch and Toniatti, 2018; Wang et al., 2018; Tuchman, 2019; Shapiro, 2020; Tsai and Honka, 2021; Yang, Lee, and Chintagunta, 2021; Shapiro, 2022).[8] The border strategy exploits discontinuities in local advertising at DMA borders, created by the delineation of TV markets. To illustrate, consider the DMA border between the Seattle-Tacoma DMA (in red) and the Portland DMA (in blue) in Figure 1.3. Two counties from the Seattle-Tacoma DMA (in light blue) and three border counties from the Portland DMA (in light red) are located along the border, and are denoted as "border counties." Since advertisers purchase local ads at the DMA level rather than the county level, their advertising decisions are likely driven by local demand conditions in the major city of each DMA (in dark blue and dark red, respectively), where a significant share of the DMA's population resides. Consequently, the border counties can be exposed to different and quasi-random levels of advertising.

[8]Although there are alternative empirical strategies, such as Thomas (2020) and Li et al. (2024), that use "preference externality" based instrumental variable approaches, we believe it is challenging to find a suitable instrumental variable in our setting, especially for our brand-level demand analysis.

Figure 1.3: DMA Border between Seattle-Tacoma and Portland



We focus on within-state borders to avoid potential state-level policy changes that could impact mortgage demand and to ensure that borrowers on both sides of DMA borders have access to similar lenders. Among the top 101 DMAs, we identify 152 within-state DMA borders, 12 of which are excluded from the analysis for various reasons. For example, the San Diego DMA includes only one county, and all counties in the Springfield-Holyoke DMA border the Boston DMA, making the definition of a border county inapplicable. Moreover, some borders, such as between the Los Angeles DMA (Esmeralda County, NV) and the Las Vegas DMA (Nye County, NV), have a small population with very few mortgage originations.

For the remaining 140 within-state DMA borders, following Tuchman (2019), we further aggregate loan originations on either side of the border to the "border market" level due to the small number of loans in some border counties. Taking Figure 1.3 as an example, loans originating from the two border counties in the Seattle-Tacoma DMA are combined as one border market, and those from the three border counties in the Portland DMA are combined as another border market. This aggregation results in a total of 280 border markets. We would be worried if a significant share of the DMA's population lives in the border markets because, in that case, advertising levels could possibly be set based on the demand within these border markets. However, this is not the case in our data: on average, a border market represents only 12.8% of the total DMA population. Thus, advertising decisions are unlikely to be driven by these border markets.

Our key assumption in the border strategy approach is that, in the absence of differences in advertising, border markets would have followed similar trends in mortgage refinancing demand. Although not directly testable, we examine the parallel trends in two important variables that may influence the refinancing demand: income and home price index (HPI). Both variables can affect refinancing eligibility through debt-to-income ratio (DTI) and loan-to-value ratio (LTV). We collect annual income and HPI growth per county each year and calculate the weighted average for each border market using the county-level population as the weight. We find that both measures are highly correlated between neighboring border markets. For annual income, the mean and median correlation coefficients are 0.85 and 0.90. For HPI growth, the mean and median correlation coefficients are 0.77 and 0.89. Further details can be found in Appendix A.4. Overall, these high correlations lend support to the assumption that refinancing eligibility (and therefore demand) would have followed similar trends without differences in advertising. With this assumption, we can estimate the causal effects of advertising by comparing how demand in border markets changes differently when exposed to different levels of advertising.

1.3.2 Category-level Demand

We start by estimating the impact of TV advertising on refinancing demand at the category level (i.e., regardless of lenders). The regression model is specified as in Equation 1.1. Let m denote the border market and t denote the year. The term $b(m)$ represents the border associated with market m. Recall that two markets from different DMAs share a border.

Leveraging the border strategy, we control for two sets of fixed effects: the border market fixed effects a_m, which capture persistent differences in demand across border markets, and border-year fixed effects $T_{b(m),t}$ to account for time-varying, local demand shocks.

$$log(Q_{m,t}) = \beta \cdot log(AD_{m,t}) + a_m + T_{b(m),t} + e_{m,t}. \tag{1.1}$$

The dependent variable, $log(Q_{m,t})$, is the logarithm of the number of refinance mortgages originating in border market m in year t. $log(AD_{m,t})$ is the logarithm of the ad spending per 100 capita in border market m in year t. The ad spending measure combines both national and local ads in the border market. To account for the time it takes to close a refinance loan, we calculate ad spending for year t with a one-month lag.[9]

We present the results of Equation 1.1 in Table 1.2 column (1), with the standard errors clustered at the border market level. Each observation represents the corresponding number from a border market in a year, with the number of observations at $280 \times 6 = 1680$. Results suggest that across DMA borders, markets with higher advertising levels tend to have more refinance mortgages. Therefore, TV advertising leads to a significant increase in the category-level demand for refinancing (i.e., market expansion effect).

One potential concern with the market expansion effect in the mortgage refinance context is the possible correlation between advertising and lenders' underwriting criteria. If lenders become more lenient in approval decisions with increased advertising, the increase in loans can come from a higher loan approval rate instead of advertising effectiveness. This could lead to an overestimation of advertising's impact. This concern, however, is unlikely to occur in our context because we focus on conventional, conforming loans with fairly standardized

[9]Since the mortgage origination data is available only at the annual level, we choose to use the advertising flow instead of building an advertising stock model, which is better suited for data with finer time granularity.

Table 1.2: Ad Effects on Category-level Demand for Refinancing

Dependent Variable:	$log(Q^{Refi})$		
	(1)	(2)	(3)
$log(AD)$	0.101**		
	(0.045)		
$log(AD^{Refi})$		0.082***	0.092***
		(0.027)	(0.032)
$log(AD^{Purch})$			-0.041
			(0.055)
Border Market FE	Y	Y	Y
DMA Border-Year FE	Y	Y	Y
N	1,680	1,680	1,680

Notes. S.E.s are clustered at the border market level; $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

loan qualification criteria. To further address this concern, we conduct a robustness analysis by estimating Equation 1.1 using data on loan applications, which include loans that were approved but not accepted, as well as denied loans, in addition to the originated loans. The results are robust to using this alternative sample (Table A.5.1 in Appendix A.5).

To contrast the effect on refinance mortgages, we run the same regression with the logarithm of the number of home purchase loans as the dependent variable. We find a small and statistically insignificant effect of TV advertising on purchase loans, at -0.0003 (s.e. 0.039). This is likely because home purchase is generally a much more deliberate and involved decision compared to refinancing. Consequently, consumers are much less likely to decide to buy a new home (and get a purchase loan) because of TV mortgage advertising. The very small coefficient also suggests that the border strategy effectively addresses the endogeneity of advertising without conflating the increased demand for purchase loans with higher advertising, both of which could be driven by low rates.

In columns (2) and (3) of Table 1.2, we further leverage the ad classification described above to distinguish between refinance and purchase loan advertising. Recall that refinance ad spending is calculated as the ad spending multiplied by the probability of the ad targeting refinance borrowers. Specifically, in column (2), the logarithm of total ad spending is replaced with the logarithm of refinance ad spending. We find that refinance advertising leads to a significant increase in refinance mortgages. Compared to column (1), the estimate is more precise (i.e., a smaller standard error), when focusing on just refinance advertising instead of overall mortgage advertising. In column (3), we further add the variable for the logarithm of purchase ad spending. We find that the coefficient on refinance advertising remains robust with the inclusion of purchase loan advertising.

The estimated TV advertising elasticity ranges from 0.082 to 0.092 using results from columns (2) and (3). We compare our estimated elasticity in mortgage refinancing with those documented in prior literature across different products. Shapiro et al. (2021) find an average elasticity of 0.023, with the 90th percentile at 0.092, among 288 consumer packaged goods. Across different industries, we see advertising elasticity at 0.027 in the health insurance market (Shapiro, 2020), 0.05-0.06 for satellite TV operator services (Yang, Lee, and Chintagunta, 2021), 0.08 for e-cigarette advertising (Tuchman, 2019), and 0.14 for statin drug advertising (Sinkinson and Starc, 2019). Overall, the estimated advertising elasticity in mortgage refinancing is comparable to those identified in the other markets.

We interpret the economic magnitude of the 0.082 advertising elasticity estimate. There are about 2.9 million refinance loans per year within the top 101 DMAs in our data. Assuming the advertising elasticity applies beyond border counties, a 10% increase in refinance advertising, which costs about $7.7 million, would lead to an additional 24,000 refinance loans. With the average net present value of savings per loan at $8,719 (Johnson et al., 2019), the increase in refinancing translates to total interest savings of around $209 million. Therefore, a $1 increase in refinance advertising will lead to $27.3 in interest savings from refinancing for mortgage holders.

The significant interest savings from relatively low advertising costs underscores the potential of TV advertising as an effective policy tool to increase refinance activities. Since refinancing provides not only private benefits through household savings, but also social benefits through increased household consumption (e.g., Abel and Fuster, 2021; Agarwal et al., 2023), the government may find TV advertising to be a useful tool to enhance the transmission of monetary policy, particularly during periods of low interest rates. Although not examined in our study, government-sponsored advertising might be more effective due to its potentially higher credibility than advertising by private lenders.

1.4 Advertising Effects on Brand-level Demand

So far, we have shown that mortgage advertising leads to a significant increase in the overall category-level demand for refinance loans. The analysis relies on aggregating the advertising spending as well as the refinance demand across lenders. In this section, we shift the focus to a brand-level analysis by estimating how the demand of a lender is impacted by its own advertising as well as advertising by other lenders. This brand-level analysis (Section 1.4.1) complements the category-level analysis above, but provide further insights into how advertising affects demand in this market. We then examine the correlation between advertising and loan price to address a potential concern that branded advertising might lead borrowers to obtain loans from more advertised, and possibly more expensive lenders in Section 1.4.2.

1.4.1 Advertising Effects on Brand-level Demand

We now proceed to a brand-level analysis by disaggregating data at the lender level. Consistent with the category-level analysis, m denotes the border market, $b(m)$ denotes the border associated with market m, and t denotes the year. For the brand-level analysis, we additionally use j to represent a lender. We control for the same two sets of fixed effects, but now applied at the lender level: the lender-market fixed effects $a_{j,m}$, which capture persistent differences in demand for lender j across border markets, and lender-border-year fixed effects $T_{j,b(m),t}$ to account for time-varying, local demand shocks specific to lender j. The regression is specified in Equation 1.2:

$$log(1 + Q_{j,m,t}) = \beta_1 \cdot log(1 + AD_{j,m,t}) + \beta_2 \cdot log(1 + AD_{-j,m,t}) + a_{j,m} + T_{j,b(m),t} + e_{j,m,t}. \tag{1.2}$$

The dependent variable, $log(1 + Q_{j,m,t})$, is the logarithm of 1 plus the number of refinance loans originated by lender j in border market m in year t. We add 1 to $Q_{j,m,t}$ to account for a small number of cases where the number of loans is zero. $log(1 + AD_{j,m,t})$ is the logarithm of 1 plus the refinance ad spending per 100 capita by lender j in border market m in year t. Similarly, $log(1+AD_{-j,m,t})$ is the logarithm of 1 plus the refinance ad spending by all lenders except lender j. We add 1 to advertising values to account for cases with zero advertising. By including other lenders' advertising, we allow for the potential spillover effect of advertising. The key parameters of interest, β_1 and β_2, capture the effects of lender $j's$ own advertising and advertising from other lenders on its demand, respectively.

To estimate the model, we use a dataset with observations at the lender, border market, and year level. We focus on the top 30 lenders as shown in Table 1.1. The total number of observations is 40,076, which is smaller than $240 \times 6 \times 30 = 43,200$, because not all lenders are available in every market. More specifically, we require a lender to originate at least one

Table 1.3: Ad Effects on Brand-level Demand for Refinancing

Dependent Variables:	$log(1+Q_j^{Refi})$		$log(s_j/s_0)$
	(1)	(2)	(3)
$log(1+AD_j^{Refi})$	0.095**	0.094**	0.074**
	(0.038)	(0.038)	(0.034)
$log(1+AD_{-j}^{Refi})$		0.033**	
		(0.015)	
Lender-Border Market FE	Y	Y	Y
Lender-DMA Border-Year FE	Y	Y	Y
N	40,076	40,076	40,076

Notes. S.E.s are clustered at the lender - border market level; $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

loan on both sides of a DMA border to be included in the data. When calculating advertising by other lenders AD_{-j}^{Refi}, we include all lenders, not just the top 30 lenders.

Table 1.3 columns (1) and (2) present the estimates from Equation 1.2. Standard errors are clustered at the lender-border market level. In column (1), we focus on only lender $j's$ refinance advertising. The result suggests that a lender's own advertising leads to a significant increase to its refinance demand. In column (2), we add advertising from other lenders as well. We find that there is a positive spillover from other lenders' refinance advertising. The magnitude of the spillover effect, as expected, is not as large as the impact of a lender's own advertising.

We calculate the advertising elasticities using the coefficient estimates, which do not directly represent elasticities due to the addition of 1 in the logarithm. The own advertising elasticity can be calculated as $\widehat{\beta}_1 \cdot \frac{AD_j^{Refi}}{1+AD_j^{Refi}} \cdot \frac{Q_j^{Refi}+1}{Q_j^{Refi}}$. Taking the sample averages for the refinance ad spending and demand, AD_j^{Refi} and Q_j^{Refi}, which are 0.727 and 72.3, respectively, the own advertising elasticity is 0.041. Similarly, the demand elasticity of others' advertising can be calculated as $\widehat{\beta}_2 \cdot \frac{AD_{-j}^{Refi}}{1+AD_{-j}^{Refi}} \cdot \frac{Q_j^{Refi}+1}{Q_j^{Refi}}$. Again using the sample averages for AD_{-j}^{Refi} and Q_j^{Refi}, which are 21.1 and 72.3, respectively, the elasticity of others' advertising is 0.032.

Calculating the own and others' advertising elasticities separately allows us to "decompose" the overall category-level advertising elasticity. In the mortgage refinance market, a lender's advertising increases the category-level demand by affecting not only its own demand but also the demand for other lenders through a substantial advertising spillover effect.

The brand-level analysis so far has focused on quantity as the dependent variable, with results showing a positive impact on the demand of both the advertised lender and other lenders. Since the elasticity is higher for the lender's own advertising than the spillover from others, advertising may have a business-stealing effect conditional on borrowers choosing to refinance. In the following analysis, we investigate whether advertising shifts the choice of lenders by using market share as the dependent variable. More specifically, we estimate the extent to which advertising changes the market share of the advertised lender.

We estimate how the advertising of lender j, $AD_{j,m,t}$, impacts its market share in border market m and year t, denoted as $s_{j,m,t}$. We continue to focus on the top 30 lenders in this analysis, and the number of observations remains the same as in the previous analysis. To calculate each lender's market share, we include all loans originated in market m in year t, including those from non-top 30 lenders. Loans from outside the top 30 are grouped into the "outside option," and their market share is denoted by $s_{0,m,t}$. The regression is specified as in Equation 1.3:

$$\log(s_{j,m,t}) - \log(s_{0,m,t}) = \beta \cdot \log(1 + AD_{j,m,t}) + a_{j,m} + T_{j,b(m),t} + e_{j,m,t}. \tag{1.3}$$

Estimating this regression is equivalent to estimating a logit model of lender choice with an individual-level utility function after applying the Berry market share inversion (Berry, 1994).

Let $\delta_{j,m,t} = \beta \cdot \log(1 + AD_{j,m,t}) + a_{j,m} + T_{j,b(m),t} + e_{j,m,t}$. The utility of consumer i refinancing with lender j can be written as

$$u_{i,j,m,t} = \delta_{j,m,t} + \epsilon_{i,j,m,t},$$

where $\delta_{j,m,t}$ is the mean utility that captures the advertising effect as well as other lender-level characteristics such as price and brand reputation. The mean utility of the outside option $\delta_{0,m,t}$ is normalized to 0.

Assuming the error term that varies across individuals, $\epsilon_{i,j,m,t}$, follows a Type I Extreme Value distribution, the market share of lender j has a closed form expression:

$$s_{j,m,t} = \frac{\exp(\delta_{j,m,t})}{1 + \Sigma_{j' \in J_{m,t}} \exp(\delta_{j',m,t})},$$

where $J_{m,t}$ is the choice set of the top 30 lenders. As described above, the choice set can vary by border and year because not all top 30 lenders are available in every market each year. The share of the outside option (non-top 30 lenders) is

$$s_{0,m,t} = \frac{1}{1 + \Sigma_{j' \in J_{m,t}} \exp(\delta_{j',m,t})}.$$

It is easy to see that taking the logarithm of the ratio of the two market shares, $log(s_{j,m,t}/s_{0,m,t})$, gives Equation 1.3, which we can estimate with linear regression.

Column (3) of Table 1.3 shows the results.[10] Standard errors are clustered at the lender-border market level. The results show that, conditional on consumers who choose to refinance, lender $j's$ refinance advertising significantly increases its market share relative to the shares of other

[10]Since the market share is zero for a small number of cases, we add a small number, e^{-5}, to the dependent variable so that the logarithm is defined. The results are robust to using other small numbers.

lenders (i.e., business-stealing effect). Similar to the category-level analysis, we conduct robustness checks using loan application data to address the potential correlation between advertising and loan approval decisions. We find that the results are robust and similar in magnitude (Table A.5.2 in Appendix A.5). We calculate the advertising elasticity based on market share as $\widehat{\beta} \cdot \frac{AD_j}{1+AD_j} \cdot (1 - s_j)$. With the sample means for AD_j and s_j at 0.73 and 0.016, respectively, the advertising elasticity is 0.031, which is comparable to the own advertising elasticity when using quantity as the dependent variable.

Taken together, the results show that refinance advertising expands the market for all lenders, including non-advertised lenders, through positive spillover. Conditional on refinancing, however, advertising increases the market share of the advertised lender. Therefore, both market expansion and business-stealing effects occur in the mortgage refinance market. The results are likely driven by the fact that advertising by private lenders provides both information that is applicable to all lenders and brand-specific information that only shifts brand choices. This finding is broadly aligned with Shapiro (2018) and Sinkinson and Starc (2019), both of which find significant market expansion and business-stealing effects of advertising in the pharmaceutical industry, specifically in the contexts of prescription antidepressants and anti-cholesterol drugs.

Using the advertising elasticity estimate, we explore the advertising return on investment (ROI) for the top 30 lenders. Using the own advertising elasticity of 0.041 from Equation 1.2, a 10% increase in refinance advertising leads to an additional 198.1 refinance loans per lender (the average number of loans per lender in the top 101 DMAs 48,312.9 $\times$ 10% $\times$ 0.041). These additional loans lead to an increase of about \$1.91 million in revenues or \$0.34 million in profits per lender,[11] which are higher than the additional advertising cost of about \$0.20

[11]To approximate margins, we collect annual revenues and profits per loan from the Mortgage Bank Association, which are \$9,310 and \$1,939.2, respectively, during our sample period. Revenues per loan include fees and net secondary marketing incomes (e.g., gains from loan sales and servicing rights). Profits per loan

million per lender. These numbers suggest that TV advertising brings a positive ROI in the mortgage refinance market, unlike consumer packaged goods, where the ROI is found to be mostly negative (Shapiro et al., 2021).

1.4.2 Impact of Branded Advertising on Mortgage Costs

Given the substantial savings from refinancing (Agarwal et al., 2023) and our findings on the market expansion effect, TV advertising could significantly increase consumer welfare. The brand-level analysis, however, shows a business-stealing effect where advertising increases the market share of the advertised lender. This could potentially hurt consumer welfare if advertising shifts consumer choices from less expensive to more expensive lenders. This is a reasonable concern, given that in the prior literature, Gurun et al. (2016) find that the reset rates (after the initial fixed rates) on adjustable-rate mortgage (ARM) loans tend to be higher among heavily advertised lenders. Following their strategy, we examine the correlation between advertising and loan expensiveness in our context.

In this analysis, we collect loan-level data from the HMDA for the period from 2018 to 2021, during which detailed information on interest rates and other key variables are available is available. In addition to the filters described in Section 1.2.1, we focus on 30-year, fixed-rate mortgages to ensure comparibility among loans. We then select loans originated by the top 30 lenders within the top 101 DMAs. We use the full dataset rather than restricting our sample to border counties for a larger sample size. We exclude a small number of loans with missing data on the key variables, such as LTV and DTI, as well as those with prepayment penalties. Lastly, we winsorize the key variables at the 0.1% and 99.9% levels, because the HMDA data includes a small number of extreme outliers (e.g., rates exceeding 100%).

are derived from revenues by deducting sales and non-sales costs (e.g., commissions), production support costs, and corporate expenses.

Table 1.4: Loan Level Summary Statistics ($N = 3,972,682$)

	Mean	Std. Dev.	25th	Median	75th
Rate spread (APR - APOR)	0.23	0.38	-0.02	0.18	0.41
Income (in $1,000)	112.24	70.86	65	96	140
Loan size (in $1,000)	311.34	149.96	195	285	405
Combined loan-to-Value (CLTV)	66.27	15.22	57.68	69.24	78.00
Debt-to-income (DTI)	33.53	10.57	25	33	42
Discount points	0.55	0.80	0	0.07	0.95
Lender credits	0.16	0.37	0	0	0.10
Refi. ad spend per 100 capita	3.99	6.81	0	0	9.02

Table 1.4 presents the loan-level summary statistics. The average rate spread, defined as the difference between the annual percentage rate (APR) and the average prime offer rates (APOR), is 0.23 percentage points or 23 basis points. This means that, on average, borrowers are offered APRs that are 23 basis points higher than the rates offered to low-risk mortgage borrowers. We use rate spread over interest rate as our measure of loan expensiveness, because APR includes interests and other costs like origination fees. Debt-to-income ratio (DTI) in the HMDA data is typically reported as integers (e.g., 40%) but can be reported as ranges when the data is sparse (e.g., 30% - 36%). For summary statistics, we calculate the mid-point whenever DTI is reported as a range. Each discount point, equal to 1% of the loan amount, typically lowers the mortgage rate by 25 basis points, while lender credits, or negative points, have the opposite effect.

Following previous work on mortgage pricing (e.g., Gurun et al., 2016; Bartlett et al., 2022), we estimate the following regression:

$$Y_{i,j,c,t} = \beta \cdot \log(1 + Ad_{j,d(c),t}) + \eta \cdot X_i + \theta_j + \theta_c + \theta_t + e_{i,j,c,t}, \tag{1.4}$$

where $Y_{i,j,c,t}$ is the rate spread for loan i originated by lender j in county c during year t. $Ad_{j,d(c),t}$ is the refinance ad spending by lender j in DMA d (to which county c belongs) in year t. X_i is a vector of loan-level characteristics that impact loan pricing, including the loan size decile, income decile, loan-to-value (LTV) decile, debt-to-income (DTI) bucket, and net mortgage points (discount points minus lender credits). Additionally, we control for the lender, county, and year fixed effects, denoted by θ_j, θ_c, and θ_t, respectively. We cluster standard errors at the lender-DMA level.

We run two versions of the regression with and without lender fixed effects and present the results in Table 1.5. Without lender fixed effects, column (1) reports a positive and statistically significant coefficient on refinance ad spending, which suggests that lenders who advertise more tend to have a higher rate spread. However, the economic magnitude of the coefficient is very small. By converting the coefficient to an advertising elasticity of 0.008,[12] we find that doubling the ad spending is associated with only a 0.8 basis point (0.008%) increase in the rate spread, or $99.6 in total interest costs for the average loan size of $311,340.[13] This increase is very small in the context of mortgage loans with total interest costs of about $185,213 (assuming an average interest rate of 3.39%) and an average of $4,350 in origination fees. After controlling for lender fixed effects, column (2) shows a negative and statistically insignificant coefficient, which suggests that increases in advertising within the same lender are not associated with higher mortgage costs.

Therefore, we conclude that TV advertising likely increases consumer welfare in the context of mortgage refinancing. More mortgage borrowers will enjoy significant interest savings from refinancing due to the market expansion effect. The business-stealing effect may lead to a decrease in consumer welfare by shifting market share to more expensive lenders, but we find the economic magnitude to be very small. Note that this discussion considers an

[12]The advertising elasticity of the rate spread is calculated with $\widehat{\beta} \cdot \frac{Ad^{Refi}}{(1+Ad^{Refi})}$.

[13]To translate changes in basis points to dollar amounts, we use the heuristic that 1 mortgage point (i.e., 1% of the loan size) is equivalent to 25 basis points, similar to Bartlett et al. (2022). Applying this heuristic to the average loan size in our sample, we calculate that 0.8 basis points corresponds to $99.6 ($311,340 $\times$ 1% $\times$ 0.8/25).

Table 1.5: Correlation between Advertising and Loan Expensiveness

	(1)	(2)
Dependent Variables:	Rate spread	
$\log(1 + Ad^{Refi.})$	0.010***	-0.009
	(0.002)	(0.006)
Lender FEs	N	Y
County FEs	Y	Y
Year FEs	Y	Y
Loan Size Decile FEs	Y	Y
Income Decile FEs	Y	Y
LTV Decile FEs	Y	Y
DTI FEs	Y	Y
Points controlled	Y	Y
N	3,972,682	3,972,682

Notes. S.E.s are clustered at the lender - DMA level; *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

average borrower and abstracts away from scenarios where some borrowers may be worse off by refinancing at a rate that is not economically desirable, after considering origination costs, or by switching to a more expensive lender after seeing their ads.

1.5 Advertising Effects by Ad Topics

Mortgage ads typically highlight different messages, such as low interest rates or the simplicity of the online application process. Understanding the effects of different ad content provides insights into the most effective message for encouraging mortgage borrowers to consider refinancing. It is also practically valuable for advertisers. In this section, we explore the heterogeneous effects of different ad topics. We begin by describing how we extract topics from advertising messages in Section 1.5.1. We then conduct demand analyses at both the category and brand levels using the extracted ad content in Section 1.5.2.

1.5.1 Topics in Advertising Messages

In order to explore the heterogeneous effects of different ad content, we need to extract topics from advertising messages. To do so, we employ zero-shot text classification, which classifies data (ad messages) into categories (ad topics) without the need for labeled data. We first describe the candidate ad topics, and then discuss how zero-shot classification works and why it is appropriate in our context compared to several alternative methods.

We seek to classify ads into topics defined based on prior literature and domain knowledge. From prior literature, Perry et al. (2016) analyzed mortgage ads from 2015 across multiple media and categorized them as informational (providing factual content like "rates as low as 3.1% APR"), transformational (evoking emotional and affective responses like "a home is more than a place you live"), or middle ground (brand-focused, trust-building like "a trusted partner at every step"). We adopt this categorization since these three topics remain relevant in our dataset. Besides these topics, we introduce a new one that highlights the ease of loan application through online processes, which reflects the recent rise of fintech mortgage lending (e.g., Fuster et al., 2019). Table 1.6 summarizes the four topics that focus on interest rate and savings, ease of application, homeownership, and brand-building, respectively.

Table 1.6: Topic Descriptions for Zero Shot Text Classification

Ad topics (1)	Topic descriptions (2)
Interest rate and savings	low mortgage rate, interest saving, save on mortgage
Interest rate and savings	online application, digital mortgage, easy, simple and quick
Homeownership	home buying, new/dream/perfect home, home ownership
Brand-building	trusted lender, customer satisfaction, expert advice

To classify text into user-defined categories, zero-shot classification leverages a pre-trained large language model to represent the semantic information of both the input documents (advertising messages in our context) and candidate topics (more specifically, the topic descriptions in column (2) of Table 1.6) using text embeddings. Zero-shot classification scores the input documents against candidate topics in terms of semantic similarity and predicts the probabilities of each document belonging to the candidate topics. To assess the sensitivity to the topic descriptions, we experiment with adding or removing a semantically similar word for each topic. The robustness of our main findings is discussed in the subsequent section.

The performance of zero-shot classification improves when a pre-trained language model is fine-tuned on Natural Language Inference (NLI) datasets (Yin et al., 2019). These NLI datasets consist of numerous "premises" and "hypotheses" pairs, annotated as "entail", "contradict", or "neutral." These annotations indicate whether the hypotheses can or cannot be logically inferred from the premises. For example, in the widely used MultiNLI dataset (Williams et al., 2017), the premise "Their mortgage payments immediately jumped $1,200 a month, to $3,290." and the hypothesis "The amount they paid for mortgage increased significantly." are coded as "entail." In contrast, the hypothesis "Their mortgage payments decreased significantly." is coded as "contradict."

We employ a zero-shot classification model based on DeBERTaV3 (He et al., 2021), fine-tuned on five NLI datasets.[14] Given the ad messages and candidate topics, the model predicts the probability of each topic for all ads. Table 1.7 reports the probabilities across all ads (Panel A), refinance-focused ads (Panel B), and purchase-focused ads (Panel C). Overall, the interest rate and savings topic occurs the most frequently in our data at 43%, followed by the homeownership topic at 28% and the brand-building topic at 22%. Although the ease of application topic is least frequent at 7% overall, it appears more frequently in ads from non-bank, fintech lenders than bank lenders (e.g., 17% for Rocket Mortgage vs. 9% for Wells

[14]This is the recommended model for zero-shot classification tasks as of the writing of this paper. https://huggingface.co/MoritzLaurer/DeBERTa-v3-large-mnli-fever-anli-ling-wanli

Fargo). Comparing Panel B and C, refinance ads tend to focus more on interest rates and savings, while purchase ads are more likely to discuss homeownership, which is intuitive. We also manually go through randomly selected ads and their corresponding topic probabilities, and find the zero-shot classification model to be highly effective. We show a few examples in Appendix A.6 as a qualitative evaluation of the model.

Table 1.7: Probability Distribution by Ad Topics

Ad topics	Mean	Std. Dev.	Min	P25	P50	P75	Max
Panel A: All ads (N=8,355)							
Interest rate and savings	0.43	0.37	0.00	0.06	0.31	0.88	1.00
Ease of application	0.07	0.15	0.00	0.00	0.01	0.07	1.00
Homeownership	0.28	0.32	0.00	0.01	0.14	0.49	1.00
Brand building	0.22	0.27	0.00	0.01	0.10	0.33	1.00
Panel B: Refinance-focused ads (N=3,200)							
Interest rate and savings	0.72	0.33	0.00	0.43	0.94	0.99	1.00
Ease of application	0.05	0.14	0.00	0.00	0.00	0.03	1.00
Homeownership	0.05	0.12	0.00	0.00	0.01	0.04	0.99
Brand building	0.17	0.25	0.00	0.00	0.02	0.27	1.00
Panel C: Purchase-focused ads (N=5,155)							
Interest rate and savings	0.24	0.27	0.00	0.02	0.14	0.35	0.99
Ease of application	0.08	0.16	0.00	0.00	0.02	0.09	1.00
Homeownership	0.43	0.33	0.00	0.13	0.35	0.73	1.00
Brand building	0.25	0.27	0.00	0.03	0.15	0.37	1.00

Notes. We classify an ad as refinance-focused if the predicted probability for refinance is larger than that for purchase loans; otherwise, it is classified as purchase-focused.

Zero-shot classification can be viewed as a semi-supervised approach that leverages rich information from document embedding and theories from prior research or domain knowledge via researcher-defined candidate topics. It is particularly well suited in our context compared to several alternative methods. One alternative method is to train a supervised machine learning model on these topics, similar to how we classify refinance vs. purchase loan ads. This method is challenging for ad topics because they tend to be subtle and nuanced, and thus, ads cannot be effectively labeled using a keyword-based approach. The second alternative method is topic modeling, such LDA (Blei et al., 2003) and its variant, guided LDA, which allows users to define topic-indicative keywords prior to estimation (Jagarlamudi et al., 2012). Compared to LDA (and guided LDA), zero-shot classification leverages a large language model to capture nuanced semantic meanings of ad messages, rather than relying on word co-occurrence for topic discovery. The third alternative method is to segment ads into clusters based on document embeddings, like BERTopic (Grootendorst, 2022). We find zero-shot classification to work better since it allows for the use of researcher-defined topics based on prior literature and domain knowledge, thereby eliminating the need for post-hoc interpretation of clusters. Empirically, we observe that unsupervised segmentation tends to group ads from a single lender into one cluster, likely because of the frequent mentions of brand names. This is less useful for our purpose, as we intend to identify the different messages that the ads focus on.

1.5.2 Heterogeneous Effects by Ad Topics

We now explore the heterogeneous effects of advertising based on the content of the ads (i.e., the identified ad topics). Recall that from the zero-shot text classification model, we obtain the probability for each topic at the ad creative level. For the demand analysis, we calculate the ad spending for each topic by multiplying the refinance ad spending of each creative by its corresponding topic probabilities.

We start with the heterogeneous effects on the category-level demand. To allow for separate coefficients for each topic, we modify our main category-level demand model (Equation 1.1). Specifically, we replace the refinance ad spending, denoted by $log(Ad^{Refi})$, with ad spending measures specific to each of the four topics: $log(Ad^{Refi} \times \text{Rate and Savings})$, $log(Ad^{Refi} \times \text{Ease of Application})$, $log(Ad^{Refi} \times \text{Homeownership})$, and $log(Ad^{Refi} \times \text{Brand-building})$. These

Table 1.8: Effectiveness by Ad Content for Category-level Demand

Dependent Variable:	$log(Q^{Refi})$	
	(1)	(2)
$log(AD^{Refi} \times$ Rate and Savings)	0.040**	0.039**
	(0.019)	(0.019)
$log(AD^{Refi} \times$ Ease of Application)	0.034	0.033
	(0.034)	(0.034)
$log(AD^{Refi} \times$ Homeownership)	0.007	0.021
	(0.036)	(0.047)
$log(AD^{Refi} \times$ Brand-building)	0.003	0.007
	(0.023)	(0.026)
$log(AD^{Purch})$		-0.038
		(0.075)
Border Market FE	Y	Y
DMA Border-Year FE	Y	Y
N	1,680	1,680

Notes. S.E.s are clustered at the border market level; $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

values are derived by aggregating the ad spending by topic from the ad creative level to the border market-year level.

Table 1.8 shows the regression results. Standard errors are clustered at the border market level. Similar to our previous analysis, we run two versions of the regression: one with and the other without ad spending on purchase loans. In both columns (1) and (2), we find that the rate and savings topic has the largest coefficient and is statistically significant. The coefficient on the ease of application topic is the second largest, but it is not statistically significant. The coefficients on the other topics are small and statistically insignificant.

These results suggest that to increase category-level demand (market expansion), ad messages that focus on low interest rates and potential savings are the most effective. This is somewhat intuitive since the primary reason borrowers choose to refinance is to take advantage of lower rates and save on interest. This topic is also the most frequently used by lenders in their refinance advertising, as shown in Table 1.7. This finding is in line with previous studies that find that price promotional ads tend to be more effective than other ad content (e.g., Morozov and Tuchman, 2022).

Ad messages that emphasize the ease of application also seem to be effective in increasing category-level demand, although the effect is not statistically significant. It is possible that for some consumers, the hassle of refinancing may be the main barrier, and thus, ads that emphasize an easy application process could lead to additional demand for refinancing. As shown in Table 1.7, the ease of application topic is the least commonly used by lenders, which may contribute to the larger standard errors in the coefficient estimates.

The homeownership topic is not effective in increasing category-level demand, which is not surprising given that the target audience for refinancing is expected to already have mortgages and own homes. Brand-building messages also have little impact on category-level demand. However, ads that focus on building brand trust could potentially influence borrowers' choice of lenders, conditional on choosing to refinance, which we examine next.

We now focus on the heterogeneous effects on the brand-level demand. From our main brand-level demand models (Equations 1.2 and 1.3), we replace the overall refinance ad spending with ad spending measures specific to each topic. These values are obtained by aggregating the ad spending by topic from the ad creative level to the lender-market-year level.

We present the results in Table 1.9. Columns (1) and (2) use loan quantity as the dependent variable, while column (3) focuses on market share, conditional on borrowers who choose to refinance. In both models, the rate and savings topic is the most effective in increasing brand-level demand, although the coefficients are not statistically significant when using quantity as the dependent variable. This finding is consistent with findings from the category-level demand, suggesting that emphasizing low rates and interest savings not only increases

Table 1.9: Effectiveness by Ad Content for Brand-level Demand

Dependent Variable:	$log(1+Q_j^{Refi})$		$log(s_j/s_0)$
	(1)	(2)	(3)
$log(1+AD_j^{Refi}\times$ Rate and Savings)	0.102	0.097	0.146**
	(0.075)	(0.075)	(0.067)
$log(1+AD_j^{Refi}\times$ Ease of Application)	-0.051	-0.046	-0.011
	(0.104)	(0.104)	(0.084)
$log(1+AD_j^{Refi}\times$ Homeownership)	0.066	0.066	-0.018
	(0.101)	(0.101)	(0.087)
$log(1+AD_j^{Refi}\times$ Brand-building)	0.089	0.088	0.055
	(0.101)	(0.100)	(0.077)
$log(1+AD_{-j}^{Refi})$		0.033**	
		(0.015)	
Lender-Border Market FE	Y	Y	Y
Lender-DMA Border-Year FE	Y	Y	Y
N	40,076	40,076	40,076

Notes. S.E.s are clustered at the lender-border market level; $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

category-level demand but also attracts potential customers to the advertised lender, especially among those who choose to refinance. Additionally, the dominant prevalence of this topic in refinance advertising suggests that lenders are reasonably well-informed about which advertising message works best.

Ad messages that focus on brand-building also appear to be somewhat effective in increasing brand-level demand, although the coefficients are not statistically significant. Together with the category-level demand analysis, it is conceivable that ads focusing on building brand trust may be more effective in shifting the choice of lenders among borrowers who decide to refinance, rather than attracting more mortgage holders to consider refinancing. The other two ad topics, ease of application and homeownership, do not seem to be effective in increasing brand-level demand.

These regression results depend on the outcomes of zero-shot text classification, which rely on the descriptions provided for each topic. As a robustness check, we experiment with alternative topic descriptions by adding or removing one semantically similar word from each topic. We re-run the zero-shot classification to obtain the topic probabilities and then re-estimate the regression models. We find that the results are largely consistent. Detailed results can be found in Appendix A.7.

Overall, our results suggest that ads that focus on low mortgage rates and savings are the most effective in both expanding the market and shifting consumers' choice of lenders. This is consistent with the main motivation for mortgage holders to refinance. This finding suggests that for government-sponsored enterprises (e.g., Freddie Mac) seeking to promote refinancing through TV advertising, emphasizing low rates and savings could be more effective than other messages. Conditional on choosing to refinance, we also find weak evidence that brand-building messages can drive consumers' lender choices, which suggests that brand equity may still play a role in this market, despite the fact that a refinance loan is arguably homogeneous across lenders.

There is an important caveat to our analysis of the heterogeneous effects based on ad content. Even though the border strategy allows us to leverage plausibly quasi-random levels of advertising, the choice of ad topics by lenders is likely endogenous. Our analysis relies on the assumption that ad topics are equally effective when used by different lenders, which may not always hold in practice. The ad topic of low interest rate and savings is likely applicable to many lenders. However, the ease of application topic might only be appropriate for lenders with the technology to support a digital loan process. Additionally, it is possible that lenders have already optimally chosen the ad messages that work best for them, in which case our estimates could be biased upward.

1.6 Discussion

In this paper, we study how TV advertising affects mortgage refinancing decisions. It is well documented that many mortgage borrowers make sub-optimal financial decisions by failing to refinance when interest rates are low, despite the substantial amount of potential savings (e.g., Keys et al., 2016; Johnson et al., 2019). This is problematic not only from a household perspective but also from a societal perspective, as mortgage borrowers could otherwise spend the interest savings towards debt repayment or other forms of consumption (e.g., Abel and Fuster, 2021; Agarwal et al., 2023).

Using six years of mortgage origination and advertising data, we find that TV advertising significantly increases the category-level demand for refinancing (i.e., market expansion effect). The economic magnitude is substantial: with an estimated ad elasticity of 0.082, we calculate that a \$1 increase in refinance advertising leads to \$27.3 in interest savings from refinancing for mortgage borrowers. Beyond the category-level analysis, we find that a lender's advertising increases not only its own demand but also the demand for other lenders through market expansion. When focusing on borrowers who choose to refinance, a lender's advertising increases its market share at the expense of other lenders' shares.

We further explore the heterogeneous effects of different ad content. We use zero-shot text classification with a pre-trained large language model to classify ad messages into four topics selected based on prior literature and domain knowledge. We find that ad messages that focus on low mortgage rate and savings are the most effective in both increasing category-level demand and shifting the choice of lenders.

Our findings offer valuable insights for both policymakers and advertisers. For policymakers, TV advertising can be an effective tool to promote refinancing, especially when doing so is particularly beneficial (e.g., during periods of low interest rates). Refinancing will not only benefit consumers in terms of savings but also help improve the effectiveness of monetary policy pass-through. For advertisers, our research provides practical guidance as to what types of messages work the best in increasing their demand.

Although we focus on the effectiveness of advertising in the specific context of mortgage refinancing, our framework can be applied to other important consumer financial decisions, especially those where information provision or reminders can make a difference. The framework can also be applied to study other categories, such as government social welfare programs, where the take-up rate is low despite being both privately and socially beneficial. For instance, health insurance advertising by the federal or state governments has been found to be effective in decreasing the fraction of the uninsured among eligible individuals (Aizer, 2007; Karaca-Mandic et al., 2017). We invite future research to study how advertising, an important marketing tool, can enhance consumer decision making in various markets.

Chapter 2: TV Advertising Effectiveness with Racial Minority Representation: Evidence from the Mortgage Market

2.1 Introduction

In recent years, there has been a growing emphasis on promoting diversity, equity, and inclusion (DEI), which has prompted companies to make commitments to advance DEI. One visible manifestation of this commitment is the increasing number of chief diversity officers (CDOs) tasked with driving DEI initiatives at the organizational level. According to an article by McKinsey & Company, over 50% of Fortune 500 firms have appointed CDOs as of 2022.[1] Efforts to promote DEI can also be observed in many other areas of society, such as inclusive hiring practices in the workplace, fostering diverse student bodies in education, and the inclusion of diverse characters and narratives in media.

From a marketing perspective, promoting diversity and minority representation in advertising holds significant importance for companies for a number of reasons. Advertising is a powerful tool for marketers to connect with consumers and convey brand values. By including minority actors in advertisements, companies can better engage with their minority customer base, as these consumers respond more positively to ads featuring actors from their own racial background (e.g., Deshpandé and Stayman, 1994; Aaker et al., 2000). Beyond the racial fit between the consumers and actors, companies can also signal their commitment

[1] https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/unlocking-the-potential-of-chief-diversity-officers. Accessed July 12, 2023

to DEI initiatives through minority representation in their ads. Recent surveys conducted by Microsoft and Facebook find that consumers are more trusting of brands that represent diversity in their ads, and ads featuring more diverse actors are associated with higher ad recall, both suggesting positive consumer attitudes towards minority representation in advertising.[2]

This paper investigates the impact of racial minority representation on the effectiveness of TV advertising in the empirical context of mortgage refinancing. Refinancing a mortgage is one of the most financially consequential decisions that a household can make. In 2020 alone, \$2.6 trillion worth of mortgage loans were refinanced.[3] Prior studies have shown that minority consumers are less likely to refinance their mortgages compared to White consumers with similar characteristics, forgoing substantial potential savings (Gerardi et al., 2021; Gerardi et al., 2023). The lower refinance take-up rate among minorities contributes to the well-documented racial disparities in the mortgage market (e.g., Bartlett et al., 2022; Bhutta et al., 2022). From this perspective, understanding the impact of minority representation in ads in this market holds particular social significance.

More specifically, we aim to answer the following research questions. First, how does the effectiveness of TV advertising change with varying degrees of minority representation? After showing the main effect that higher minority representation increases advertising effectiveness, we further investigate potential heterogeneous effects. In particular, how do the effects differ based on the borrower's own race and political leanings? Finally, taking together results from both observational and experimental studies, we discuss several potential mechanisms that are consistent with our findings.

[2]https://about.ads.microsoft.com/en-us/insights/inclusive-marketing-whitepaper; https://www.facebook.com/business/news/insights/the-difference-diversity-makes-in-online-advertising. Accessed July 12, 2023.

[3]https://www.statista.com/statistics/205946/us-refinance-mortage-originations-since-1990/. Accessed July 12, 2023.

To answer these research questions, we obtain loan-level mortgage origination data from 2018 to 2021. This dataset includes information on the borrower's race and census tract of the property, which allows us to link the home location to census tract-level political voting data. We then merge this loan origination data with TV mortgage advertising data obtained from Kantar Media, covering the same time period. In addition to advertising spending data, we also obtain the video files of the TV ad creatives. Using these videos, we utilize computer vision techniques to determine the race of each actor and construct a measure of minority representation in advertising.

There are two main challenges in estimating the impact of minority representation on TV advertising effectiveness using observational data. First, lenders may advertise more towards individuals who are more likely to be responsive or during time periods when the return on advertising is expected to be high. Second, the level of minority representation in ads may be correlated with video features, such as visual elements or advertising messages, which can also affect the effectiveness of the ads. Ignoring these potential correlations can introduce omitted variable bias in the estimates.

We address these concerns using a Double Machine Learning (Double ML) model (Chernozhukov et al., 2018). In the Double ML model, we explicitly allow the advertising levels and minority representation in ads to depend on a very flexible functional form on high-level interactions between lender, location (Designated Market Area, or DMA), and time (year) fixed effects, as well as high-dimensional video attributes. The video attributes include image embeddings which we obtain by training a variational autoencoder (VAE), as well as ad transcript embeddings, which we obtain from a pre-trained embedding model from OpenAI. Because the Double ML model allows for interactions between each of the control variables, it includes the benefits one would get from a fixed effects regression with a rich set of interactive fixed effects. As a benchmark, we also present the results of a fixed effects regression with lender-DMA and lender-time fixed effects. This model, however, does not control for the possibility that advertisements with more minority actors might also have different messages or video imagery. The Double ML approach controls for this correlation by estimating how the advertising levels and minority representation are correlated with the video and message attributes, and then using only the residual variation to estimate the causal effect. We find that the estimates from the fixed effect model are mostly similar to results from the Double ML model.

Our results show that increased minority representation leads to an increase in the effectiveness of the ads. Specifically, we find that as the minority share in ads increases from 15% (representing the median value at the lender-DMA level in our data) to 25%, the advertising elasticity increases from 0.037 to 0.042 (a 14% relative increase). Moreover, we observe that the impact of minority representation is more pronounced among minority borrowers than White borrowers, but is also positive among White borrowers. Across the political spectrum, we find a stronger effect among liberal-leaning consumers compared to conservative-leaning consumers, suggesting that support for diversity and minority representation, or attitudes about race in general, may play an important role.

We further complement our observational study with a pre-registered lab experiment ($N = 2,796$) where we directly manipulate the race of the families in advertisements using generative AI technology. The results from our experimental study confirm the main effect observed in our observational study: Participants who are randomly assigned to advertisements featuring minority families report a higher likelihood of applying for refinancing from the advertised lender and recommending the lender compared to those randomly assigned to ads featuring White families. Furthermore, the experimental results show consistent patterns on the relative impact of minority presence in ads based on the participants' self-reported race and political orientation. The converging results from the experiment with random assignment help assure us that our estimated effects from our observational study are not simply the result of a subtle endogeneity or omitted variable biases.

There are several possible mechanisms that are consistent with our results. First, consistent with our heterogeneous effects, prior results have documented that minority consumers have a stronger preference for racial homophily than White consumers (e.g., Deshpandé and Stayman, 1994; Aaker et al., 2000; Mollica et al., 2003), and liberal-leaning individuals have greater support for racial diversity and equity than conservatives (e.g., Agarwal and Sen, 2022; Aneja et al., 2023; Babar et al., 2023). By asking follow-up questions in the experiment, we find that ads featuring minority actors lead to favorable brand perceptions, such as perceptions of broad loan options, fair lending practices, and inclusiveness toward individuals of all backgrounds, all of which can contribute to higher ad effectiveness. Lastly, ads featuring minority actors can be more effective simply because they are less common, making them stand out to consumers (Pieters et al., 2002; Rosengren et al., 2020). Indeed, our experimental participants perceive ads featuring minority actors as new, fresh, and attention-grabbing compared to ads with only White actors.

Our findings offer valuable insights for brands in shaping their advertising strategies. By featuring minority actors, brands can not only signal their commitment to DEI, but can also increase the effectiveness of their advertising efforts. Our results also have important policy implications. Given that featuring minority actors is particularly effective in reaching minority consumers, it has the potential to be a useful strategy to provide these consumers with information about refinancing opportunities and encourage them to refinance, especially in times when interest rates are low. This, in turn, could help reduce racial disparities in the refinance take-up rate.

Our research contributes to several streams of literature. First, our paper is closely related to the nascent literature on the impact of DEI initiatives and social equity movements on both society and business. In terms of societal impact, Agarwal and Sen (2022) find a significant increase in demand for anti-racist books requested by public school teachers following the killing of George Floyd. From a managerial perspective, Balakrishnan et al. (2022) and Khan and Kalra (2022) demonstrate that signals of diversity at the corporate level have a positive impact on consumer attitudes. Furthermore, signaling racial identity can increase demand for minority-owned businesses on platforms like Yelp and improve the success rate of requests for help (Kirgios et al., 2022; Aneja et al., 2023; Babar et al., 2023). Beyond racial diversity, Goli and Mummalaneni (2023) find that an increase in women's screen time positively impacts the viewership of cable news shows. However, it is important to note that DEI initiatives may not always receive favorable responses, underscoring the need for careful assessment of the potential benefits and risks associated with such initiatives. For example, Wang et al. (2022) find that firms' social media posts related to the Black Lives Matter movement reduce consumer engagement on social media platforms. Our paper contributes to this literature by studying how minority representation in TV ads impacts advertising effectiveness.

Within this domain, the studies closest to ours are two concurrent working papers by Hartmann et al. (2023) and Overgoor et al. (2023). Hartmann et al. (2023) find that online display advertisements featuring minority actors achieve higher click-through rates than those with White actors. Overgoor et al. (2023) study the impact of Black actor share in TV ads on consumers' purchase intentions and find that the effect depends on the processing route. Our study differs from these sets of papers in a few important ways. First, we study advertising effectiveness with actual consumer demand rather than relying on clicking behaviors or self-reported purchase intentions. Second, we account for the potential correlation between the race of the actors and visual and text features in ads in order to minimize omitted variable biases. Thus, we have a much richer set of controls in our study. Third, by leveraging detailed information on each consumer's race and (census tract level) political leaning, we estimate heterogeneous effects along these dimensions. This, in conjunction with the experimental results, allows us to discuss potential mechanisms at play.

Our paper also adds to the literature on the impact of advertising content. Beyond lab experiments, ad content has received less attention than the examination of ad quantity in economics and marketing. Bertrand et al. (2010) measure the effect of informational content (e.g., interest rates) and non-informational content (e.g., a photo featuring an attractive woman) in direct mail ads for consumer loans through a large-scale field experiment, and find strong evidence of the significant impact of ad content. Since then, marketers have utilized either field experiments (e.g., Sudhir et al., 2016; Sahni et al., 2018; Morozov and Tuchman, 2022) or observational data (e.g., Liaukonyte et al., 2015; Lee et al., 2018; Tsai and Honka, 2021; Fossen et al., 2022) to gain deeper insights into the impacts of both informational and non-informational ad content. Building on this literature, our study explores how racial representation in ads, as a form of non-informational content, can influence consumer demand for the advertised lender.

The rest of the paper is organized as follows. Section 2.2 describes our main data and provides descriptive statistics. Section 2.3 describes the multi-modal features we extract from the video data. Section 2.4 presents our empirical strategies. Section 2.5 documents the empirical results. Section 2.6 describes the online experiment and documents the experimental results. Section 2.7 discusses potential mechanisms. Finally, Section 2.8 concludes.

2.2 Data Source and Descriptive Statistics

In this section, we describe the two main datasets used in our study: the mortgage origination data (Section 2.2.1) and the TV advertising data (Section 2.2.2). We also discuss how we extract race information from advertising video files and present descriptive statistics in Section 2.2.3.

2.2.1 Mortgage Origination Data

We obtain loan-level mortgage origination data from the Home Mortgage Disclosure Act (HMDA) database for the period from 2018 to 2021. The HMDA is a U.S. federal law that mandates mortgage lenders to disclose detailed information about their mortgage lending activities. The data collected under this law cover approximately 90% of total mortgage originations with reporting exemptions for small lenders Bhutta et al. (2017). This dataset provides comprehensive loan-level information, including the originating lender, year of origination, loan size, and loan type. Additionally, the HMDA data include borrower characteristics, including race, which allow us to study how consumers from different racial backgrounds respond to minority representation in ads. We also use the property's census tract to collect information about political voting patterns.[4]

A new mortgage loan can be originated for the purpose of a home purchase or refinancing. We focus on refinances because consumers often rely on real estate agents or mortgage brokers when choosing a lender during the home purchase process, which can lead to limited advertising effectiveness at that stage. In contrast, refinancing decisions are typically made independently by consumers.

[4]We obtain census block group-level estimates for the 2020 presidential election from Bryan (2022), which uses the methods described in Amos et al. (2017). This data has been used in academic studies, such as Babar et al. (2023).

We focus on conventional mortgage loans that follow standard underwriting guidelines, such as having a minimum credit score of 620 or above and a debt-to-income ratio below 50%. Consequently, we exclude non-conventional, government-backed loans, including VA loans for veterans or active-duty service members, FHA loans for low-income and low-credit score consumers, and USDA loans for rural areas. Conventional mortgage loans represent approximately 80% of the total mortgage originations during our sample period (F. Liu et al., 2022). Following previous studies on household finance, we impose additional inclusion criteria (e.g., Bartlett et al., 2022; Bhutta et al., 2022; Gerardi et al., 2023): These loans must be first-lien mortgages for owner-occupied, site-built, single-family residential homes with a minimum loan size of $100,000. Additionally, we exclude jumbo loans that exceed conventional loan size limits and other unconventional loan types, such as reverse mortgages, interest-only loans, balloon payment loans, and negatively amortizing loans.

Since the TV advertising data (discussed in Section 2.2.2) covers the top 101 Designated Market Areas (DMAs), we include loans originating from these markets. Further, as we seek to analyze the heterogeneous responses of consumers from different racial backgrounds, we exclude loans with missing or mixed joint race information.[5] These data cleaning procedures lead to a sample of 9.7 million loans for our analysis. We provide further details of the data cleaning process and the number of excluded loans at each step in Appendix B.1. For ease of estimation, we randomly select 2.89 million borrowers, which account for 30% of the full sample, as our estimation sample.

When studying the choice of lenders, we narrow our focus to the top 30 lenders, which collectively represent over 50% of all refinancing mortgage originations, and categorize the remaining smaller lenders as the "outside" option. Column 1 of Table 2.1 presents the top

[5]We compare borrowers with missing or joint ethnic/racial information to those with complete information on loan size, income, and age and find similar distributions. Further details are provided in Appendix B.2.

Table 2.1: Market Share, Ad Spending, and Minority Share in Ads by Lender per Year

Lender	Market Share	Ad Spending (per 1,000 Capita)	Minority Share in Ads
Rocket Mortgage	10.87%	$1147.90	34.69%
United Wholesale Mortgage	5.69%	$7.62	39.04%
Wells Fargo	4.03%	$341.13	20.48%
JP Morgan Chase	3.77%	$0.67	21.95%
LoanDepot	2.78%	$98.17	13.98%
Nationstar	1.90%	$26.97	13.33%
Bank of America	1.72%	$0.00	–
Caliber Home Loans	1.52%	$0.00	–
US Bank	1.30%	$6.68	18.25%
Fairway Independent	1.28%	$197.05	8.08%
PennyMac	1.18%	$0.00	–
Guaranteed Rate	1.14%	$80.50	23.53%
Flagstar Bank	1.12%	$16.81	0.00%
Home Point Financial	1.08%	$0.00	–
Freedom Mortgage	1.01%	$1.45	25.15%
Newrez Mortgage	0.86%	$0.00	–
Provident Funding	0.85%	$0.00	–
AmeriSave Mortgage	0.77%	$105.42	4.29%
Citizens Bank	0.74%	$6.07	34.88%
Better Mortgage	0.74%	$0.20	0.00%
CrossCountry Mortgage	0.74%	$75.59	23.52%
PNC Bank	0.72%	$0.06	36.28%
Broker Solution Bank	0.70%	$0.31	32.52%
Cardinal Financial	0.69%	$6.08	8.02%
Finance of America	0.68%	$10.62	6.42%
Guild Mortgage	0.60%	$13.94	6.03%
Fifth Third Bank	0.52%	$0.00	–
Huntington Natl. Bank	0.49%	$0.00	–
Movement Mortgage	0.49%	$0.26	25.00%
American Financing Corp.	0.42%	$941.54	5.59%

Note. Ad spending denotes the total ad spending per 1,000 capita, including both national ads and local ads across the top 101 DMAs.

30 lenders ranked by their average market share per year, as reported in column 2.[6] Rocket Mortgage (formerly known as Quicken Loans) has the largest market share, followed by United Wholesale Mortgage, Wells Fargo, and JP Morgan Chase. These top four lenders

[6]In this list, we exclude BB&T and SunTrust, which merged into Truist in December 2019. Because of the merger, there was a time period where they had separate ad campaigns but reported to HMDA under the new name Truist, which creates challenges in matching the advertising data with the loan origination data.

account for 24% of the market share. Mortgage lending is a much less concentrated market than many other markets, such as airlines or breakfast cereals, where the top 4 companies have market shares of 67% and 85%, respectively.[7] We define the variables in columns 3 and 4 of Table 2.1 in the sections below.

2.2.2 Mortgage TV Advertising Data

We obtain TV advertising data from Kantar Media for the same sample period as in the mortgage data. This data includes monthly advertising spending at the lender-DMA-ad creative level for both national and local ads, covering the top 101 Designated Market Areas (DMAs).[8] To account for population differences across DMAs, we scale the local advertising spending using the population of the corresponding DMA to obtain ad spending per capita, following previous research in TV advertising (e.g., Shapiro, 2018; Tsai and Honka, 2021). Similarly, we scale the national advertising spending using the national population. The total ad spending for a specific lender within a specific DMA is defined as the lender's national ad spending per capita plus their local ad spending per capita within the DMA.

Using the ad spending data, we show the average ad spending per 1,000 capita per year for each lender across all of the 101 DMAs in Column 3 of Table 2.1. We observe significant variations in total advertising spending among lenders. Rocket Mortgage is the largest advertising spender during the sample period, followed by American Financing and Wells Fargo. However, some major lenders, such as JP Morgan Chase and Bank of America, allocate little or no budget to TV advertising.

[7] https://www.statista.com/statistics/250577/domestic-market-share-of-leading-us-airlines/; https://www.statista.com/statistics/858562/cereal-company-market-share-us/. Accessed July 12, 2023.

[8] In the U.S., TV markets, known as DMAs, are defined by the Nielsen Company to measure ratings across different geographic regions. Each DMA typically consists of multiple counties, with a major city at its center, along with surrounding smaller counties. Advertisers have the option to purchase national ads that are broadcasted across all 210 DMAs or local ads that are limited to specific DMAs (e.g., Boston DMA).

Besides the advertising spending data, we also collect the video files of the TV ad creatives. In our data, there are a total of 1,441 unique ad creatives aired by the top 30 lenders.[9] For each ad creative, we observe the total ad spending at the DMA-month level. We utilize these ad videos to determine if and to what extent they feature minority actors. In addition to race, we extract visual and textual features from these video ads, as detailed in Section 2.2.3.

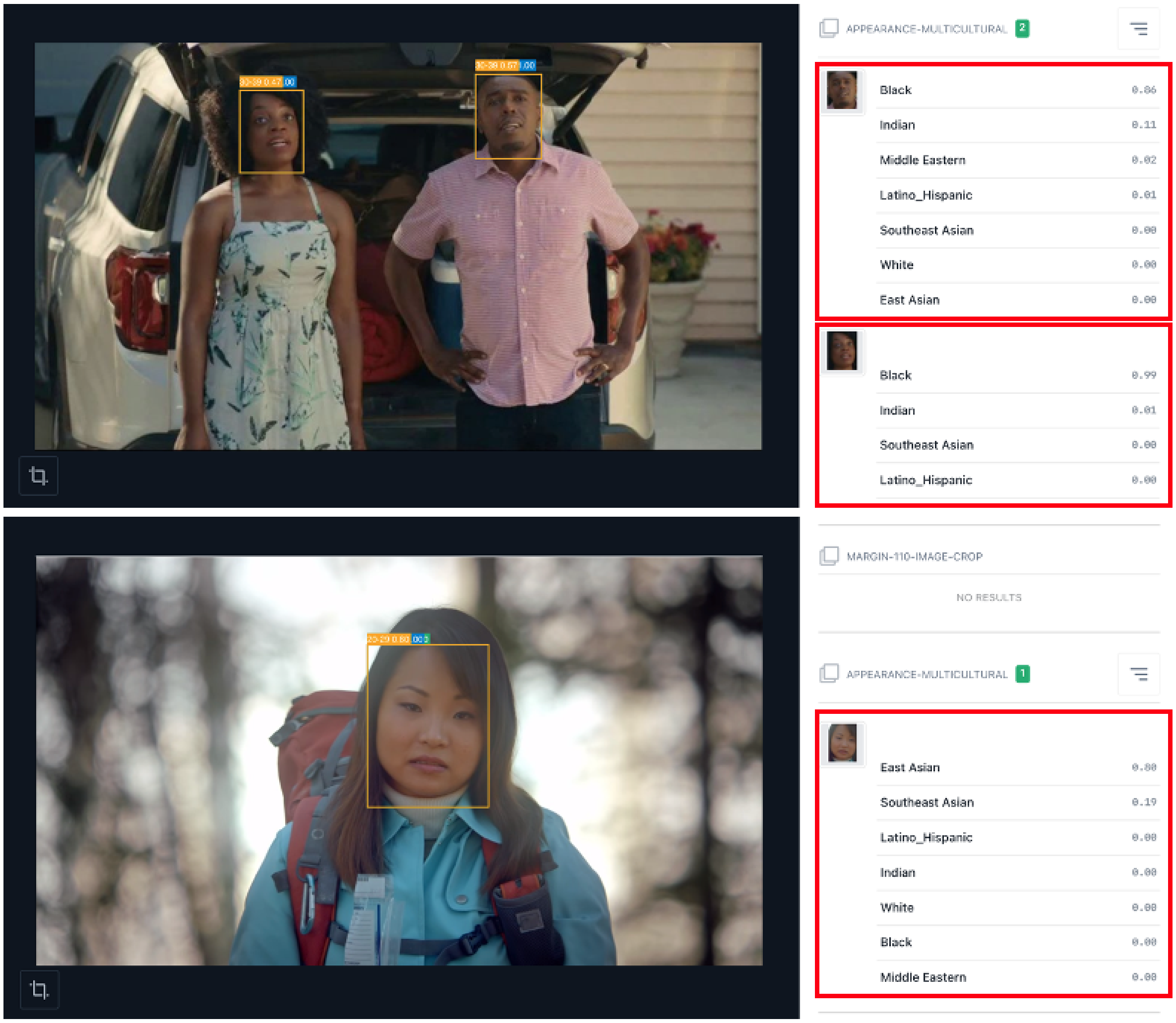
2.2.3 Race Detection and Minority Representation Measure

To determine the race of each actor in advertisements, we leverage pre-trained computer vision algorithms from Clarifai Inc. rather than training our own model. Pre-trained models, which are trained on large datasets, generally outperform models trained by researchers on smaller data. Several prior studies in marketing and management have also used Clarifai's pre-trained models (e.g., Dzyabura and Peres, 2021; Khern-am-nuai et al., 2021; M. Zhang and Luo, 2023; Hartmann et al., 2023). We build a customized "workflow" on Clarifai: We first detect any faces in a given frame, then crop an image for each detected face, and finally predict the probability that the face belongs to each racial group. We apply this workflow to frame-level image data, where we sample one frame per second from each ad video. Figure 2.1 shows two examples. On the top panel, Clarifai detects two faces and predicts that the male actor is Black with a probability of 0.86 and the female actor is Black with a 0.99 probability. On the bottom panel, one face is detected, and the actor is predicted to be East Asian with a probability of 0.80. We have manually checked a number of predictions and found the Clarifai algorithms to be highly accurate.

Before constructing our measure of racial representation using the predicted race information, we conduct two additional data processing steps. First, we group certain racial categories

[9] This number is based on the ad creative names reported in the Kantar data. We exclude a small number of ads specifically targeting reverse mortgage loans.

Figure 2.1: Clarifai Examples



from Clarifai to align with the categories in the HMDA data. Specifically, we combine "White" and "Middle Eastern" into the White category and group "East Asian", "Southeast Asian", and "Indian" into the Asian category. This results in four racial categories: White, Black, Hispanic, and Asian.[10] Second, we exclude a small number of predictions where the probability of the most likely race falls below 70%, similar to previous studies (e.g., An and Kwak, 2019; Gunarathne et al., 2022). This is to ensure that the detected race variable contains minimal measurement errors.

[10]In our study, we use race to refer to both ethnicity and race, and these classifications are based on the appearance of actors in the ads. Thus, we treat Hispanic as a racial category.

To measure the level of racial representation in videos, we take into account both the duration of time that each race appears on the screen and the extent of screen sharing when a video features multiple actors. Suppressing the subscript for each ad video for brevity, let $f = 1, ..., F$ denote the frame with human faces in the video. Let J^f denote the number of actors in frame f and $R_j^f \in \{W, B, H, A\}$ denote whether the race of the j^{th} actor in frame f is White (W), Black (B), Hispanic (H) or Asian (A). When a frame contains multiple actors ($J^f \geq 2$), we divide the screen share for each actor by the number of actors present in the frame ($\frac{1}{J^f}$). The measure of racial representation for each video is calculated as follows:

$$Share^{Race} = \frac{1}{F} \sum_{f=1}^{F} \left(\frac{1}{J^f} \sum_{j=1}^{J^f} \mathbb{1}\{R_j^f = Race\} \right), \tag{2.1}$$

where $Race \in \{W, B, H, A\}$ denotes each racial group, and $\mathbb{1}\{R_j^f = Race\}$ is an indicator function that takes the value of 1 if the race of the j^{th} actor in frame f matches the given $Race$, and 0 otherwise.

Let's consider a 30-second ad as an example. Suppose a White actor appears in the first 14 seconds, and then two actors, one White and one Black, appear together during the next 14 seconds, and the lender's logo is displayed in the last 2 seconds. In this example, the total number of frames with human faces F is 28. The share of White actors is calculated as $Share^W = \frac{1}{28}(14 \cdot 1 + 14 \cdot 0.5) = 0.75$ because the first 14 seconds only have a White actor (screen share of 1), and the next 14 seconds have both a White actor and a Black actor (screen share of $\frac{1}{2}$ for each). Similarly, the share of Black actors is $Share^B = \frac{1}{28}(14 \cdot 0 + 14 \cdot 0.5) = 0.25$ because the Black actor appears in the latter 14 seconds together with the White actor. Table 2.2 presents the average share for each racial group at the ad creative level. We observe that the average share of White representation is approximately 0.8. Among the minority groups, the share of Black representation is the highest at 0.13, indicating that a significant portion of the minority representation in our data comes from Black actors. In contrast, the shares of Hispanic and Asian representation are relatively low at 0.01 and 0.06, respectively.[11]

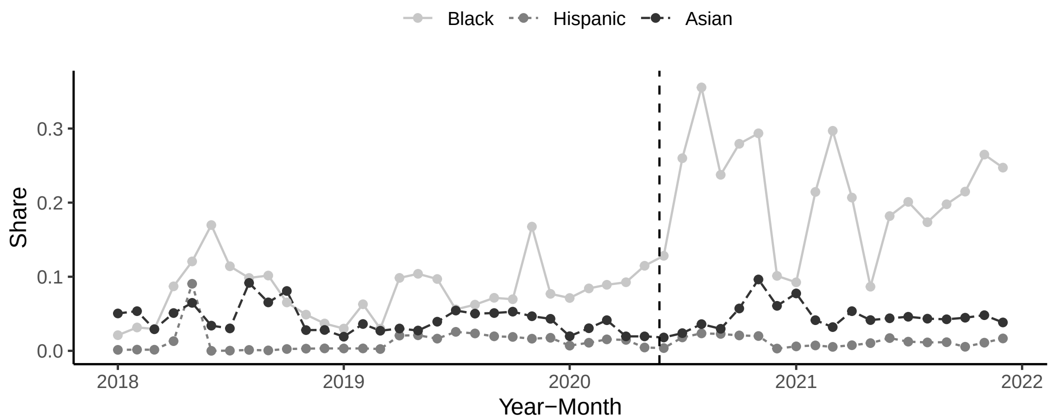
Table 2.2: Racial Representation in Ads at the Ad Creative Level ($N = 1,441$)

Race	Mean (SD)	Min	Median	Max
White	0.797 (0.262)	0	0.902	1
Black	0.130 (0.229)	0	0	1
Hispanic	0.009 (0.042)	0	0	0.545
Asian	0.064 (0.124)	0	0	1

Figure 2.2 plots the monthly shares of minority representation for Black, Hispanic, and Asian actors from 2018 to 2021. We calculate the weighted average share of racial representation for each minority group within a given month, considering both national and local ads across all of the 101 DMAs and using ad spending per capita as the weight. The shares of Hispanic and Asian representation stay relatively stable over time. However, there is a significant increase in the share of Black representation in ads in the second half of 2020. The dashed line in the figure corresponds to the time of George Floyd's murder. While it is difficult to pin down the exact reasons, the increase in Black representation in ads aligns with lenders responding to the social movement advocating for greater diversity, equity, and inclusion in ads in response to George Floyd's murder. A similar upward trend in the representation of minority actors, primarily driven by an increase in Black actors, is also observed in online display advertising (Hartmann et al., 2023).

[11]We observe that the Hispanic representation is the lowest. Clarifai's race detection algorithm can only identify individuals with more indigenous Hispanic features. As a result, many Hispanic individuals may be classified as White. A similar challenge is discussed in a related study by Davis et al., 2019.

Figure 2.2: Racial Representation in Ads over Time



Given that the majority of minority representation in ads comes from Black actors, we combine the shares of Black, Hispanic, and Asian representations to obtain the overall share of minority representation in each advertisement, a:

$$Share_a^{Minority} = Share_a^B + Share_a^H + Share_a^A. \tag{2.2}$$

Because the HMDA data is only available at the annual level, we ultimately aggregate the advertising spending and minority share variables to the annual level as well by lender and DMA. The advertising spending is calculated by taking the sum of the spending per capita, and the minority share is the weighted average of minority share across different advertising creatives, where the weight is the spending for each advertisement in the DMA.

We present the minority share in ads for each lender in Table 2.1 column 3. While most lenders that advertise use minority actors in at least some of their ads, the minority share in ads varies significantly across lenders. For example, Rocket Mortgage and United Wholesale Mortgage have relatively higher minority shares, while other lenders, such as Fairway Independent, AmeriSave Mortgage, and American Financing, are significantly less likely to feature minority actors during our sample period, despite their significant advertising expenditures.

Table 2.3 presents summary statistics of the key variables at the borrower-lender level. In Panel A, we see that, on average, borrowers have access to 28 lenders in their respective DMA, including the outside option. While the top 30 lenders are generally available in most of the 101 DMAs, there are some exceptions, such as Fifth Third Bank, which primarily focuses its loan originations in the Midwest. We also find that 23% of borrowers belong to racial minority groups, with Black, Hispanic, and Asian borrowers accounting for 4.52%, 8.44%, and 9.85% of the market, respectively. The difference between Democratic and Republic vote shares is calculated as the difference between the number of votes for the Democratic and Republican candidates divided by the sum of their votes in the 2020 presidential election at the census tract level. Panel B of Table 2.3 presents the summary statistics for the ad spending per capita and minority share variables. On average, 16.8% of the screen time is occupied by minority actors in mortgage ads, which is slightly lower than the fraction of the minority consumers in this market, as shown in Panel A.

Table 2.3: Summary Statistics at the Borrower-Lender Level ($N = 81.2$M)

Panel A: Borrower Characteristics	Mean (SD)	Min	Median	Max
Number of Available Lenders	28.06 (3.52)	8	29	31
Minority Borrower	0.23 (0.42)	0	0	1
Dem.−Rep. Vote Shares	0.08 (0.36)	-0.97	0.07	0.97
Panel B: Advertising Characteristics	**Mean (SD)**	**Min**	**Median**	**Max**
Ad Spending per 1,000 Capita	31.11 (125.09)	0	0	1,071.33
Minority Share in Ads	0.168 (0.149)	0	0.151	1
Black Share in Ads	0.108 (0.128)	0	0.040	1
Hispanic Share in Ads	0.012 (0.025)	0	0	0.500
Asian Share in Ads	0.048 (0.046)	0	0.028	0.413

Note. Minority share in ads, as well as the racial breakouts, are conditional on positive ad spending.

2.3 Other Video Ad Features

2.3.1 Visual Features

While our main goal is to estimate the impact of minority representation in advertising on consumers' choice of lenders, the presence of different races in ads may be correlated with other confounding factors, such as visual elements or advertising messages. For instance, advertisements featuring minority actors might hypothetically place a greater emphasis on the ease of loan application. This potential difference in messaging could introduce omitted variables bias if it is not properly addressed. Therefore, we seek to account for the possibility of such correlation by controlling for a large number of video features. In this section, we describe how we extract high-dimensional visual and textual features from ads.

To extract visual features from the video data, we first pre-process the video data and select a smaller number of images per video. A typical video in our data has 1.37 billion pixel values (960 pixels for width $\times$ 540 pixels for height $\times$ 30 seconds $\times$ 30 frames per second $\times$ 3 color channels).[12] To reduce the computational burden, we sample one frame every 5 seconds. This results in an average of 6 frames per video since most video ads are 30 seconds long. Following the standard practices in computer vision, we then resize the frame-level data to have 150,528 pixel values per frame (224p for width $\times$ 224p for height $\times$ 3 color channels).

There are two broad approaches that one can use to extract features from image data. The first approach is to extract a number of researcher-defined features, which are often interpretable. For example, S. Zhang et al. (2022) examine how 12 image features, such as brightness and situation, impact the demand for Airbnb properties. This approach can offer

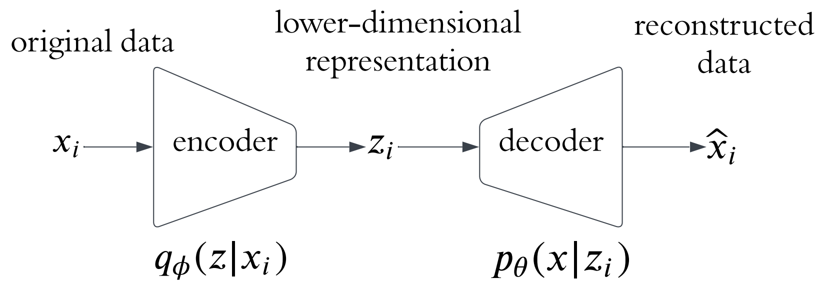
[12] The vast majority of the video ads have a resolution of 960p $\times$ 540p with a frame rate of 30 frames per second.

interpretable insights into what exactly is captured in the image data. However, it is often not straightforward to determine which features to extract. Furthermore, unless pre-trained algorithms are available, researchers would need to train their own models to extract the desired features, which are more likely to be subject to measurement errors.

The second approach is to obtain embeddings that represent image data. These embeddings are numeric vectors that may not be easily interpretable. However, unlike the first approach, researchers do not need to manually define a list of features to extract. Instead, they rely on a data-driven approach to capture as much relevant information as possible from the data. Since the primary purpose of extracting these features is to use them as control variables to address potential confounding factors in our application, we prioritize comprehensiveness over interpretability. Therefore, we use the second approach to extract image features.

We use a variational autoencoder (VAE) to obtain a lower-dimensional representation of the high-dimensional image data (Kingma and Welling, 2013; Rezende et al., 2014). VAEs have been applied in several recent marketing studies (e.g., Dew et al., 2022; Burnap et al., 2023, Tian et al., 2023). The structure of a VAE is illustrated in Figure 2.3. In our context, x_i denotes the i-th image from our frame-level video data, represented by high-dimensional pixel-level values ($x_i \in \mathbb{R}^d$), where d is equal to 150,528. This input x_i is passed through an encoder network denoted as $q_\phi(z|x_i)$, which generates a lower-dimensional latent representation denoted as z_i. More specifically, the encoder network generates a stochastic representation by outputting the parameters (mean and variance) of the distribution $q_\phi(z|x_i)$, which is a Gaussian probability density. $z_i \in \mathbb{R}^k$ is sampled from this distribution and has a significantly smaller dimension compared to x_i, with $k << d$. In our case, we set $k = 100$. The low-dimensional representation z_i is then fed into the decoder network denoted as $p_\theta(x|z_i)$, which generates the reconstructed image $\widehat{x}_i \in R^d$.

Figure 2.3: Illustration of Variational Autoencoder



Intuitively, a VAE is a semi-supervised machine learning model that aims to reconstruct the original image using a low-dimensional latent representation that captures the necessary information to minimize the reconstruction error. More formally, a VAE is trained by minimizing the following loss function:

$$L_{\phi,\theta}(x) = -E_{z \sim q_\phi(z|x)} \left[log(p_\theta(x|z))\right] + D_{KL}\left(q_\phi(z|x)||p(z)\right), \tag{2.3}$$

where the first term is the expected log-likelihood of the data x given the latent representation z. The expectation is taken over the encoder's distribution over the representation z, so that it depends on both the encoder parameters ϕ and the decoder parameters θ. This term is typically known as the reconstruction loss since it encourages the decoder to accurately reconstruct the original data by maximizing the likelihood. The D_{KL} term is the regularization loss. This is the Kullback–Leibler divergence between the encoder's distribution $q_\phi(z|x)$, which is the variational approximation to the posterior distribution, and the prior distribution $p(z)$, which is assumed to be a standard normal distribution. This regularization term ensures that the latent representation follows a smooth distribution. We refer interested readers to Kingma, Welling, et al. (2019) for further details on the model and estimation.

We implement a VAE using the pre-processed frame-level video data. We train a convolutional encoder-decoder neural network that includes fully connected layers in between to generate a 100-dimensional latent representation z. For more detailed information about the network structure and the training process, please refer to Appendix B.3.

Once the model is trained, we use the trained encoder to generate a lower-dimensional representation z for each frame. We then aggregate the z's at the video level by taking the average across all frames within the video. To align with the unit of observation in our mortgage origination data, we further aggregate these video-level image embeddings to the lender-DMA-year level by calculating the weighted average across both national and local ads, where the ad spending per capita is the weight.

2.3.2 Text Features

To extract features that capture the messages conveyed in the ads, we first obtain the video transcripts using the Amazon Transcribe API, a speech recognition service that converts the audio content of the ad into text. Similar to image analysis, there are two broad approaches to extracting features from textual data. One could seek to identify a set of interpretable topics or sentiments within the ads (e.g., ease of application or a low mortgage rate). Alternatively, one could use a pre-trained large language model to obtain text embeddings that may not be directly interpretable but contain more comprehensive information. As the text features will serve as control variables in our application, we choose the second approach and use a text embedding model to represent the content of ads.

To represent the transcript of each ad in a low-dimensional vector, we utilize a pre-trained embedding model from OpenAI. Specifically, we use the "text-embedding-ada-002" model from OpenAI. While OpenAI offers multiple embedding models, such as "davinci," "curie," and “baggage” that are better suited for different tasks (e.g., clustering or search), they recommend “ada-002” for most use cases.[13] The ada-002 model generates 1,536-dimensional embeddings per document. Since this dimensionality is still relatively high to be used as control variables in our causal inference model, we further compress the embeddings into 100 dimensions using another VAE model. Further details on the network structure and training process can be found in Appendix B.3. Similarly as the image embeddings, we aggregate these text embeddings to the lender-DMA-year level by taking the weighted average across both national and local ads, where the ad spending per capita is the weight.

2.4 Empirical Strategy

After establishing that lenders have increased the minority share in their ads, we seek to understand how consumers respond to minority representation in ads. In this section, we describe our empirical strategy for estimating the impact of minority share in ads on consumers’ choice of lenders. We begin by describing a benchmark regression model with fixed effects in Section 2.4.1, which can account for many concerns related to advertisers targeting ads with greater minority representation over specific DMAs and time periods. However, these fixed effects regressions could be vulnerable to potential confounding variables. To address this concern, we use the double machine learning (Double ML) estimator as our main empirical strategy, which is described in Section 2.4.2. The Double ML estimator allows us to account for the high-dimensional video features described in Section 2.3 in a flexible functional form and obtain consistent estimates for the main parameters.

[13]For more information, see https://platform.openai.com/docs/guides/embeddings. Accessed July 13, 2023.

2.4.1 Benchmark: Regression with Fixed Effects

We start by describing a benchmark regression model with a large number of fixed effects to estimate the impact of ad spending and minority representation in the ads on consumer choices:[14]

$$y_{i,j} = \beta_1 \cdot log(1 + Ad_{i,j}) + \beta_2 \cdot log(1 + Ad_{i,j}) \cdot MS_{i,j} + \delta_{j,m(i)} + T_{j,t(i)} + e_{i,j}, \qquad (2.4)$$

where $y_{i,j}$ is a binary variable that equals 1 if consumer i chooses lender j and 0 otherwise. $Ad_{i,j}$ represents the total ad spending per capita by lender j in DMA $m(i)$ in year $t(i)$. Here, $m(i)$ denotes the DMA where consumer i resides, and $t(i)$ denotes the year when the consumer obtains a refinance loan. Note that the $m(i)$ and $t(i)$ subscripts are suppressed whenever the i subscript is included because each customer only considers a refinance in a specific market during a specific year, making the extra two subscripts superfluous. $MS_{i,j}$ indicates the corresponding minority share in the ads lender j has in market $m(i)$ in year $t(i)$, as defined in Section 2.2.3.

This regression includes lender-DMA fixed effects $\delta_{j,m(i)}$ and lender-year fixed effects $T_{j,t(i)}$. The lender-market fixed effects account for local, time-invariant confounding factors, such as lenders consistently advertising more in certain DMAs with higher demand or certain consumer characteristics. These fixed effects can also account for higher demand due to a particular lender having more offices or a longer history in a particular market. The lender-year fixed effects capture global, time-varying confounding factors, such as lenders

[14] Because multinomial logit or probit models do not handle large numbers of fixed effects well, for the ease of computation, we use a linear probability model as a reasonable approximation for the more micro-founded multinomial logit or probit models, similar to Tsai and Honka (2021). This also allows us to be parallel to our Double ML model, where it is challenging to run a multinomial logit or probit. To account for potential correlations within individuals, we bootstrap the standard errors, and further details are provided in Section 2.4.2.

choosing to advertise more, or including more minorities in advertising, in certain time periods, such as the increase in minority representation in advertising that occurred after the murder of George Floyd. The two key parameters of interest, β_1 (which measures the baseline level of advertising effectiveness) and β_2 (which captures how advertising effectiveness varies based on different levels of minority share), would then be identified from the variation in advertising by specific lenders within given markets across time. We show in Appendix B.4 that there is sufficient residual variation in advertising with fixed effects that allows for the estimation of these parameters. If this level of variation were to be approximately random, then the fixed effects model would yield causal estimates.

Besides the main effects, we could also estimate the results vary based on each consumer's race. Let $\mathbb{1}\{M_i\}$ be an indicator variable that takes the value of 1 if individual i is from a racial minority group and 0 otherwise (i.e., non-Hispanic White). We extend the model in Equation 2.4 as follows:

$$\begin{aligned} y_{i,j} = &(\beta_1 + \beta_2 \cdot \mathbb{1}\{M_i\}) \cdot \log(1 + Ad_{i,j}) + (\beta_3 + \beta_4 \cdot \mathbb{1}\{M_i\}) \cdot \log(1 + Ad_{i,j}) \cdot MS_{i,j} \\ &+ \delta_{j,m(i)} + T_{j,t(i)} + e_{i,j}, \end{aligned} \tag{2.5}$$

where β_2 captures the difference in baseline advertising effectiveness for minority consumers compared to White consumers, and β_4 estimates the difference in the impact of minority share in ads on minority consumers compared to White consumers. If the estimated β_4 is positive and statistically significant, it indicates that minority representation in ads has a stronger impact on minority consumers compared to White consumers.

Similarly, we examine the heterogeneous effects based on borrowers' political ideology. We use the census tract-level voting outcomes from the 2020 presidential election to represent the political ideology of each consumer. We create a new variable DEM_i that represents the difference in the vote shares between Biden and Trump in 2020 within the census tract where consumer i's property is located. More specifically, DEM_i is defined as $\frac{(\text{number of Biden votes}_i - \text{number of Trump votes}_i)}{(\text{number of Biden votes}_i + \text{number of Trump votes}_i)}$. Third party votes are discarded in this calculation. DEM_i is bounded between 1 and -1, where 1 (-1) means that 100% of votes went to Biden (Trump). We then estimate the following regression:

$$
\begin{aligned}
y_{i,j} = & (\beta_1 + \beta_2 \cdot DEM_i) \cdot \log(1 + Ad_{i,j}) + (\beta_3 + \beta_4 \cdot DEM_i) \cdot \log(1 + Ad_{i,j}) \cdot MS_{i,j} \\
& + \delta_{j,m(i)} + T_{j,t(i)} + e_{i,j}.
\end{aligned}
\tag{2.6}
$$

Similar to Equation 2.5, β_2 captures the different baseline advertising effectiveness based on consumers' political leanings and β_4 captures how the impact of minority share in ads varies with political leanings. If the estimated β_4 is positive and statistically significant, it indicates that ads featuring minority actors have a stronger impact on liberal-leaning consumers compared to conservative-leaning consumers.

2.4.2 Double Machine Learning

While the benchmark fixed effects regressions are likely to account for the largest sources of endogeneity, it is still possible that advertisements with higher minority representation may also differ in terms of their messaging or other video features. We can account for these effects using the Double ML estimator (Chernozhukov et al., 2018). Double ML allows us to estimate causal effects in the presence of high-dimensional covariates. It has recently gained increasing popularity in economics and marketing for causal inference using observational data with high-dimensional control variables (e.g., Dube et al., 2020; Ellickson et al., 2022; Gershon and Jiang, 2022; Gordon et al., 2022). The high-level intuition behind Double ML is to leverage machine learning models to remove or "partial out" the influences of high-dimensional control variables from both the outcome and treatment variables. By doing so, we obtain orthogonalized residuals, which are then used to estimate the causal parameters. We start by describing the Double ML estimator in our application. We specify the outcome model as a partial linear model:

$$y_{i,j} = \boldsymbol{\beta} D_{i,j} + g\left(X_{i,j}\right) + e_{i,j}, \tag{2.7}$$

where $y_{i,j}$ denotes the binary choice variable as defined previously, and $D_{i,j}$ denote the key causal variables of interest: ad spending per capita, $log\left(1 + Ad_{i,j}\right)$, and the interaction term of ad spending and minority share, $log\left(1 + Ad_{i,j}\right) \cdot MS_{i,j}$.[15] $X_{i,j}$ denotes our high-dimensional control variables: the visual and textual features of the ads as described in Section 2.3, as well as lender, DMA, and year fixed effects. All the variables in $X_{i,j}$ can be thought of as nuisance variables that need to be accounted for in the model but are not the main variables of interest. The impact of the high-dimensional control variables $X_{i,j}$ on the outcome $y_{i,j}$ is captured through a flexible function denoted by $g\left(\cdot\right)$.

One naive approach to estimating Equation 2.7 would be to simply fit a machine learning model to obtain the estimate of the flexible function $\widehat{g\left(X\right)}$ and plug it into the regression model in Equation 2.7. However, under such an approach, the estimates for the main coefficients of interest $\widehat{\boldsymbol{\beta}}$ will be biased. The reason behind this bias can be understood through the regularization in machine learning models, which results in $\mathbb{E}(\widehat{g\left(X\right)}) \neq g(X))$, and introduces regularization bias. While this bias diminishes as the sample size (n) increases,

[15]To estimate heterogeneous effects, $D_{i,j}$ can further include the interaction terms with consumers' own race $\mathbb{1}\{M_i\}$ as in Equation 2.5, and interaction terms with political leaning DEM_i as in Equation 2.6.

it does so at a rate slower than $n^{-1/2}$. Furthermore, the machine learning model can also overfit the training data because of the flexible functional form, resulting in overfitting bias.

The Double ML estimator solves both issues of regularization bias and overfitting bias through orthogonalization and sample-splitting. For orthogonalization, we estimate a second equation to predict the key causal variables $D_{i,j}$ given control variables $X_{i,j}$ (video features and fixed effects) through another flexible function denoted by $h(\cdot)$:

$$D_{i,j} = h\left(X_{i,j}\right) + \epsilon_{i,j} \tag{2.8}$$

To implement Double ML in our application, we use a Random Forest as the machine learning model to obtain the estimates of the conditional expectations $\widehat{l(X)} = E\widehat{(y \mid X)}$ and $\widehat{h(X)} = E\widehat{(D \mid X)}$. Using neural networks as the machine learning model gives similar results. Since we have multiple treatment variables, we fit a separate machine learning model for each treatment variable. After fitting the model, the Double ML estimate is obtained through a residuals-on-residuals regression. Using vector notation, we use the residuals of the outcome variable as $\widehat{\boldsymbol{e}} = \boldsymbol{y} - \widehat{\boldsymbol{l}(\boldsymbol{X})}$, and the residuals of the treatment variables as $\widehat{\boldsymbol{\epsilon}} = \boldsymbol{D} - \widehat{\boldsymbol{h}(\boldsymbol{X})}$. These residuals can be thought of as the variations in the dependent and key causal variables (e.g., advertising spending, minority share in ads) after controlling for, or "partialling-out" the effects of the control variables. The main parameters of interest can then be estimated as:

$$\widehat{\boldsymbol{\beta}} = \left(\widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{\epsilon}}\right)^{-1}\left(\widehat{\boldsymbol{\epsilon}}'\widehat{\boldsymbol{e}}\right) \tag{2.9}$$

There is one more consideration. The procedure above deals with regularization bias, but the estimated $\widehat{\boldsymbol{\beta}}$ may still be biased due to overfitting bias. This issue is solved by sample-splitting, where we randomly partition the data into K subsets, called folds. For each fold k, we fit the machine learning models to obtain $\widehat{l(\cdot)}$ and $\widehat{h(\cdot)}$ using all folds except the k-th fold, take the fitted models, and estimate $\widehat{\boldsymbol{\beta}^{\boldsymbol{k}}}$ using the k-th fold. The key is that the observations used to estimate $\widehat{\boldsymbol{\beta}^{\boldsymbol{k}}}$ are different from those used to fit the machine learning models. Doing so avoids bias that can arise due to overfitting. After iterating through all K folds, we compute the final Double ML estimate by averaging the K estimates. In our application, we have experimented with different numbers of folds, ranging from two to four, and obtained similar results. We opt for two folds for computational efficiency. We refer interested readers to Chernozhukov et al. (2018) and references wherein for more technical details.

The conventional approach of obtaining standard errors from the Double ML estimator does not directly apply in our setting. This is because the conventional approach assumes that the error terms are independent and identically distributed across all observations. This assumption is violated in our multinomial linear probability model because the observations for each individual are correlated since each consumer chooses one of the lenders for refinancing. To properly account for the correlation structure, we calculate the standard errors at the individual level using the block bootstrapping technique (Cameron and Miller, 2015). More specifically, we resample data on an individual level with replacement and compute the parameter estimate for each bootstrap sample. The bootstrap standard error is then calculated as the standard deviation of the 200 bootstrap estimates.

2.5 Empirical Results

In this section, we present results from both the fixed effects model and the Double ML model. We start by discussing the main effect of including more minority actors on the overall effectiveness of advertising in Section 2.5.1. We then describe the heterogeneous effects based on consumer characteristics, including race and political leanings, in Section 2.5.2.

2.5.1 Main Effect

As discussed in Section 2.2.1, we estimate both models on a random subset of 30% of consumers (or 2.89 million consumers). Table 2.4 presents the results of estimating the impact of minority representation in ads on consumers' choice of lenders. The fixed effects model corresponds to Equation 2.4, while the Double ML model corresponds to Equations 2.7 and 2.8, where the treatment variables D include ad spending and the interaction term of ad spending and minority share in ads. The results from the two models are not statistically different. The baseline effect of ad spending on lender choice (β_1) is positive and statistically significant. Moreover, the main parameter of interest, the interaction term of ad spending and minority share (β_2), is positive and statistically significant. These results indicate that a higher representation of minority actors increases the overall effectiveness of advertising.

Table 2.4: Effects of Minority Share in Ads on Consumer Choices

	Lender choice	
Model	F.E. Reg.	Double ML
	(1)	(2)
$\beta_1 : log(1 + Ad)$	0.026***	0.035***
	(0.006)	(0.007)
$\beta_2 : log(1 + Ad) \cdot$ MS	0.056**	0.060***
	(0.022)	(0.020)
N	81,203,548	81,203,548

Notes. S.E.s are clustered at the candidate level; $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

To interpret the effect size of the estimates, we calculate the implied advertising elasticity of demand, which is commonly used as a measure of advertising effectiveness. With simple algebra, the advertising elasticity can be derived as: $\frac{\partial y}{\partial Ad}\frac{Ad}{y} = \beta_1 \cdot \frac{Ad}{1+Ad}\frac{1}{y} + \beta_2 \cdot MS \cdot \frac{Ad}{1+Ad}\frac{1}{y}$, where the notations are defined the same way as in Section 2.4. The first term represents the elasticity of ad spending in the absence of minority representation, while the second term represents the incremental effect of the minority share in ads on ad elasticity. Using the sample averages for all the variables, we calculate the average advertising elasticity of demand to be 0.030 in the fixed effects model and 0.038 in the Double ML model. The effect sizes are broadly in-line with previous studies: the average elasticity is 0.023 for 288 consumer packaged goods (Shapiro et al., 2021), 0.026 for cigarette product placement on TV (Goli et al., 2022), 0.030 for auto insurance (Tsai and Honka, 2021), 0.031 for antidepressant (Shapiro, 2022), 0.05−0.06 for satellite TV operators (Yang, Lee, and Chintagunta, 2021), and 0.08 for e-cigarettes (Tuchman, 2019).

To examine the impact of minority share on advertising effectiveness, we calculate the average advertising elasticity under different levels of minority share in ads while keeping the total level of TV advertising spending constant. Specifically, we present the advertising elasticity for two levels of minority share: 15% (close to the median minority share at the lender-DMA level) and 25%, using our preferred Double ML model. The results are presented in Table 2.5. As the minority share increases from 15% to 25%, the estimated elasticity increases from 0.037 to 0.042, representing a 13.6% increase in relative terms. This result suggests that increasing the minority share in advertising can increase the effectiveness of advertising at an economically meaningful level.

Table 2.5: Advertising Elasticities with Minority Representation

Minority share	Advertising elasticity	
	Mean	95% C.I.
	(1)	(2)
15%	0.0373	[0.0207, 0.0534]
25%	0.0424	[0.0224, 0.0623]

Note. The range in brackets [ ] denotes the 95% confidence interval.

2.5.2 Heterogeneous Effects

In this section, we investigate how the impact of including more minorities in advertising varies with the borrowers' characteristics. We start by examining how the results vary with the borrower's race. The results are presented in Table 2.6. Across both models, the coefficient of minority share in ads for minority borrowers (β_4) is positive and significant. This indicates that higher minority representation in ads has a larger impact on minority borrowers compared to White borrowers. The effect of minority share for White borrowers (β_3) is also positive and significant, although smaller than that for minority borrowers.

Table 2.6: Heterogeneous Effects based on Consumers' Race

	Lender choice	
Model	F.E. Reg.	Double ML
	(1)	(2)
$\beta_1 : log(1 + Ad)$	0.030***	0.032***
	(0.006)	(0.007)
$\beta_2 : log(1 + Ad) \cdot \mathbb{1}\{M\}$	-0.014***	0.0003
	(0.002)	(0.002)
$\beta_3 : log(1 + Ad) \cdot \text{MS}$	0.043*	0.043**
	(0.022)	(0.020)
$\beta_4 : log(1 + Ad) \cdot \text{MS} \cdot \mathbb{1}\{M\}$	0.108***	0.061***
	(0.006)	(0.004)
N	81,203,548	81,203,548

Notes. MS denotes the minority share in ads; $\mathbb{1}\{M\} = 1$ for minority borrowers. Standard errors, clustered at individual, in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Further comparing the results from the fixed effects model and the Double ML model, we observe that although the results are generally in-line with each other, there are some differences. In particular, we see that the baseline impact of ad spending with no minority actors ($\beta_1+\beta_2$) is much smaller in the fixed effects regression than in the Double ML model, while the impact of including minority actors in the advertising ($\beta_3+\beta_4$) is larger for minority customers in the fixed effects regression. These differences suggest that ads featuring minority actors may contain some video features or messaging attributes that can particularly influence minority borrowers' choice of lenders, and that a failure to account for these features in a simply fixed effects regression may overstate the impact of featuring minority actors in ads on minority borrowers.

To interpret the effect size of the estimates, we examine the advertising elasticities for both White and minority borrowers at two different levels of minority share (15% and 25%), using our preferred model of Double ML. The results are presented in Table 2.7. Among White borrowers, when the minority share increases from 15% to 25%, the advertising elasticity increases by 11.0%. The effect is even more pronounced among minority borrowers. With the same change in minority share, the advertising elasticity increases by 21.8%. These results confirm that increased minority representation in ads has a larger impact on minority borrowers compared to White borrowers.

Table 2.7: Advertising Elasticities based on Consumers' Race

Minority share	White consumers		Minority consumers	
	Mean (1)	95% C.I. (2)	Mean (3)	95% C.I. (4)
15%	0.0326	[0.0171, 0.0481]	0.0404	[0.0196, 0.0615]
25%	0.0362	[0.0171, 0.0554]	0.0492	[0.0244, 0.0743]

Note. The range in brackets [ ] denotes the 95% confidence interval.

The stronger impact observed on minority borrowers aligns with previous behavioral research indicating that minority borrowers are more sensitive to racial cues and show stronger racial homophily effects (e.g., Deshpandé and Stayman, 1994; Aaker et al., 2000; Mollica et al., 2003). However, the positive impact on White borrowers suggests that other mechanisms beyond racial homophily are likely contributing to the observed effects.

We also examine the heterogeneous effects based on the borrower's political ideology. The results are presented in Table 2.8. Recall that DEM is defined as the difference in the vote shares between the Democratic and Republican candidates in the 2020 election, divided by the sum of the vote shares for the Democratic and Republican candidates. Both models show that the coefficient on the interaction term between the minority share in ads and the vote share difference (β_4) is positive and significant, indicating that increased minority representation in ads has a larger impact on liberal-leaning borrowers than on conservative-leaning borrowers. However, the magnitudes are again quite different between the two models. Ultimately, we believe that the Double ML estimates represent the more robust estimates because this model accounts for the higher-level video and transcript attributes.

Table 2.8: Heterogeneous Effects based on Political Leaning

	Lender choice	
Model	F.E. Reg.	Double ML
	(1)	(2)
$\beta_1 : log(1 + Ad)$	0.032***	0.035***
	(0.006)	(0.007)
$\beta_2 : log(1 + Ad) \cdot$ DEM	-0.035***	-0.020***
	(0.003)	(0.002)
$\beta_3 : log(1 + Ad) \cdot$ MS	0.044**	0.067***
	(0.022)	(0.020)
$\beta_4 : log(1 + Ad) \cdot$ DEM $\cdot$ MS	0.050***	0.020**
	(0.007)	(0.005)
N	81,203,548	81,203,548

Notes. MS denotes the minority share in ads; DEM denotes the difference in vote shares between the Democratic and Republican candidates. Standard errors, clustered at individual, in parentheses. $^{**}p < 0.05$; $^{***}p < 0.01$.

To interpret the effect size of the estimates, we examine the advertising elasticities for two groups of borrowers: moderately liberal and moderately conservative. Specifically, we consider a moderately liberal-leaning group with a 25% advantage for the Democratic candidate (i.e., $DEM = 0.25$) and a moderately conservative-leaning group with a 25% advantage for the Republican candidate (i.e., $DEM = -0.25$). For each group, we consider the impact of changing the minority share in ads from 15% to 25%. Results are presented in Table 2.9. Among liberal-leaning borrowers, the advertising elasticity increases by 17.6%, while the advertising elasticity increases by 15.7% among conservative-leaning borrowers, which is slightly smaller.

These results are consistent with related studies that indicate that individuals with liberal-leaning political ideologies are more likely to be more supportive of racial diversity and related social movements in multiple contexts (e.g., Agarwal and Sen, 2022; Aneja et al., 2023; Babar et al., 2023). We do not find a strong negative impact of minority representation on conservative-leaning borrowers. Even when we extrapolate DEM to an extreme value of -1 (i.e., 100% of votes going for the Republican candidate), the total impact of minority share in ads is positive in the double ML model and close to 0 in the fixed effects model. One caveat with this analysis is that the political-leaning data is only observed at the aggregate census tract level. We will revisit this point when discussing results from our experimental study in Section 2.6.

Table 2.9: Advertising Elasticities based on Political Leaning

	Liberal consumers		Conservative consumers	
Minority share	Mean (1)	95% C.I. (2)	Mean (3)	95% C.I. (4)
15%	0.0346	[0.0164, 0.0528]	0.0418	[0.0267, 0.0569]
25%	0.0407	[0.0190, 0.0624]	0.0471	[0.0289, 0.0652]

Note. The range in brackets [ ] denotes the 95% confidence interval.

2.6 Experimental Study

We complement the analysis with observational data with an experimental study, where we directly manipulate the race of actors using generative AI technology. We describe the experimental design in Section 2.6.1 and discuss the results and implications in Section 2.6.2. The experimental study serves two main purposes. First, while we believe that the Double Machine Learning estimator provides causal estimates, there is always the hypothetical concern of endogeneity or omitted variable biases with observational data. With the experiment, we are able to measure a clean causal relationship with random assignment of ads with different racial compositions. The fact that our results from the observational study match those from the experiment adds confidence that our empirical findings are not driven by some subtle endogeneity story. Second, we use the experiment to inform us about the potential mechanisms by asking participants a number of attitudinal questions about the ads they see, which we discuss in Section 2.7.

2.6.1 Experimental Design

As outlined in our preregistered research plan (https://aspredicted.org/B7C_P49), we aim to recruit a total of 2,800 participants from CloudResearch. We plan to recruit 2,000 participants who are either mortgage borrowers or homeowners, and due to the limited available pool of participants, supplement with 800 general population participants. The purpose of prioritizing homeowners is to ensure that the sample is comparable to the borrowers in the observational data. We ended up with a sample of 1,903 participants who were either mortgage borrowers or homeowners and 902 general population participants. After excluding participants who failed the attention check (n = 9), our final sample consists of 2,796 participants with an average age of 42 years and 51.3% female.[16]

Participants were presented with the following text: "In this survey, we would like you to imagine that you currently have a mortgage loan on your home and you are considering refinancing the mortgage to reduce interest rates. You come across a refinance advertisement from AnchorPoint Refi. Please consider the ad as you are thinking about your refinancing decisions." They were then shown an advertisement stating: "Refinance with AnchorPoint Refi and start saving today! Whether you're looking to lower your monthly payments, consolidate debt, or fund home improvements, AnchorPoint Refi has got you covered. Our customized solutions are tailored to your unique financial situation, and with our online application process, refinancing has never been easier. We'll help you get the best possible rate and save money with no stress or hassle. Check out the people we've helped. Apply today at AnchorPointRefi.com and become our next success story!" Below the text, the ad also featured customers who had recently refinanced with AnchorPoint Refi. Participants were randomly assigned to one of seven conditions, with each condition featuring actors from different racial groups in the ad.

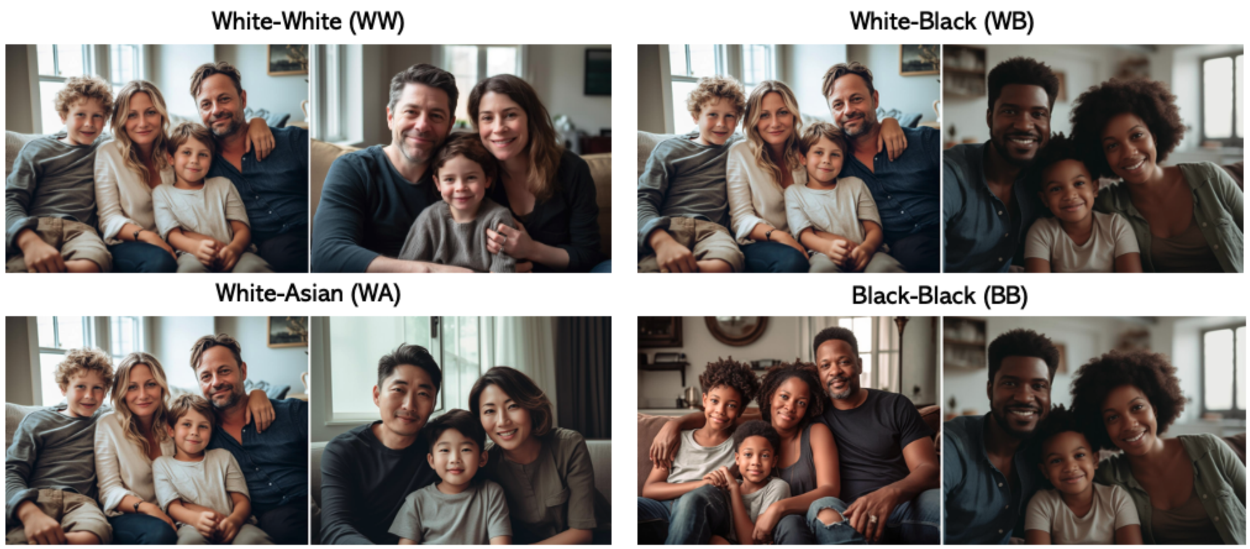
We use Midjourney V5, a generative AI technology that generates highly realistic images based on text prompts, to experimentally manipulate the race of the actors in the ads. To ensure similarity among the generated images, we provided explicit and detailed instructions to Midjourney, including specifications for the number of children and their genders. This is to ensure that other aspects of the images do not differ systematically across conditions. In four out of the seven conditions, participants were presented with an image featuring two families. As shown in Figure 2.4(a), these four conditions include two White families (WW), one White and one Black family (WB), one White and one Asian family (WA), or two Black

[16]We exclude two outliers in the reported age (660 and 677) when calculating the average age.

families (BB). The remaining three conditions, shown in Figure 2.4(b), featured an image with one White family (W), one Black family (B), or one Asian family (A). The purpose of having both the two-family conditions and the single-family conditions is to explore whether the positive consumer responses observed in our observational study hold when ads feature racially diverse representation (the two-family conditions) as well as minority representation (the single-family conditions).

Figure 2.4: Advertising Images Featuring Different Races

(a) Two-family conditions



(b) Single-family conditions

After showing one of the seven advertisements, we measure two key dependent variables (DVs): the likelihood of submitting a loan application with the advertised lender and the likelihood of recommending the advertised lender to a friend who is looking to refinance. Both variables are measured on a scale of 1 to 7, where 1 indicates "Not at all likely" and 7 indicates "Very likely." While the DV of the likelihood to submit an application is better aligned with our dependent variable in the observational analysis, we also included the DV of the likelihood to recommend the lender, which may better capture the overall brand attitude or impression. The two measures are positively correlated with the correlation coefficient $\rho = 0.771$.

To explore the potential mechanisms behind our results, we also ask participants six attitudinal questions after they respond to the key outcome variables: "To what extent would you agree or disagree with the following statements?" on a 1 to 7 scale where 1 indicates "Strongly disagree," and 7 indicates "Strongly agree." The statements cover the breadth of product offerings ("The advertised lender has broad, flexible loan options that fit different financial situations and needs"), whether the respondent felt included ("The advertised lender caters to people like me"), financial inclusiveness ("The advertised lender is inclusive towards individuals of all backgrounds"), fair lending practices ("The advertised lender has fair lending practices without predatory interest rates and hidden fees"), the freshness of the ad ("The advertisement feels new and fresh"), and whether the ad garners attention ("The advertisement is attention-grabbing").

Lastly, participants were asked to provide their demographic information, including age, gender identity, and race and ethnicity (White/Caucasian, Black/African American, Hispanic/Latinx, Asian, Native American/Alaska Native, Mixed race/multiracial, Others, and Prefer not to disclose). They were also asked to describe their political orientation on a 1 to 7 scale, where 1 indicates "Very liberal" and 7 indicates "Very conservative."

2.6.2 Experimental Results

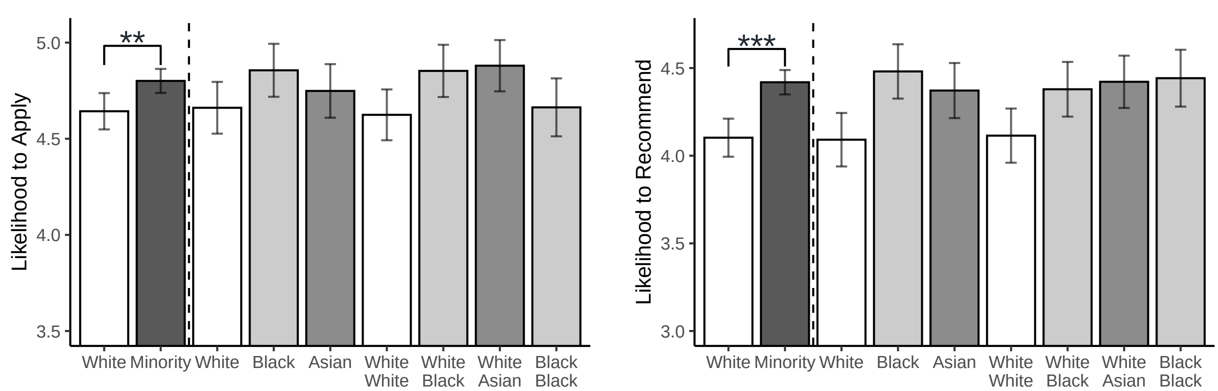
Main Effect

We start by presenting the results of the key outcome measures. Figure 2.5(a) shows the likelihood-to-apply measure and Figure 2.5(b) shows the likelihood-to-recommend measure. We first compare the conditions featuring minority families (WB, WA, BB, B, and A) with those featuring only White families (WW and W). Overall, participants who are randomly assigned to advertisements featuring minority families report a higher likelihood of submitting an application with the advertised lender compared to those who see White families (4.80 vs. 4.64, $p = 0.006$). The gap is larger for the likelihood to recommend: Participants who see minority families report a higher likelihood to recommend the advertised lender compared to those who see White families (4.42 vs. 4.10, $p < 0.001$). The experimental results are consistent with the observational study where we find that ads with a higher minority share are more effective.

Figure 2.5: Likelihood to Apply and Recommend

(a) Likelihood to Apply

(b) Likelihood to Recommend



Notes. Error bars denote the 95% confidence interval; $^{**}p < 0.01$; $^{***}p < 0.001$.

Figure 2.5 further shows the results separately for each of the 7 experimental conditions. In the single-family conditions, participants in the Black condition report a higher likelihood of applying with the lender (4.86 vs. 4.66, $p = 0.048$) and recommending the lender (4.48 vs. 4.09, $p < 0.001$) compared to those in the White condition. The Asian condition shows a slightly higher likelihood to apply (4.75 vs. 4.66, $p = 0.376$) and a higher likelihood to recommend (4.37 vs. 4.09, $p = 0.012$) compared to the White condition.

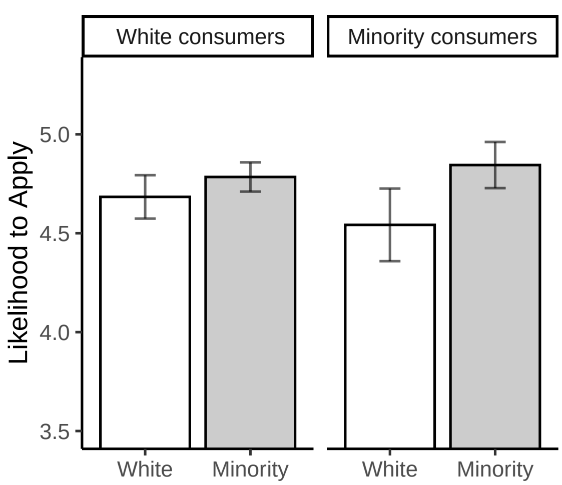
Among the two-family conditions, we use the condition with two White families as the benchmark. Participants in the White-Black condition are more likely to apply to the lender (4.85 vs. 4.62, $p = 0.018$) and recommend the lender (4.38 vs. 4.11, $p = 0.018$). Similarly, participants in the White-Asian condition are more likely to apply (4.88 vs. 4.62, $p = 0.008$) and recommend (4.42 vs. 4.11, $p = 0.005$). When participants see an ad featuring two Black families, the likelihood of applying is just slightly higher (4.66 vs. 4.62, $p = 0.703$) and the likelihood of recommending is significantly higher (4.44 vs. 4.11, $p = 0.004$). The results with two Black families point to the possibility of a boundary condition in our observational study: The measured effect may not fully extrapolate to a region where the minority share is substantially higher than what was observed in our observational data. With that said, despite a possible boundary condition, the minority share in ads observed in the data is still lower than the actual share of minority borrowers in the mortgage refinancing market. Overall, while not all comparisons show statistical differences, we find converging results that featuring minority actors increases advertising effectiveness.

Heterogeneous Effects

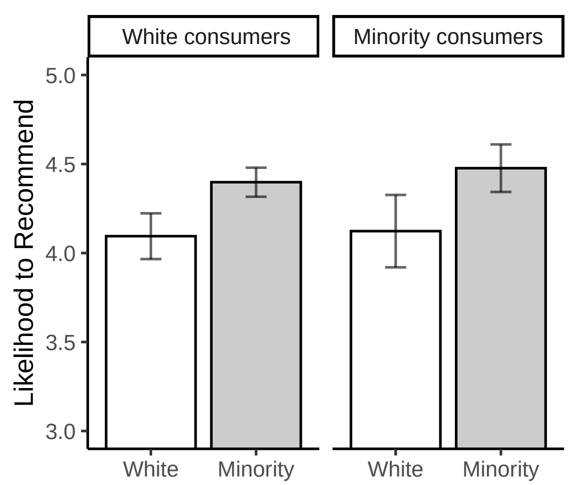
We examine the heterogeneous effects based on the self-reported racial/ethnic information of the participants. After excluding 10 participants who chose not to disclose their race, there are 2,025 participants who identified as "White/Caucasian," and we group the rest 761 participants as "minority consumers." Figure 2.6 presents the likelihood of making a loan application and a recommendation for these two groups under the conditions that displayed only White families (WW or W) and those with minority families (the other 5 conditions). Figure 2.6(a) shows that the presence of minorities in the ads increases the likelihood of applying much more for minority consumers (4.84 vs. 4.53, $p = 0.007$) than White consumers (4.78 vs. 4.68, $p = 0.136$). Figure 2.6(b) shows that both White consumers (4.40 vs. 4.09, $p < 0.001$) and minority consumers (4.48 vs. 4.12, $p = 0.005$) report a significantly higher likelihood of recommending the lender when they are assigned to conditions featuring minority families compared to conditions featuring White families. We report the results for each condition separately in Appendix B.5. These results are broadly consistent with those from our observational study, where we find that while ads with a higher minority share are more effective for both White and minority groups, the impact of minority share is stronger among minority consumers.

Figure 2.6: Heterogeneous Effects By Race

(a) Likelihood to Apply



(b) Likelihood to Recommend



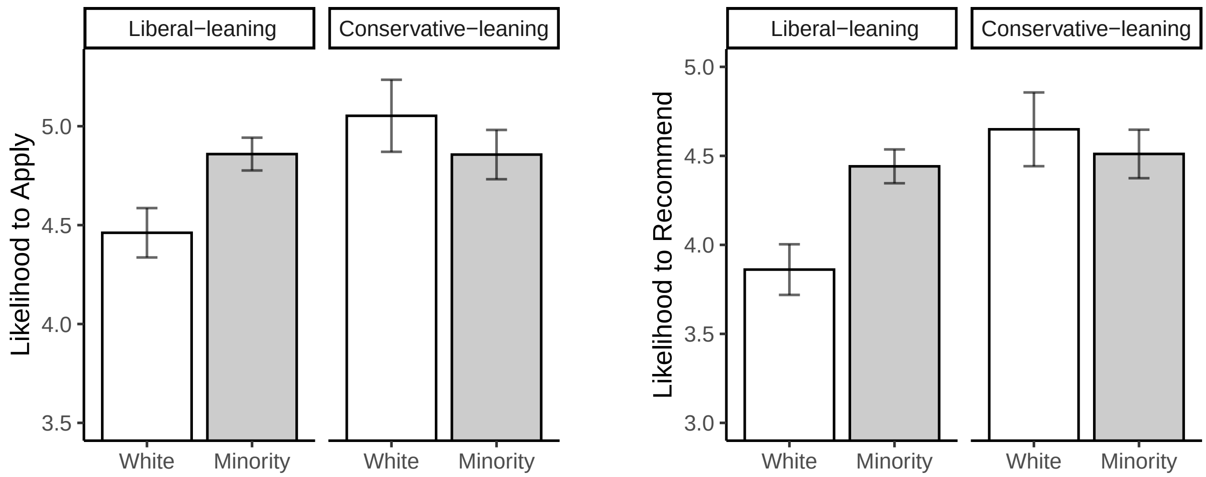
Note. Error bars denote the 95% confidence interval

Next, we investigate how the results vary based on the self-reported political leanings of participants. As a reminder, our political scale ranged from 1 to 7, where 1 represents "Very liberal," and 7 represents "Very conservative." There are 1,454 "liberal-leaning" participants who chose 1 to 3 on this scale, and 786 "conservative-leaning" participants who chose 5 to 7. As shown in Figure 2.7(a), liberal-leaning consumers report a significantly higher likelihood to submit an application when they are assigned to a condition featuring minority families compared to a condition featuring only White families (4.86 vs. 4.46, $p <= 0.001$). Conservative-leaning consumers, on the other hand, are less likely to apply when assigned to conditions that feature minority families compared to only White families (4.86 vs. 5.05, $p = 0.083$), although this effect is borderline insignificant. This pattern also holds when we look at the likelihood to recommend, as shown in Figure 2.7(b). Liberal-leaning consumers are significantly more likely to recommend the advertised lender when assigned to the conditions featuring minority families (4.44 vs 3.86, $p < 0.001$), while conservative-leaning consumers do not have a significant difference in their likelihood of recommending (4.51 vs 4.65, $p = 0.275$). These experimental results are broadly consistent with the observational study, where we find that the impact of minority representation is larger for liberal-leaning consumers compared to conservative-leaning consumers, although they cast doubt as to whether conservative customers respond positively or negatively to minority representation in an ad.

Comparing Figures 2.6 and 2.7, the difference in consumer responses to ads featuring White versus minority families appears even larger when participants are grouped based on their political leaning rather than their own race. This suggests that the effectiveness of minority representation in ads may be better predicted by political ideology than race. Recall that our observational study indicates that race plays a larger role than political leaning (Tables 2.7 and 2.9). This could be because while we observe individual-level race information in the observational data, we rely on the census tract level data to proxy political leaning

Figure 2.7: Heterogeneous Effects By Political Leaning

(a) Likelihood to Apply (b) Likelihood to Recommend



Note. Error bars denote the 95% confidence interval

for each consumer, which inherently introduces measurement errors. The experimental results suggest that we would likely expect an even greater difference between liberal- and conservative-leaning consumers based on individual-level political ideology.

2.7 Potential Mechanisms

So far, we have shown converging evidence from both the observational and experimental studies. In this section, we discuss several potential mechanisms that could explain our findings. To do so, we draw on related prior literature, our empirical results, and the follow-up questions in the experiment that measure participants' perceptions of the advertised lender and the advertisement (see Section 2.6.1). Our goal here is not to pin down a single definitive mechanism; rather we show evidence of several possible explanations that are consistent with our findings. Indeed, the effect that ads with minority representation are more effective is likely to be multi-determined.

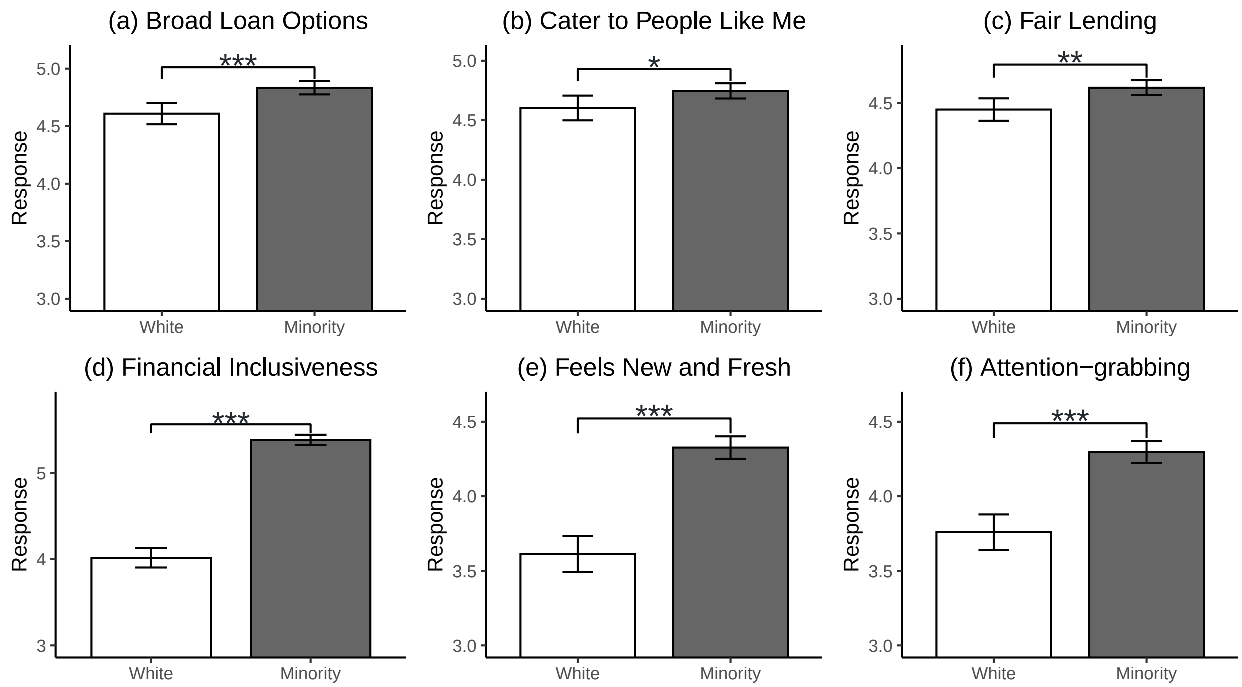
One potential explanation for our findings is that customers care about the racial match between themselves and the race shown in the ads, with the importance of match being especially high for minority customers. Past behavioral literature has found that racial fit plays a significant role in advertising effectiveness, particularly for minority consumers (e.g., Deshpandé and Stayman, 1994; Aaker et al., 2000). In our study, we find that minority consumers are significantly more likely to believe that "the advertised lender caters to people like me" when they see ads featuring minority actors, as opposed to ads with White actors (4.79 vs. 4.13, $p < 0.001$). White consumers, on the other hand, show only a marginal increase in the belief that ads featuring White actors cater to people like them compared to ads with minority actors (4.79 vs. 4.73, $p = 0.354$). Taking these two findings together, placing minority actors in advertisements should be more effective because of the positive effect this has on minority consumers and non-negative response from White consumers in terms of racial fit. This is consistent with our finding that minority representation has a stronger impact on minority consumers from both observational and experimental studies (Table 2.6 and Figure 2.6). Racial fit is unlikely to be the only mechanism, however, since we find ads featuring minority consumers are also more effective among White consumers, although at a smaller magnitude. There are a number of mechanisms that are consistent with this result, as well.

One possible mechanism is that the presence of minority actors in advertisements can speak to consumers' support for diversity and minority representation. Several recent studies have documented greater support for DEI initiatives among liberal-leaning individuals (Agarwal and Sen, 2022; Aneja et al., 2023; Babar et al., 2023). Consistent with these studies, we find that liberal-leaning consumers, including White customers, respond more positively toward ads featuring minority actors compared to conservative-leaning consumers. The heterogeneity of the effect across the political spectrum suggests that the extent of support for diversity and minority representation, or attitude about race in general, likely plays an important role in determining how consumers react to ads featuring minority consumers.

Moreover, the presence of minority actors in ads can affect consumers' perceptions of the advertised brand. In other words, even when presented with the same ad copy for a fictional brand, consumers may perceive the brand differently based on the race of the actors. This aligns with prior research that finds that firms with diverse workforces are perceived to be more moral Khan and Kalra (2022). We examine brand perceptions using the attitudinal measures collected in the experimental study. Figure 2.8 compares the results for participants who are randomly assigned to see ads featuring White vs. minority families. Detailed results by each condition can be found in Appendix B.5. After seeing ads featuring minority consumers, participants are more likely to perceive the advertised lender to have broad and flexible loan options (4.83 vs. 4.61, $p < 0.001$), have fair lending practices without predatory pricing and hidden fees (4.62 vs. 4.45, $p = 0.001$), and be inclusive towards individuals of all backgrounds (5.38 vs. 4.01, $p < 0.001$). All these brand perceptions are positively correlated with the likelihood of applying for a loan and recommending the lender (see Appendix B.5). Minority representation can increase the ads effectiveness through these favorable brand perceptions.

Lastly, ads featuring minority actors can be more effective simply because they are less common, making them stand out and appear more salient to consumers. Having minority representation in ads, therefore, can be an effective way for firms to differentiate their advertisements from others and increase their effectiveness (Pieters et al., 2002; Rosengren et al., 2020). In our experimental study, we find that participants are more likely to perceive the ads featuring minority actors as new and fresh (4.33 vs. 3.61, $p < 0.001$) and attention-grabbing (3.88 vs. 3.64, $p < 0.001$) compared to ads with White actors. Therefore, ads with

Figure 2.8: Consumer Perceptions of the Advertised Lender and Advertisement



Notes. Error bars denote the 95% confidence interval; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

minority actors are more effective because of these visual cues, which are positively correlated with the likelihood of loan application and recommendation (see Appendix B.5).

2.8 Conclusion

Given the growing emphasis on diversity and minority representation, it is crucial for brands to understand the impact of including racial minority actors in ads on advertising effectiveness. In this paper, we find that greater minority representation in TV ads increases the effectiveness of advertising in the mortgage refinancing market. The impact of minority representation is stronger among minority consumers as well as liberal-leaning consumers. A pre-registered experimental study where we directly manipulate the race of the actors in ads shows consistent results with the observational study. Leveraging the heterogeneous results from our observational study and attitudinal questions from the experimental study, we discuss several potential mechanisms that are consistent with our findings.

Our research offers valuable insights for brands seeking to promote racially diverse and inclusive representation in their advertising strategies. Our results suggest that featuring minority actors not only achieves the social goal of increasing minority representation but also leads to higher advertising effectiveness, given on the current level of advertising spending. Our results also have important policy implications. Since featuring minority actors in advertising is an effective strategy for reaching minority consumers, it has the potential to contribute to improving financial inclusion in the mortgage market. In particular, minority consumers have been shown to be less likely to refinance when it is beneficial to do so (Gerardi et al., 2021; Gerardi et al., 2023). Increasing minority representation in TV advertising can encourage refinancing among minority consumers and help reduce racial disparities in the mortgage market.

Chapter 3: The Role of Slant and Message Consistency in Political Advertising Effectiveness: Evidence from the 2016 Presidential Election

3.1 Introduction

Political advertising is typically the largest expense for political campaigns, with the vast majority of this spending going toward television advertising (Baer and Sinagoga, 2018; Adgate, 2019; OpenSecretes, 2020). In the 2016 U.S. election, candidates, parties, and political action committees (PACs) spent \$6 billion on television advertising, accounting for 8% of U.S. television ad revenue in 2016 (Kaye, 2017). With many online sites like Twitter and Google restricting online political advertising (Kaye, 2019b; Kaye, 2019a), political expenditures on television advertising rose to even higher levels – \$8.5 billion – in the 2020 U.S. election (Passwaiter, 2020). Numerous studies justify this heavy investment in political television advertising, which has significant impacts on election outcomes (e.g., Huber and Arceneaux, 2007; Gordon and Hartmann, 2013; Spenkuch and Toniatti, 2018; Wang et al., 2018). Indeed, changes in political television ad spending strategies in 2000 could have resulted in a different U.S. president (Gordon and Hartmann, 2013).

While there is a large literature that studies the effectiveness of political advertising, limited research has examined how the content of political advertisements (beyond tone) changes the effectiveness of these ads. Given recent advances in text analytic methods (see Berger et al., 2020), this offers an opportunity to further examine the way in which the content of political advertising shapes voter behavior. More broadly, text analytic methods offer a means of examining how messages from human brands (e.g., Thomson, 2006) affect consumers' perceptions.

We apply text analysis to national political television ads from 2016 by the two main presidential candidates, Hillary Clinton and Donald Trump, to assess two key features of the advertising messages: slant and consistency. Slant refers to the extent to which each candidate's messages are extreme versus centrist, while consistency refers to how much the candidate's messages remain consistent between the primary and the general election.

We focus on these aspects of ad measures for two key reasons. First, politically there is a conventional wisdom that after candidates win their nominations, they should moderate their positions and become more centrist over the course of a campaign (e.g., Hummel, 2010). However, changing one's positions and tone can demotivate the base of voters who supported the candidate in the primary election, and reduce the creation of brand associations that inform voters as to what the brand stands for. We aim to empirically investigate this tension between slant and consistency in our analyses. Second, both slant and consistency are two vital dimensions related to the branding of political candidates, with slant representing what the product (the candidate) stands for, and consistency representing the extent to which the candidate creates a clear and repeated message of what they represent.

Keller and Lehmann (2006) consider the importance of being consistent in the brand message, asking whether differing messages to distinct segments may cause confusion of what the brand stands for. Kotler and Keller (2016) similarly argue that message consistency is a fundamental element of the brand, playing a vital role in brand building and brand equity creation and reinforcement. As in the political domain, research on brand extensions links the branding dimensions of slant and consistency. Specifically, work on brand extensions has long studied how the similarity and dissimilarity of associations between a brand variant and the schema of that brand impacts the success of the brand variant and its ad effectiveness (e.g., Boush and Loken, 1991; Park et al., 1991; Broniarczyk and Alba, 1994; Y. Liu et al., 2017). Overall, research on branding underscores the vital importance of slant and consistency in impacting the effectiveness of brand messages.

In our investigation, we examine how these two aspects of advertising content affect two distinct measures of voter engagement, online word-of-mouth (WOM) about the candidate (e.g., Bermingham and Smeaton, 2011; Jahanbakhsh and Moon, 2014) and voter preference captured through daily polling polls (e.g., Jennings and Wlezien, 2018; Kennedy et al., 2018; Silver, 2018), both of which have been shown to predict the candidate's vote share. Furthermore, online WOM and consumers' related online behaviors have been widely leveraged as a means to study the effectiveness of television advertising (e.g., Lewis and Reiley, 2013; Joo et al., 2014; Liaukonyte et al., 2015; Fossen and Schweidel, 2017; Tirunillai and Tellis, 2017; Fossen et al., 2021).

Utilizing data on over 800 television ad airings during the 2016 U.S. presidential election, we find that ad messages that use language that is more (1) centrist and (2) consistent with a candidate's primary-election platform are associated with increases in online WOM volume and voter preference for the candidate. We also see some evidence that these attributes are more important early in the campaign. Overall, the results support that candidates should favor more moderate messaging for the general election as well as messaging that is consistent with their primary election communications in the early stages of the campaign. This pattern demonstrates a nuance to the conventional wisdom that candidates should move to more centrist messaging after winning their nomination. While this strategy is beneficial, consistency with the candidate's primary election messaging is particularly important early in the campaign. These results suggest that it would be advantageous for candidates to adhere close to their primary campaign messages in the early stage of the general election and emphasize moderate messages as the election further develops. Further, our results suggest that the rising use of extremist messages in political advertising (e.g., Bartels, 2016; Wells and Seetharaman, 2018) may be a flawed strategy for candidates that could decrease candidate-related WOM volume and voter preference for the candidate.

With this investigation, we acknowledge the challenge of establishing a causal relationship in a study using historical data. Our main results are based on using an event study in which we compare the volume of WOM on Twitter just before versus just after an advertisement. As we discuss later, while using such an approach removes most sources of coincident timing, there may remain some feasible confounds. For example, our measures of slant and consistency may be correlated to another attribute of the ad that is really the driver of our results. Further, an anonymous reviewer noted that an advertisement could prompt a viewer to tweet about the campaign at the time of the ad, when instead the viewer might have made the same tweet at a later time. In our analyses, we control for several observable ad attributes to reduce potential confounds. Ultimately, we discuss the assumptions under which the identified relationships are likely causal and allow the reader to decide whether these conditions are likely to be met. We also present confirmatory regression analysis using voter preference. This secondary analysis has more potential endogeneity concerns but is broadly consistent with the WOM analysis, presenting confirmatory correlational evidence of our key findings.

Our findings contribute to the political marketing literature by considering the role of message content in political advertising. Though prior research has examined the differential impact of tone (e.g., Lovett and Shachar, 2011; Wang et al., 2018) and message source (Wang et al., 2018; L. Zhang and Chung, 2020), we are among the first studies to consider other aspects of political ad content, namely the similarity of messaging over the course of the campaign with regard to the party (slant) and the candidate's earlier messaging (consistency). Our findings demonstrate that these dimensions of political advertising content affect voter behaviors beyond conventionally studied advertising dimensions in the political domain (e.g., tone, source, volume of advertising). Our results also contribute to the broader streams of research on advertising and text analysis. While extant research has examined the role of different themes in advertising campaigns (e.g., Bass et al., 2007), to the best of our knowledge, our research is among the first in marketing to use automated text analysis to derive message-related metrics that are linked to the performance of television commercials. This has implications for brands as they develop and launch new marketing campaigns.

The remainder of this paper proceeds as follows. Section 3.2 describes the data and our key measures and presents descriptive evidence of our findings. Section 3.3 discusses our identification strategy and model specifications, and the results are reported in Section 3.4. Section 3.5 concludes with a discussion of the implications of our findings, both within the political domain and more broadly for marketers.

3.2 Data Description

We combine four data sources: national political television advertising, political communications from campaign speeches and congressional records, daily poll data on voter preferences, and online chatter about the candidates from Twitter. We detail each of these in turn.

3.2.1 Television Advertising Data

We collect data on political advertisements that aired on national primetime television during the 2016 presidential general election (June 30th to November 8th) from Kantar Media's Stradegy database. We utilize national television advertising because it is very challenging to obtain daily (or more granular) ad outcomes (i.e., polling and social media measures) at the DMA level. National ad buys typically account for more than 25% of a presidential candidate's television advertising spending (Miller, 2015). This share is expected to increase, as the rising costs of local ad inventory in battleground markets during elections make national ad buys more economical (Passwaiter, 2018). Given that national ads have significantly higher reach and that our outcome measures of voter behavior are collected at the national level, we believe it is reasonable to focus on national ads in our investigation of how slant and consistency relate to their effectiveness.[1][2] The Stradegy data contains the date and time each ad airs, the program and network in which the ad airs, ad length, ad position, and the advertiser's name. We supplement the Stradegy data by collecting information on the number of viewers for each ad from Comscore's TV Essentials database. In total, our data include 824 ad airings for 60 unique ad creatives aired by 11 political advertisers, which include the campaigns for the two main candidates, two political party entities, and seven PACs.

We present descriptive information about the political ads in Table 3.1. Although the number of unique ad creatives are similar across the candidates, Clinton has more ad airings, while ads supporting Trump have larger audience sizes. Ads supporting the two candidates are comparable in tone, length, and ad position. Figure 3.1 illustrates the number of ad airings over time by candidate.

[1]The WOM measures are available only nationally. While there are state-level polls, these tend to be conducted very irregularly, and by a number of different pollsters, making the creation of a high-quality state-specific panel very difficult (and nearly impossible to do so at the even more granular DMA level).

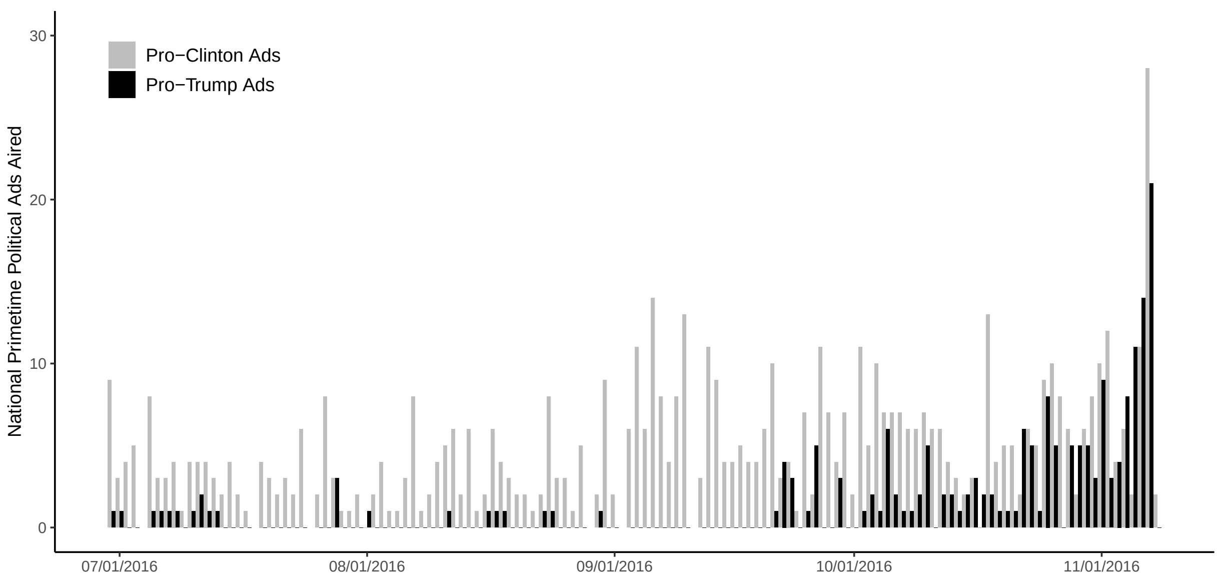
[2]We probe the overlap of national versus local primetime political ads in our data using the name of the TV creative and find significant overlap. This suggest that the content of national ads likely reflect the campaigns' overall message. Specifically, 100% of both Clinton and Trump's national primetime ad creatives in our data were also aired locally in primetime. That said, the candidates do employ more ad creatives for their primetime local ad airings. Yet, the majority of their local primetime airtime consists of ad creatives that are aired both locally and nationally. Specifically, only 3.1% (9.2%) of Clinton's (Trump's) primetime local ad airings consist of ad creatives that were only aired locally. Additionally, we find that Clinton and Trump's national primetime ads exhibit very similar airing patterns as their local primetime ads in terms of time, day of the week, and month aired (see Appendix C.7).

Table 3.1: Descriptive Statistics for Candidates' Advertisements

Candidate	Unique ad creatives	Ad airings	% attack ad airings	Ad length in seconds (SD)	Relative ad position in break (SD)	Break position in program (SD)	Audience in 1,000s (SD)
Clinton	28	634	71.50%	30.30 (4.12)	0.48 (0.21)	3.63 (3.02)	1,396.56 (3,181.85)
Trump	32	190	75.30%	31.70 (10.30)	0.48 (0.19)	3.96 (3.44)	2,403.34 (3.444.20)
Total	60	824	73.33%	30.62 (6.16)	0.48 (0.20)	3.71 (3.12)	1,634.34 (3.271.20)

Note. Relative ad position is defined as the ad's position in the break divided by the number of ads in the break

Figure 3.1: National Political Television Ads from June 30th to Election day in 2016



3.2.2 Using Text Analysis to Derive Measures of Slant and Message Consistency

We define two dimensions of ad messaging for our analysis: political slant and message consistency. For political slant, we evaluate each ad in a purely political domain and measure the extent to which the ad message is extreme versus centrist. For message consistency, we evaluate each ad in the domain of semantic meanings and calculate the extent to which the ad message has similar content to the candidate's platform, which is derived from their speeches during the primary campaign. We detail each in turn.

Slant

We operationalize our slant measure using a technique introduced by Gentzkow and Shapiro (2010). For this, we utilize multiple datasets: the 114th Congressional Record, which includes all speeches and debates made on the congress floor during early 2015 to early 2017 (collected from Stanford Social Science Data and Software), public speech data during the primary elections (collected from American Presidency Project), ad creative transcriptions, and vote share data (collected from https://dailykos.com). We pre-process our texts – including punctuation, stop word and number removal, tokenization, and stemming – before constructing slant.

Following Gentzkow and Shapiro (2010), we first derive the mapping between a vector of word counts a politician used[3] and the political leanings of their district, as measured by the

[3] Among the words in the congressional records that remain after pre-processing, we focus on words that appear at least 2 times and fewer than 100 times in the candidates' public speech documents and transcribed ad texts. We then select 1,000 words for analysis that are found to be most asymmetrically used by the parties. We conduct sensitivity analysis and find that our results are robust to the number of words in the analysis.

Republican vote share in their district.[4] Specifically, we estimate the extent to which each word is associated with the two parties. Table 3.2 presents a subset of the words that are most associated with a Democratic or Republican slant.[5] We observe that these words match the conventional wisdom of how they would map onto the conservative-liberal continuum, confirming the validity of the approach.

Table 3.2: Words that are Highly Associated with Slant by Party

Republican	Democratic
senat, administr, regul, govern, rule, iran, law, spend, energi, busi, land, time, washington, balance, power, author, west, servic, move, obamacare, defens, north, life, obama, appreci, nuclear, deal, execut, plan, legisl, produc, materi, cost, debt, forc, process, account, dollar, appropri, manufactur, office, purpos, veteran, secretari, compani, provid, babi, illeg, price, suspend, defend, freedom, oil, pass, militari, south, sell, sponsor, alli, industri, entitl, islam, arm, bank, iranian, natur, document, radic, grow, control, enemi, innov, amend, faith, misson, alien, agreement, terror, missile	gun, act, vote, women, violence, hous, health, student, educ, commun, right, children, york, african, invest, citi, school, background, public, caucu, check, cut, live, climat, address, major, worker, prevent, justice, fund, water, child, civil, equal, pay, repress, afford, infrastructur, immigr, polic, colleg, action, homeland, girl, moment, progress, silenc, oppos, system, polit, kill, crisi, clean, access, democrat, care, wage, effort, discrimin, undermin, improv, corpor, secur, join, democraci, danger, shoot, crimin, famili, trade, incarcer, voter, fair, loan, kid, research, lgbt, inequ

Note. All words are stemmed (e.g., business amd businesses are stemmed to busi).

After we estimate the slant of each word, we compute the slant of each television ad as the extent to which the words in the ad would be associated with being Republican (strongly Democratic messages have negative values). We then subtract 0.5, which corresponds to an even Democratic and Republican ideology. Finally, we multiply this score by -1 for Clinton's ads, so that a greater (lower) number corresponds to more extreme (more centrist) messages for both candidates compared with their party. We provide further technical details

[4] We consider the Republican candidate's vote share in the 2012 presidential election in the state for senator and electoral district for representative.

[5] We consider unigrams for our implementation. Although higher-order n-grams capture richer information in general, this severely reduces both the total number and variety of n-grams that we can extract from ads due to their shortness.

in Appendix C.1. Table 3.3 shows descriptive statistics for the slant measure. We observe that while most of the ads are fairly centrist, Trump's ads tend to be somewhat more centrist than Clinton's ads. Yet, we find that some ads from Trump appear to lean towards Democratic ideology, and some ads from Clinton appear to lean towards Republican ideology. We probe them and find that such ads from Trump highlight his promised support for gender equality, including equal pay and support for childcare. The Republican-leaning ads for Clinton discuss threats from nuclear weapons or the Islamic state. Therefore, these outliers seem to confirm the validity of our slant measure.

Table 3.3: Descriptive Statistics for Political Slant and Message Consistency

Panel A: Political Slant						
Candidate	Mean (SD)	Min	10th	50th	90th	Max
Clinton	0.148 (0.514)	-0.625	-0.406	0.201	0.906	1.108
Trump	-0.084 (0.353)	-0.795	-0.626	0.002	0.294	0.539
Panel B: Message Consistency						
Candidate	Mean (SD)	Min	10th	50th	90th	Max
Clinton	0.389 (0.067)	0.295	0.303	0.390	0.510	0.510
Trump	0.353 (0.066)	0.209	0.217	0.359	0.407	0.452

Message Consistency

We compute message consistency using doc2vec (Le and Mikolov, 2014) to measure the semantic similarity between the contents in a candidate's ads and their platform as articulated in their speeches during the primary election (collected from American Presidency Project). To understand doc2vec, we first discuss its predecessor, word2vec (Mikolov et al., 2013). Word2vec is a word embedding methodology that projects each word into a low dimensional vector space such that words that frequently co-occur in similar contexts are closely located in the space. That is, word2vec considers co-occurrences of words – a given target word and its surrounding/context words – and then assigns those words coordinates that are close to each other. Doc2vec extends this idea to allow the relationship to depend on the paragraph in which the words appear, allowing the context of the words to matter more broadly. Documents themselves are represented as high-dimensional vectors. In such a context, the cosine between the vectors for a pair of documents is the typical measure used to calculate textual similarities (e.g., Berman et al., 2019; Feng, 2020), which we use here.

We train the model using 62 documents as inputs: the 60 ad creatives plus two documents, each consisting of the text of the aggregated primary campaign speeches given by one of the candidates. We use the primary election speeches as the benchmark against which their ads are compared because this is the time when the candidates define their campaign's message. After training, we compute the cosine similarity between each candidate's ads and their aggregated speech document to assess how similar the ad message is to the candidate's messaging during the primary. We note that, in addition to capturing similarity in political issues, the similarity measure also captures other dimensions, such as frequently used attack phrases by each candidate and other phrases that set the tenor of their campaign, which is confirmed by our validation checks.[6]

In our implementation, we first pre-process the corpus, including punctuation removal, tokenization, and stemming, and then train the Distributed Bag of Words (DBOW) model using the gensim module in Python.[7] Our results under different dimensions of the vector space are reported in Appendix C.3. Table 3.3 shows descriptive statistics for our message

[6]Similar to Wu et al. (2019), we conduct two types of validation checks. First, we treat the training documents (i.e., 62 documents that we use as inputs) as if they were new, infer vectors for the documents using the trained model, and see whether the inferred documents are found to be most similar to themselves via cosine similarity. We find that all documents are most similar to themselves. Second, in order to showcase the face validity of our approach to measuring consistency, we select a few documents and show documents that are most and least similar in terms of the cosine similarity. See Appendix C.2.

[7]For hyperparameters, we set vector size = 200, epochs = 300, window = 5, sub-sampling = 10^{-2}, and negative = 5.

consistency measure for the ads in our data. We observe that Clinton ads have a higher level of consistency than Trump, while both candidates have similar standard deviations.

Our analysis approach of using automated text analysis tools to extract measures of political advertising content (specifically, slant and consistency) offer advantages over commonly used survey-based measures of political ad content, such as those from the Wesleyan Media Project, which uses the same Kantar Media Stradegy raw data we use. For example, the Wesleyan Media Project codes tone (i.e., positive vs. negative), topic, and certain emotions (e.g., fear, anger, sadness, etc....), but not any measures of slant and consistency, nor do they code other linguistic measures that have been derived using automated text analysis such as language sophistication, concreteness and familiarity, arousal and dominance, and objectivity and subjectivity. The automated extraction and analysis of the ad transcript, in contrast, does not impose such restrictions and enables the identification of textual features of interest.[8] Additionally, through the use of an embedding space, our analysis takes into account the context in which words appear. This is particularly relevant in the case of political advertisements, as political campaigns may seek to link emotions with a particular topic. While there may be alternative methods of measuring these constructs, the use of automated text analysis offers an objective method to derive these variables that is both scalable and reproducible. Our research is among the first in marketing to use text analysis to derive message-related metrics that are linked to the performance of television commercials. Similar approaches could be used outside of political marketing, for example to assess the importance of message consistency (perhaps along specific dimensions) in messaging used by product and service marketers. Overall, our analyses present an efficient, automated approach to construct measures of ad content.

[8]While our approach offers many benefits, one caveat is that the ad texts are very short. This may lead to some measurement error in our measures of slant and consistency. It is likely that any coding mechanism would also have some measurement error. Given that slant and consistency are independent variables, whatever measurement error does exist would be expected, on average, to lead to attenuation of coefficients.

3.2.3 Outcome Measures of Political ad Effectiveness

Online WOM

We collect online WOM data from Twitter about Clinton and Trump during the 2016 general election from Crimson Hexagon (now Brandwatch), a certified Twitter partner. This data includes information on the volume of Tweets about the two candidates in one-second increments of time. We then examine the WOM activity about the candidates for the five minutes before and after the television ad airings. Table 3.4 presents summary statistics of the percentage change in WOM around the airings of the ads. On average, both candidates experience about a 30% average increase in WOM between the five-minute window before the ad is shown and the five-minute window after the ad is shown.

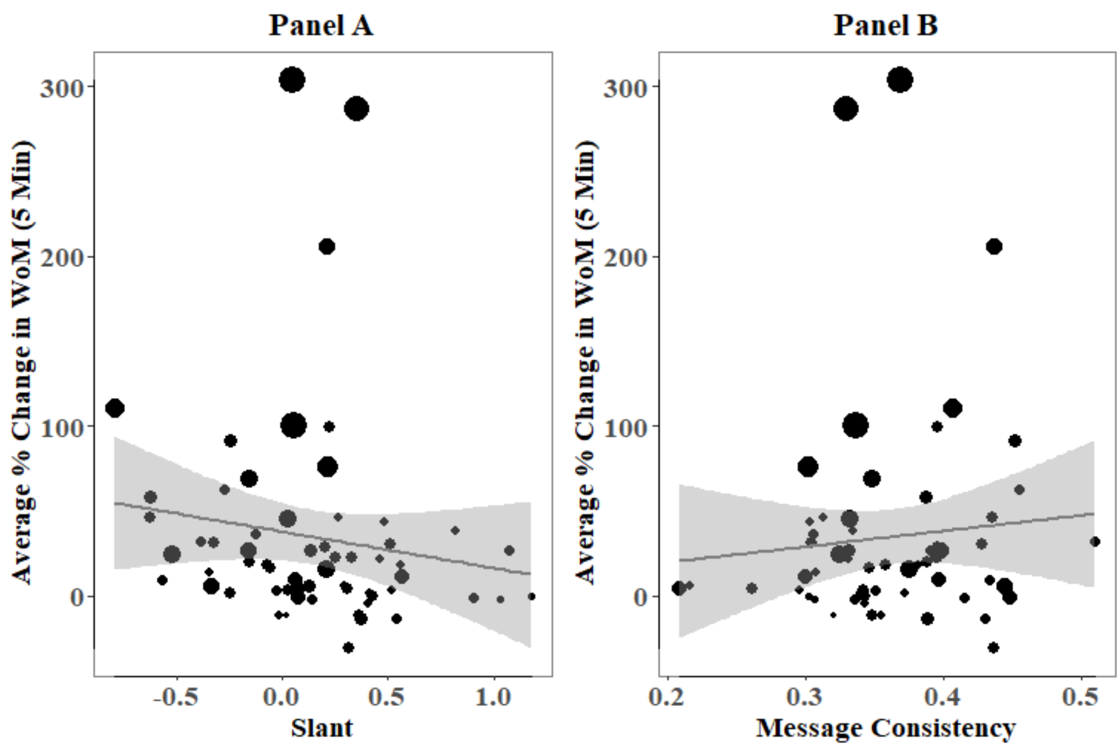
Table 3.4: Descriptive Statistics for Percentage Change in WOM and Voter Preference

Panel A: Percent change in WOM						
Candidate	Mean (SD)	Min	25th	50th	75th	Max
Clinton	35.5% (117.0)	−83.3%	−20.0%	6.1%	50.0%	930.0%
Trump	28.3% (96.3)	−60.3%	−15.5%	6.8%	30.9%	643.0%
Panel B: Voter preference						
Candidate	Mean (SD)	Min	25th	50th	75th	Max
Clinton	43.1 (1.4)	40.0	42.3	43.3	44.2	46.3
Trump	45.0 (1.6)	41.6	43.8	44.8	46.3	48.2

Note. Percentage change in WOM is calculated as $\frac{WOM^{post} - WOM^{pre}}{WOM^{pre}+1} \cdot 100$ at each ad each airing.

Figure 3.2 provides descriptive support for the idea that the changes in WOM may be correlated with political slant and message consistency of each ad. Panel A shows an overall negative relationship between the advertisement's slant and the average percentage change in WOM, indicating that extreme (centrist) messages may be associated with decreased (increased) WOM. Panel B suggests that message consistency is positively related with increases in the volume of WOM following the ads. These relationships are consistent with our empirical results.

Figure 3.2: Relationship between Slant, Message Consistency, and Percentage Change in WOM



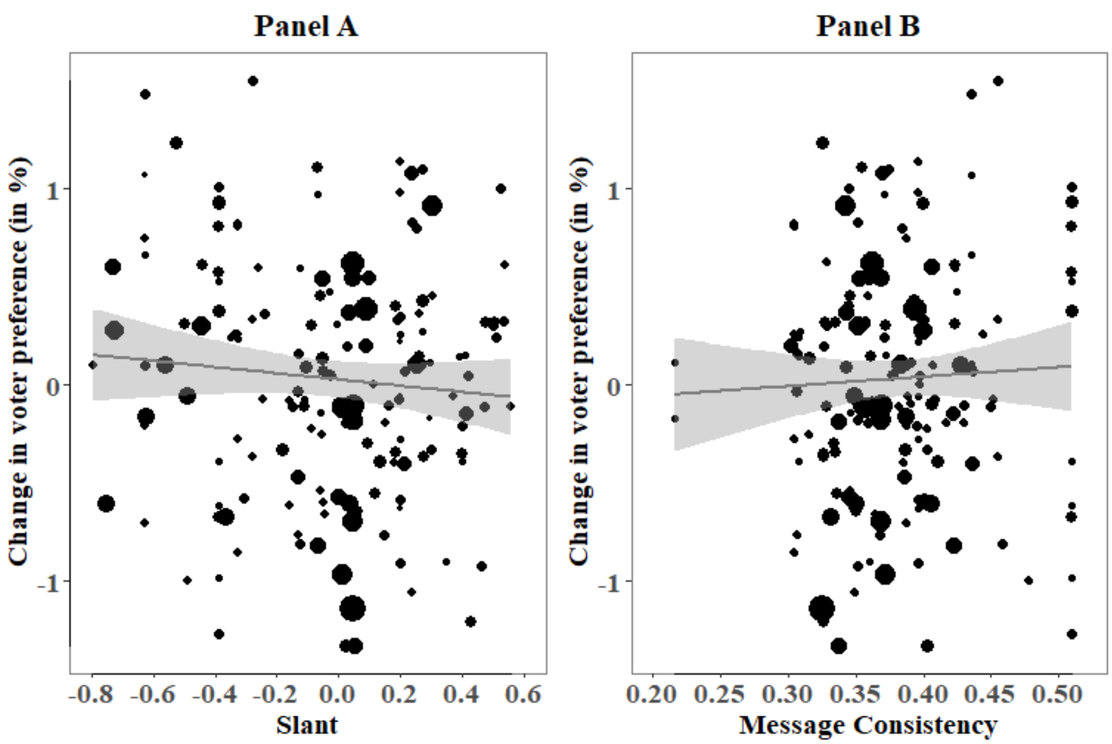
Note. The size of each dot represents the average ratings per each of the unique ad creatives (in millions).

Voter Preference Data

Our voter preference data is inferred from daily polls from the USC Dornsife/Los Angeles Times Poll. This poll surveyed voter preferences from about 3000 registered voters on a daily basis during the 2016 general election. We use the USC Dornsife/Los Angeles Times Poll data because this poll is one of the few daily tracking polls that surveyed a fixed set of voters, which has the advantage of revealing changes in respondents' preferences among the same panelists over time. We present summary statistics for the voter preference data in Table 3.4.

Figure 3.3 illustrates the relationship between the change in voter preferences, which is calculated as the difference between the two consecutive polling numbers, and different levels of slant and message consistency. Panel A shows that more extreme messages are correlated with decreased voter preferences, while Panel B shows that message consistency is positively related to changes in voter preference. These patterns match those of our empirical results.

Figure 3.3: Relationship between Slant, Message Consistency, and Change in Voter Preference



Note. The size of each dot represents the average ratings per each of the unique ad creatives (in millions).

3.3 Models

We model the impacts of political slant and message consistency on Twitter WOM in a manner consistent with prior literature on the effects of television advertising on online consumer behaviors (e.g., Liaukonyte et al., 2015; Fossen and Schweidel, 2017). These approaches leverage granular time windows around an ad's airing to study the impact of the ad. Using short time windows makes it unlikely that outside variables will impact the outcome measure. As such, past work has argued that such identification strategies that analyze changes in behaviors in narrow time windows around television advertisements are effective at investigating the causal impact of television advertising and can produce results similar to a randomized experiment (e.g., Lewis and Reiley, 2013; Joo et al., 2014; Liaukonyte et al., 2015; Fossen and Schweidel, 2017).

An important element for such identification strategies is the exogenous nature of ad positioning (e.g., Wilbur et al., 2013; Fossen and Schweidel, 2017; Deng and Mela, 2018). Advertiser-network contracts rarely write-in the airing time of an ad or even state the specific show in which an ad will air (Liaukonyte et al., 2015). Television networks decide the sequence of ads and commonly use a random order within ad breaks, an assertion further verified in more recent data sets on television advertising (Deng and Mela, 2018; McGranaghan et al., 2019; Fossen et al., 2021). If the selected time-slots deliver insufficient ratings, networks compensate the advertisers by rerunning the ad in a comparable spot on the same or a similar show to make up the remaining rating points (Katz, 2013, p. 200). Consequently, advertisers currently have limited control over selecting a specific program, let alone the specific ad break, in which to air an ad to affect immediate online WOM.

The main threat to this identification strategy would occur if other campaign activities (e.g., social media positing) coincide with national TV ads, and these other activities drive changes in the outcome. Our use of granular time windows, and the random positioning of ads, make such coordination very unlikely. Nevertheless, to further probe this concern, we assessed Clinton and Trump's tweet activity in our data window. Clinton (Trump) tweeted 3,732 (2,688) times during our data window. Only 196 (60) of these were posted five minutes before or after the airing of one of their national ads. We test whether this is over-representative of what would be expected in a benchmark case without any coordination between advertising and social media. Specifically, we generate empirical probability distributions for ad airing and Twitter posting times for each candidate and draw randomized ad airing and posting times at the minute level, holding dates fixed. From the simulated data, we find that, on average, one would get 191 (69) tweets from Clinton (Trump) that would overlap the 5-min pre- and post-ad windows if these were set in an uncoordinated manner. Statistical tests on whether the coincident probabilities are different between the two conditions are rejected, suggesting that we do not see evidence for coordination. Additionally, we find that none of these coincident tweets in the observed data seem to be coordinated with the television advertising (e.g., mention the ad in order to amplify its impact). Thus, we don't see evidence that campaigns were coinciding their social media positing activities with their national television advertising activity to amplify their reach.

We measure the volume of online WOM mentioning the candidates using a five- minute window before and after the ad airs.[9] Specifically, for each ad airing i, we model the volume of post-ad WOM as follows:

$$\begin{aligned}\log(1+WOM_i^{post}) = \beta_0 + \beta_1 \cdot \log(1+WOM_i^{pre}) + \beta_2 \cdot Slant_i + \beta_3 \cdot Consistency_i \\ + \beta_4 \cdot Attack_i + X_i^{'}\rho + \epsilon_i,\end{aligned} \tag{3.1}$$

where WOM_i^{post} is the online WOM volume about the candidate that occurs from the beginining of ad i until five minutes later. Similarly, WOM_i^{pre} is the volume of online WOM about the candidate that occurs from five minutes before an ad starts airing to the beginning of ad i. $Slant_i$ and $Consistency_i$ are the political slant and message consistency measures discussed in Section 3.2, respectively. $Attack_i$ is a dummy variable indicating that the ad

[9]For robustness, we also run the analysis using 2-min and 3-min time windows. See Appendix C.4.

is an attack ad or not. X_i is a vector of control variables that have been shown to impact political ad effectiveness and/or WOM activity following television ads (e.g., Fossen and Schweidel, 2017; Wang et al., 2018). These variables include a dummy variable for which candidate the ad supports,[10] the log of audience size, a dummy variable of ad length (which is equal to 1 if the ad is greater than 30 seconds and 0 otherwise), the relative ad position in the ad break, program genre fixed effects, network fixed effects,[11] day of the week, time of the day, and week of the data window in which the ad airs.

Additionally, we test whether slant and message consistency have different effects on WOM across time as follows:

$$
\begin{aligned}
\log(1 + WOM_i^{post}) = \beta_0 &+ \beta_1 \cdot \log(1 + WOM_i^{pre}) \\
&+ \beta_2 \cdot PreOct1st_i \cdot Slant_i + \beta_3 \cdot PostOct1st_i \cdot Slant_i \\
&+ \beta_4 \cdot PreOct1st_i \cdot Consistency_i + \beta_5 \cdot PreOct1st_i \cdot Consistency_i \\
&+ \beta_6 \cdot PreOct1st_i \cdot Attack_i + \beta_7 \cdot PreOct1st_i \cdot Attach_i \\
&+ X_i' \rho + \epsilon_i,
\end{aligned}
\tag{3.2}
$$

where $PreOct1st_i$ ($PostOct1st_i$) are equal to 1 if ad i airs before Oct. 1st (on or after Oct. 1st) and 0 otherwise. We also interact tone ($Attack_i$), a vital component of political ad content, with the time division in order to test whether attack vs. non-attack ads differentially

[10]This is done through the inclusion of a Clinton dummy variable. We consider only whether the ad is pro-Clinton or pro-Trump, and not whether the ad was sponsored by the candidate or by a supporting PAC because the vast majority of Trump's ads were PAC ads while the vast majority of Clinton's ads were run by the campaign. Thus, the Clinton coefficient captures both the difference between Clinton and Trump as well as the difference between candidate ads and PAC ads.

[11]As there are many networks in the data, we group networks with fewer than 7 ad airings together as "Other Networks."

affects WOM activities over the course of the election campaign. We estimate Eq. (3.1) and (3.2) with clustered standard errors at the candidate level.[12]

To add credence to the WOM analysis, we additionally run confirmatory regressions of voter preferences on political slant and message consistency. In modeling daily voter preferences, we control for lagged voter preference and include a rich set of controls and fixed effects to help us isolate impacts of slant and consistency on voter preference. Nevertheless, it is impossible to control for all potential confounds. We model voter preference for candidate c at day t as:

$$\begin{aligned} VP_{c,t} = \alpha_c + \gamma_0 \cdot VP_{c,t-1} + \gamma_1 \cdot Slant_{c,t-1} + \gamma_2 \cdot Consistency_{c,t-1} \\ + \gamma_3 \cdot NoAds_{c,t-1} + X^{'}_{c,t-1}\beta + T_{w_t} + \epsilon_{c,t}, \end{aligned} \tag{3.3}$$

where $VP_{c,t}$ is voter preference of candidate c at day t (in %), α_c is a candidate fixed effect, and T_{w_t} is a weekly fixed effect. $Slant_{c,t-1}$ and $Consistency_{c,t-1}$ are the daily measures of slant and message consistency at $t-1$, respectively, which are measured as the weighted averages of these variables across all ads aired in a given day, where each ad is weighted by its audience size. $NoAds_{c,t-1}$ is a dummy variable equal to 1 if there are no ads supporting the candidate c on date $t-1$ and 0 otherwise. $X_{c,t-1}$ is a vector of control variables that includes the number of ads aired by the candidate, the audience size for the candidate as well as the rival's ads, the weighted average of attack tone, relative ad positions in breaks, ad length (as defined above), the share of ads that aired on different program genres (for

[12]We cluster at the candidate level to allow for the WOM residuals to correlate within candidates. However, we could instead cluster at the ad creative level, which is the level at which the slant and consistency measures vary. To guide our decision, we conduct statistical tests for the appropriate level of clustering (MacKinnon et al., 2023) and find support for clustering at the candidate level. This matches the recommendation of Cameron and Miller (2015) to cluster at a more aggregate level rather than a more disaggregate level. In Appendix C.5, we document the results of the statistical tests for clustering.

example, comedy, drama, news, etc....), and the share of ads that aired on major networks.[13] All of these control variables are calculated as the audience size weighted averages across all ads that aired in a given day, except for the number of ads aired and audience size.

Similar to Equation (3.2), we estimate a variation of Equation (3.3) where we interact $Slant$ and $Consistency$ (as well as $Attack$) with $PreOct1st$ and $PostOct1st$, to see if the two variables of interest have differing impacts on voter preference over the course of the election campaign. In both equations, we cluster standard errors at the candidate level.

3.4 Results

Table 3.5 presents the estimates for the WOM model. Column 1 presents the results from Equation 3.1. We first observe that politically extreme messages decrease the volume of candidate-related WOM. In contrast, consistent messages increase the volume of WOM. We also see that attack ads decrease the volume of WOM. Column 2 reports the results from Equation 3.2, which includes the time interactions on the slant, consistency, and attack tone variables. We observe that the impact of slant appears to be larger in the early stages of the campaign, although the differences in the coefficients in each time period are not statistically significantly different. Next, we find that consistent messages are positively associated with WOM, although its impacts are much stronger and statistically significant early in the campaign. Attack ads have a slight positive impact on WOM in the early stages of the campaign, but the effects become negative in the later stage of the campaign. To summarize, our results show that politically centrist and consistent messages are beneficial for the candidates in terms of spurring a larger volume of online WOM.[14]

[13] As there are many small networks with only a few ad airings, we select the top 10 networks by the number of ad airings and group the rest as "Other Networks."

[14] As noted in the introduction, it is also possible that the ads prompt a flurry of tweets that would have been sent at a later time in the absence of the ad.

Table 3.5: Effects of Political Slant and Message Consistency on WOM

	(1)	(2)
	$\log(1 + WOM^{post})$	
$\log(1 + WOM^{pre})$	0.649***	0.650***
	(0.002)	(0.0004)
Slant	−0.027***	
	(0.006)	
Consistency	0.214***	
	(0.039)	
Attack ads	−0.065***	
	(0.014)	
Slant × PreOct1st		−0.078
		(0.098)
Slant × PostOct1st		−0.002
		(0.017)
Consistency × PreOct1st		0.377***
		(0.053)
Consistency × PostOct1st		0.167
		(0.338)
Attack ads × PreOct1st		0.038*
		(0.021)
Attack ads × PostOct1st		−0.128**
		(0.057)
Pro-Clinton ads	−0.179**	−0.179**
	(0.091)	(0.084)
log(Audience size)	0.127***	0.129***
	(0.048)	(0.048)
Ad length	0.319*	0.316*
	(0.170)	(0.157)
Ad position in break	−0.125**	−0.105***
	(0.054)	(0.027)
Week, Day, Hourly F.E.s	Yes	Yes
Program Genre and Network F.E.s	Yes	Yes
N	824	824
R^2	0.777	0.779

Notes. S.E.s are clustered at the candidate level; *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

While we interpret the impacts measured in Table 3.5 as being causal, it is worth clarifying the assumptions under which we consider this effect. First, as we note in Section 3.3, the impact of the advertisements on WOM could be affected by coincident events with the advertisements. The institutional details about how television advertisements are sold makes such coordination to the level of a 5-min interval very difficult. We also note in Section 3.3 that we examined whether either candidate tweeted at times that coincided with the ads. We find that there were very few such tweets and that these numbers are not statistically different than those that one would expect with randomness. In Appendix C.6, we probe the influence of outliers on the results. Specifically, we consider (1) winsorizing (i.e., replacing outliers with certain percentiles of the data) and (2) trimming the post-WOM volume at the 1% and 99%, and show that the results from these alternative analyses are very similar to our main analysis.

Another concern may be that these ads are not placed on random networks, but on optimal networks, some of which may match the audience with the message of the ad. While some readers may consider this to be a form of endogeneity, a more accurate interpretation of our results, given that we account for the network in which the ads airs, would be that our estimates reflect a causal effect of slant and consistency on WOM after accounting for the campaign's matching mechanism between the ad and the best media on which to place the ad. That is, our estimates account for the total effect consisting of both the direct effect of a change in slant or consistency, as well as the indirect effect of how slant and consistency would change the medium used to deliver such a message.

While the results suggest that it is best for campaigns to present ad messages that are both politically centrist and consistent with their primary election messaging, these message characteristics may be at odds with each other. Thus, we consider whether slant or message consistency has a stronger impact on WOM behavior. Since the variables are on different scales, we compare the relative importance of these two measures using two different metrics. We first multiply the estimated coefficient with the difference between the 90th percentile and 10th percentile values for slant and consistency, respectively. As a robustness check, we also multiply the coefficients with the standard deviation (SD) of each variable. The results in Table 3.6 show that consistency has a stronger impact on WOM than slant. Specifically, an increase from the 10th to 90th percentile values of consistency (one SD) is associated with an increase in WOM by about 4.4% (1.5%), while a similar change in slant is associated with a decrease in WOM by about 3.8% (1.3%). As was the case above, we find that both slant and consistency matter more in the early stages of the campaign, with the contrast being especially large for consistency. This seems to suggest that people may be more responsive to a candidate's messages in political ads early in the campaign.

Table 3.6: Relative Importance of Slant and Consistency on WOM

	(1)	(2)
	Change in WOM	
	90th vs 10th percentile	Standard Deviation
Slant	−0.038	−0.013
Consistency	0.044	0.015
Slant × PreOct1st	−0.132	−0.039
Consistency × PreOct1st	0.077	0.025
Slant × PostOct1st	−0.002	−0.001
Consistency × PostOct1st	0.023	0.011

While WOM is a means to generate attention and spread the candidate's message, which has been linked to increased vote share (e.g., Bermingham and Smeaton, 2011; Jahanbakhsh and Moon, 2014), voter preference offers a more direct measure of voting behavior that accurately predicts election results (e.g., Kennedy et al., 2018; Silver, 2018). Thus, we examine the relationship between each candidate's voter preference and the slant and message consistency of the candidate's ads to lend further support to and confirm the relevance of the WOM analysis.

The results for our voter preference model appear in Table 3.7. Consistent with the WOM analysis, we observe from Column 1 that voter preference for a candidate is lower when the candidate's ads are more slanted, while it is higher when the candidate's ads are more consistent. The impact of whether the ad is an attack ad is very small. Column 2 reports the results when we interact slant, consistency, and tone with the time period (pre-Oct. 1st vs. post-Oct. 1st). We find that more consistent messages are strongly and positively associated with increases in voter preferences in the early stages of the election, but its effects in the later stages of the election are much smaller and not statistically significant. In contrast, the coefficients on slant are both negative and statistically significant. However, we again see that the impact of slant is stronger in the early stage of the general election. Additionally, the coefficient on attack ads in the late stage is negative, though only marginally significant. In sum, our analysis on voter preference suggests that a candidate would be better off by airing ads containing politically centrist and consistent messages, especially in the early stages of the campaign. While these results provide confirmatory evidence of the key findings from the WOM analysis, we note that we interpret these results as correlational, because it is possible that the strategy of ads being used (attack versus inconsistent new message versus slant) could be correlated with underlying trends or events not observed in our data that affect both the type of advertising and the voter preferences.

Many other papers that have sought to measure the impact of political advertising (e.g., Gordon and Hartmann, 2013; Spenkuch and Toniatti, 2018) have noted that measuring its impact is difficult because advertising intensity is often confounded with other campaign activities. However, the problem of measuring the impact on the quantity of advertising is somewhat different than the confounds we face in trying to understand how the message

Table 3.7: Effects of Political Slant and Message Consistency on Voter Preference

	(1)	(2)
	Voter preference (in %)	
Lagged voter preference	0.865***	0.855***
	(0.012)	(0.015)
Slant	−0.027***	
	(0.022)	
Consistency	0.352***	
	(0.072)	
Attack ads	−0.006	
	(0.082)	
Slant × PreOct1st		−0.323***
		(0.081)
Slant × PostOct1st		−0.231***
		(0.017)
Consistency × PreOct1st		0.688**
		(0.339)
Consistency × PostOct1st		0.130
		(0.087)
Attack ads × PreOct1st		0.148
		(0.140)
Attack ads × PostOct1st		−0.326*
		(0.175)
No ads	−0.196	0.272
	(0.338)	(0.376)
log(Audience size: Own ads)	−0.039	−0.026
	(0.031)	(0.035)
log(Audience size: Rival's ads)	0.001	0.005***
	(0.001)	(0.001)
Number of ads	0.010	0.014
	(0.028)	(0.029)
Ad position	−0.102	−0.085
	(0.291)	(0.321)
Ad length	0.173	0.171
	(0.941)	(0.964)
Candidate and Week F.E.s	Yes	Yes
Program Genre and Network Controls	Yes	Yes
N	240	240
R^2	0.852	0.855

Notes. S.E.s are clustered at the candidate level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

content of advertising affects voter preferences. In particular, if one is trying to calculate the impact of TV advertising spending on campaign outcomes, then one would over-estimate the impact of this advertising if one did not account for the fact that the TV advertising may coincide with non-TV advertising, campaign events, or earned media. However, our goal is not to measure the effectiveness of the advertising, per se, but of the impact of the messaging of the campaign as a whole.

Even with the set of fixed effects and controls we use, the key assumption we make is that the national campaign ads serve as central messaging devices used to amplify the same message that the campaign is delivering in other activities it engages in. This assumption may be reasonable partially because we know that it is hard to create brands with multiple conflicting messages (Kotler and Keller, 2016), so campaigns need to focus on the main message of the day. To support that national advertising campaigns reflect the message that the campaigns seek to get out, we note that all of Clinton's and Trump's national primetime TV ad creatives in our data were also aired as local TV ads in primetime, demonstrating the intended broad appeal of these messages. That said, the main purpose of this regression is to replicate the results of the WOM analysis. Also, as was the case with the WOM analysis, any estimated effects would have to be interpreted as being the effect of the message after the campaign's decision of how to match the message to the appropriate TV audience.

3.5 Conclusion

Using data on political advertising, online WOM, and daily polls for the 2016 Presidential Election, we find that ad messages that are politically centrist and consistent with the candidate's primary election messaging are associated with increases in online WOM. The voter preference analysis confirms these results. By comparing the relative impact of these two variables of interest, we find that consistent messages have a slightly larger impact on WOM. The result is somewhat reversed for the voter preference analysis, which may reflect significant differences or may reflect that the voter preference estimation is less able to control for confounds. Both measures are found to have greater impacts on voter preferences in the early stage of the campaign. Our results add nuance to the conventional wisdom that candidates should focus on taking stances that appeal to their party's primary electorate in the primary elections but then shift their message to be more centrist for the general election. While this strategy is beneficial, our results also show the value of consistent political branding, especially in the early stages of a campaign. Our results further suggest that the rising use of extremist messages in political advertising (e.g., Bartels, 2016; Wells and Seetharaman, 2018) may be a flawed strategy for candidates as more extreme messages are associated with decreased candidate-related WOM volume and decreased voter preference for the candidate.

Our study is not without limitations. Our analysis is limited to one presidential election. Future research is needed to study whether our findings generalize to other presidential elections or other types of elections such as senatorial elections. Another limitation of the data we use is that we only focus on national television advertisements. This is done because it is very difficult to link local television advertising and online advertisements to our outcomes, as such ads are distributed widely throughout the day with a low density at any particular time. In contrast, national television advertising occurs at specific times, allowing us to detect the changes in our outcomes. Additionally, our consistency measure assumes that what would constitute a consistent message remains static. This reflects a natural breaking point that arises as candidates finish the primary season and enter the general election. However, a dynamic model that allows for additional break points throughout the campaign, or allows for a continuously updating measure of consistency, may prove to be fruitful. Finally, as noted in the introduction, there are several assumptions that are needed to make our findings causal. Unfortunately, we do not have exogenous data variation to exploit to further probe causal relationships. As such, it may be prudent to restrain the causal interpretations of the results until further studies confirm the results with other methods.

Beyond the political domain, our findings demonstrate the potential for marketers to make use of automated text analytic methods to evaluate their advertisements. We find that consumers react more favorably in terms of the volume of WOM and preference to advertising content that is consistent with the established brand image of the political candidates. Future research may consider applying measures derived from text analysis to traditional brands. Evaluating the impact of ad campaigns on WOM and preferences may reveal insights for these traditional firms to increase the effectiveness of advertising. Such a relationship may vary across categories, potentially due to the nature of the products or the competitiveness of the category.

References

Aaker, J. L., Brumbaugh, A. M., & Grier, S. A. (2000). Nontarget markets and viewer distinctiveness: The impact of target marketing on advertising attitudes. *Journal of Consumer Psychology*, *9*(3), 127–140.

Abel, J., & Fuster, A. (2021). How do mortgage refinances affect debt, default, and spending? evidence from harp. *American Economic Journal: Macroeconomics*, *13*(2), 254–91.

Adgate, B. (2019). *The 2020 elections will set (another) ad spending record.* Retrieved October 9, 2020, from https://www.forbes.com/sites/bradadgate/2019/09/03/the-2020-elections-will-set-another-ad-spending-record/?sh=1a7aa6fc1836

Agarwal, S., & Sen, A. (2022). Antiracist curriculum and digital platforms: Evidence from black lives matter. *Management Science*, *68*(4), 2932–2948.

Agarwal, S., Amromin, G., Chomsisengphet, S., Landvoigt, T., Piskorski, T., Seru, A., & Yao, V. (2023). Mortgage refinancing, consumer spending, and competition: Evidence from the home affordable refinance program. *The Review of Economic Studies*, *90*(2), 499–537.

Aizawa, N., & Kim, Y. S. (2021). Public and private provision of information in market-based public programs: Evidence from advertising in health insurance marketplaces. *Available at SSRN 3434450*.

Aizer, A. (2007). Public health insurance, program take-up, and child health. *The Review of Economics and Statistics*, *89*(3), 400–415.

Amos, B., McDonald, M. P., & Watkins, R. (2017). When boundaries collide: Constructing a national database of demographic and voting statistics. *Public Opinion Quarterly*, *81*(S1), 385–400.

An, J., & Kwak, H. (2019). Gender and racial diversity in commercial brands' advertising images on social media. *Social Informatics: 11th International Conference, SocInfo 2019, Doha, Qatar, November 18–21, 2019, Proceedings 11*, 79–94.

Andersen, S., Campbell, J. Y., Nielsen, K. M., & Ramadorai, T. (2020). Sources of inaction in household finance: Evidence from the danish mortgage market. *American Economic Review*, *110*(10), 3184–3230.

Aneja, A., Luca, M., & Reshef, O. (2023). The benefits of revealing race: Evidence from minority-owned local businesses. *Available at SSRN 4346915*.

Babar, Y., Mahdavi Adeli, A., & Burtch, G. (2023). The effects of online social identity signals on retailer demand. *Management Science*, *69*(12), 7335–7346.

Baer, J., & Sinagoga, D. (2018). *Why TV totally dominates political advertising*. Retrieved October 9, 2020, from https://adage.com/article/cmo-strategy/why-tv-totally-dominates-political-advertising/315498

Balakrishnan, M., Nam, J., & Buell, R. W. (2022). Differentiating on diversity: How disclosing workforce diversity improves brand attitudes. *Available at SSRN 4369758*.

Bartels, L. (2016). *The rise of presidential extremists*. Retrieved October 9, 2020, from https://www.nytimes.com/2016/09/12/opinion/campaign-stops/the-rise-of-presidential-extemists.html

Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics*, *143*(1), 30–56.

Bass, F. M., Bruce, N., Majumdar, S., & Murthi, B. (2007). Wearout effects of different advertising themes: A dynamic bayesian model of the advertising-sales relationship. *Marketing Science*, *26*(2), 179–195.

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the tribes: Using text for marketing insight. *Journal of Marketing*, *84*(1), 1–25.

Berman, R., Melumad, S., Humphrey, C., & Meyer, R. (2019). A tale of two twitterspheres: Political microblogging during and after the 2016 primary and presidential debates. *Journal of Marketing Research*, *56*(6), 895–917.

Bermingham, A., & Smeaton, A. F. (2011). On using twitter to monitor political sentiment and predict election results. *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011)*, 2–10.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., & Zinman, J. (2010). What's advertising content worth? evidence from a consumer credit marketing field experiment. *The Quarterly Journal of Economics*, *125*(1), 263–306.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Milkman, K. L. (2015). The effect of providing peer information on retirement savings decisions. *The Journal of Finance*, *70*(3), 1161–1201.

Bhutta, N., Hizmo, A., & Ringo, D. (2022). How much does racial bias affect mortgage lending? evidence from human and algorithmic credit decisions. *Available at SSRN 3887663*.

Bhutta, N., Laufer, S., & Ringo, D. R. (2017). Residential mortgage lending in 2016: Evidence from the home mortgage disclosure act data. *Fed. Res. Bull.*, *103*, 1.

Blaufus, K., & Milde, M. (2021). Tax misperceptions and the effect of informational tax nudges on retirement savings. *Management Science*, *67*(8), 5011–5031.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Boush, D. M., & Loken, B. (1991). A process-tracing study of brand extension evaluation. *Journal of Marketing Research*, *28*(1), 16–28.

Broniarczyk, S. M., & Alba, J. W. (1994). The importance of the brand in brand extension. *Journal of Marketing Research*, *31*(2), 214–228.

Bryan, M. (2022). *U.S. Voting by Census Block Groups*. https://doi.org/10.7910/DVN/NKNWBX

Burnap, A., Hauser, J. R., & Timoshenko, A. (2023). Product aesthetic design: A machine learning augmentation. *Marketing Science*.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of human resources*, *50*(2), 317–372.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometric Journal*, *21*(1), C1–C68.

Davis, D. R., Dingel, J. I., Monras, J., & Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, *127*(4), 1684–1738.

DeFusco, A. A., & Mondragon, J. (2020). No job, no money, no refi: Frictions to refinancing in a recession. *The Journal of Finance*, *75*(5), 2327–2376.

Deng, Y., & Mela, C. F. (2018). Tv viewing and advertising targeting. *Journal of Marketing Research*, *55*(1), 99–118.

Deshpandé, R., & Stayman, D. M. (1994). A tale of two cities: Distinctiveness theory and advertising effectiveness. *Journal of Marketing Research*, *31*(1), 57–64.

Dew, R., Ansari, A., & Toubia, O. (2022). Letting logos speak: Leveraging multiview representation learning for data-driven branding and logo design. *Marketing Science*, *41*(2), 401–425.

Dube, A., Jacobs, J., Naidu, S., & Suri, S. (2020). Monopsony in online labor markets. *American Economic Review: Insights*, *2*(1), 33–46.

Dzyabura, D., & Peres, R. (2021). Visual elicitation of brand perception. *Journal of Marketing*, *85*(4), 44–66.

Ellickson, P. B., Kar, W., & Reeder III, J. C. (2022). Estimating marketing component effects: Double machine learning from targeted digital promotions. *Marketing Science*.

Feng, S. (2020). The proximity of ideas: An analysis of patent text using machine learning. *PloS one*, *15*(7), e0234880.

Fossen, B. L., Kim, D., Schweidel, D. A., & Thomadsen, R. (2022). The role of slant and message consistency in political advertising effectiveness: Evidence from the 2016 presidential election. *Quantitative Marketing and Economics*, *20*(1), 1–37.

Fossen, B. L., Mallapragada, G., & De, A. (2021). Impact of political television advertisements on viewers' response to subsequent advertisements. *Marketing Science*, *40*(2), 305–324.

Fossen, B. L., & Schweidel, D. A. (2017). Television advertising and online word-of-mouth: An empirical investigation of social tv activity. *Marketing Science*, *36*(1), 105–123.

Freddie Mac. (2021). Refinance trends in the first half of 2021 [Accessed: insert-date-here].

Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The role of technology in mortgage lending. *The Review of Financial Studies*, *32*(5), 1854–1899.

Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? evidence from us daily newspapers. *Econometrica*, *78*(1), 35–71.

Gerardi, K., Lambie-Hanson, L., & Willen, P. (2021). *Racial differences in mortgage refinancing, distress, and housing wealth accumulation during covid-19* (Discussion Paper No. 21-02). Federal Reserve Bank of Philadelphia Consumer Finance Institute.

Gerardi, K., Willen, P. S., & Zhang, D. H. (2023). Mortgage prepayment, race, and monetary policy. *Journal of Financial Economics*, *147*(3), 498–524.

Gershon, R., & Jiang, Z. (2022). Referral contagion: Downstream benefits of customer referrals. *Available at SSRN 4214781*.

Goli, A., & Mummalaneni, S. (2023). Gender diversity on cable news: An analysis of on-screen talent and viewership. *Available at SSRN 4462592*.

Goli, A., Mummalaneni, S., Chintagunta, P. K., & Dhar, S. K. (2022). Show and sell: Studying the effects of branded cigarette product placement in tv shows on cigarette sales. *Marketing Science*, *41*(6), 1163–1180.

Gordon, B. R., & Hartmann, W. R. (2013). Advertising effects in presidential elections. *Marketing Science*, *32*(1), 19–35.

Gordon, B. R., Moakler, R., & Zettelmeyer, F. (2022). Close enough? a large-scale exploration of non-experimental approaches to advertising measurement. *Marketing Science*.

Grootendorst, M. (2022). Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Guitart, I. A., & Stremersch, S. (2021). The impact of informational and emotional television ad content on online search and sales. *Journal of Marketing Research*, *58*(2), 299–320.

Gunarathne, P., Rui, H., & Seidmann, A. (2022). Racial bias in customer service: Evidence from twitter. *Information Systems Research*, *33*(1), 43–54.

Gurun, U. G., Matvos, G., & Seru, A. (2016). Advertising expensive mortgages. *The Journal of Finance*, *71*(5), 2371–2416.

Hartmann, J., Netzer, O., & Zalta, R. (2023). Diversity in advertising in times of racial unrest. *Available at SSRN 4466378*.

He, P., Gao, J., & Chen, W. (2021). Debertav3: Improving deberta using electra-style pre-training with gradient-disentangled embedding sharing. *arXiv preprint arXiv:2111.09543*.

Hu, L., Li, K., Ngo, P. T., & Sosyura, D. (2023). Media as a money doctor: Evidence from refinancing decisions. *Available at SSRN 3679875*.

Huber, G. A., & Arceneaux, K. (2007). Identifying the persuasive effects of presidential advertising. *American Journal of Political Science*, *51*(4), 957–977.

Hummel, P. (2010). Flip-flopping from primaries to general elections. *Journal of Public Economics*, *94*(11-12), 1020–1027.

Jagarlamudi, J., Daumé III, H., & Udupa, R. (2012). Incorporating lexical priors into topic models. *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 204–213.

Jahanbakhsh, K., & Moon, Y. (2014). The predictive power of social media: On the predictability of us presidential elections using twitter. *arXiv preprint arXiv:1407.0622*.

Jennings, W., & Wlezien, C. (2018). Election polling errors across time and space. *Nature Human Behaviour*, *2*(4), 276–283.

Jiang, Z., Zhang, D. J., & Chan, T. (2021). How do bonus payments affect the demand for auto loans and their delinquency? *Journal of Marketing Research*, *58*(3), 476–496.

Johnson, E. J., Meier, S., & Toubia, O. (2019). What's the catch? suspicion of bank motives and sluggish refinancing. *The Review of Financial Studies*, *32*(2), 467–495.

Joo, M., Wilbur, K. C., Cowgill, B., & Zhu, Y. (2014). Television advertising and online search. *Management Science*, *60*(1), 56–73.

Karaca-Mandic, P., Wilcock, A., Baum, L., Barry, C. L., Fowler, E. F., Niederdeppe, J., & Gollust, S. E. (2017). The volume of tv advertisements during the aca's first enrollment period was associated with increased insurance coverage. *Health Affairs*, *36*(4), 747–754.

Katz, H. (2013). *The media handbook: A complete guide to advertising media selection, planning, research, and buying*. Routledge.

Kaye, K. (2017). *Data-driven targeting creates huge 2016 political ad shift*. Retrieved October 9, 2020, from https://adage.com/article/media/2016-political-broadcast-tv-spend-20-cable-52/307346

Kaye, K. (2019a). *Google to restrict political ad targeting on its platforms*. Retrieved October 9, 2020, from https://www.wsj.com/articles/google-to-restrict-political-ad-targeting-on-its-platforms-11574293253

Kaye, K. (2019b). *Twitter will ban all political ads, C.E.O. Jack Dorsey says*. Retrieved October 9, 2020, from https://www.nytimes.com/2019/10/30/technology/twitter-political-ads-ban.html

Keller, K. L., & Lehmann, D. R. (2006). Brands and branding: Research findings and future priorities. *Marketing Science*, *25*(6), 740–759.

Kennedy, C., Blumenthal, M., Clement, S., Clinton, J. D., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., et al. (2018). An evaluation of the 2016 election polls in the united states. *Public Opinion Quarterly*, *82*(1), 1–33.

Keys, B. J., Pope, D. G., & Pope, J. C. (2016). Failure to refinance. *Journal of Financial Economics*, *122*(3), 482–499.

Khan, U., & Kalra, A. (2022). It's good to be different: How diversity impacts judgments of moral behavior. *Journal of Consumer Research*, *49*(2), 177–201.

Khern-am-nuai, W., So, H., Cohen, M. C., & Adulyasak, Y. (2021). Selecting cover images for restaurant reviews: Ai vs. wisdom of the crowd. *Available at SSRN 3911469*.

Kim, D., Jiang, Z., & Thomadsen, R. (2023). Tv advertising effectiveness with racial minority representation: Evidence from the mortgage market. *Available at SSRN 4521178*.

Kim, T., & Kc, D. (2020). The impact of hospital advertising on patient demand and health outcomes. *Marketing Science*, *39*(3), 612–635.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Kingma, D. P., Welling, M., et al. (2019). An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, *12*(4), 307–392.

Kirgios, E. L., Rai, A., Chang, E. H., & Milkman, K. L. (2022). When seeking help, women and racial/ethnic minorities benefit from explicitly stating their identity. *Nature Human Behaviour*, *6*(3), 383–391.

Kotler, P., & Keller, K. L. (2016). *Marketing management* (15th ed.). Pearson.

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. *International Conference on Machine Learning*, 1188–1196.

Lee, D., Hosanagar, K., & Nair, H. S. (2018). Advertising content and consumer engagement on social media: Evidence from facebook. *Management Science*, *64*(11), 5105–5131.

Lewis, R. A., & Reiley, D. H. (2013). Down-to-the-minute effects of super bowl advertising on online search behavior. *Proceedings of the fourteenth ACM conference on Electronic commerce*, 639–656.

Li, X., Hartmann, W. R., & Amano, T. (2024). Preference externality estimators: A comparison of border approaches and ivs [Forthcoming]. *Management Science*, *0*(0).

Liaukonyte, J., Teixeira, T., & Wilbur, K. C. (2015). Television advertising and online shopping. *Marketing Science*, *34*(3), 311–330.

Liu, F., Jo, Y., & Chen, E. (2022). Data point: 2021 mortgage market activity and trends. *Consumer Financial Protection Bureau Office of Research Reports Series*, (22-7).

Liu, Y., Li, K. J., Chen, H., & Balachander, S. (2017). The effects of products' aesthetic design on demand and marketing-mix effectiveness: The role of segment prototypicality and brand consistency. *Journal of Marketing*, *81*(1), 83–102.

Lovett, M. J., & Shachar, R. (2011). The seeds of negativity: Knowledge and money. *Marketing Science*, *30*(3), 430–446.

MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2023). Testing for the appropriate level of clustering in linear regression models. *Journal of Econometrics*, *235*(2), 2027–2056.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly Journal of Economics*, *116*(4), 1149–1187.

McGranaghan, M., Liaukonyte, J., Fisher, G., & Wilbur, K. C. (2019). Lead offer spillovers. *Marketing Science*, *38*(4), 643–668.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, *26*.

Miller, J. (2015). *Report: TV ad spending in 2016 election to reach $4.4 billion*. Retrieved October 9, 2020, from https://www.cbsnews.com/news/report-tv-ad-spending-in-2016-election-to-reach-4-4-billion/

Mollica, K. A., Gray, B., & Trevino, L. K. (2003). Racial homophily and its persistence in newcomers' social networks. *Organization Science*, *14*(2), 123–136.

Morozov, I., & Tuchman, A. (2022). Where does advertising content lead you? we created a bookstore to find out. *Available at SSRN 4308400*.

OpenSecretes. (2020). *Expenditures*. Retrieved October 9, 2020, from https://www.opensecrets.org/campaign-expenditures

Overgoor, G., Yildirim, G., Bart, Y., & Pauwels, K. (2023). From representation to reception: Evaluating the impact of diversity in tv advertising on consumer purchase intention. *Available at SSRN 4471248*.

Park, C. W., Milberg, S., & Lawson, R. (1991). Evaluation of brand extensions: The role of product feature similarity and brand concept consistency. *Journal of Consumer Research*, *18*(2), 185–193.

Passwaiter, S. (2018). *2018 campaign ad spend will be in the billions*. Retrieved October 9, 2020, from https://www.cookpolitical.com/analysis/national/political-advertising/2018-campaign-ad-spend-will-be-billions

Passwaiter, S. (2020). *Political ad spending this year reached a whopping $8.5 billion*. Retrieved October 9, 2020, from https://adage.com/article/campaign-trail/political-ad-spending-year-reached-whopping-85-billion/2295646

Perry, V. G., Motley, C. M., & Adams Jr, R. L. (2016). What's the point (s)? information content and messaging strategies in mortgage loan advertisements. *Cityscape*, *18*(2), 35–54.

Pieters, R., Warlop, L., & Wedel, M. (2002). Breaking through the clutter: Benefits of advertisement originality and familiarity for brand attention and memory. *Management Science*, *48*(6), 765–781.

Rezende, D. J., Mohamed, S., & Wierstra, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. *International conference on machine learning*, 1278–1286.

Rosengren, S., Eisend, M., Koslow, S., & Dahlen, M. (2020). A meta-analysis of when and how advertising creativity works. *Journal of Marketing*, *84*(6), 39–56.

Sahni, N. S., Wheeler, S. C., & Chintagunta, P. (2018). Personalization in email marketing: The role of noninformative advertising content. *Marketing Science*, *37*(2), 236–258.

Shapiro, B. T. (2018). Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy*, *126*(1), 381–437.

Shapiro, B. T. (2020). Advertising in health insurance markets. *Marketing Science*, *39*(3), 587–611.

Shapiro, B. T. (2022). Promoting wellness or waste? evidence from antidepressant advertising. *American Economic Journal: Microeconomics*, *14*(2), 439–77.

Shapiro, B. T., Hitsch, G. J., & Tuchman, A. E. (2021). Tv advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica*, *89*(4), 1855–1879.

Silver, N. (2018). *The Polls Are All Right*. Retrieved October 9, 2020, from https://fivethirtyeight.com/features/the-polls-are-all-right/

Sinkinson, M., & Starc, A. (2019). Ask your doctor? direct-to-consumer advertising of pharmaceuticals. *The Review of Economic Studies*, *86*(2), 836–881.

Spenkuch, J. L., & Toniatti, D. (2018). Political advertising and election results. *The Quarterly Journal of Economics*, *133*(4), 1981–2036.

Sudhir, K., Roy, S., & Cherian, M. (2016). Do sympathy biases induce charitable giving? the effects of advertising content. *Marketing Science*, *35*(6), 849–869.

Sussman, A. B., & Alter, A. L. (2012). The exception is the rule: Underestimating and overspending on exceptional expenses. *Journal of consumer research*, *39*(4), 800–814.

Thomas, M. (2020). Spillovers from mass advertising: An identification strategy. *Marketing Science*, *39*(4), 807–826.

Thomson, M. (2006). Human brands: Investigating antecedents to consumers' strong attachments to celebrities. *Journal of Marketing*, *70*(3), 104–119.

Tian, Z., Dew, R., & Iyengar, R. (2023). Mega or micro? influencer selection using follower elasticity. *Journal of Marketing Research*, 00222437231210267.

Tirunillai, S., & Tellis, G. J. (2017). Does offline tv advertising affect online chatter? quasi-experimental analysis using synthetic control. *Marketing Science*, *36*(6), 862–878.

Tsai, Y.-L., & Honka, E. (2021). Informational and noninformational advertising content. *Marketing Science*, *40*(6), 1030–1058.

Tuchman, A. E. (2019). Advertising and demand for addictive goods: The effects of e-cigarette advertising. *Marketing Science*, *38*(6), 994–1022.

Van Rooij, M. C., Lusardi, A., & Alessie, R. J. (2011). Financial literacy and retirement planning in the netherlands. *Journal of Economic Psychology*, *32*(4), 593–608.

Wang, Y., Qin, M. S., Luo, X., & Kou, Y. (2022). Frontiers: How support for black lives matter impacts consumer responses on social media. *Marketing Science*, *41*(6), 1029–1044.

Wang, Y., Lewis, M., & Schweidel, D. A. (2018). A border strategy analysis of ad source and message tone in senatorial campaigns. *Marketing Science*, *37*(3), 333–355.

Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., Du, N., Dai, A. M., & Le, Q. V. (2021). Finetuned language models are zero-shot learners. *arXiv preprint arXiv:2109.01652*.

Wells, G., & Seetharaman, D. (2018). *Coming soon to Facebook: Lots of extreme political ads*. Retrieved October 9, 2020, from https://www.wsj.com/articles/coming-soon-to-facebook-lots-of-extreme-political-ads-1535707802

Wilbur, K. C., Xu, L., & Kempe, D. (2013). Correcting audience externalities in television advertising. *Marketing Science*, *32*(6), 892–912.

Williams, A., Nangia, N., & Bowman, S. R. (2017). A broad-coverage challenge corpus for sentence understanding through inference. *arXiv preprint arXiv:1704.05426*.

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378–382.

Yang, J., Lee, J. Y., & Chintagunta, P. K. (2021). Commercial success through commercials? advertising and pay-tv operators. *Journal of Marketing Research*, *58*(5), 925–947.

Yang, J., Xie, Y., Krishnamurthi, L., & Papatla, P. (2021). High-energy ad content: A large-scale investigation of tv commercials. *Journal of Marketing Research*, 00222437211067802.

Yin, W., Hay, J., & Roth, D. (2019). Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach. *arXiv preprint arXiv:1909.00161*.

Zhang, L., & Chung, D. J. (2020). The air war vs. the ground game: An analysis of multichannel marketing in us presidential elections. *Marketing Science*, *39*(5), 872–892.

Zhang, M., & Luo, L. (2023). Can consumer-posted photos serve as a leading indicator of restaurant survival? evidence from yelp. *Management Science*, *69*(1), 25–50.

Zhang, S., Lee, D., Singh, P. V., & Srinivasan, K. (2022). What makes a good image? airbnb demand analytics leveraging interpretable image features. *Management Science*, *68*(8), 5644–5666.

Appendix A: The Effects of TV Advertising and Ad Content on Consumer Financial Decisions: Evidence from Mortgage Refinancing

A.1 Sample Selection Criteria

In this appendix, we outline our sample selection criteria and present the resulting number of observations. While most of the variables used for our sample selection criteria are available throughout the entire sample period, some additional variables, such as indicators for reverse mortgages, interest-only payment loans, and negatively amortizing loans, become available beginning in 2018. Therefore, we apply these variables solely for the reported years (i.e., 2018 - 2021). The selection process is as follows:

1. We collect loans directly originating from lenders (not purchased from other lenders) for home purchase or refinance within our sample period from 2016 to 2027, resulting in a total of 57,484,148 loans.

2. We remove loans with missing state or county codes, as well as those originating outside of the 50 states and Washington, DC, resulting in 57,229,968 loans.

3. Among these loans, we select 43,846,062 conventional loans.

4. We limit our focus to first-lien mortgage loans for owner-occupied, single-family, site-built, residential homes. After applying these filters, we are left with 36,111,078 loans.

5. We remove a small number of loans with Home Ownership and Equity Protection Act (HOEPA) status, indicating unusually high interest rates. Additionally, we remove a few loans with the dummy variable "Application Date $<$ 01-01-2004" equal to 1 in the 2016 HMDA data. This step results in 36,099,341 loans.

6. We condition that the loan size is greater than or equal to \$50,000 and below the annual county-level loan size limit (set by the FHFA). This process yields a final sample of 33,622,202 loans.

7. Finally, we exclude reverse mortgages and mortgages with not fully amortizing or negatively amortizing loans, including balloon and interest only payment loans. After applying these additional filters, we end up with 33,124,105 loans.

A.2 Constructing Training Labels for Ad Classification

As described in Section 1.2.3, we use a keyword approach to identify whether an ad focuses on refinancing or purchasing to construct labels. We create two dummy variables, one for refinance and the other for purchase loans. For each type, the dummy variable takes the value of 1 if the ad contains at least one of the key phrases in the category.

We provide the full list of phrases for refinancing and purchasing below. Parentheses are used to denote alternative wording within phrases. For example, "your current (rate, interest rate)" means either "your current rate" or "your current interest rate." We categorize the extensive list of phrases by grouping similar ones together.

- "refi," "if you're a homeowner."
- "cash out," "cash back," "take out cash," "home equity," "home's equity," "use the new equity in your home," "tap into the equity," "it's your home's turn to work for you,"

"equity to work," "let the equity work," "use the (equity, new equity) in your home," "use the equity in my home," "use equity from your current home," "how much is the equity in your home," "make the equity in your home work for you," "using some of the equity built up in your home," "need cash from the equity in your home," "how much has the equity in your home increased," "your increased equity means extra cash," "turn equity into cash," "unlock that equity," "convert some of the equity in your home."

- "paying too much," "costing you way too much money," "your current (rate, interest rate, mortgage rate)," "if the interest rate on your home loan is higher than," "lower your (payment, current monthly payment, house payment, monthly payment, monthly mortgage payment)," "lower your (rate, interest rate, mortgage rate)," "lower the (rate, interest rate)," "lower the monthly (payment, house payment)," "lower their (rate, interest rate, monthly bills, monthly payment)," "lower my (payment, mortgage payment)," "lowering your (current mortgage payment, mortgage payment, interest rate)," "save you (hundreds, thousands)," "saving money every month on your mortgage," "reduce your (payment, monthly payment, monthly mortgage payment)," "reduce your interest rate," "reducing their (current interest rate, monthly payment)," "homeowners could benefit from current interest rates," "one third of homeowners isn't aware what their current interest rate is," "homeowners mortgage rates continue to drop," "homeowners ready for some good news interest rates continue to drop."

In addition, we create a dummy variable for ads focused on home purchasing. This variable equal to 1 if the ad contains at least one of the purchase-related phrases listed below, and 0 otherwise.

- "buy (home, first home, house)," "buy a (condo, home, house, place, property)," "buy a (larger, new, nice, second) home," "buy a (new house, bigger place, second property)," "buy the (home, best home, house, ideal house)," "buy your (home, house, property)," "buy your (first, new, next, own, very first, very own) home," "buy your (first, new) house," "buy your first condo," "buy my (first home, house, own house)," "buy our (first home, first house, own home)," "buy this home," "buy that new home," "buy their (home, first home, house, own)," "buys (a home, their first house)," "ready to buy," "get approved to buy," "what it means to buy," "better to buy," "planning to buy," "looking to buy," "buy somewhere a place," "property you want to buy," "home you want to buy," "whether you want to buy," "if you want to buy," "make it simple for you to buy," "time to buy," "when you buy," "how simple it was to buy," "buy or build," "build buy," "buy and build."

- "buying (home, house)," "buying for the first time," "home buying," "buying a (home, house)," "buying a (bigger, first, new) home," "buying a (bigger, first, new) house," "buying an investment house," "buying the (home, house)," "buying the (first, perfect, new) home," "buying the perfect house," "buying your (home, house, property)," "buying your (first, new, next, own) home," "buying your (first, own) house," "buying our first home," "buying (this house, that dream second home, that new home)," "buying their (home, first home, new home)," "if you're buying," "whether you're buying," "ease the process of buying," "buying or building," "thinking about buying," "when it comes to buying," "home you're buying," "buying the home you want."

- "bought a (condo, home, house)," "bought a new (home, house)," "bought the (home, house)," "bought this house," "bought their (home, first home)," "bought (her/his) first home," "bought my first house."

- "homebuyer," "home buyer," "first (time buyer, timer, buyer)."

- "purchase," "purchasing."

- “own a (home, house, piece of the country place),” “own the (house, place),” “own your (home, place),” “own your (new, own) home,” “own my own (home, ranch),” “own our (home, own home),” “own their (home, own home).”
- “owning a (home, new home, house),” “owning their (first, own) home,” “owning your (home, own home, dream),” “owning (her, my, our) own home,” “dream of owning.”
- “your own (home, house, place),” “our own (home, house, place),” “their own (home, house, place),” “home of (your, our, their) own,” “house of your own,” “place of (your, our, my, their) own,” “make it your own,” “home you want to own,” “home we call our own,” “place to call (your, my) own,” “place to finally call your own,” “place we can call our own,” “place that we could call our own.”
- “home ownership,” “homeownership,” “become a homeowner,” “become a (happy, hawaii, detroit) homeowner,” “become the (homeowner, owner, proud owner),” “become homeowners,” “become proud owner,” “be (a homeowner, homeonwers),” “transform renters into owners,” “first time homeowner.”
- “looking for a (first, second, new, brand new, bigger) home,” “looking for a house,” “looking for a new (house, place, space),” “looking for an investment property,” “looking for your (home, first home, new home),” “looking for your perfect house,” “looking for their perfect home,” “home you’ve been looking for,” “looking to get a mortgage loan to buy.”
- “shopping for a (home, new home, house, new house),” “shopping for your (new home, next home, house),” “shopping for that new home,” “shop around for a new house,” “shop for your new home,” “house hunt,” “(house, home) shopping,” “new home shoppers,” “in the market for a new home,” “home search,” “searching for your new home.”

- “find a (house, place, new home),” “find a home (that fits, you love),” “find the home (you want, you love, of your dream, you’ve been dreaming),” “find the perfect (home, house, place),” “find the way to a new home,” “find the forever home,” “find your (first, perfect, new) home,” “find your perfect place,” “find your way (home, to a new home),” “find their (new home, place),” “find that perfect place.”

- “finding a (new house, place),” “finding your (forever, new) home,” “finding the home (that fits, of your dreams, that’s perfect),” “finding the perfect (home, house),” “finding that special home,” “finding own home,” “finding the right neighborhood.”

- “found the perfect (home, place),” “found that perfect house,” “found your (first, new, perfect) home,” “found it the right house the right home.”

- “first (home, house, mortgage).”

- “new (house, place),” “(door, key, way) to your new home,” “(get, get you, moving) into your new home,” “get into that new home,” “in your new home,” “couple look for new home.”

- “perfect (home, house, place, piece of property),” “this (place, house) is perfect,” “perfect future home.”

- “keys to your (first home, new home, house, doors),” “keys to their new home,” “just got the keys,” “key to your new home,” “handing keys to families,” “keys handed over to them.”

- “tired of renting,” “stop renting,” “you might be a renter,” “are you still renting,” “stop paying someone else’s mortgage,” “throw in the towel on renting,” “turn renters into homeowners,” “transform renters into owners,” “invest in a house or a condo.”

- “american dream,” “dream (home, house),” “(home, house) of dream,” “(home, house) of your dream,” “home of (my, their) dream,” “your dream kitchen,” “dream (come, comes) true,” “dreams come true,” “dream is coming true,” “dreaming of a (bigger, new) home,” “the home you’re dreaming of,” “achieve (my, the, your) dream,” “achieving your big dream,” “(finance, realize) your dream,” “(chase, realize) their dream,” “home is where dreams begin,” “get into homes they never dreamed,” “always dreamed about this house.”

- “(dream, dreams) into a reality,” “(dream, dreams) a reality,” “(dream, dreams) reality,” “(dream, dreams) become reality,” “(dream, dreams) becoming reality,” “home a reality,” “becomes a reality,” “make your (vision, next home) a reality,” “make that dream kitchen a reality,” “make them a reality,” “it can be a reality,” “it’s finally a reality,” “transform your vision into a reality,” “from dreams to reality,” “reality of actually coming home to this house,” “dream is now a reality,” “dreams are becoming a reality.”

- “help you (buy, finance a home),” “help you get (into a home, that home),” “help you get the home you (want, deserve),” “help financing my home.”

- “make (a house, your next house) a home,” “(make, making) it a home,” “house becomes home,” “make your house (feel like a home, your home),” “make their house feel like home,” “make that house your home,” “place (they, to, we can all) call home,” “feel (at, right at, like a) home,” “feels like home,” “feeling of home,” “make yourself at home,” “where you can be yourself,” “where i can be myself,” “home is (where, more than),” “home it’s more than,” “more than (four walls, a house, a place, just a place),” “turn (a, that) house into your home,” “what your home could be like,” “nothing (better, is better) than being at home,” “home sweet home,” “home is (my sanctuary, your happy place).”

- “getting into the home,” “get into my own home,” “the house (you want, you’ve always wanted),” “you find one you love,” “house that you fall in love,” “you love this town this neighborhood this house,” “place (in our hearts, where families can feel safe),” “we want to be that house,” “always wanted (a house, to own my own),” “move into your own home,” “we’ll get you home,” “get a home with more space,” “need a bigger place,” “acquire a larger dwelling,” “the right house the right home,” “nice house,” “the home you love,” “a lovely new home,” “signing for your new home,” “our forever home,” “to finance our next home,” “your next home fits,” “making it yours is what we do,” “ready to make home happen,” “cross that line from wishing you had your own slice.”

A.3 Examples of Unlabeled Ads

In this appendix, we show a few ads from the unlabeled set with the highest predicted probabilities of targeting refinance or purchase loan borrowers. In cases of nearly identical ads (e.g., ads with the same transcribed texts but different phone numbers), we include the ad with the next highest probability.

Refinance Ads

- Refinance probability: 98.25%

  “call to save thousands. homeowners as states begin to open back up and the economy rebounds interest rates are still at near all time lows call rate plus now and take advantage of these historically low rates rate plus is now offering a 15 year fixed at just 1.875% rate 2.23% apr or a 30 year fixed at just 2.5% rate 2.665% apr and we’ll pay your title escrow and appraisal fees call 800 785 9045 or visit rate plus dot com.”

- Refinance probability: 94.01%

"get the lowest rate j. ruedy. this is jason rudy with the home loan arranger seized the moment with these historically low mortgage rates 30 year fixed rates 2.25% unbelievable fast closings skip up to two mortgage payments call today or go to the home loan arranger dot com."

Purchase Loan Ads

- Purchase loan probability: 94.08%

  "providing best services b.daniels. this is the house home to future can't miss moments this is you with your mortgage team getting your questions answered and details handled so you can think about new paint colors capcom's mortgage subsidiary homeowners advantage offers affordable home loans to match your budget partnering with you every step of the way now here's you going online to get started find out what it's like to bank where you matter most."

- Purchase loan probability: 93.81%

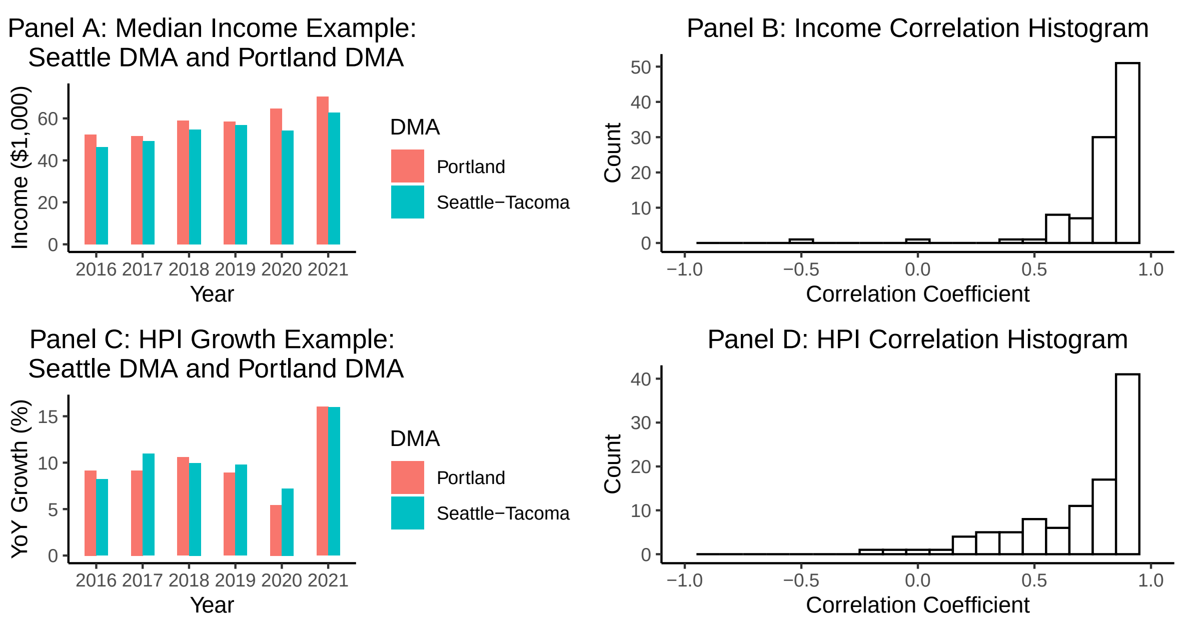
  "help your future family nate & roni. it was a lot of fun designing a home to cater to what we wanted with our future family going into the financing process we weren't sure what to expect we wanted something that would be explained to us and benchmark mortgage really did that for us it was simple once all the documents were uploaded then were approved and we kind of move forward with everything when the house was done and we walked in for the first time i felt like this was our home."

A.4 Parallel Trends in Refinancing Eligibility

In this appendix, we provide additional details on our analysis to investigate parallel trends in income and HPI growth between neighbouring border markets. We collect two data points per county each year, one for income and the other for the HPI. In cases where multiple border counties are present on either side of a DMA border, we calculate a weighted average, using the county-level population as the weight.

In Figure A4, Panel A and C present the annual median household income and the year-over-year change in the HPI for the border counties along the Seattle-Tacoma DMA and Portland DMA border. We observe similar time-series patterns in both income and HPI. Panel B and D present histograms of correlation coefficients for income and HPI across all DMA borders. The mean and median correlation coefficients for income are 0.85 and 0.90, respectively, and for HPI, they are 0.77 and 0.89, respectively. These results suggest that refinancing eligibility has not evolved in systematically different ways at the DMA borders. Consequently, any observed differences in refinancing demand are likely attributable to variations in advertising levels rather than differences in economic trends.

Figure A4: Suggestive Evidence for Parallel Trends in Income and HPI



A.5 Robustness Checks Using Loan Applications

The main sample includes approved and originated loans. In this appendix, we examine the robustness of our results using loan applications. Specifically, we expand the data to include loans that were approved but not accepted, as well as denied loans, in addition to the originated loans.

First, we present the results corresponding to Equation 1.1 in Table A.5.1. The coefficients are slightly smaller but statistically indistinguishable from those in Table 1.2.

Table A.5.1: Category-level Demand with Loan Applications

Dependent Variable:	$log(Q^{Refi})$		
	(1)	(2)	(3)
$log(AD)$	0.096**		
	(0.040)		
$log(AD^{Refi})$		0.065**	0.063**
		(0.024)	(0.028)
$log(AD^{Purch})$			0.008
			(0.051)
Border Market FE	Y	Y	Y
DMA Border-Year FE	Y	Y	Y
N	1,680	1,680	1,680

Notes. S.E.s are clustered at the border market level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Next, we examine the robustness of our brand-level demand analysis. Note that the expanded dataset contains a slightly larger number of observations. The results, corresponding to Equation 1.2 and 1.3, are presented in Table A.5.2. The coefficients across columns (1) through (3) are similar in magnitude and mostly consistent in statistical significance, compared to those in Table 1.3.

Table A.5.2: Brand-level Demand with Loan Applications

Dependent Variable:	$log(1+Q_j^{Refi})$		$log(\hat{s_j}/\hat{s}_0)$
	(1)	(2)	(3)
$log(1+AD_j^{Refi})$	0.083**	0.083**	0.051
	(0.040)	(0.040)	(0.033)
$log(1+AD_{-j}^{Refi})$		0.040***	
		(0.015)	
Lender-Border Market FE	Y	Y	Y
Lender-DMA Border-Year FE	Y	Y	Y
N	41,198	41,198	41,198

Notes. S.E.s are clustered at the border market level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

A.6 Topic Distributions on Selected Ads

In this appendix, we show a few examples of ads from the largest lender, Rocket Mortgage, and the corresponding probabilities for the four topics.

- Interest rate and savings: 79%; Ease of application: 1%; Homeownership: 0%; Brand-building: 20%.

  "benefits of yourgage 791 3577, rates. some good news you can control your mortgage as long as it's a yourgage the fixed rate mortgage exclusively from quicken loans that conforms to your needs with a yourgage you control your home loan term anywhere from 8 to 30 years that custom term can help you finish paying off your mortgage in a time frame that works best for you and by taking years off the loan you'll potentially save thousands of dollars in interest over the life of your mortgage right now could be a great time to take some positive financial steps with a cash out refinance from quicken loans you can use the cash from your home's equity to make some much needed home improvements or reduce nagging credit card debt a great way to take cash out is with our most popular loan a quicken loans 30 year fixed rate mortgage here's one more reason you'll want to work with us we're number one in the nation according to j d power to learn more about how you can leverage the value in your home call us today or go to rocket mortgage dot com."

- Interest rate and savings: 0%; Ease of application: 99%; Homeownership: 0%; Brand-building: 0%.

  "push button & get mortgage 68. yeah you could spend the next few days weeding through w two's pay stubs and bank statements to refinance your home or you could push that button skip the bank skip the paperwork and go completely online securely share your financial info and confidently get an accurate mortgage solution in minutes lift the burden of getting a home loan with rocket mortgage by quicken loans."

- Interest rate and savings: 0%; Ease of application: 6%; Homeownership: 93%; Brand-building: 1%.

  "rocket mortgage has helps millions. home it's so much more than a house it's your own little slice of heaven for over 30 years rocket mortgage has helped millions of americans finance the home of their dreams something we've got to go guys so you can spend your time making your house rocket mortgage push button get mortgage."

A.7 Robustness Checks with Topic Descriptions

In this appendix, we explore the robustness of our main findings using different topic descriptions. For this analysis, we update the topic descriptions (Section 1.5.1) by adding or removing about one semantically similar word from each topic. Table A.7.1 presents the revised topic descriptions, which are largely semantically consistent.

Table A.7.2 presents the heterogeneous effects by topic on the category-level demand using the updated topic descriptions. These results are qualitatively and statistically similar to those in Table 1.8.

Table A.7.1: Alternative Topic Descriptions for Zero shot Classification

Ad topics (1)	Topic descriptions (2)
Panel A. Short Topic Descriptions	
Interest rate and savings	low mortgage rate, save money
Ease of application	online application, easy and simple
Homeownership	home buying, new/dream home
Brand-building	trusted lender, customer satisfaction
Panel B. Long Topic Descriptions	
Interest rate and savings	low mortgage rate, interest saving, lower payments, save on mortgage
Ease of application	online application, digital mortgage, easy, simple and quick process
Homeownership	home buying, first/new home, dream home, perfect home, home ownership
Brand-building	trusted lender, customer satisfaction, expert advice, professional

Table A.7.2: Heterogeneous Effects on Category-level Demand for Refinancing

Dependent Variable:	$log(Q^{Refi})$			
	Short descriptions		Long descriptions	
	(1)	(2)	(3)	(4)
$log(AD^{Refi} \times$ Rate and Savings)	0.036**	0.035**	0.038*	0.039*
	(0.017)	(0.017)	(0.020)	(0.020)
$log(AD^{Refi} \times$ Ease of Application)	0.024	0.023	0.042	0.041
	(0.020)	(0.021)	(0.036)	(0.037)
$log(AD^{Refi} \times$ Homeownership)	0.016	0.029	0.003	0.008
	(0.036)	(0.045)	(0.034)	(0.038)
$log(AD^{Refi} \times$ Brand-building)	-0.005	0.001	0.007	0.017
	(0.023)	(0.029)	(0.026)	(0.027)
$log(AD^{Purch})$		-0.039		-0.044
		(0.072)		(0.066)
Border Market FE	Y	Y	Y	Y
DMA Border-Year FE	Y	Y	Y	Y
N	1,680	1,680	1,680	1,680

Notes. S.E.s are clustered at the border market level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Next, we examine the robustness of our brand-level demand analysis using alternative topic descriptions. Table A.7.3 presents the results, which are both qualitatively and statistically similar to those in Table 1.9.

Table A.7.3: Heterogeneous Effects on Brand-level Demand for Refinancing

Dependent Variable:	$log(1+Q_j^{Refi})$				$log(\hat{s_j}/\hat{s}_0)$	
	Short descriptions		Long descriptions		Short	Long
	(1)	(2)	(3)	(4)	(5)	(6)
$log(AD_j^{Refi} \times$ Rate and Savings)	0.094	0.092	0.098	0.095	0.118*	0.119*
	(0.069)	(0.069)	(0.076)	(0.076)	(0.063)	(0.069)
$log(AD_j^{Refi} \times$ Ease of application)	-0.022	-0.019	-0.003	-0.003	-0.008	0.022
	(0.065)	(0.065)	(0.098)	(0.098)	(0.008)	(0.098)
$log(AD_j^{Refi} \times$ Homeownership)	0.017	0.014	-0.010	-0.009	-0.042	-0.089
	(0.121)	(0.121)	(0.184)	(0.184)	(0.117)	(0.159)
$log(AD_j^{Refi} \times$ Brand-building)	0.136	0.136	0.134	0.132	0.094	0.093
	(0.087)	(0.087)	(0.099)	(0.099)	(0.094)	(0.073)
$log(AD_{-j}^{Refi})$		0.033**		0.033**		
		(0.015)		(0.015)		
Border Market FE	Y	Y	Y	Y	Y	Y
DMA Border-Year FE	Y	Y	Y	Y	Y	Y
N	40,076	40,076	40,076	40,076	40,076	40,076

Notes. S.E.s are clustered at the lender - border market level; $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

Appendix B: TV Advertising Effectiveness with Racial Minority Representation: Evidence from the Mortgage Market

B.1 Sample Selection Criteria

In this appendix, we outline our sample selection criteria and provide the resulting number of observations. The selection process is as follows:

1. We collect loans that originated directly from lenders for home purchase or refinance purposes during our sample period, resulting in a total of 43,878,666 loans.

2. Among these loans, we select 34,510,763 conventional loans.

3. We then select 19,523,098 refinance loans, which account for about 57% of all conventional loans.

4. Our selection criteria focus on first-lien mortgage loans for owner-occupied, single-family, site-built residential homes. Additionally, we apply two conditions: the loan length must fall within the range of 10 to 30 years, and the loan size should exceed $100,000 but be less than $1,000,000. After applying these filters, we are left with 14,849,323 loans.

5. Next, we remove jumbo loans, which are loans where the loan amount exceeds the limit for conforming loans. This step results in a remaining total of 14,408,940 loans.

6. We exclude "exotic loans" with the following characteristics: reverse, open-end, interest-only, non- or negatively amortizing, balloon payment, or those with prepayment penalties. Additionally, we remove a small number of loans with zero or over 20% APRs. After applying these additional filters, we end up with 14,075,414 loans.

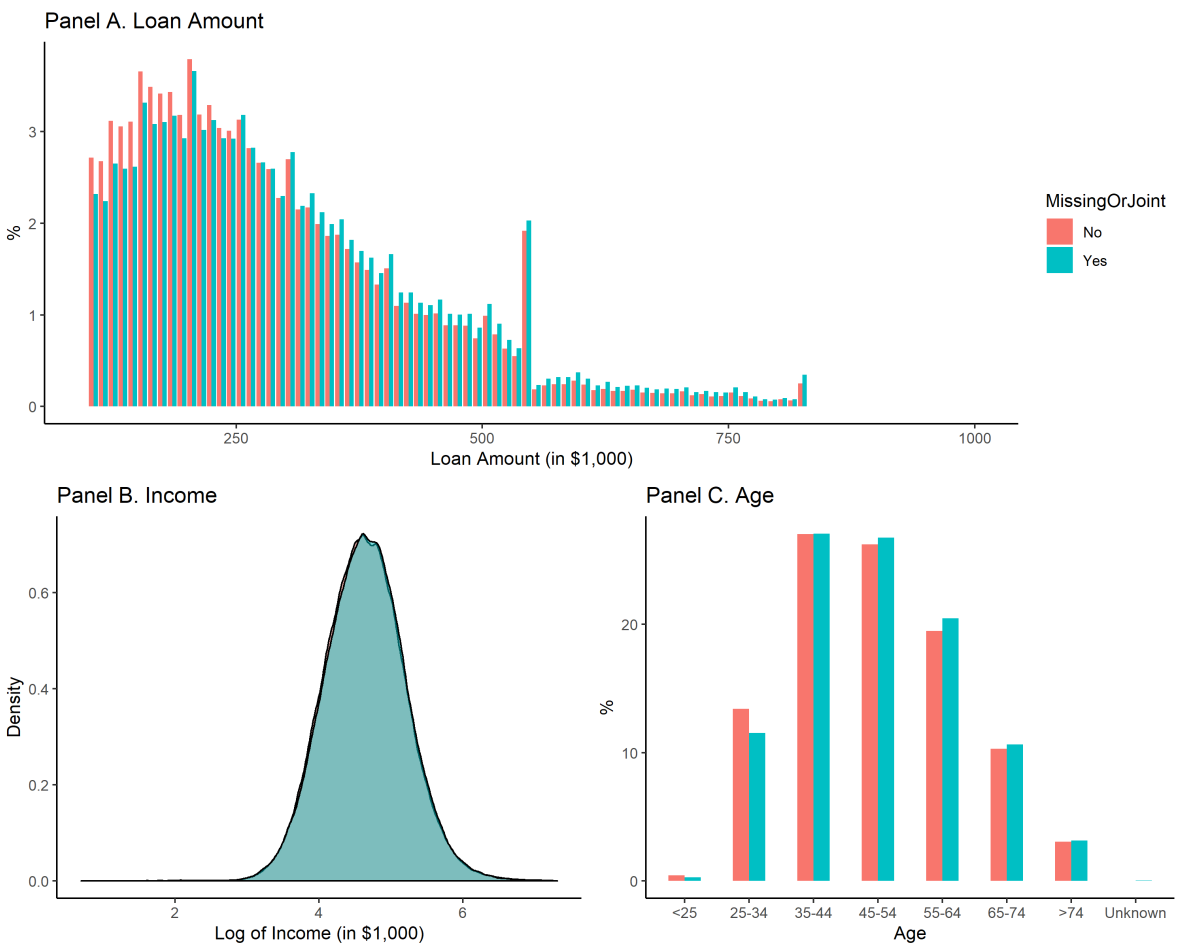
7. We further refine our selection by choosing loans whose underlying property locations belong to the top 101 Designated Market Areas (DMAs) we consider, resulting in a remaining sample of 12,814,253 loans.

8. Finally, we exclude loans with missing ethnic/racial information or cases where the borrowers on the same loan belong to different ethnic or racial backgrounds (i.e., joint). Additionally, we remove a small number of loans originating to American Indian or Alaska Native and Native Hawaiian or Other Pacific Islander borrowers. After applying these criteria, we are left with a total of 9,781,509 loans.

B.2 Selection on Observables

In this appendix, we present suggestive evidence of similar observable characteristics between individuals with missing or joint ethnic/racial information (Group A) and those with complete or single ethnic/racial information (Group B) using the 2021 data. Figure B2 shows these comparisons. In Panel A, we observe a slight difference in loan amounts: borrowers from Group A, on average, borrow $15,266 more than those from Group B. However, considering the average loan amount of $290,854, the difference appears relatively small. Panel B shows nearly identical income distributions, indicating that the disparity in loan size is not driven by income. In Panel C, borrowers from Group A have a slightly higher average age than those from Group B. Overall, the observed differences in these characteristics are relatively small, suggesting that any potential selection bias is likely small.

Figure B2: Distribution of Loan Size, Income, and Age by Ethnic/Racial Information Status

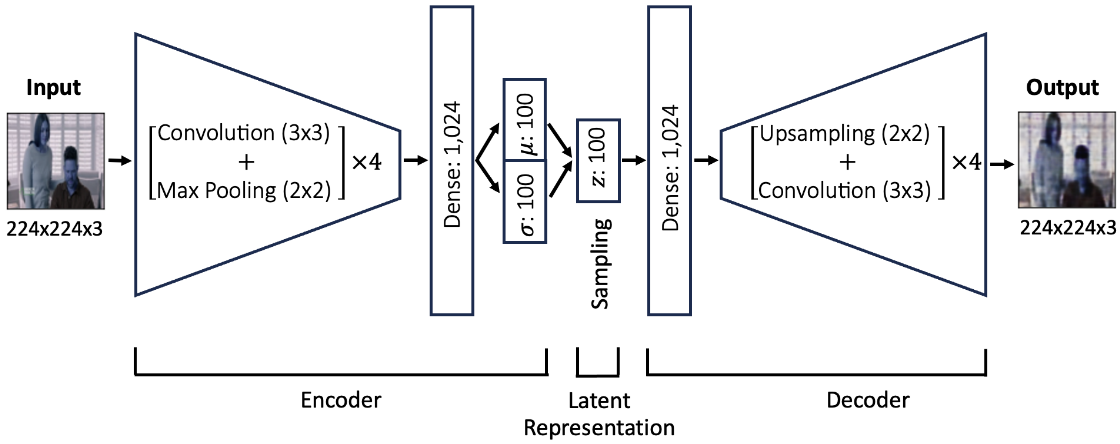


B.3 VAE Architecture and Estimation Details

Figure B3 shows the Variational Autoencoders (VAEs) network structure employed to represent image data in a lower-dimensional vector space, along with an example of an input image and the corresponding reconstructed image. As discussed in Section 2.3.1, an input image is represented as a 150,528-dimensional vector ($224 \times 224 \times 3$). This input is passed through the encoder network, which consists of four convolutional layers and two fully-connected/dense layers. In the convolutional layers, we apply commonly used 3×3 filters and 2×2 max pooling operations, gradually increasing the number of channels to 32, 32, 64, and 64. The purpose of this process is to extract meaningful features from the image while reducing its dimensionality. The resulting output from the convolutional layers is then flattened to a 12,544-dimensional vector ($14 \times 14 \times 64$) and fed into two fully-connected/dense layers. We use the Rectified Linear Unit (ReLU) as the activation function in these two layers. The output of the encoder network is the latent vector, denoted as z, which is sampled from a multivariate Gaussian distribution with parameters μ and σ. This latent vector has a dimensionality of 100 and thus captures a compact representation of the input image. The decoder network is the inverse of the encoder network. It takes the latent vector z as input and gradually increases its dimensionality until it matches the original vector space of the input image. The purpose of the decoder network is to reconstruct an image close to the input image based on the compact representation z. The model is trained with a batch size of 64 for 100 epochs. We employ adaptive learning rates, starting with an initial rate of 0.005. We implement an early stopping rule to prevent overfitting.

We implement another VAE to represent the 1,536-dimensional output of the OpenAI's text embedding model ("text-embedding-ada-002") into a 100-dimensional latent vector space. Since the input data does not contain spatial information like images, we use a simpler network structure that does not require convolutional neural networks. Specifically, we

Figure B3: VAE Network Structure



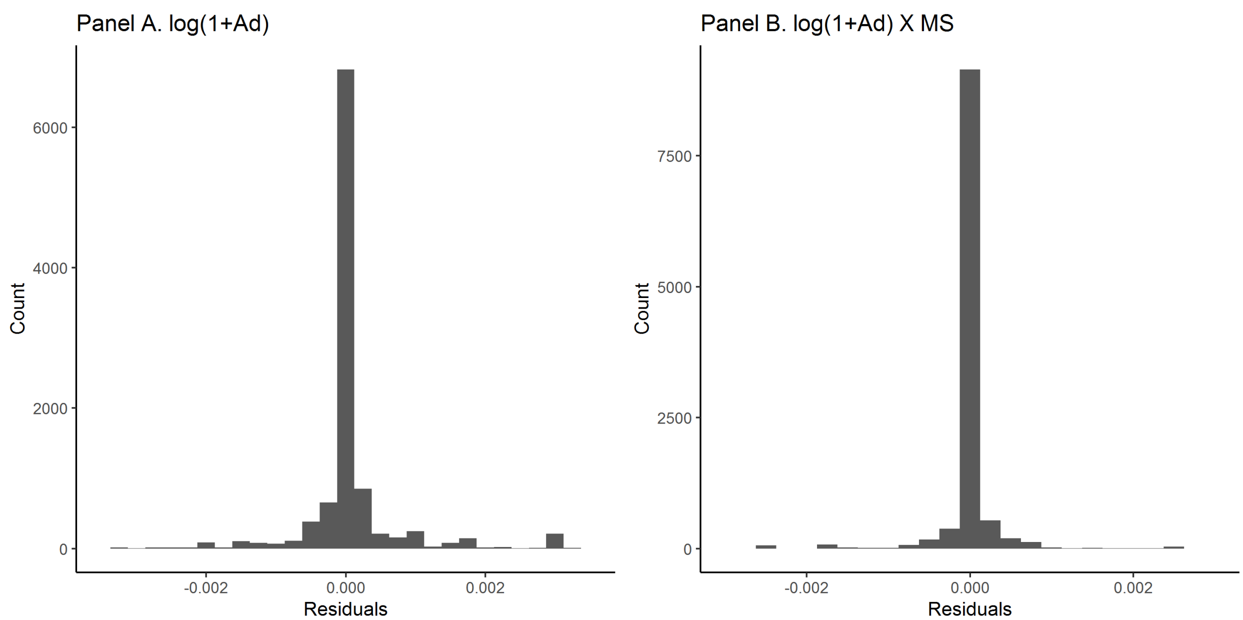
use the fully connected layers for training. In the encoder network, we use a sequence of fully-connected layers with dimensions of 1024, 512, 256, and 100. In the decoder network, we perform the inverse operations, gradually increasing the dimensionality of the latent vector until it matches the original input space. The estimation procedure follows a similar approach as before. We train the model using a batch size of 64 for 100 epochs with adaptive learning rates that start with an initial rate of 0.005. We implement an early stopping rule to prevent overfitting.

B.4 Residual Variation in Advertising

One potential concern about controlling for a large set of fixed effects is that there may be little residual variation in advertising. We explore whether we have sufficient variation in advertising after accounting for the fixed effects. Following previous studies (e.g., Shapiro et al., 2021; Tsai and Honka, 2021), we regress $log(1 + Ad)$ on the lender-DMA and lender-year fixed effects, where Ad represents the total ad spending per capita, including both national and local ad spending. The unit of observation is the lender-DMA-year level. To assess the extent of variation in advertising not explained by the fixed effects, we calculate the ratio of the standard deviation of the residuals to the unconditional mean of ad spending. Additionally, we regress the interaction between $log(1 + Ad)$ and MS on the same set of fixed effects, where MS denotes the corresponding minority share in ads at the lender-DMA-year level. We then calculate the ratio of the standard deviation of the residuals to the unconditional mean of the interaction term.

Figure B4 shows the distributions of the residual variations. We observe a significant level of residual variation in both Panel A and B. Moreover, the calculated ratios of the standard deviation to the unconditional mean are 0.179 for Panel A and 0.171 for Panel B, suggesting sufficient variation in the data.

Figure B4: Residual Variations in Advertising Variables



B.5 Additional Experimental Findings

In this appendix, we present additional findings from the experiment. We first examine the heterogeneous effects by race. Specifically, we regress the likelihood of loan application and recommendation on the seven conditions, using the WW condition as the baseline (intercept), among White consumers and minority consumers. Table B.5.1 presents the results. In columns 1 and 2, we observe that the conditions featuring minority families generally have positive impacts on both DVs among White consumers, although some of the effects are statistically insignificant. On the other hand, in columns 3 and 4, we find that the impacts are greater among minority consumers, and the coefficients are mostly significant. These results align with the findings from our observational study.

Table B.5.1: Heterogeneous Effects based on Each Consumer's Race

	White Consumers		Minority Consumers	
	(1)	(2)	(3)	(4)
	Application	Recommend	Application	Recommend
White-Black (WB)	0.157	0.200	0.421**	0.437**
	(0.118)	(0.132)	(0.184)	(0.209)
White-Asian (WA)	0.230**	0.310**	0.300	0.295
	(0.138)	(0.153)	(0.187)	(0.212)
Black-Black (BB)	-0.074	0.275**	0.324*	0.468**
	(0.117)	(0.131)	(0.184)	(0.209)
Single White (W)	-0.019	-0.070	0.168	0.078
	(0.117)	(0.131)	(0.182)	(0.207)
Single Black (B)	0.103	0.341***	0.560***	0.407*
	(0.117)	(0.132)	(0.193)	(0.219)
Single Asian (A)	0.025	0.205	0.389**	0.418**
	(0.119)	(0.134)	(0.186)	(0.211)
Intercept	4.694***	4.130***	4.462***	4.077***
	(0.084)	(0.094)	(0.129)	(0.146)
N	2,025	2,025	761	761
Adj. R^2	0.002	0.006	0.007	0.005

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Next, we investigate the heterogeneous effects by each consumer's political ideology. To do so, we divide the sample based on the self-reported political leanings, as described in Section 2.6.2. The first group consists of individuals who lean toward a liberal ideology, while the second group consists of those who lean toward a conservative ideology. Within each group, we regress the likelihood of loan application and recommendation on the seven conditions, using the WW condition as the baseline (intercept). The results are presented in Table B.5.2. In columns 1 and 2, we observe that the conditions featuring minority families consistently have positive and statistically significant effects among liberal consumers. In contrast, when considering columns 3 and 4, which correspond to conservative-leaning consumers, the effects are generally negative, although not statistically significant in many cases. Among other conditions, the BB condition stands out with a large and statistically significant negative impact. Overall, these results show a stronger impact of minority representation among liberal consumers, which aligns with the findings from our observational study.

Table B.5.2: Heterogeneous Effects based on Each Consumer's Political Leaning

	Liberal Consumers		Conservative Consumers	
	(1)	(2)	(3)	(4)
	Application	Recommend	Application	Recommend
White-Black (WB)	0.344***	0.436***	-0.061	-0.074
	(0.132)	(0.151)	(0.194)	(0.214)
White-Asian (WA)	0.481***	0.547***	-0.096	0.096
	(0.128)	(0.147)	(0.194)	(0.215)
Black-Black (BB)	0.516***	0.824***	-0.517***	-0.354
	(0.133)	(0.151)	(0.194)	(0.215)
Single White (W)	0.044	0.003	-0.069	0.029
	(0.131)	(0.149)	(0.195)	(0.216)
Single Black (B)	0.381***	0.586***	-0.148	-0.117
	(0.133)	(0.152)	(0.194)	(0.215)
Single Asian (A)	0.371***	0.524***	-0.350*	-0.180
	(0.131)	(0.150)	(0.202)	(0.223)
Intercept	4.439***	3.860***	5.087***	4.635***
	(0.092)	(0.105)	(0.137)	(0.152)
N	1,454	1,454	786	786
Adj. R^2	0.016	0.029	0.006	-0.0003

Note. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

We further examine consumers' perceptions of the advertised lender and advertisements across the seven conditions. Specifically, we regress each of the six attitudinal measures collected in our experiment on the seven experimental conditions, using the WW condition as the baseline (intercept). The results are presented in Table B.5.3. Overall, we observe that conditions featuring minority families lead to more positive perceptions across all dimensions we consider.

Table B.5.3: Perceptions of the Advertised Lender and Advertisement

	(1)	(2)	(3)	(4)	(5)	(6)
	Broad Options	Cater to Me	Fair Lending	Inclusive	Fresh&New	Attention
White-Black (WB)	0.278***	0.258**	0.200**	1.857***	0.774***	0.658***
	(0.094)	(0.104)	(0.090)	(0.100)	(0.122)	(0.118)
White-Asian (WA)	0.320***	0.350***	0.276***	1.716***	0.708***	0.610***
	(0.093)	(0.103)	(0.090)	(0.099)	(0.121)	(0.118)
Black-Black (BB)	0.184**	-0.030	0.099	1.421***	0.864***	0.612***
	(0.094)	(0.104)	(0.091)	(0.100)	(0.122)	(0.118)
Single White (W)	0.126	0.080	0.158*	0.694***	0.187	0.182
	(0.093)	(0.103)	(0.090)	(0.099)	(0.121)	(0.117)
Single Asian (A)	0.272***	0.189*	0.365***	1.823***	0.723***	0.451***
	(0.095)	(0.105)	(0.092)	(0.101)	(0.123)	(0.120)
Single Black (B)	0.391***	0.147	0.303***	1.793***	0.979***	0.814***
	(0.094)	(0.104)	(0.091)	(0.101)	(0.123)	(0.119)
Intercept	4.545***	4.562***	4.368***	3.662***	3.517***	3.667***
	(0.066)	(0.073)	(0.064)	(0.071)	(0.086)	(0.084)
N	2,796	2,796	2,796	2,796	2,796	2,796
Adj. R^2	0.006	0.005	0.006	0.179	0.035	0.023

Note. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Lastly, we consider how the six attitudinal measures towards the lender and the advertisement correlates with the likelihood of loan application and recommendation. The results are presented in Table B.5.4. In both Panel A and B, we observe that each of the six attitudinal measures is positively correlated with the likelihood to apply for the advertised lender as well as recommend it.

Table B.5.4: Impact of Consumer Perceptions on the Likelihood of Application and Recommendation

Panel A: Likelihood of Loan Application						
	(1)	(2)	(3)	(4)	(5)	(6)
	Application					
Broad Options	0.598***					
	(0.017)					
Cater to Me		0.591***				
		(0.014)				
Fair Lending			0.585***			
			(0.017)			
Inclusive				0.405***		
				(0.015)		
Fresh & New					0.453***	
					(0.013)	
Attention						0.452***
						(0.013)
Intercept	1.906***	1.975***	2.081***	2.738***	2.888***	2.883***
	(0.082)	(0.070)	(0.083)	(0.080)	(0.056)	(0.059)
N	2,796	2,796	2,796	2,796	2,796	2,796
Adj. R^2	0.318	0.381	0.286	0.202	0.320	0.296

Panel B: Likelihood of Recommendation						
	(1)	(2)	(3)	(4)	(5)	(6)
	Recommendation					
Broad Options	0.651***					
	(0.019)					
Cater to Me		0.632***				
		(0.017)				
Fair Lending			0.671***			
			(0.020)			
Inclusive				0.456***		
				(0.017)		
Fresh & New					0.538***	
					(0.014)	
Attention						0.552***
						(0.014)
Intercept	1.223***	1.355***	1.263***	2.057***	2.111***	2.042***
	(0.094)	(0.082)	(0.093)	(0.090)	(0.062)	(0.064)
N	2,796	2,796	2,796	2,796	2,796	2,796
Adj. R^2	0.297	0.342	0.295	0.200	0.353	0.347

Note. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix C: The Role of Slant and Message Consistency in Political Advertising Effectiveness: Evidence from the 2016 Presidential Election

C.1 Slant Variable Details

In this appendix, we present details on the calculations of the slant index.

We first pre-process texts using the NLTK module in Python: we make words lowercase, and remove stop words,[1] punctuations, and numbers. We then tokenize each of the texts and stem words in the text.

Following Gentzkow and Shapiro (2010) and their notations for consistency, we first derive the mapping between a vector of word counts a congressperson used and the political leanings of their district. For each word p in the 114th Congressional Record, we count the number of times the word p is used by each of the two parties and calculate a chi-square statistic, χ^2_p. We restrict our focus to words that occur at least two times but fewer than 100 times in the candidates' public speeches and transcribed ad texts.[2] This removes some of the most common and least common words, which are not useful for the analysis. We then select the 1,000 words with the highest values of χ^2_p.

[1]We add a handful of words to the existing list of stop words from NLTK in Python, such as madam, speaker, and thank. These words appear frequently but are not informative of one's political ideology.

[2]Contrary to Gentzkow and Shapiro (2010) who used newspaper articles, ad texts are typically short. Therefore, to overcome the scarcity of words, we use both the candidates' speeches and ad texts to select words to consider.

Among the 1,000 selected words, we regress congressperson c's relative frequency of word p, $\tilde{f}_{pc}$, on their ideology, $ideology_c$, measured by the Republican vote share in the district[3] from the 2012 presidential election (collected from https://dailykos.com) and estimate an intercept parameter α_p and slope parameter β_p. A positive (negative) slope estimate suggests that the word p is associated with the Republican (Democratic) party.

We then compute the political slant for each of the ad creatives by applying the same mapping between the relative word frequencies and political slants of those words used in the ad. Specifically, the political slant of ad creative n is computed as $\tilde{y}_n = \frac{\sum_{p=1}^{p=1,000} \beta_p \cdot (\tilde{f}_{pn} - \alpha_p)}{\sum_{p=1}^{p=1,000} \beta_p^2}$.

Finally, we re-index the estimated political slant of ad creative n, $\tilde{y}_n$ to $\tilde{y}_{cn}$ to denote the candidate c that ad creative n supports. Our slant measure is then calculated as $\hat{y}_{cn} = -(\tilde{y}_{cn} - 0.5)$ if c is Clinton and $\hat{y}_{cn} = (\tilde{y}_{cn} - 0.5)$ if c is Trump. Thus, a greater (lower) slant measure always corresponds to a more politically extreme (centrist) message for both candidates.

C.2 Doc2Vec Validation

In this appendix, we provide additional evidence supporting the doc2vec algorithm's performance. Specifically, we present two ad creatives and compare them with other ads that are identified as the most and least similar. Table C.2.1 presents selected ads supporting Clinton and Trump in Panel A and B, respectively.

[3]We use the Republican candidate's vote share in the state for senators and congressional district for representatives.

Table C.2.1: Examples of Ads that Support Clinton or Trump

Selected Ad	Candidate (tone)	Ad content
Panel A. Pro-Clinton Advertisement		
Focal ad	Clinton (neg.)	I spent many years as a nuclear missile launch officer. If the president gave the order, we had to launch the missiles. That would be it. I prayed that call would never come. Self-control may be all that keeps these missiles from firing. [Trump speaking] I would bomb the F out of them I want to be unpredictable. I love war. The thought of Donald Trump with nuclear weapons scares me to death. Should scare everyone.
Most similar	Clinton (neg.)	If he governs consistent with some of the things he said as a candidate, I would be very frightened. He's been talking about the option of using a nuclear weapon against our Western European allies. This is not somebody who should be handed the nuclear codes. You have to ask yourself, do I want a person of that temperament control the nuclear codes? And as of now, I have to say no.
Least similar	Trump (pos.)	The American moment is here, two choices, two Americans decided by you. Hillary Clinton will keep us on the road to stagnation. Fewer jobs, rising crime, America diminished at home and abroad. Donald Trump will bring the change we are waiting for. America better, stronger, more prosperous. For everyone, a plan for a future brighter than our past. The choice is yours.
Panel B. Pro-Trump Advertisement		
Selected Ad	Candidate (tone)	Ad content
Focal ad	Trump (pos.)	The most important job any woman can have is being a mother, and it shouldn't mean taking a pay cut. I'm Ivanka Trump, a mother, a wife and an entrepreneur. Donald Trump understands the needs of the modern work force. My father will change outdated labor laws so that they support women and American families. He will provide tax credits for childcare, paid maternity leave and dependent care savings accounts. This will allow women to support their families and further their careers.
Most similar	Clinton (pos.)	Far too many families today don't earn what they need and don't have the opportunities they deserve. I believe families deserve quality education for their kids. Childcare they can trust and afford. Equal pay for women and jobs they can really live on. People ask me what'll be different if I'm president? Well, kids and families have been the passion of my life, and they will be the heart of my presidency.
Least similar	Clinton (pos.)	What does showing up when it's time to vote actually mean? You care about protecting his legacy and our progress. You care about moving forward, united as one, because when we show up in full force and when we refuse to stand by quietly, we show what it means to be stronger together.

C.3 Sensitivity of the Results to Alternative Size of Vector Space in dov2vec

In this appendix, we show the robustness of our main results to variations in the dimension of the vector space, which is one of the most important hyper-parameters in the doc2vec model. Table C.3.1 presents the effects on WOM volume, and Table C.3.2 presents the effects on voter preference. In all cases, standard errors are clustered at the candidate level.

Table C.3.1: Effects of Political Slant and Message Consistency on WOM

	(1)	(2)	(3)	(4)
	Vector size = 150 $\log(1 + WOM^{post})$		Vector size = 300 $\log(1 + WOM^{post})$	
$\log(1 + WOM^{pre})$	0.649***	0.650***	0.649***	0.650***
	(0.001)	(0.0004)	(0.001)	(0.0001)
Slant	−0.030***		−0.030***	
	(0.009)		(0.008)	
Consistency	0.173**		0.186***	
	(0.076)		(0.072)	
Attack ads	−0.064***		−0.065***	
	(0.015)		(0.014)	
Slant × PreOct1st		−0.082		−0.083
		(0.098)		(0.096)
Slant × PostOct1st		−0.007		−0.005
		(0.018)		(0.017)
Consistency × PreOct1st		0.208***		0.305***
		(0.036)		(0.035)
Consistency × PostOct1st		0.162		0.151
		(0.389)		(0.349)
Attack ads × PreOct1st		0.045*		0.039
		(0.024)		(0.026)
Attack ads × PostOct1st		−0.131**		−0.129**
		(0.059)		(0.060)
Pro-Clinton ads	−0.178*	−0.183**	−0.177*	−0.179**
	(0.093)	(0.088)	(0.092)	(0.082)
log(Audience size)	0.128***	0.129***	0.127***	0.129***
	(0.048)	(0.048)	(0.048)	(0.048)
Ad length	0.317*	0.316*	0.319*	0.316*
	(0.168)	(0.154)	(0.169)	(0.156)
Ad position in break	−0.125**	−0.107***	−0.125**	−0.106***
	(0.054)	(0.028)	(0.054)	(0.028)
Week, Day, Hourly F.E.s	Yes	Yes	Yes	Yes
Program Genre and Network F.E.s	Yes	Yes	Yes	Yes
N	824	824	824	824
R^2	0.777	0.779	0.777	0.779

Notes. S.E.s are clustered at the candidate level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table C.3.2: Effects of Political Slant and Message Consistency on Voter Preference

	(1)	(2)	(3)	(4)
	Vector size = 150 Voter preference (in %)		Vector size = 300 Voter preference (in %)	
Lagged voter preference	0.865***	0.855***	0.866***	0.856***
	(0.012)	(0.015)	(0.011)	(0.015)
Slant	−0.299***		−0.299***	
	(0.023)		(0.023)	
Consistency	0.279***		0.279***	
	(0.046)		(0.046)	
Attack ads	−0.005		−0.005	
	(0.079)		(0.079)	
Slant × PreOct1st		−0.343***		−0.340***
		(0.084)		(0.083)
Slant × PostOct1st		−0.228***		−0.226***
		(0.010)		(0.012)
Consistency × PreOct1st		0.521		0.615
		(0.317)		(0.435)
Consistency × PostOct1st		−0.223**		−0.159
		(0.105)		(0.234)
Attack ads × PreOct1st		0.129		0.132
		(0.140)		(0.142)
Attack ads × PostOct1st		−0.290*		−0.286*
		(0.174)		(0.166)
No ads	−0.189	0.300	−0.193	0.305
	(0.341)	(0.367)	(0.337)	(0.362)
log(Audience size: Own Ads)	−0.037	−0.019	−0.037	−0.020
	(0.031)	(0.034)	(0.028)	(0.032)
log(Audience size: Rival's Ads)	0.001	0.005***	0.001	0.005***
	(0.001)	(0.0004)	(0.001)	(0.0004)
Number of ads	0.010	0.012	0.010	0.012
	(0.028)	(0.029)	(0.027)	(0.028)
Ad position	−0.103	−0.094	−0.103	−0.096
	(0.291)	(0.320)	(0.294)	(0.320)
Ad length	0.173	0.136	0.171	0.141
	(0.946)	(0.965)	(0.936)	(0.951)
Candidate and Week F.E.s	Yes	Yes	Yes	Yes
Program Genre and Network Controls	Yes	Yes	Yes	Yes
N	240	240	240	240
R^2	0.852	0.855	0.852	0.855

Notes. S.E.s are clustered at the candidate level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

C.4 Sensitivity of the WOM Results to Time Windows

In this appendix, we show that our main findings are robust to different time windows. Table C.4.1 presents the results results for two- and three-minute windows.

Table C.4.1: Effects on WOM - Alternative Time Windows

	(1)	(2)	(3)	(4)
	Two-min. window $\log(1+WOM^{post})$		Three-min. window $\log(1+WOM^{post})$	
$\log(1+WOM^{pre})$	0.529***	0.530***	0.603***	0.603***
	(0.002)	(0.004)	(0.002)	(0.001)
Slant	−0.032		−0.084***	
	(0.039)		(0.027)	
Consistency	0.446***		0.158	
	(0.079)		(0.118)	
Attack ads	−0.059		−0.085***	
	(0.041)		(0.021)	
Slant × PreOct1st		0.027		−0.062
		(0.077)		(0.075)
Slant × PostOct1st		−0.068**		−0.099***
		(0.028)		(0.008)
Consistency × PreOct1st		0.135		0.072
		(0.092)		(0.068)
Consistency × PostOct1st		0.632*		0.227
		(0.347)		(0.444)
Attack ads × PreOct1st		0.196***		0.106**
		(0.045)		(0.047)
Attack ads × PostOct1st		−0.209*		−0.196**
		(0.126)		(0.093)
Pro-Clinton ads	−0.270***	−0.270***	−0.205***	−0.209**
	(0.061)	(0.056)	(0.074)	(0.068)
log(Audience size)	0.176***	0.172***	0.180***	0.178***
	(0.046)	(0.040)	(0.042)	(0.039)
Ad length	0.131	0.148	0.220	0.228
	(0.176)	(0.170)	(0.175)	(0.167)
Ad position in break	−0.131**	−0.116	−0.178	−0.161
	(0.199)	(0.162)	(0.140)	(0.167)
Week, Day, Hourly F.E.s	Yes	Yes	Yes	Yes
Program Genre and Network F.E.s	Yes	Yes	Yes	Yes
N	824	824	824	824
R^2	0.675	0.680	0.731	0.734

Notes. S.E.s are clustered at the candidate level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

C.5 Tests on the Level of Clustering

In this appendix, we conduct the statistical test for the appropriate level of clustering proposed by MacKinnon et al. (2023). MacKinnon et al. (2023) test the null hypothesis of a finer clustering level against the alternative hypothesis of a coarser clustering level. In our context, we can test whether clustering at the ad creative level (null hypothesis) against clustering at the candidate level (alternative hypothesis). The results, shown in Table C.5.1, reveal that the no clustering case (the finest case) is rejected for both the ad creative level and the candidate level, but clustering at the ad creative level is rejected against the candidate level. Taken together, these results suggest that we cluster standard errors at the candidate level.

Table C.5.1: Statistical Tests on the Level of Clustering

	Estimated Model: Equation (2)	
Test	Statistic	Bootstrapped p-value
N vs. A	76.63	0.000
N vs. C	29.35	0.014
N vs. C	19.08	0.031

Notes. N denotes no clustering; A denotes clustering at the ad creative; C denotes clustering at the candidate level.

C.6 Influence of Outliers on the WOM Analysis

In this appendix, we show that our results remain very similar to removal of outliers. Specifically, we run our WOM analysis after both winsorizing and trimming the post-WOM volume at the 1% and 99% level.

Table C.6.1: Effects on WOM - Winsorizing and Trimming Outliers

	(1)	(2)	(3)	(4)
	Winsorized WOM $\log(1+WOM^{post})$		Trimmed WOM $\log(1+WOM^{post})$	
$\log(1+WOM^{pre})$	0.633***	0.634***	0.629***	0.629***
	(0.006)	(0.008)	(0.013)	(0.015)
Slant	−0.095***		−0.029***	
	(0.007)		(0.005)	
Consistency	0.215***		0.140***	
	(0.079)		(0.044)	
Attack ads	−0.068***		−0.063***	
	(0.017)		(0.002)	
Slant × PreOct1st		−0.073		−0.074
		(0.100)		(0.109)
Slant × PostOct1st		0.005		−0.015***
		(0.014)		(0.002)
Consistency × PreOct1st		0.361***		0.393***
		(0.013)		(0.066)
Consistency × PostOct1st		0.173		−0.018
		(0.359)		(0.278)
Attack ads × PreOct1st		0.031**		0.010
		(0.015)		(0.010)
Attack ads × PostOct1st		−0.130**		−0.133***
		(0.057)		(0.037)
Pro-Clinton ads	−0.193***	−0.196***	−0.178**	−0.176**
	(0.086)	(0.079)	(0.080)	(0.073)
log(Audience size)	0.117**	0.118**	0.106*	0.107*
	(0.053)	(0.052)	(0.056)	(0.073)
Ad length	0.247***	0.245***	0.239***	0.237***
	(0.079)	(0.067)	(0.058)	(0.044)
Ad position in break	−0.148***	−0.129***	−0.150***	−0.133***
	(0.065)	(0.040)	(0.043)	(0.022)
Week, Day, Hourly F.E.s	Yes	Yes	Yes	Yes
Program Genre and Network F.E.s	Yes	Yes	Yes	Yes
N	824	824	809	809
R^2	0.779	0.781	0.769	0.771

Notes. S.E.s are clustered at the candidate level; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

C.7 Comparison of National Primetime Ads versus Local Primetime Ads

In this appendix, we show that Clinton's and Trump's national primetime ads exhibit very similar airing patterns to their local primetime ads in terms of time, day of the week, and month aired. Note that the summary statistics for the local prime ads are generated from the raw Stradegy data on the candidate's primetime advertising and have not been cleaned for data errors.

Table C.7.1: National versus Local Primetime Ad Airings by Time, Day of the Week, and Month

Panel A: By time								
	Clinton				Trump			
Time Slot	Local ad airings		National ad airings		Local ad airings		National ad airings	
7:00–7:59 PM	12,522	4.2%	10	1.6%	2,576	6.9%	5	2.6%
8:00–8:59 PM	108,931	36.8%	178	28.1%	13,808	37.1%	57	30.0%
9:00–9:59 PM	99,866	33.7%	232	36.6%	11,930	32.0%	81	42.6%
10:00–10:59 PM	74,887	25.3%	214	33.8%	8,928	24.0%	47	24.7%
Panel B: By day of the week								
	Clinton				Trump			
Day	Local ad airings		National ad airings		Local ad airings		National ad airings	
Sun	50,782	17.1%	104	15.4%	8,195	22.0%	32	16.8%
Mon	33,262	11.2%	75	11.1%	6,061	16.3%	42	22.1%
Tue	42,763	14.4%	120	17.8%	4,011	10.8%	25	13.2%
Wed	44,725	15.1%	121	17.9%	4,468	12.0%	22	11.6%
Thu	41,944	14.2%	77	11.4%	5,388	14.5%	23	12.1%
Fri	44,247	14.9%	78	11.5%	4,420	11.9%	21	11.1%
Sat	38,483	13.0%	59	8.7%	4,699	12.6%	25	13.2%
Panel C: By month								
	Clinton				Trump			
month	Local ad airings		National ad airings		Local ad airings		National ad airings	
Jun.	3,679	1.2%	9	1.4%	0	0.0%	1	0.5%
Jul.	52,326	17.7%	92	14.5%	206	0.6%	13	6.8%
Aug.	35,314	11.9%	99	15.6%	172	0.5%	8	4.2%
Sep.	85,014	28.7%	178	21.8%	2,275	6.1%	17	8.9%
Oct.	87,039	29.4%	181	28.5%	23,618	63.4%	81	42.6%
Nov.	32,834	11.1%	75	11.8%	10,971	29.5%	70	36.8%