

©Copyright 2023

Mariia Mitkina

Impact of Social Media on Prosocial Offline Outcomes

Mariia Mitkina

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Yong Tan, Chair

Ming Fan

Mingwen Yang

Program Authorized to Offer Degree:
Business

University of Washington

Abstract

Impact of Social Media on Prosocial Offline Outcomes

Mariia Mitkina

Chair of the Supervisory Committee:

Yong Tan

Information Systems and Operations Management

This dissertation consists of two essays studying how specific artifacts of social media platforms create conditions for improved offline outcomes for their users. The contemporary landscape of social media is marked by a burgeoning number of platforms and an ever-expanding user base, giving rise to a multitude of novel social media artifacts that have the potential to significantly influence the outcomes experienced by users. Within this context, our research aims to shed light on two distinct and understudied social media phenomena and assess their causal impact on offline outcomes.

In the first essay, the focus is on the potential aid of User Generated Content (UGC) in the rising of political consumerism. In recent years, there have been multiple social movements that resulted in consumer choice shifts of the engaged social movement participants. For example, during the recent Black Lives Matter protests, the movement participants were actively looking for Black-owned businesses to support. In the absence of detailed lists for such and similar businesses, UGC mentioning the ownership status has become instrumental in the minority-owned business search. In this essay, we estimate the impact of increased digital visibility for Black-owned restaurants on Yelp through user-generated reviews mentioning Black ownership. We also study the heterogeneity of this effect across different consumer and business categories.

The second essay focuses on the role of creators on social media and how their content

and growth strategies can benefit content consumers. We use the context of the physical exercising genre posted on social media and study whether the popular free fitness content is able to provide a setting for consistent exercising for a wide audience. Namely, we study the effect of fitness challenges – a common content posting strategy utilized by fitness content creators – on the engagement dynamics of content consumers. We demonstrate that structured fitness programs improve immediate and future creator performance and provide improved individual user engagement as well as habit formation, a sense of progress, and a feeling of community among content consumers.

TABLE OF CONTENTS

	Page
List of Figures	iii
List of Tables	iv
Chapter 1: Introduction	1
Chapter 2: Literature Review	4
2.1 Literature Review on Online Platforms and Social Movements	4
2.2 Literature on Exercise Behaviors and Health Management	5
Chapter 3: Effect of Online Visibility on Minority Business Performance	9
3.1 Introduction	9
3.2 Research Context & Identification Strategy	14
3.3 Data	21
3.4 Empirical Analysis and Results	28
3.5 Heterogeneous Effects	34
3.6 Implications and Conclusion	42
Chapter 4: Effect of Online Fitness Challenges on User' Exercising Behavior	46
4.1 Introduction	46
4.2 Empirical Setting and Data	51
4.3 Empirical Strategy and Results	57
4.4 Discussion and Conclusion	71
Chapter 5: Concluding Remarks	75
Bibliography	77

Appendix A	89
Appendix B	100

LIST OF FIGURES

Figure Number	Page
3.1 Results For Generic Search	16
3.2 Results For Black-Owned Search	18
3.3 Difference between restaurants with reviews mentioning Black ownership and restaurants that have voluntarily adopted Black-owned badges in the search interface for Black-owned searches	19
3.4 Black-owned restaurants with reviews mentioning Black-ownership and no voluntary badges appearing in generic and Black-owned searches	20
3.5 Google Trends for “Black-Owned Business”	24
3.6 Event Study Analysis	30
4.1 Example of Fitness Challenge on YouTube	48
4.2 Parallel Trends Check: Lagged Coefficients Estimates	70
A.1 Dynamic Treatment Effects	91

LIST OF TABLES

Table Number	Page
3.1 Summary Statistics of Monthly Visits	26
3.2 Effect of Reviews Mentioning Black-Ownership on Visits to Black-Owned Restaurants	31
3.3 Impact of City Level Segregation on Political Consumerism	37
3.4 Effect of Yelp Reviews by Neighborhood Characteristics	38
3.5 Effect on Exposure	41
4.1 Comment Labeling Examples	55
4.2 Labeling Performance	56
4.3 Summary Statistics	58
4.4 Effect of Challenges on Aggregate Daily Performance	61
4.5 Effect of Challenges on Daily Performance - Counterfactual Estimates	63
4.6 Mechanisms	65
4.7 Effect of Yoga Challenges on Individual Commenting Behavior	68
4.8 Robustness Check: Effect of Challenges on Daily Performance	71
A.1 Effect of Yelp Reviews Robustness Checks - CEM	89
A.2 Effect of Yelp Reviews Robustness Checks - CE	90
A.3 Effect of Yelp Reviews on Dine-In Visits	92
A.4 Effect of Yelp Reviews on Black-Owned Restaurant Visits After Removing Restaurants That Received Their First Black-Owned Review During Months Of High Black Interest	93
A.5 Effect of Yelp Reviews on Black-Owned Restaurant Visits After Removing Restaurants with Offline Labels	94
A.6 Effect of Yelp Reviews on Black-Owned Restaurant Visits - Placebo Test . .	95
A.7 Effect of Yelp Reviews on Black-Owned Restaurant Visits	96
B.1 Comment Labeling Keywords	100
B.2 List of Yoga Channels Used in Analysis	103

B.3 Effect of Yoga Challenges on Individual Users - Probability Margins 104

ACKNOWLEDGMENTS

I would like to express my profound gratitude to the people who played a crucial role in shaping my Ph.D. journey. First and foremost, my sincere appreciation goes to my advisor, Yong Tan, whose unwavering support has been instrumental throughout this expedition. His commitment to fostering intellectual freedom, providing a robust network for knowledge exchange, and offering unconditional support to all his students, regardless of their chosen paths, is truly commendable and deeply valued.

I extend heartfelt thanks to the distinguished professors with whom I had the privilege to collaborate during my studies: Uttara Ananthkrishnan, Stephanie Lee, and Yingda Lu. Your contributions of knowledge, guidance, and inspiration have left an indelible mark on my academic journey, and for that, I am truly grateful.

To the faculty of the ISOM Department, I express my gratitude for the wealth of knowledge you shared through your exceptional courses, seminars, talks, and day-to-day interactions. Your dedication to academic excellence has significantly enriched my learning experience.

My heartfelt appreciation also extends to the friends I made during my Ph.D. studies at the Foster School of Business. Your friendship has made this period truly special and full of joyous moments. I am also grateful for all the thoughtful discussions, advice, and support I had, thanks to them.

A special acknowledgment is reserved for Nuzulita Budhiari, Shawna Reimers, and Beau Kirkeby. Your unwavering support and willingness to assist in every situation, irrespective of deadlines, have been invaluable.

Lastly, I thank my husband, Mohamed Hussein, for his unwavering support, boundless

love, and remarkable patience. His encouragement has been a constant source of strength throughout this academic endeavor.

DEDICATION

to my husband, Mohamed, and our families

Chapter 1

INTRODUCTION

The contemporary landscape of social media is marked by a burgeoning number of platforms and an ever-expanding user base, giving rise to a multitude of novel social media artifacts that have the potential to significantly influence the outcomes experienced by users. Within this context, our research aims to shed light on two distinct and understudied social media phenomena, with the primary purpose of estimating the causal impact of these phenomena on offline outcomes.

In Chapter 3, we turn to studying the effect of the User-Generated Content (UGC) on the performance of minority-owned businesses at times of increased societal interest. Although UGC has been a popular and widely researched area (Chevalier and Mayzlin, 2006; Erkan and Evans, 2016; Verma and Yadav, 2021), there are few studies focusing on the impact of UGC on the performance of the underperforming products and its potential to decrease the earning gap between businesses. Below is a short summary of the research questions discussed in Chapter 3.

Consumers are increasingly supporting businesses and brands that align closely with their values in an act of political consumerism. Despite the increasing demand to support minority-owned businesses, discovering these businesses through traditional channels remains challenging. We examine whether and how online reviews enable the discovery of Black-owned businesses, particularly during increased interest in supporting Black-owned businesses. We leverage GPS-enabled foot traffic data, hand-collected data from a directory of Black-owned restaurants, business details from Yelp.com, and zipcode-level social capital data to understand the potential mechanisms driving this behavior. Our results suggest that Black-owned restaurants that receive reviews mentioning Black ownership experience a

10.8% increase in foot traffic compared to Black-owned restaurants that do not receive these reviews. Further, reviews that mention black ownership provide a 4.6% increase in foot traffic to these restaurants during increased consumer interest in supporting Black causes. We find that reviews mentioning Black ownership aid Black-owned restaurants located in cities with lower segregation and racial isolation and in neighborhoods with stronger ties across socioeconomic status. Our results also suggest online reviews mentioning Black ownership increase the foot traffic to Black-owned restaurants by increasing interracial exposure - non-White consumers who patronize Black-owned restaurants experience an increase in the likelihood of encountering White consumers after the restaurants receive a review mentioning Black ownership. Together, our results demonstrate the value of user-generated content in enabling political consumerism even without targeted interventions from digital platforms or voluntary disclosure of race by business owners.

After that, we shift the focus to another emerging phenomenon of social media platforms. With the increasing time that customers spend online, video content creation has been a constantly increasing market. In Chapter 4, we aim to understand whether specific content creation and posting strategies, and their effect on the content performance as well as individual consumer response effects. We focus on the physical fitness genre, where the improved content performance might lead to improved physical activity and outcomes for the audience as well. Understanding of the content creation strategies and their effect on content performance is an emerging field of study, and Chapter 4 contributes to this literature. We summarize our approach and findings below.

Despite the boom of the fitness industry and recent technological developments that enhance physical activity, individuals still experience barriers to consistent exercise. Recently, the online fitness market has expanded through the growing popularity of fitness apps and online workouts posted on social media. One way that these fitness creators have kept their audience engaged is fitness challenges, that is, a workout plan published by a creator with defined start and end dates. We use the context of YouTube yoga channels to estimate the effect of such challenges on the performance of the hosting channels and their followers'

activity during and after the challenges. We collect data for 21 popular yoga channels on YouTube and estimate the impact of a total of 37 fitness challenges. We demonstrate that fitness challenges improve aggregate channel performance, such as daily views, engagements, new followers, and views per follower, both during and after a challenge. Further, we analyze user comments content and demonstrate that habit formation, community support, and improved physical performance might be the mechanisms behind the effectiveness of the fitness challenges. Finally, we show that individual users are more likely to comment on the channel during and after the challenge. Overall, our findings suggest that fitness challenges can be an efficient strategy for fitness workout creators to improve performance and for content consumers to increase engagement. The findings have several important implications for fitness content creators, consumers, sponsors, and policymakers.

Chapter 2

LITERATURE REVIEW

In this chapter, we discuss the literature related to both essays. The first part of this chapter relates to the first essay and reviews the only recently emerging literature on 1) the role of social platforms in decreasing racial disparities and 2) the role of social media in social movements and societal outcomes. The second part of this chapter relates to the second essay and covers the various related streams of literature, such as the role of online platforms in health management and the factors affecting the performance of video content on social media. We first discuss the prior literature related to online exercise platforms. After that, we review behavioral theories that can explain the effect of the fitness challenges on individual users' exercise patterns.

2.1 Literature Review on Online Platforms and Social Movements

First, Chapter 3 is related to the literature studying the impact of digital platforms and policies on mitigating racial disparities. For example, Agarwal and Sen (2022) find that a crowdfunding platform for school teachers can promote more racially-inclusive curricula in schools during a time of increased national interest in race-related topics. Additionally, (Zhang et al., 2021) find that a smart-pricing feature on the Airbnb platform can help decrease the earnings gap between the Black and the White hosts, conditional on adoption. Studies have shown how technology adoption might be able to close the gap between access to capital between Black and White Americans Howell et al. (2021). Our work examines the impact of features designed to increase the visibility of minority business owners from a new angle. We analyze how online platforms can assist minority business owners and what kinds of businesses benefit from these features.

Second, Chapter 3 is related to the literature studying the outcomes of social movements amplified by social media activity and coordination (Mundt et al., 2018). In a recent study, (Luo and Zhang, 2021) find that the #MeToo movement led to an increase in the number of women on the production board in the film-making industry. Agarwal and Sen (2022) conclude that the 2020 BLM protests increased the demand for inclusive literature in schools. In our study, we leverage a platform’s introduction of digital badges and user-generated content identifying Black-owned businesses in response to the 2020 BLM protests and study the impact of these platform changes on demand for these businesses.

2.2 Literature on Exercise Behaviors and Health Management

First, Chapter 4 relates to the studies of electronic systems in exercise and healthcare, which have been actively studied in the information systems literature. In particular, one of the prominent directions has been gamification systems with game-like features such as competitions or earning rewards (see Liu et al. (2017) for a conceptual framework). The prior literature has also examined the effect of peer effects such as support groups or competitions on health behaviors and outcomes (Bojd et al., 2022; Uetake and Yang, 2020; Jurgens et al., 2015). We extend the literature and analyze the potential of online platforms to provide nudges for healthier behaviors. In particular, we focus on workout programs available to people for free on a popular video hosting platform. Content posted by popular creators on social media, even though increasingly prevalent in the past years, has been relatively understudied in the information systems field. Specifically, many content providers hold challenge programs in which the users are offered a fitness program with clear outlines, easy online access, and potential for interaction with the host and other participants. Note that, although these programs are referred to as challenges, they do not commonly have typical gamified features such as point or leaderboard systems. Since self-paced programs are becoming more prevalent on online platforms, it is important to understand whether they are able to create sufficient levels of user engagement.

Chapter 4 is related to the literature that examines incentives to exercise. A common type

of incentive studied in behavioral economics and sports medicine literature has been the effect of monetary incentives to get people to start and sustain exercising (Charness and Gneezy, 2009; Acland and Levy, 2015; Royer et al., 2015; Strohacker et al., 2014; Beshears et al., 2021; Carrera et al., 2020). In this domain, studies find that monetary rewards help people start exercising and stay more active for weeks after receiving the reward. Conversely, Royer et al. (2015) and Strohacker et al. (2014) note that the effect of these incentives vanishes after the incentives are removed. Apart from the monetary incentives, previous research has studied the effect of non-monetary rewards for attending the gym (Milkman et al., 2014) as well as reminders from wearable devices and gain-loss incentives (Adjerid et al., 2022). As in the papers discussed above, fitness challenges provide incentives for viewers to participate. At the same time, fitness challenges occur in the competitive market structure without the necessity of monetary interventions. Therefore, fitness challenges may be a useful tool to help people increase their exercise activity that does not require additional healthcare policies. Our study complements the previous research by studying potential mechanisms to incentivize people to exercise without policymaker interventions.

In addition, our study relates to the emerging stream of literature studying YouTube and other social media content consumption. A few studies have aimed to understand the performance of individual videos in different contexts (Park et al., 2023; Liu et al., 2019; Kristjorac et al., 2019). To the best of our knowledge, the studies on exercising behaviors via YouTube or other social media platforms have been limited, with some exceptions in other disciplines. Some consumer behavior studies have examined fitness content on YouTube. Sokolova and Perez (2021) conduct surveys and use data collected from 306 respondents to find that watching videos about exercising posted by fitness influencers might help individuals sustain physical activity. Additionally, Kim (2022) conducts surveys and collects data from 379 respondents to find that interaction between the fitness channel and viewer characteristics may explain the viewer's loyalty to the channel. Unlike these studies, we focus on these channels' common strategy of posting fitness challenges in order to improve channel performance.

Furthermore, our study differs from the previous research on YouTube videos, in that we examine the aggregate performance of the channel in addition to the individual video performance. Since creators commonly aim to build a loyal community, our study enriches understanding of the strategies content creators use to build a sustainably performing social media page. In addition, because few quantitative studies have studied exercising with the help of workout videos posted on YouTube, we also contribute to the quantitative estimation of the effect of the strategies employed by the fitness channels.

Now, we extend the existing behavioral theories to examine the effects of fitness challenges on video channel performance and audience activities during and after fitness challenges. We hypothesize that the video channels that host fitness challenges will improve the overall channel performance, which can be measured by an increase in the channel's total number of views, engagements, new followers, and views per follower. Furthermore, we posit that the positive effects of fitness challenges on channel performance would be present not only during the challenge period, but also after the challenge ends. Specifically, we extend the prior literature to posit that the following three potential mechanisms underlie the immediate and persistent effects of fitness challenges: habit formation, communal support, and a feeling of competence.

First, prior studies of exercise incentives use habit as the driving mechanism behind regular exercise (Charness and Gneezy, 2009; Acland and Levy, 2015; Royer et al., 2015; Strohacker et al., 2014; Beshears et al., 2021). This idea is conceptualized as follows: exercising is a type of good with current negative and future positive utility, and the disutility from exercising becomes less over time as the person gets used to regular exercising (Becker and Murphy, 1988). In this light, participants of the fitness challenges should exercise more after completing them, as regular exercise during the challenge forms a habit and reduces the future disutility from the physical activity. Fitness challenges may help establish exercise habits by nudging people to work out regularly for several weeks and make exercise more enjoyable and easier to incorporate into their schedules.

The other potential mechanism explaining the effectiveness of fitness challenges might

be the support coming from the fitness host and other participants during the challenge. Previous research has found that social influence might be an explanatory factor in behaviors (Fulk et al., 1990), including online communities (Butler, 2001; Yan and Tan, 2014). Many channels encourage comment posting during the challenges and talking about participants' experiences and commitments during the challenge. This increased information and support exchange may assist more sustainable exercise behaviors among the viewers participating in fitness challenges.

The third potential mechanism that underlies the positive effect of challenges is that challenges can increase a participant's feeling of competence. Namely, after performing exercises regularly, the users may see improvements in their physical abilities and will be motivated to keep exercising in order to see more progress. According to the Cognitive Evaluation Theory (CET, Deci and Ryan (1985)), intrinsic motivation for behaviors may be driven by autonomy, competence, and relatedness. Studies have found that the perception of competence may drive exercise attendance (Richard et al., 1997; Dyrland and Wininger, 2006). Fitness challenges provide an environment where participants perform the same type of physical activity repeatedly, therefore potentially improving their performance of the suggested exercise and increasing their perception of competence. Besides an increased probability to keep exercising during the challenge, participants might want to progress to more challenging videos posted by the channel host. This may in turn increase the demand for more advanced-level workouts posted on the channel.

The feeling of competence and habit formation can create a potential for keeping up the exercising behaviors. Relatedly, Acland and Levy (2015) find evidence for habit formation when experiment participants are asked to attend gym repeatedly over several weeks. Also, Wilson et al. (2008) find that the sense of confidence helps people to maintain regular exercise. Fitness challenges may serve as a starting point for building exercise habit and building more confidence toward exercising through physical progress and peer support. Together, these mechanisms can lead to increased exercise activity even after the fitness challenge is over and the incentive of a fitness challenge is removed.

Chapter 3

EFFECT OF ONLINE VISIBILITY ON MINORITY BUSINESS PERFORMANCE

3.1 Introduction

One of the most significant social changes in the last few years has been the rise of political consumerism. Consumers increasingly support firms closely aligned with their ethical, political, and moral values. Empirical research on political consumption is nascent, with studies examining the impact of political polarization and brands adopting political stances on consumption (Liaukonytė et al., 2023; Schoenmueller et al., 2023). However, the pivotal role of digital platforms via User Generated Content (UGC) in enabling political consumerism has not yet been studied. In this chapter, we study how UGC, in the form of online reviews, enables the discovery of Black-owned businesses, particularly during polarized discussions about race on the national stage. Although increased awareness of national events can lead to an increase in political consumerism (Liaukonytė et al., 2023; Schoenmueller et al., 2023; Ananthakrishnan et al., 2023), it is not clear *how* consumers find businesses that align with their values when they decide to support these businesses.

Black Americans comprise 14.2% of the U.S. population, but Black-owned businesses only constitute 2.3% of all firms. Studies have highlighted the challenges Black business owners face in owning and operating their own firms (Butler, 2012; Fairlie and Robb, 2007), particularly accessing capital, the lack of generational wealth transfers, and experience running a small business, as well as other opportunities crucial to success in entrepreneurship (Bates and Robb, 2013; García and Darity Jr, 2021; Fairlie and Robb, 2007).

The interest in supporting Black causes ebbs and flows based on the national conversation around the disparities between White and Black communities. According to a recent indus-

try report, over 45% of Americans believe in supporting Black businesses, although Black businesses are vastly underrepresented among traditional retailers (McKinsey, 2022). Retailers are beginning to address the lack of diversity in brands and have started highlighting Black businesses during certain times of the year, such as Black History Month, to allow the discovery of Black-led enterprises¹. However, Black-owned businesses remain severely underrepresented among major retailers and contribute only 1.5% of retail sales. Thus, there exists a significant gap between the consumers' increasing need to support Black-owned businesses and the lack of avenues to discover them through traditional channels.

These limited paths to discover Black-owned businesses through traditional channels have created an opportunity for digital platforms dealing primarily with user-generated content like Yelp to fill the void. Studies in the IS literature have shown how online platforms have reduced information asymmetry and friction (Brynjolfsson and Smith, 2000) and have democratized access to markets for a large segment of society. The importance of online reviews in decreasing uncertainty in product quality and influencing purchasing decisions has been extensively studied in the IS and marketing literature (Dellarocas, 2003; Chatterjee, 2001; Chevalier and Mayzlin, 2006; Forman et al., 2008; Archak et al., 2011; Zervas et al., 2021). Consumers increasingly turn to the Internet to discover the businesses they want to support. For example, in the days following the nationwide Black Lives Matter (BLM) protests in June 2020, there was significant interest in supporting Black-owned businesses. According to Yelp's official blog, the platform experienced more than 35 times more searches for "Black-owned businesses" in the first week of the BLM protests.²

In this chapter, we examine whether online reviews that identify Black-owned businesses on Yelp.com affected the demand for these businesses during increased national support for Black causes. We measure how Black-owned firms, identified as such by reviewers on Yelp, performed compared to other Black-owned businesses not identified as Black-owned on Yelp

¹Retailers are also pledging to increase their retail shelf-space to Black-owned businesses <https://15percentpledge.org/>

²<https://blog.yelp.com/news/yelp-teams-up-with-my-black-receipt-to-support-black-owned-businesses/>

by reviewers. We employ a novel data set of GPS-enabled foot traffic data, hand-collected data from a directory of Black-owned restaurants (as ground truth), restaurant details from Yelp, and zip code-level social capital data to understand the potential mechanisms driving this change. We aim to answer the following questions: *First*, what is the impact of reviews that mention Black ownership for Black-owned businesses? *Second*, how do businesses identified as Black-owned by consumer reviews fare during increased national attention on Black communities? *Third*, what factors contribute to the heterogeneity in the results?

We are primarily focused on user-generated content in disclosing Black ownership. Consequently, we do not include restaurants with voluntary disclosures of being Black-owned in our analysis, especially since identifying the causal impact of voluntary disclosure is challenging without knowing the time of adoption of badges. Aneja et al. (2023) study voluntary disclosure through badges by obtaining proprietary information from Yelp, which provides them with the exact date of adoption. Our analysis studies the causal impact of user-generated content identifying a business owner’s ethnicity, a distinct and different context, and fortunately, this data information is readily available on Yelp.

We leverage three important points that enable us with an identification strategy to tease out the causal impact of UGC on political consumerism. First, searches for “Black-owned business” are a proxy for political consumerism. Consumers are likely motivated by offline events and use online platforms to discover businesses to support. Second, restaurants become visible on Yelp searches for “Black-owned” businesses as soon as they receive reviews that contain these words. Finally, we know the exact time of the first review mentioning “Black-owned business”. Together, we can estimate the time a Black-owned restaurant becomes visible on Black-owned searches and compare the performance of Black-owned restaurants that have at least one review mentioning Black ownership (and thus becoming visible on Black-owned searches on Yelp) with those that don’t have these reviews in a staggered Difference-in-Difference framework.

We find that Black-owned restaurants that receive a review mentioning Black ownership get a 10.8% increase in visits compared to the Black-owned restaurants that don’t have

such reviews. Further, Black restaurants that receive reviews mentioning Black ownership get an additional 4.6% increase in foot traffic during times of increased awareness of Black causes. Our results also suggest significant heterogeneity. We find that the location of the restaurants plays a significant role in moderating the impact of online reviews that mention Black ownership. First, Black-owned restaurants in cities that have lower levels of segregated communities experience a higher increase in foot traffic after they receive reviews mentioning Black ownership. However, Black-owned restaurants in cities with higher levels of segregation experience an increase in foot traffic only during times of increased national interest. Further, we study the moderating role of a restaurant’s location at a micro-geographical level. Black-owned restaurants in White neighborhoods with reviews mentioning Black ownership experience an increase in foot traffic only during times of increased national interest in Black causes. In other words, consumers in predominantly White neighborhoods search for Black-owned restaurants to support during times of increased national interest and patronize restaurants in predominantly White neighborhoods. Thus, our results highlight the segregation of consumption in American cities due to residential segregation and social frictions (Davis et al., 2019; Waldfogel, 2007) even during times of increased political consumerism.

Additionally, we find that the effect is higher in the neighborhoods with higher social capital; these are the geographies with stronger ties between users across income categories. Studies in the social capital literature (Chetty et al., 2022a,b) find economic connectedness as a significant predictor of social mobility, and our results show that economic connectedness plays a major role as an antecedent to political consumerism. Finally, recent studies in the economics literature study interracial interactions at the micro-geographical level (Athey et al., 2021; Massenkoff and Wilmers, 2023) using GPS-enabled location data. We contribute to this literature by leveraging similar data and studying how online reviews disclosing minority ownership affect interracial interactions. Our foot traffic analysis reveals that Black-owned restaurants that receive at least one review mentioning Black ownership not only increase the foot traffic to these restaurants, but the increase bears out through the increase in exposure of non-White patrons to White patrons. This effect is also particularly

high in White neighborhoods. This implies that while significant consumption segregation occurs in American communities (Davis et al., 2019), online reviews not only enable the discovery of Black-owned businesses but also potentially mitigate segregation occurring in everyday activities when consumers choose to engage in political consumerism.

We contribute to two main streams of literature. First, we contribute to the literature studying political consumerism, which addresses how consumers attempt to exert their purchasing power and enact changes in politics (Micheletti et al., 2004; Stolle et al., 2005; De Zúñiga et al., 2014), society (Stolle and Micheletti, 2013; Jacobsen and Dulsrud, 2007) and business practices (Neilson, 2010; Levi and Linton, 2003). More recently, studies in this literature have studied the impact of brands or CEOs of firms adopting political stances on consumption (Liaukonytė et al., 2023; Schoenmueller et al., 2023; Hou and Poliquin, 2023). We complement this literature by focusing on how online reviews enable political consumerism, especially when the businesses in question are not easily discoverable via traditional channels. Additionally, we demonstrate the mechanisms, such as the location of the business, the social capital of the community, and the exposure, that affect these changes.

Second, we contribute to the literature by studying the impact of digital platforms on mitigating racial disparities. Recently, Agarwal and Sen (2022) find that a crowdfunding platform for school teachers can promote more inclusive curricula in schools. Mayya et al. (2020) find the economic gap in forgoing screening on Airbnb platforms and how this aids Black and female hosts. Additionally, (Zhang et al., 2021) find that a smart-pricing feature on the Airbnb platform can help decrease the earnings gap between the Black and the White hosts, conditional on adoption. More recently, Babar et al. (2022) demonstrated the impact of the voluntary disclosure of business owners' ethnicity on consumers through lab experiments. In this chapter, we find strong evidence that consumer reviews play a pivotal role in helping consumers identify Black-owned businesses, even without the voluntary adoption of labels by these businesses or the introduction of unique features by the platform.

Our results indicate the significant positive externality of online reviews in the context of empowering minority-owned businesses, an important and thus far unstudied aspect of

UGC. Online reviews are publicly available information, are generated at an unprecedented scale, and thus can impact change as long as they are available for other users to peruse and search. UGC does not require any special intervention from the platform or the adoption of badges among business owners. This is important because platform-level features designed to improve the visibility of minority-owned businesses still serve only a small percentage of these businesses, due to a lack of adoption among Black-owned businesses and lack of awareness among consumers (Zhang et al., 2021).

3.2 Research Context & Identification Strategy

3.2.1 Black-owned Restaurants

We first consider the different mechanisms through which Black-owned restaurants get increased visibility on Yelp. First, some restaurants have user-generated reviews mentioning Black ownership. Next, we have voluntary adoption of Black-owned badges by some businesses after the introduction of these badges in June 2020. These two mechanisms are distinctly different — the reviewers drive the former in the form of UGC. The platform (as a feature introduction) and the restaurants on the platform (adoption of these badges) drive the latter. Each of these mechanisms confers different avenues of increased digital visibility, as described below. In our analysis, we want to study the impact of UGC in enabling political consumerism. More specifically, we want to measure the impact of online reviews in enabling the discovery process when consumers want to support Black-owned restaurants.

3.2.2 Black-Owned Restaurants on Yelp Search

There are two ways users on Yelp can encounter Black-owned restaurants: generic search and “Black-owned” search. Important differences in how Black-owned restaurants feature in each of these search results are critical to our identification strategy.

Generic Search : We refer to searches that most customers would perform on Yelp as generic searches. For example, we categorize searches such as “Restaurants”, “Takeout”,

“Italian Restaurants” under generic search. Such a search would display all restaurants based on Yelp’s search algorithm (this is a function of the user, location, restaurant type, etc.).

After June 2020, Black-owned restaurants that have voluntarily adopted a Black-owned badge displayed on the search results with a gem icon indicating Black ownership, as shown in the first restaurant in the top half of Figure 3.3. In other words, if a Black-owned restaurant that has adopted a badge were going to be featured in a generic search result, this restaurant would be shown with increased visibility with an icon denoting Black ownership³. Black-owned restaurants that have not voluntarily adopted a badge would be displayed according to Yelp’s search algorithm but featured without a badge. Thus, the increased visibility afforded by the Black-owned badge in a generic search depends on whether a Black-owned restaurant has adopted a badge.

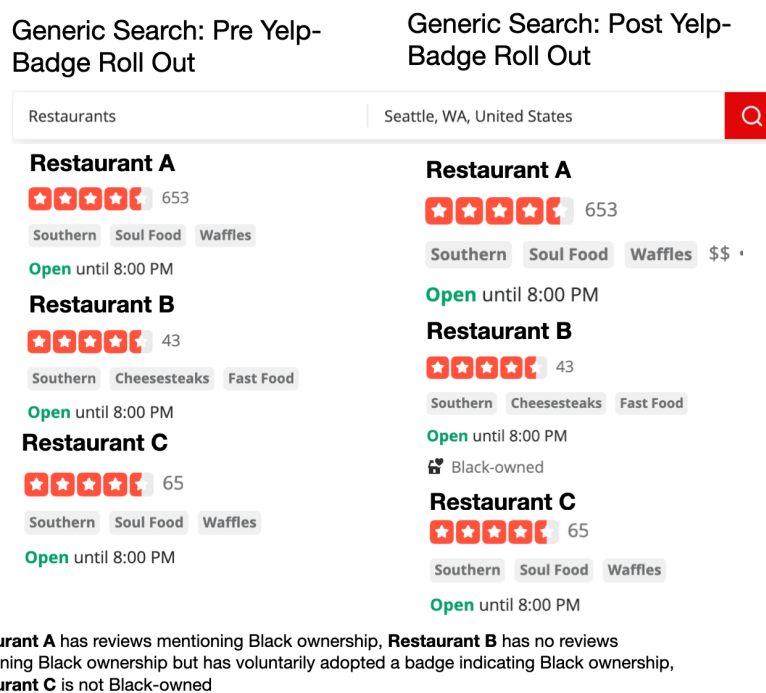
Importantly, in our context, Black-owned restaurants that have not adopted a Black-owned badge but have a review mentioning Black ownership are *not treated* in any unique way in a generic search result (they don’t have any icons to display Black-owned reviews).

Figure 3.1 is a stylized example of changes to the search algorithm when the user performs a generic search (such as “restaurants”). In this example, Restaurant A has reviews mentioning Black ownership, Restaurant B has no reviews mentioning Black ownership but has voluntarily adopted a badge indicating Black ownership, and Restaurant C is not Black-owned. Before the introduction of badges and the changes on Yelp, Restaurants A, B, and C would look similar on the search result page. After the introduction of badges, Restaurant B would have a gem icon indicating Black-owned since it has voluntarily adopted a badge indicating Black ownership. There is no difference between Restaurants A and C since the user did not specifically search for a particular term.

Increase in political consumerism – “Black-owned” search: Next, we consider the searches that are more relevant to our analysis. Consumers conduct specific searches due

³Yelp’s official blog details how Black-ownership (voluntary adoption of badges) became a searchable attribute - “If you choose to opt in, [...] when a user searches for Black-owned businesses, a gem will appear in the search results of your business page highlighting the attribute.”. Yelp does not verify Black ownership.

Figure 3.1: Results For Generic Search



to increased political consumerism. In other words, we have consumers explicitly looking to support Black-owned businesses due to offline events or an increase in national attention to Black causes. It is reasonable to assume that a consumer who searches for a “Black-owned” business does so to discover these businesses. It is crucial to note that such searches are not a rarity. Customers searched for “Black-owned” restaurants in an unprecedented manner in the aftermath of BLM protests. In fact, according to Yelp’s official blog, this search increased by 3600%

“Black-owned” searches are essential in our analysis because of two reasons: First, they allow us to delineate political consumerism clearly. Second, the presence or absence of a restaurant in the results of a Black-owned search depends on whether the restaurant has received at least one review mentioning Black ownership at the time of the search. In other words, “Black-owned” searches help us measure the value of UGC in enabling political consumerism.

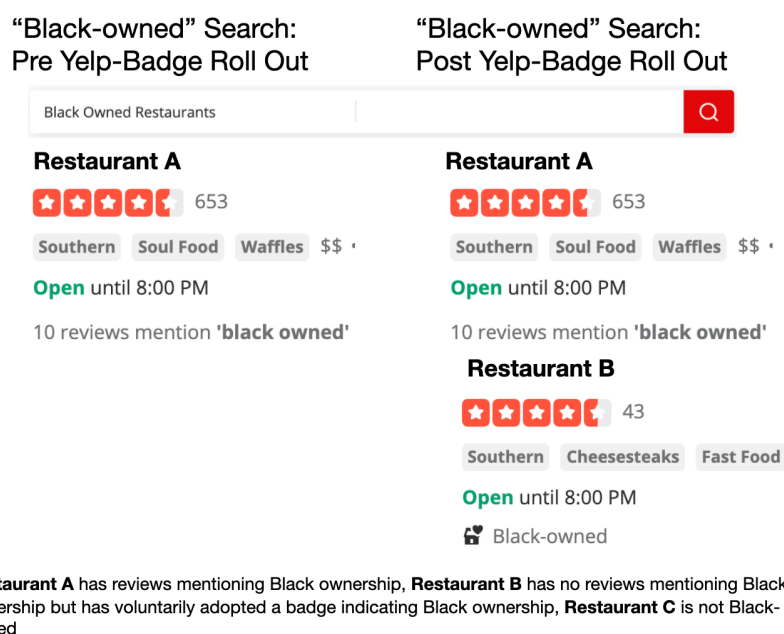
Specifically, when customers searched for Black-owned restaurants, the search results would display all restaurants with at least one review mentioning “Black-owned” as a simple string match. A lot of restaurants obtained reviews mentioning Black ownership much before 2020. Before June 2020, search results for the term “Black-owned business” would only include restaurants with at least one review mentioning “Black-owned” in the review text. After June 2020, search results for the term “Black-owned business” would include both restaurants with at least one mention of “Black-owned” in the reviews and restaurants that have voluntarily adopted a badge *even* if users had not mentioned Black-ownership in the text of their review. Further, during our study period, Yelp’s search results mentioned why a restaurant was displayed when users searched for “Black-owned”. For example, the search would say - restaurant A has ten reviews mentioning “Black-owned” or the search would have a gem icon indicating that restaurant B has adopted a badge as shown in the restaurant in Figure 3.3.

Therefore, Black-owned restaurants with at least one review mentioning Black-ownership would appear in Black-owned searches if they had at least one review mentioning Black ownership at the time of search, during the entire period of our analysis from January 2019 to June 2021. Black-owned restaurants that have voluntarily adopted a badge declaring Black-ownership and *not* have any Black-owned review would appear in Black-owned searches only after they adopt the badge. This feature was released after June 2020. Note that only about 27 percent of all the Black-owned restaurants in our sample had voluntarily adopted a badge even by September 2021, so the majority of the restaurants still became visible to users who searched for Black-owned restaurants only through reviews that mentioned minority ownership. In fact, these restaurants with reviews were the only ones displayed on the “Black-owned” search before June 2020. We isolate restaurants that only have user-generated reviews mentioning Black ownership in our analysis and discuss more in our identification section below.

Figure 3.2 is a stylized example of changes to the search algorithm when the user performs a “Black-owned” search. As before, Restaurant A has reviews mentioning Black ownership,

Restaurant B has no reviews mentioning Black ownership but has voluntarily adopted a badge indicating Black ownership, and Restaurant C is not Black-owned. Before the introduction of badges and the changes on Yelp, a search for “Black-owned” restaurant would only display Restaurant A with an explanation of why Restaurant A was being displayed. Restaurants B and C would not be displayed in the Black-owned search since they don’t have any reviews mentioning Black ownership. After the introduction of badges, the voluntary Black-owned badges became a searchable attribute, placing them in the “Black-owned” searches along with Restaurant A. Restaurant C would not be displayed since they do not have a review mentioning Black ownership nor have they voluntarily adopted a badge.

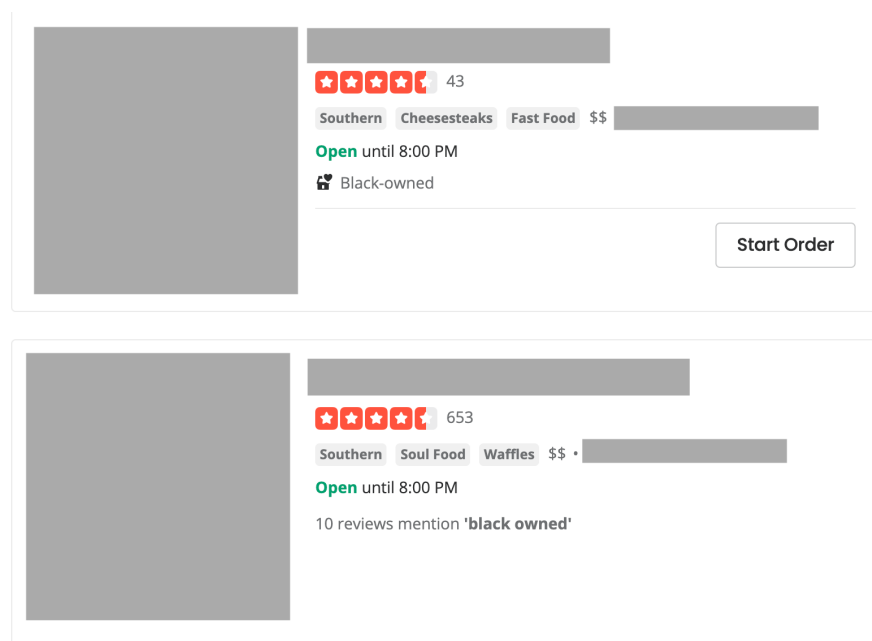
Figure 3.2: Results For Black-Owned Search



3.2.3 Identification Strategy

How do we measure the impact of coordinating offline movement with online reviews that promote digital visibility? In other words, how do we measure the impact of user-generated content in enabling political consumerism? In our analysis, we choose to study the digital

Figure 3.3: Difference between restaurants with reviews mentioning Black ownership and restaurants that have voluntarily adopted Black-owned badges in the search interface for Black-owned searches

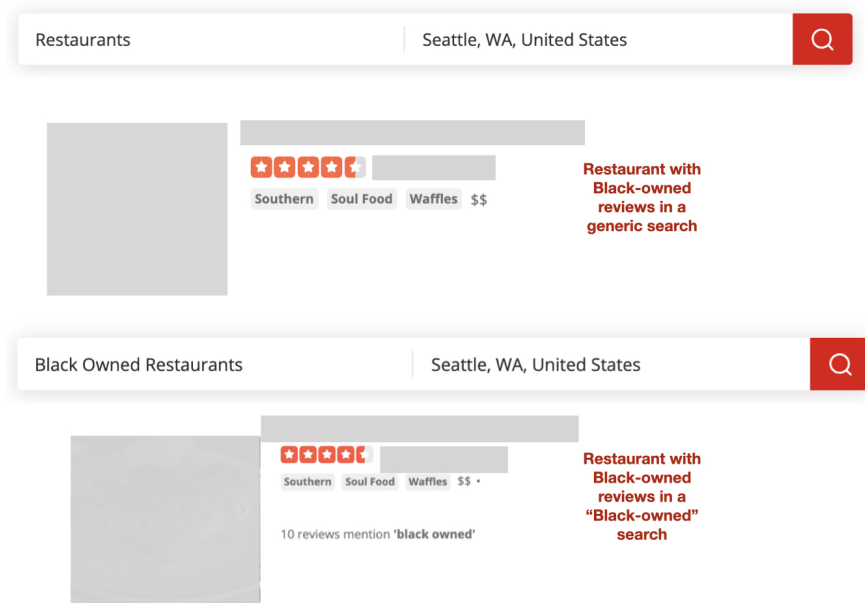


visibility afforded by *reviews* when users search to support Black-owned businesses. When a consumer searches for a specific term on Yelp, Yelp’s search algorithm displays restaurants with at least one review containing the term that the consumer searches for that term. Restaurants have no control over whether and when consumers mention “Black-owned” in their reviews. Different Black-owned restaurants obtain reviews mentioning Black ownership at different times, if at all, which determines their presence in the “Black-owned” search results. This provides us with a staggered DiD framework.

Black-owned restaurants only with reviews in the generic search result without any information about being Black-owned and would look the same way as any other restaurant on the search result but would be highlighted as Black-owned in the “Black-owned” search as shown in Figure 3.4. Therefore, having at least one review mentioning Black-ownership would ensure a position in the “Black-owned” search results (searches proxying for political

consumerism).

Figure 3.4: Black-owned restaurants with reviews mentioning Black-ownership and no voluntary badges appearing in generic and Black-owned searches



Some nuances are worth clarifying. When users search for “Black-owned” restaurants, search results contain Black-owned restaurants that obtain at least one review mentioning Black ownership and those that have adopted a “Black-owned” badge (if the search happened after June 2020 after Yelp introduced these features and after the voluntary adoption of the badges by a restaurant). Our analysis only focuses on political consumerism and how consumers choosing to support Black-owned businesses discover these businesses through online reviews. We do not consider accidental encounters in generic search. Hence, we remove the restaurants that have voluntarily adopted badges for four reasons: First, such restaurants could get increased digital visibility on searches through the Black gem icon unrelated to offline political movements (generic search) and could confound our results. Second, these restaurants could be part of both generic and “Black-owned” searches, and

since the time of adoption is unknown, we would not know when they became part of the “Black-owned” search result. Third, Black-owned restaurants adopting badges could self-select and be very different from other Black-owned restaurants. Fourth, badges require restaurant owners to adopt these badges. Technology adoption and increased visibility can increase revenue (Ghose and Yang, 2009; Ghose et al., 2014; Agarwal et al., 2011). However, technology adoption among minority business owners is low, even if technology adoption reduces the revenue gap between Black and White business owners (Zhang et al., 2021). The number of such restaurants is less than 27% even during our data collection in September 2021. Other studies estimate this number as around 10% on all of Yelp one year after the introduction of these badges (Aneja et al., 2023). Fortunately, user-generated content does not require technical savvy among minority business owners or any other particular interventions from the platform.

This setting isolates political consumerism (with Black-owned searches as the proxy). Further, our modeling choice enables us to infer the precise time a restaurant would be featured in a Black-owned search based on the time of the first review mentioning Black ownership. Further, our decision to only consider Black-owned searches and the Black-owned restaurants with and without reviews mentioning Black-ownership would enable us to keep the treated and the control group among Black-owned restaurants.

3.3 Data

To conduct our empirical analysis, we use data from ten large US cities (Los Angeles, Chicago, Houston, Philadelphia, Dallas, Seattle, Boston, Washington, D.C., Denver, and Portland) and scraped restaurant-level information from multiple sources.

3.3.1 List of Black-Owned Restaurants on Yelp

Since we focus on political consumerism and the search for Black-owned restaurants, we query for “Black-owned restaurant” in each city neighborhood in our analysis and combine the search results to obtain the list of all Black-owned restaurants in each city. As men-

tioned earlier, search results display Black-owned restaurants with at least one mention of “Black-owned” in the reviews or have voluntarily adopted a Black-owned badge. Further, we obtain all observable restaurant-level characteristics for each restaurant in our database that restaurant owners can claim. In addition, we also obtain Yelp reviews for the Black-owned restaurants in our data set. We collected all this information from Yelp in September 2021, and all the restaurant-level characteristics remain time-invariant for our analysis.

3.3.2 Directory of Black-Owned Restaurants

The list of Black-owned restaurants on Yelp is not comprehensive. We require a list of all Black-owned restaurants in the cities in our sample to identify ground truth to measure the value of digital visibility afforded by online reviews - this is especially important since technology adoption is relatively low among the Black business community (Zhang et al., 2021).⁴ Other studies in this literature address the issue of ground truth by performing image analytics on business owners’ images (Zhang et al., 2021). However, Yelp restaurants rarely have photos of business owners, making this issue even more difficult. Fortunately, we obtain this information from the website “EatOkra”⁵. Currently, EatOkra lists 9,000+ Black-owned businesses across the US and is considered one of the largest directories of black restaurants.⁶ We compare the list of restaurants in EatOkra and the restaurants listed on Yelp and find that a substantial number of restaurants (335 - around 33%) that are Black-owned are not displayed as Black-owned on Yelp (they neither have user-generated reviews nor have claimed a “Black-owned” badge on Yelp). We then collect all information available

⁴We contacted city officials and Black-owned business community leaders in all the cities in our sample to obtain this information and discovered that such a comprehensive list was unavailable

⁵www.eatokra.com. EatOkra initially used various online resources to list Black-owned restaurants. After the BLM protests in 2020, the company received significant media attention and support from businesses such as UberEats, Apple, and Pepsi.

⁶We manually verified the information on EatOkra for over 50 restaurants listed on the website. We found them listed as Black-owned in our Yelp data or as Black-owned in the independent user-generated available on the Internet.

on Yelp for Black-owned restaurants that are not identified as such on Yelp ⁷

After this step, we have 1012 Black-owned restaurants in all the cities. Of these 1012, 677 were displayed as Black-owned on Yelp. Among these, 274 had the Black-owned badge on their page, and the remaining 403 were shown in “Black-owned” search as Black-owned based on user reviews. Our analysis focuses primarily on the value of UGC that describes the ethnicity of the business owner. Therefore, we do not consider restaurants that have adopted the Black-owned badges. This leaves us with 738 Black-owned restaurants with reviews mentioning Black ownership or no indication that they were Black-owned on Yelp. Table 3.1 provides the number of restaurants in each category.

3.3.3 *Reviews Mentioning Black Ownership*

The period for our analysis is between January 2019 and April 2021. We collect user-generated reviews for all restaurants labeled as “Black-owned” on Yelp. When restaurants receive their first review mentioning Black ownership, they are featured in the Black-owned searches and are considered treated. There is a significant variation in the time restaurants obtain their first review mentioning Black ownership. This is the source of variation that we exploit in our DiD. We consider a restaurant to be treated once it receives its first “Black-Owned” review. We define our treatment variable as $AfterBlackOwnedReview_{it}$, 1 for all months after the restaurant receives the first review mentioning Black ownership. All restaurants that received a Black-owned review before January 2019 are considered to be treated at the beginning of our study period.

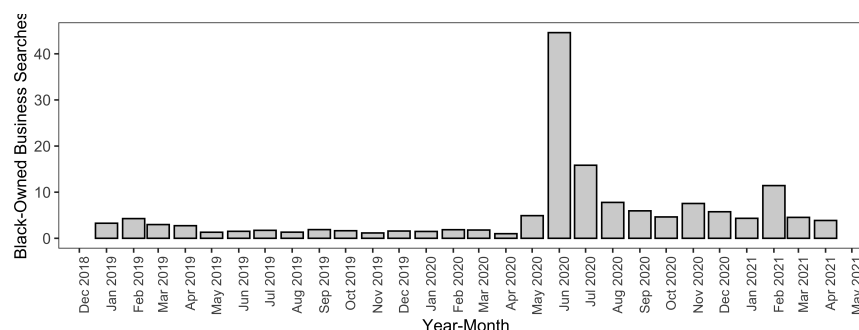
3.3.4 *Increase in demand to support Black-owned businesses*

There was a drastic increase in national interest in supporting Black-owned businesses after the nationwide BLM protests. Figure 3.5 shows the Google Trends data for the term “Black-owned business.” In this figure, we observe a dramatic increase in the searches for Black-

⁷We discuss more about EatOkra in the Appendix 9

owned businesses in the first months after the nationwide BLM protests. Yelp’s official blog cited this meteoric rise in searches for Black-owned businesses for the release of a feature that would enable businesses to disclose ethnic ownership voluntarily.⁸

Figure 3.5: Google Trends for “Black-Owned Business”



The vertical axis represents the normalized number of Google searches for the term “Black-owned business.” Normalization is done with respect to the lowest number of searches in the observed periods. Source: Google Trends

In our analysis, we investigate the effects of online reviews that mention Black ownership, particularly in times of heightened consumer interest in supporting Black businesses. We define the periods of heightened support of Black businesses after the nationwide BLM protests, as apparent in the Google searches data. Additionally, we also include Black History Month (February each year) as a time of high interest in supporting Black causes (we observe a small spike in February 2019 and 2021 in the search results, which is explained by Black History Month). In particular, we define a binary variable *BlackInterest*, which is 1 during Black History Month in February 2019, 2020, 2021, and during nationwide BLM protests in June and June 2020.

⁸<https://blog.yelp.com/news/laying-the-groundwork-for-change/>, <https://blog.yelp.com/news/yelp-teams-up-with-my-black-receipt-to-support-black-owned-businesses/>

3.3.5 *SafeGraph Data*

We obtain our outcome variable, $\text{Log}(\text{Visits})$, from SafeGraph. SafeGraph provides foot traffic data for over 3.6 million points of interest (POIs) throughout the United States. According to the firm’s documentation, SafeGraph covers over 45 million mobile devices around the United States.⁹ Users of these devices provide permission to track their location via different mobile apps. SafeGraph then matches the users’ location with various POIs and aggregates the visits at the POI-month level. The users’ traffic data is anonymized and aggregated at the census block level. This provides a unique peek into different retailers’ traffic patterns and consumers’ behavioral patterns across the United States. IS, Economics, and Marketing researchers have used SafeGraph data to study various social phenomena (Chiou and Tucker, 2020; Athey et al., 2018)

SafeGraph’s directory of POIs provides extensive information about each POI, including descriptive details such as the store’s name, brands, address, geo-coordinates, and census block group. The variable of interest in the SafeGraph data is the monthly number of visitors to the establishments of interest. We match the restaurants in our Yelp data with those in the SafeGraph data and obtain the foot traffic for all these restaurants from January 1, 2019, to June 30, 2021¹⁰. SafeGraph data provides us with the census tract where restaurants are located and the census tract level foot traffic data. We provide the distribution of foot traffic across the different types of restaurants in Table 3.1.

3.3.6 *Social Capital Atlas*

Social capital refers to the strength of a person’s ties to their community. We are interested in investigating social capital as a potential driving mechanism for political consumerism (Stolle and Micheletti, 2013; Neilson, 2010; Neilson and Paxton, 2010). Measuring social capital and interactions at the community level is challenging, and this information is unavailable

⁹<https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset>

¹⁰Note that for the regression analysis, we exclude all the observations for March and April 2020, because of the business lockdowns due to the COVID-19 pandemic.

Table 3.1: Summary Statistics of Monthly Visits

	Visits				Number of Restaurants
	Mean	Median	SD	N	
RestaurantType					
BlackOwnedReviews	178.6	132	171.8	10,462	403
BlackOwnedNotIdentified	150.5	107	152.8	8,700	335

in census data. Most of the studies in this stream of literature employ surveys or aggregate measures at the country level to quantify social capital.

However, more recently, Chetty et al. (2022a) use data from 21 billion friendships on Facebook and investigate the associations between social capital and economic mobility. In particular, they find that economic connectedness, the extent of friendships across high and low socio-economic status in a given community, positively affects upward mobility. We leverage this data made available to researchers through the Social Capital Atlas ¹¹ (Chetty et al., 2022a,b) and obtain zip code-level social capital characteristics. Next, we match this data with the restaurant addresses and split the restaurants into restaurants in low and high social capital areas (social capital measured as economic connectedness).

3.3.7 Exposure, Segregation Measures & Socioeconomic Data

We obtain demographic data at the census tract (CT) and zip-code levels from the American Community Survey (ACS). This data also enables us to measure the levels of residential segregation. Specifically, Massey and Denton (1988) in their seminal work created various measures of segregation in American communities. This includes measures of evenness (the degree to which neighborhoods within a city differ in terms of race) and “exposure” (the

¹¹<https://socialcapital.org>

degree of contact or interaction between members of different races). These measures are widely used in economics research (Athey et al., 2021; Gentzkow and Shapiro, 2011) to study the impact of societal integration on various outcomes. We mainly use the dissimilarity - a measure of the minority population that needs to change in a specific neighborhood so that the neighborhood has the same percentage of minority population as the larger city. This varies at the city level and takes a value between 0 to 1, with 0 being the least segregated city and 1 being the most segregated city¹². We also calculate the isolation index at the city level. We obtained all these measures using the 2020 census data at the city level from the Diversities and Disparities project (Bischoff and Reardon, 2014) — a joint collaboration between Russell Sage Foundation and Brown University.

We delve into the exposure metrics at the restaurant level. In a recent study, Massenkoff and Wilmers (2023) use SafeGraph data to measure segregation and social isolation in daily activities between members of different race. We define exposure or the likelihood of a chance encounter between people of different races at each restaurant using the procedure from Massenkoff and Wilmers (2023).¹³ We find that, on average, 43% of the visitors to Black-owned restaurants are from White neighborhoods from January 2019 to June 2021. Further, 55% of the visitors are White for Black-owned restaurants in White neighborhoods, and 31% of the visitors are White for Black-owned restaurants in Non-White neighborhoods. We use the exposure at the restaurant level in our analysis as a dependent variable to explain the share of interracial encounters at each Black-owned restaurant before and after the reception of reviews mentioning Black ownership. Finally, we use SafeGraph’s foot traffic data for each POI aggregated at the CBG level. We combine this data with the American Community Survey (ACS) data and obtain the racial and income distribution for each CBG (Davis et al., 2019; Athey et al., 2021).

¹²<https://www.census.gov/topics/housing/housing-patterns/guidance/appendix-b.html>

¹³We explain in detail our calculation in Appendix 8

3.4 Empirical Analysis and Results

In this section, we estimate the impact of online reviews that identify Black-owned restaurants on the restaurant’s performance, particularly during increased political consumerism. Finally, we explore the heterogeneity and mechanisms that drive our results.

We aim to estimate two primary parameters of interest: 1) the effect of online reviews mentioning Black ownership on foot traffic to Black-owned restaurants compared to other Black-owned restaurants that have not received these reviews, and 2) the moderating impact of increased interest in supporting Black causes due to offline movements.

We estimate these parameters using the DiD model as shown in our main specification in Model 4.1. To ensure that the treated and the control groups are comparable, we only consider Black-owned restaurants in our primary analysis (those identified by reviews and those not identified as Black-owned on Yelp). Further, we do not consider restaurants that have voluntarily adopted a “Black-owned” badge since our study focuses on the impact of UGC on political consumerism and not the voluntary adoption of badges by businesses, as explained in detail in the previous section.¹⁴

$$\begin{aligned} \ln(Visits_{it}) = & \beta_0 + \beta_1 AfterBlackOwnedReview_{it} \\ & + \beta_2 AfterBlackOwnedReview_{it} \times BlackInterest_t + \alpha_i + \gamma_t + \mu_{ct} + \epsilon_{it} \end{aligned} \quad (3.1)$$

$AfterBlackOwnedReview_{it}$ is an indicator variable equal to 1 if the restaurant had at least one review indicating Black ownership at month t . Having at least one review mentioning Black ownership would ensure that the restaurant would be featured in Black-owned search. $BlackInterest_t$ is a dummy variable that equals 1 at times of increased interest in Black causes, as defined in Section 3. β_1 is the DiD estimator that captures the average impact of reviews mentioning “Black-owned” business on foot traffic to Black-owned restaurants. β_2 is the coefficient that corresponds to the impact of online reviews on the

¹⁴Note that some restaurants with “Black-owned” badges have reviews that mention Black ownership. By excluding all restaurants that voluntarily adopted the badges, our estimate of the impact of receiving online reviews mentioning Black ownership is conservative.

performance of Black restaurants at times of increased interest in Black causes. In addition, we control for the time-invariant restaurant-level characteristics using restaurant fixed effects α_i . We also account for time shocks common to all the restaurants using year-month fixed effects γ_m . To account for the city-level time trends impacting restaurant performance, we include μ_{ct} , a city-specific time fixed effect.

3.4.1 Results

We first estimate the effect of receiving a review mentioning Black ownership on Yelp search compared with other Black-owned restaurants without these reviews. In other words, we estimate Model 4.1 without the interaction term $AfterBlackOwnedReview_{it} \times BlackInterest_t$ and display the results in column 1 of Table 3.2. We observe the coefficient corresponding to $AfterBlackOwnedReview_{it}$ is positive and significant. In other words, online reviews provide an 11.7% increase in visits to restaurants that receive at least one review mentioning Black ownership compared to Black-owned restaurants without such reviews.

Next, we add the interaction term corresponding to increased interest in supporting Black causes as shown in specification 4.1 and present the results in Column 2 of Table 3.2. This result demonstrates that, in addition to the 10.8% increase in visits after getting a Black-owned review on Yelp, these reviews provide additional foot traffic of 4.6% during times of increased interest in supporting Black causes (such as Black History Month and a few months around BLM). Together, our results demonstrate the role of UGC in enabling political consumerism.

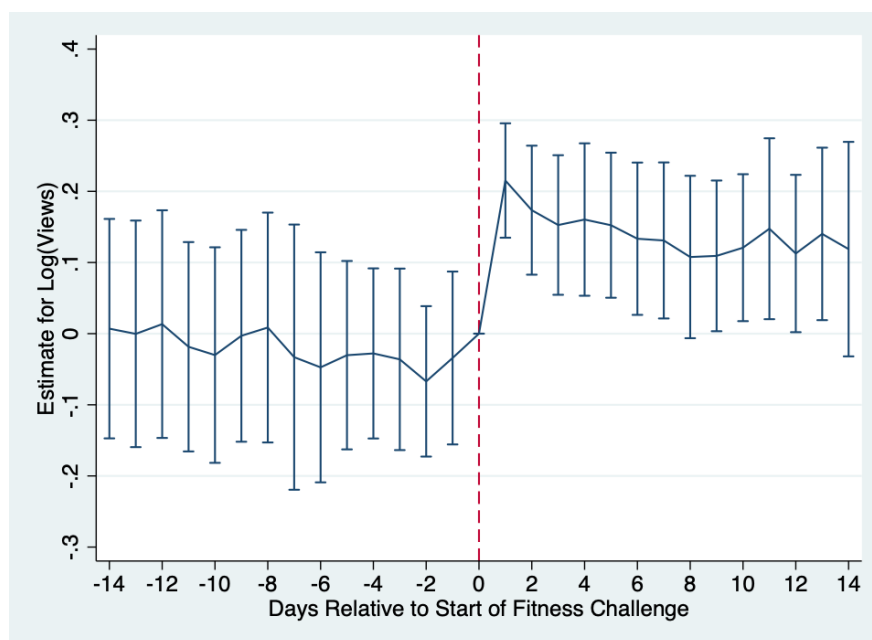
We verify the parallel trends assumption using a Relative Time Model (Autor, 2003; Agarwal and Sen, 2022). To do so, we modify 4.1 to a full set of leads and lags dummies.

$$LogVisits_{ct} = \sum_{j=-6}^6 \beta_j Review_{ct}(t = k + j) + \lambda_c + \gamma_t + \epsilon_{ct}. \quad (3.2)$$

We estimate the treatment coefficients for six periods before getting the first black-owned review and present them in Figure 3.6. In the equation, each coefficient β_j ($j =$

$-6, -5, \dots, 5, 6$) represents the treatment coefficient for the treated units j months before or after the first review. All the point estimates to the left of the dashed line in Figure 3.6 present the foot traffic to restaurants before they received reviews mentioning Black ownership, compared to Black-owned restaurants with no such reviews. We find that these estimates are statistically insignificant, verifying the parallel trends assumption. All the point estimates after the restaurant receives its first Black-owned review are to the right of the dashed line.

Figure 3.6: Event Study Analysis



The figure shows relative time model estimates using the TWFE estimator with log of visits as the dependent variable. The estimates are shown with a 95% confidence intervals.

In this analysis, our control group consists of restaurants that are not yet treated (the restaurant did not receive a review mentioning “Black-owned” at time t) and restaurants that are never treated (Black-owned restaurants that don’t get any reviews mentioning the minority ownership at any point in our analysis). We run the same analysis as a robustness check, removing restaurants that never get treated. We present this analysis in Appendix 7

and, reassuringly, find very similar results.

Table 3.2: Effect of Reviews Mentioning Black-Ownership on Visits to Black-Owned Restaurants

	(1)	(2)
	LogVisits	LogVisits
AfterBlackOwnedReview=1	0.117*** (0.0347)	0.108*** (0.0345)
AfterBlackOwnedReview=1 × BlackInterest=1		0.0457*** (0.0130)
Observations	19162	19162
R ²	0.866	0.866

Note: The dependent variable is log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.* p<0.10, ** p<0.05, *** p<0.01

Overall, this analysis demonstrates how the increased digital visibility provided by online reviews for Black-owned businesses leads to improved performance for minority businesses, especially during increased national interest in race-related issues. Given this setting, our results demonstrate the power of the crowd’s wisdom. Online reviews are generated each day at an unprecedented scale. Digital platforms and consumers use the information from these reviews to identify and support businesses that closely align with their ethical values, even without any intervention from the business or the platform. This is particularly important given the low rate of technology adoption among minority business owners even though technology adoption increases revenue (Mayya et al., 2020; Zhang et al., 2021). In

fact, platform-level features designed to ameliorate bias among minority groups often do not work or exacerbate bias because they require adoption among the minority groups (Zhang et al., 2021). Our results demonstrate that online reviews serve as a powerful tool in enabling political consumerism — when consumers want to support Black-owned businesses, they can utilize information from other consumers to discover Black-owned businesses even without any intervention or voluntary disclosure from the business. As long as a digital platform collects and categorizes information from UGC, consumers can readily leverage this information when they choose to support these businesses.

3.4.2 Robustness Checks

In this section, we present several robustness checks to verify the validity of our main results and rule out other explanations.

Matching

Unobserved differences may exist between restaurants that receive reviews mentioning Black ownership and those that don't. To test this possibility, we verify our main specification using coarsened exact matching (CEM) (Blackwell et al., 2009; Pattabhiramaiah et al., 2021). We match Black-owned restaurants with similar characteristics and find consistent results, as shown in Table A.1 in Appendix 1.

Counterfactual Estimates (Synthetic Control)

Next, we apply a Counterfactual Estimates (CE) (Liu et al., 2021; Pattabhiramaiah et al., 2021) method for treatment effect estimations under weaker assumptions than DiD. The CE method extends the DiD framework and allows for unobservable time-varying confounders with parametric restrictions on these unobservable components. Accounting for unobservable

components eliminates some estimation concerns common to the DiD method, such as pre-trends and other unobserved channel-level differences. We present the results in Table A.2 of the Appendix 2 and find that our results are consistent.

Additional Robustness Checks

We conduct a series of additional robustness checks to rule out alternate explanations. First, we verify whether our results are valid if we change the definition of foot traffic by considering the time spent in the restaurant (Li and Wang, 2020) and find consistent results as shown in Appendix 4. Second, we consider the possibility that there might be a higher number of reviews mentioning Black ownership for restaurants with high demand. This might be correlated with the outcome variable. However, we consider only the time of the first review mentioning Black-ownership as our treatment. While more reviews are associated with higher demand, we are only concerned with the timing of the first review mentioning Black ownership, which typically happens much later than the first review. We find a significant difference in days (mean of 1356 days and a median of 791 days) between the first review and the first Black review mentioning Black ownership. Therefore, the treatment (the first time a restaurant receives a review mentioning Black ownership) is not directly correlated with the restaurant’s popularity. We also find that the number of reviews mentioning Black ownership is considerably smaller than other reviews — on average, restaurants that get at least one review mentioning Black ownership have less than 3% of the reviews that mention Black ownership. Third, it might be possible that restaurants receive more reviews mentioning Black ownership during the time of increased national causes. However, we are only concerned with the timing of the first review, which determines their appearance on “Black-owned” searches. In fact, we find that most restaurants received their first Black-owned review not during increased national attention. We remove all the restaurants that received their first review mentioning Black ownership during the heightened interest in Black causes and re-run our analysis. Our results are consistent, as shown in Appendix 5. This demon-

strates that the restaurants treated during increased national interests are not driving our results. Fourth, we study the potential role of offline cues that might increase visits to Black-owned restaurants. We remove restaurants with labels, posters, or other visual cues that indicate Black ownership (offline indicators voluntarily adopted) from our sample and again find consistent results as shown in Appendix 6. Finally, we conduct a falsification analysis to verify the robustness of our results, also discussed in detail in Appendix 6.

3.5 Heterogeneous Effects

In this section, we study the heterogeneous effects that can affect the impact of online reviews that disclose minority ownership.

3.5.1 Restaurant Location

First, we study the role of a restaurant’s location in driving foot traffic. We study the effects of location at the city and community levels. Prior studies have found that the performance of local businesses is correlated with community characteristics such as the neighborhood’s average income, education, and racial composition, irrespective of the business owner’s race (Farrell et al., 2019; Perry et al., 2020).

First, we study the housing segregation metrics at the city level. Massey and Denton (1988) compiled over twenty different similarity and exposure metrics widely used in the Economics literature to analyze the nature of cities and urban segregation. For example, the dissimilarity index measures the distribution of any racial group across a city. Cities with high values of dissimilarity (over 60) are highly segregated. Studies in sociology consider values of 40 or more to be moderate segregation. In other words, the higher the dissimilarity index, the more concentrated the members of one racial group in a particular census tract and are not distributed equally in the city. The median dissimilarity is 36 for cities in our sample, indicating that cities have high and low segregation. We split our sample into cities with higher-than-median and lower-than-median dissimilarities. We then estimate the model in 4.1 and present the results in Columns (1) and (2) of Table 3.3.

Interestingly, the coefficient corresponding to $AfterBlackOwnedReview_{it}$ is insignificant in cities that experience higher levels of neighborhood segregation. However, the coefficient corresponding to $BlackInterest_t$ is significant and positive, indicating a higher interest in supporting Black-owned businesses during increased interest in Black-owned causes. This is not the case for cities with lesser residential segregation. We find that both the coefficients $AfterBlackOwnedReview_{it}$ and $BlackInterest_t$ are significant and positive, indicating that Black-owned restaurants with reviews mentioning Black ownership in these cities experience higher patronage at all times compared to those restaurants that don't have these reviews.

We repeat the same analysis with the isolation index — the percentage of the same-group population in the areas where an average member of a particular racial group resides in Table 3.3. We compute these estimates for above-median and below-median Black-Black and White-White isolation. The results are consistent with the dissimilarity index. Black-owned restaurants located in cities that have higher interracial isolation experience an increase in traffic only during times of increased national interest after they receive a review mentioning Black ownership.

While city-level measures are informative about the mechanisms at a high level, much of the everyday encounters happen at a neighborhood level. We investigate whether the demographics of the census tract where the restaurant is located moderate the effect of increased digital visibility on foot traffic. Using ACS data, we obtain the census tract (CT) where each restaurant in our data is located and classify the restaurant's neighborhood as a Black or a non-Black neighborhood. We run Specification (4.1) and display the results in Table 3.4. Column (1) refers to restaurants in non-Black (White) neighborhoods, and Column (2) refers to restaurants in Black neighborhoods.

We find that restaurants in Black-majority and non-Black-majority neighborhoods benefit from reviews mentioning Black ownership after receiving these reviews. Specifically, restaurants in non-Black (Black) areas gain 15.4% (11.2%) more visits compared to other Black-owned restaurants without reviews disclosing ethnic ownership. However, during increased interest in supporting minority groups, only restaurants in non-Black neighborhoods

enjoy additional foot traffic. In other words, Black-owned restaurants in non-Black neighborhoods experience a more significant increase in foot traffic during increased political consumerism.

The consumption segregation theory (Davis et al., 2019) explains our results. Davis et al. (2019) show how urban consumption is greatly influenced by residential segregation and social frictions. More specifically, Davis et al. (2019) find that spatial frictions play a significant role in consumption, and consumers are less likely to visit restaurants far away from their homes or work due to increased transit times. Other studies have documented location preferences in consumption (Shi et al., 2023). Davis et al. (2019) also find that social friction exists and that consumers are more likely to visit restaurants in neighborhoods with large populations of their race. Our result indicates that consumption segregation exists even in the presence of political consumerism. Black-owned restaurants that receive at least one review mentioning Black ownership in White neighborhoods experience an increase in foot traffic only during times of increased national interest. In other words, consumers in predominantly White neighborhoods discover Black-owned businesses due to the increased visibility afforded by reviews mentioning Black-ownership, but patronize those Black-owned businesses in White neighborhoods. This aligns with the antecedents of consumption segregation – spatial segregation and social friction.

3.5.2 Social Capital

Studies predict that areas with higher social capital, such as volunteering in organizations, can facilitate networking and political participation (Copeland and Boulianne, 2022; Giugni et al., 2017). This can, in turn, spur collective action among community members to use their purchasing power to influence corporate practices (Stolle and Micheletti, 2013; Neilson, 2010; Neilson and Paxton, 2010). We use the definitions of areas with high and low social capital (Chetty et al., 2022a,b) as described in the data section, estimate our main specification for areas with high and low social capital, and present our results in Table 3.4.

We find a significant difference between the impact of online reviews in high- and low-

Table 3.3: Impact of City Level Segregation on Political Consumerism

	(1)	(2)	(3)	(4)
	HiDissimilarity	LowDissimilarity	HiIsolationWhite	LowIsolationWhite
AfterBlackOwnedReview=1	0.0516 (0.0843)	0.126*** (0.0365)	0.0374 (0.0805)	0.132*** (0.0372)
AfterBlackOwnedReview=1 × BlackInterest=1	0.0803*** (0.0263)	0.0357** (0.0149)	0.0536** (0.0221)	0.0439*** (0.0155)
Observations	5225	13937	5174	13988
R ²	0.862	0.864	0.851	0.870

Note:

The dependent variable is log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.* p<0.10, ** p<0.05, *** p<0.01

social capital areas. Specifically, in neighborhoods with higher economic connectedness, Black-owned restaurants experience a 13% increase in foot traffic due to online reviews, and the impact of online reviews on foot traffic is 5% more during periods of increased political consumerism. However, we don't observe any effect due to reviews mentioning Black ownership in areas with low economic connectedness with weak ties between individuals of high and low socio-economic status.

This result is consistent with the studies in the political consumerism research – areas with higher economic connectedness are more sensitive to the calls for social equality and, therefore, support Black-owned businesses identified as such by other consumers, particularly during periods of increased attention to Black causes (Copeland and Boulianne, 2022; Stolle and Micheletti, 2013). Higher economic connectedness implies that consumers are more aware of issues around inequality because they have social ties with those who might be affected by such disparities. These consumers are more open to supporting minority-owned businesses.

Table 3.4: Effect of Yelp Reviews by Neighborhood Characteristics

	(1)	(2)	(3)	(4)
	Non-Black	Black	LowEC	HighEC
AfterBlackOwnedReview	0.0963*	0.120***	0.0537	0.130**
	(0.0569)	(0.0407)	(0.0493)	(0.0536)
AfterBlackOwnedReview \times BlackInterest	0.0595***	0.0199	0.00797	0.0509**
	(0.0197)	(0.0170)	(0.0175)	(0.0234)
Observations	9879	9283	8746	7608
R ²	0.870	0.867	0.876	0.857

Note: The dependent variable is log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.5.3 *Exposure of Patrons to Black-Owned Restaurants to Patrons of Other Racial Groups*

A nascent literature in economics examines isolation and interaction in public spaces. Studies in this literature argue how people spend their leisure time could impact the segregation they face in their everyday lives based on their socioeconomic status (Massenkoff and Wilmers, 2023; Athey et al., 2021). While traditional measures in sociology use information about where people live to construct rigid definitions of segregation (Massey and Denton, 1988; Cutler and Glaeser, 1997), recent studies in this literature use GPS-enabled location tracking data to quantify exposure or interracial contact in the locations individuals frequent in their everyday life (Athey et al., 2021; Zhang et al., 2019; Phillips et al., 2021). These measures can be highly informative in our context. Our main results demonstrate how UGC enables political consumerism — Black-owned restaurants experience increased visits after receiving a review indicating Black ownership. Our secondary results suggest that the restaurant’s location plays a vital role in moderating this effect. However, it is unclear if online reviews disclosing minority ownership change how people of different races intermingle. In other words, does the nature of exposure and intermingling among patrons change when Black-owned restaurants receive reviews mentioning minority ownership?

Our secondary analysis reveals that Black-owned restaurants in White neighborhoods receive increased foot traffic after they receive a review mentioning Black ownership during times of increased national interest. While a simple analysis of the nature of the foot traffic could demonstrate the increase or decrease in patronage by members of different races, a more nuanced approach would be to measure how online reviews change the nature of interracial interactions in these restaurants. This also informs the discussion around consumption segregation — do White patrons frequent Black-owned restaurants that are already popular among White patrons? Or do White patrons discover Black-owned restaurants usually frequented by non-White consumers?

The challenge in conducting this analysis would be reliable data. Fortunately, recent studies in this literature leverage census block level information from Safegraph to calculate

interracial interactions (Massenkoff and Wilmers, 2023; Athey et al., 2021). This measure is termed exposure, and Massenkoff and Wilmers (2023) use SafeGraph foot traffic data at the census tract level to document segregation in American life, not just in residential spaces but also in public and commercial spaces. The higher the exposure, the greater the probability of chance encounters for individuals of a particular race with other races. The average exposure of Non-White individuals to White individuals in Black-owned restaurants is 57%. Interestingly, the average exposure varies in tandem with the city-level dissimilarity metrics. For example, the average Non-White to White exposure in Black-owned restaurants is only 32% in cities with higher levels of segregation (dissimilarity > sample median) compared to the average Non-White to White exposure of 62% at Black-owned restaurants in cities that have lower levels of segregation.

To study if online reviews disclosing minority ownership change the nature of exposure in restaurants, we modify the specification in (4.1) by changing the dependent variable to *ExposureNonWhiteWhite*. We further interact *NonBlackNeighborhood* with the neighborhood with *AfterBlackOwnedReview* to study the impact of a restaurant’s location on the exposure of its patrons after the reception of Black-owned reviews.

$$\begin{aligned}
 \textit{ExposureNonWhiteWhite} = & \beta_0 + \beta_1 \textit{AfterBlackOwnedReview}_{it} \\
 & + \beta_2 \textit{AfterBlackOwnedReview}_{it} \times \textit{BlackNeighborhood}_i \quad (3.3) \\
 & + \alpha_i + \gamma_t + \mu_{ct} + \epsilon_{it}
 \end{aligned}$$

Here, *ExposureNonWhiteWhite* refers to the likelihood of an encounter between a non-White resident and a White resident at a restaurant. *BlackNeighborhood_i* refers to the restaurant’s location and can take the value 1 or 0 depending on whether the location is Black or White. All the other fixed effects remain the same. We display the results in Table 3.5. Column (1) shows that the exposure of non-white patrons to patrons increases significantly for Black-owned restaurants after they receive reviews indicating Black ownership. Column (2) shows that the location of the restaurant moderates the increase in exposure. Restaurants

in Non-Black neighborhoods see a significantly higher exposure increase than restaurants in Black neighborhoods.

Our main results and the results in Table 3.5 indicate online reviews disclosing minority ownership aid the discovery of Black-owned restaurants and increase interracial intermingling. More specifically, the increase in foot traffic is due to a higher probability of accidental encounters between Non-White and White patrons in Black-owned restaurants. This suggests that Black-owned restaurants that were previously visited by non-White consumers are now seeing increasing levels of White users after they receive reviews mentioning Black ownership. These results are in line with anecdotal evidence. Unlike the civil rights movement of the 1960s, the majority of protesters in 95% of the participating counties were White (Buchanan et al., 2020). Media reports that there was increasing interest among White consumers to support Black businesses in the immediate aftermath of BLM protests (Williams, 2021; Horowitz, 2021). We find the same effect and demonstrate that the White population demonstrated the most prominent behavioral change due to external events.

Table 3.5: Effect on Exposure

	(1)	(2)
	ExposureNonBlackBlack	ExposureNonBlackBlack
AfterBlackOwnedReview	0.0187*** (0.00711)	0.0374*** (0.0103)
AfterBlackOwnedReview × BlackNeighborhood		-0.0378*** (0.0131)
Observations	17095	17095
R ²	0.811	0.811

Note:

The dependent variable is exposure. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month, and city-year-month (time trends at the city level) fixed effects.* p<0.10, ** p<0.05, *** p<0.01

3.6 Implications and Conclusion

In this chapter, we study the effect of online reviews disclosing minority-owned businesses and how the heightened national conversation around race can increase political consumerism.

Our results have many important implications. First, we find strong evidence that online reviews enable political consumerism. Consumers tend to support businesses closely aligned with their values. Minority-owned businesses are often overlooked and are under-represented in traditional channels of discovery. User-generated content can significantly increase the discovery of minority-owned businesses, and ours is the first study demonstrating the value of online reviews in enabling political consumerism. We find that the restaurants identified as Black-owned by reviews on Yelp receive about an 11% increase in foot traffic compared to Black-owned restaurants that do not have these reviews. Further, we observe that the impact of these reviews is much more salient during times of increased national attention to Black causes, with an additional increase of about 4.6% during increased interest in Black causes such as Black History Month or nationwide BLM protests.

Second, the restaurants' location is essential in determining their ability to attract politically conscious consumers. Residential segregation appears to impact political consumerism, which aligns with the consumption segregation theory. Reviews disclosing Black ownership have the most significant impact on restaurants located in cities with lower residential segregation. Restaurants in cities with higher residential segregation experience the benefits of a review mentioning Black ownership only during increased interest in Black causes. At the micro-geography level, Black-owned restaurants in White census tracts observe increased foot traffic after receiving reviews disclosing Black ownership during increased interest in Black-causes. Thus, consumers in predominantly White census tracts display political consumerism and visit Black-owned restaurants in their neighborhoods. We also observe a consistent result for Black-owned restaurants located in areas with higher social capital in the form of higher economic connectedness with stronger ties across income lines. Studies in the sociology literature term this as the "Advantaged Group Allies" (Hong and Peoples,

2021; Droogendyk et al., 2016). In other words, the majority group can make a meaningful impact when they understand they can empathize with the members of the minority group. Our results demonstrate how Advantaged Group Allies drove the increase in revenue to Black-owned establishments, with reviews enabling the discovery of these restaurants.

Third, we observe how online reviews can mitigate consumption segregation. Our results suggest that the exposure or the likelihood of non-White patrons encountering White patrons at the Black-owned restaurants changed after these restaurants received reviews mentioning Black ownership. The restaurants' location also drives this effect — consumers in White neighborhoods are more likely to visit Black-owned restaurants frequented by non-White patrons after these restaurants obtain reviews mentioning Black ownership.

While our analysis focuses on Black-owned businesses, we address the broader question of how technology platforms can enable political consumerism. Our work speaks to the need and the benefits of higher digital visibility for minority-owned businesses. Clearly, media awareness around supporting minority-owned businesses drives up consumer interest. In fact, our results show how Black History Month and BLM protests increase political consumerism, particularly among consumers in predominantly White neighborhoods who, in turn, patronize Black-owned restaurants that they can discover on Yelp through consumer reviews. One could make the same argument for Veterans Day, Pride Month, etc., and how veteran-owned or LGBTQ+-owned businesses could attain higher visibility during times of increased interest in supporting them.

However, technology adoption remains low across minority-owned businesses. Even though Yelp introduced a voluntary self-disclosure label, only about 10% of the Black-owned restaurants adopted badges one year after the launch of these badges (Aneja et al., 2023). Low adoption among minority owners is a significant challenge, even if such features can be beneficial *conditional* on adoption (Mayya et al., 2020). For example, Zhang et al. (2021) demonstrate how smart-pricing algorithms on Airbnb widen the gap between Black and White hosts because of low adoption among Black hosts. There is much discussion about building socially inclusive products to help underrepresented minorities on digital platforms.

Socially inclusive digital products are essential in helping underrepresented minorities overcome systemic biases. However, the role of digital platforms does not stop at implementing these features. Underrepresented minorities realize the full benefit of technological capabilities only when platforms solve the last-mile problem in technology adoption, which is particularly low among minority-owned business owners.

When interest in supporting any minority group increases, minority business owners with any form of increased digital visibility benefit tremendously. Badges and other platform-level features are introduced in response to the increase in demand to support minority-owned businesses, which take time for adoption, and by the time business owners adopt these badges, the interest in supporting these businesses might have waned. This is where user-generated content comes into play. Online reviews are real-time, and large-scale, and have provided tremendous benefits to consumers and suppliers alike in reducing information asymmetry. When users of a digital platform search to support a minority-owned business, they are more likely to encounter those businesses identified as such by other users of the platform.

In addition to these efforts that require business owners to identify as minority-owned businesses voluntarily, technology platforms should generate the vast resources at their disposal through user-generated content and foster the growth of this collective knowledge database. Where adoption fails, UGC can fill the gap. For example, 55% of Black-owned restaurants highlighted in Yelp's search results as Black-owned feature on the Black-owned search queries because of user-generated content mentioning Black ownership. Digital platforms should incentivize users to engage and provide altruistic motivation for information about minority business owners to narrow the digital visibility gap that minority business owners experience.

The role of UGC in providing digital visibility to minority business owners cannot be understated. In our analysis, we interviewed and interacted with the Black-business leaders in the cities to obtain a potential list of Black-owned businesses. Interestingly, we found no unified approach to identifying Black businesses, except only via user-generated lists on

forums and social media ¹⁵. Therefore, digital platforms should leverage the wisdom of the crowd and foster the growth of UGC in their quest to address the issues stemming from digital divide.

Our analysis is not without potential weaknesses. We only observe an increase in foot traffic without accounting for growth in online delivery. Further, our analysis is conditional on the user's search for "Black-owned" business. Finally, we do not study the impact of digital visibility on online ratings and reviews. These are interesting directions for future research.

¹⁵<https://blog.yelp.com/news/yelp-teams-up-with-my-black-receipt-to-support-black-owned-businesses>

Chapter 4

EFFECT OF ONLINE FITNESS CHALLENGES ON USER' EXERCISING BEHAVIOR

4.1 *Introduction*

Despite the proven health benefits of exercising, several studies consistently find the lack of physical activity among the population, as well as failure to exercise consistently by 80 percent of Americans (Piercy et al., 2018). The development of the digital economy has generated new tools to promote exercise, including, among others, websites, social media content, and mobile apps. The online fitness industry has become increasingly attractive to consumers and is projected to reach \$79 billion by 2026.¹ Likewise, exercise with YouTube follow-along workouts has been on the rise, too. For example, in August 2022, YouTube hosted 714 channels in English that posted videos about physical fitness, with a subscriber count of over 100,000 followers.²

Fitness workouts posted on online platforms such as apps and social media attract viewers for several reasons. First, viewers do not have to depend on a gym schedule; second, they can follow these workouts from any location. Additionally, online platforms provide access to a wide variety of videos and channels, where users can search for exercise type, length, and instructions that fit their needs most. Finally, unlike gym memberships, many of the online workouts are available for free—for example, through YouTube, Instagram, Reddit, and mobile apps.³ Nevertheless, as our data suggests, 67% of users who left a comment

¹<https://www.globenewswire.com/news-release/2022/03/22/2407884/0/en/The-Online-Virtual-Fitness-Market-Is-Expected-To-Reach-79-Billion-By-2026-With-The-Rising-Penetration-Of-Smart-Devices-As-Per-The-Business-Research-Company-s-Online-Virtual-Fitness.html>

²Source: Tubular Labs

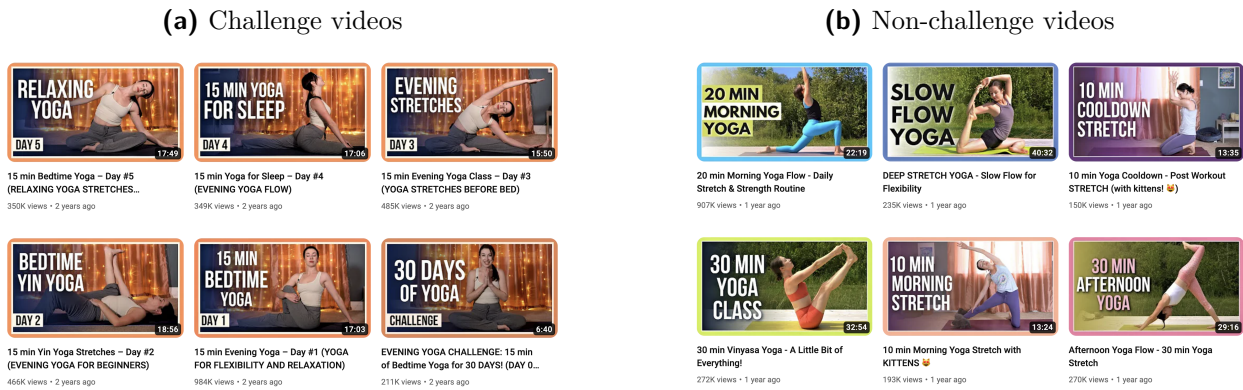
³<https://www.nytimes.com/wirecutter/blog/free-home-workouts/>

under a workout video on YouTube never commented on that channel again, suggesting that many viewers of online workout videos fail to exercise consistently with the same instructor. Therefore, it is important to understand what strategies for audience engagement are effective at helping users exercise consistently. In this chapter, we explore the effects of fitness challenges—a common strategy that fitness creators incorporate in order to improve their performance.

A fitness challenge is typically an exercise program that has a defined start and end date, provides a set of workouts to follow, and sometimes includes additional tools such as nutrition programs or community support. YouTube channels usually upload a brand-new workout video every day during a fitness challenge. Recently, an increasing number of YouTube creators have been hosting such challenges on their channels, potentially hoping to increase their audience loyalty and increase the number of views and followers. In this wellness market, it is particularly important to understand what creator strategies help increase channel performance and increase followers' loyalty, because more active viewership of workout videos is not only profitable for the creator, but also beneficial to the customers' health. In this chapter, we demonstrate that fitness challenges are a valuable strategy for fitness content creators because they help them to gain more views and engagements, attract new followers, and help followers watch more videos (and, therefore, exercise) in the future.

Figure 4.1A shows an example of a fitness challenge posted on YouTube with daily uploads of new workouts. Figure 1A presents a screenshot of one of the YouTube channels in our sample with uploads during a fitness challenge. Figure 1B shows typically uploaded videos on the same channel outside of the fitness challenge frame; Figure 1B shows a variety of different yoga workouts without a specific theme. In Figure 1A, challenge videos have a similar thumbnail and are marked in the same way. Specifically, the last video in the bottom row is a challenge announcement (marked as “challenge” at the left bottom corner of the thumbnail), and all the following videos are marked as a specific day of the challenge (day 1, day 2, etc.) In this chapter, we examine the effect of hosting such fitness challenges on the channels' performance.

Figure 4.1: Example of Fitness Challenge on YouTube



On the one hand, fitness challenges might improve channel performance. First, challenges are usually posted relatively infrequently and differ from the rest of the content posted by the channel, which might signal as “novel” to the followers and therefore be more attractive (Hirschman, 1980). Second, fitness challenges are frequently complemented with additional support from the coach such as daily emails and check-ins with the followers. Additionally, we find that the followers’ commenting activity also increases. These factors might provide additional support and motivation for the participants that has been shown to increase exercise participation (Rackow et al., 2015). Third, fitness challenges provide an exercise schedule for the participants and therefore remove the task of picking the workouts. Studies have found that the process of choice-making leads to decision fatigue and impairs self-control (Chernev et al., 2015; Baumeister, 2018).

On the other hand, not all channel viewers might want to commit to a fixed-timeframe workout program. The viewers might also be weary of following a specific instructor regularly, and after completing the challenge, they might switch to different channels. This phenomenon has been studied as the satiation effect in the consumer behavior literature (Sevilla et al., 2019). Additionally, the fitness challenges suggest that participants watch a specific video on a given day. Some people might prefer to choose their own workouts instead of following a fixed schedule. Finally, fixed workouts might not work for users with

certain physical restrictions or users with varying levels of physical conditions (e.g., workouts might be too hard for beginners and too easy for more advanced users.) Therefore, whether the effect of challenges on the channels' and viewers' activity is positive or negative is an empirical question.

We address the following three primary research questions. First, we examine whether fitness challenges improve the hosting channel performance. Second, we focus on the driving mechanisms and test whether users mention habit formation, physical fitness progress, and community support more often during and after the fitness challenges. Finally, we study whether the fitness challenges increase individual users' activity on the hosting channel.

To do so, we collect a detailed posting activity and performance dataset incorporating 21 popular yoga channels that publish follow-along workouts on YouTube that, in total, hosted 37 challenges during the period of our study. We obtain channel-day data on the aggregate channel views, engagements, and new followers, as well as all the comments for all the channels. We study the effect of the challenges on the total daily views, engagements, followers, and views per follower. Moreover, we scrape all the comments under the videos from these channels and, using the detailed comments data, we estimate the effect of the fitness challenges on the individual users' propensity to leave comments on the channel.

Finally, it is important to understand what behavioral phenomena make fitness challenges successful among participants. In order to explore this, we extend potential mechanisms previously studied in relation to exercise. More specifically, we propose that habit formation, community support, and physical performance improvements coming from a period of regular exercising might be driving the effectiveness of challenges. One mechanism, habit formation, has been frequently studied in the behavioral economics literature (Becker and Murphy, 1988) in the context of exercise behaviors. In addition, we hypothesize that community support might bolster consistent exercise because of the coach's support and because many users are following the program simultaneously. Finally, we test whether the physical progress from a program completion might motivate users to continue exercising. We utilize the textual content of user comments as well as their unique IDs and timestamps to analyze the dynamics

of mentions of these mechanisms before, during, and after fitness challenges.

We apply the Difference-in-Differences framework to estimate the effects of fitness challenges. Additionally, in order to ensure the robustness of our findings, we use the novel Counterfactual Estimates (Liu et al., 2021) framework that relaxes some of the Difference-in-Differences assumptions. We find that during the challenge, the numbers of daily views, daily engagements, new followers per day, and views per follower that the channel receives increase by 22.8, 42.9, 51.2, and 3.2 percent, correspondingly. We also demonstrate the presence of carryover effects of challenges on channel performance. Namely, we find that the positive effect of the challenges persists at least as long as the length of the challenge after it is over.⁴ Even after the challenge, the views, engagements, new followers, and views per follower of challenge-hosting channels remain up by 14.3, 18.6, 68, and 2.6 percent, respectively. Further, we study the content of the individual comments and conclude that the three proposed mechanisms—habit formation, community support, and physical performance—might be driving the effects. These findings together propose that the fitness challenges are an effective tool to both grow a YouTube channel and keep the viewers exercising more often. Finally, we find that individual viewers are more likely to engage with the channel during and after the challenge.

Our contributions to the literature are as follows. First, we study the causal effects of a creator’s strategy on the channel’s subsequent performance. Despite the burgeoning influencer economy, few studies (Zhao et al., 2022), have quantified the effects of strategies utilized by popular accounts on UGC platforms. Furthermore, we contribute to the health and wellness literature by studying the effects of tools that can help people exercise more consistently. Although online exercise platforms and fitness challenges are growing in popularity, their effects have been understudied. We utilize a unique dataset that contains text user feedback data that allows us to explore the mechanisms that help users exercise more frequently. Our findings are beneficial to content creators who want their channels to per-

⁴For example, if the challenge was 30 days long, we find the presence of carryover effects for 30 days after the end of the challenge.

form better, to sponsors that may want to support content creators, policymakers and health practitioners who may look for new ways to keep the population physically fit, and customers who may want to exercise more consistently.

The remainder of the chapter is organized as follows. We introduce the empirical context and the data in Section 4.2. Further, in Section 4.3 we present our estimation strategy and the results. We discuss our findings and conclude in Section 4.4.

4.2 Empirical Setting and Data

In this section, we describe our data sources and data collection process. We study the case of YouTube, the largest video hosting platform—some of its popular fitness channels have over 1.2 billion views.⁵ In our analysis, we first estimate the effects of fitness challenges on aggregate channel performance. After that, we conduct a more detailed analysis and estimate the effect of challenges on individual video performance and viewer activity. To study these effects, we collect and construct datasets on (1) channel performance and (2) individual user activity.

4.2.1 Channel Performance Data

We obtain the data from two sources: Tubular Labs, an analytics platform that tracks performance of social media video content creators,⁶ and Google YouTube Data API⁷, which allows the collection of publicly available data from YouTube. To ensure channel homogeneity, we focus on only the channels that post yoga workouts. Yoga has been one of the popular types of physical exercise, well suitable for at-home practice.⁸ For the channel set selection, we select the channels that produce content in the English language, with a follower count above 200,000 at the moment of the data collection, and whose content falls into the cate-

⁵<https://www.youtube.com/@yogawithadriene/about>

⁶<https://servicesdirectory.withyoutube.com/directory/tubular-labs/>

⁷<https://developers.google.com/youtube/v3>

⁸<https://disturbmenot.co/yoga-statistics/>

gory “Health, Fitness, and Self-help.”⁹ Additionally, we manually check that the channels on our list receive most of their views from their follow-along workout videos. Such filtering is important since we aim to estimate the effect of fitness challenges on the exercise activities of their followers. The resulting data cover 21 channels with a total of over 10,000 videos posted on their pages. Next, we collect information on whether and when these channels hosted fitness challenges. For that, we check the content posted by the channels (i.e., videos and community posts) as well as their websites and other social media pages to determine whether they posted information about challenges. We find that 11 channels in our sample have had at least one challenge. We track each challenge’s length, start date, and end date. We identify 37 challenges of varying lengths (from 5 to 30 days) across all the channels. We list all the channels used in our analysis, their follower count, and fitness challenge frequency in Web Appendix A.

After the channel selection, we collect the performance data for each channel. Tubular Labs provides detailed daily performance metrics for YouTube channels, such as the total daily views received by all videos posted in the channel, daily engagements (i.e., the combined number of likes and comments), the channel’s new followers, and the number of videos the channel uploaded on a given day. We obtain the daily channel performance data from January 2019 to January 2022. In the video performance literature, daily performance data are challenging to observe, but our data provide a unique opportunity to examine the performance of channels at a granular daily level. Furthermore, although the previous literature has often employed video-level metrics (Park et al., 2023; Krijestorac et al., 2019; Zhao et al., 2022), we examine channel-level daily performance metrics because a channel’s daily performance is one of the most important metrics to content creators¹⁰ and is more representative of the total consumption of the fitness content by a channel’s audience.

By using the information about the daily views of the channel and the daily number of followers for each channel, we additionally introduce the metric for views per follower as the

⁹The category information is extracted from the Tubular Labs intelligence tools.

¹⁰<https://databox.com/youtube-kpis>

ratio between the number of views on a given day and the number of followers on that same day: $ViewsPerFollower_{ct} = Views_{ct}/Followers_{ct}$, where c is channel identifier, and t is the day. This metric is commonly used among practitioners as an indicator of user engagement.¹¹ The views per follower metric estimates the average activity of the channel’s followers, and in the context of the exercise videos, it can also allow us to estimate whether an average follower of the channel manages to exercise more frequently.

4.2.2 User Comments Dataset

In addition to assessing the effect of fitness challenges on the hosting channel’s performance, we are interested in estimating the effects on individual user behaviors. Since the data about individual users’ viewing activity are unavailable to the channel hosts and researchers, we utilize a rich dataset of user comments. For that, we collect user comments posted on each of the channels. Customer comments are a valuable source of information about a product’s performance, frequently used in the literature (Yang et al., 2019). Comment content analysis allows us deeper understanding of the performance, in addition to tracking performance metrics (Zhang and Moe, 2021).

On YouTube, the audience frequently leaves comments under videos to share their thoughts and interact with other audience members and content creators. We use YouTube Data API to collect the comment information under all the videos posted by each channel on our list. Along with collecting a comment’s text, we compile the data on the comment timestamp, and, importantly, the unique ID of the commenter who wrote the comment. Commenters’ IDs allow us to track and examine returning commenters in the dataset. In total, we collected 1,601,981 comments written by 610,019 unique users.

We utilize user comment data for several purposes. First, the timestamp of the comments and unique commenter ID information let us estimate the effect of fitness challenges on an individual viewer’s propensity to return to the channel. Additionally, we use the comment

¹¹<https://tubularlabs.com/blog/3-metrics-youtube-success/>

content to explore the potential underlying mechanisms driving the effect. In the next section, we describe the text processing methods that we utilize to extract the mechanism information from the unstructured comment text.

4.2.3 Comment Labeling

To analyze the changes in comment content during and after the challenges, we identify the comments that capture (1) users' formation of habit to exercise, (2) community support, and (3) progress in physical performance. As a baseline for comparing the occurrence of these comments, we also identify the comments that compliment that specific video.

The importance of UGC analytics has been frequently highlighted in the information systems and marketing literature (Archak et al., 2011; Netzer et al., 2012). However, most of this work has been done in the context of online reviews with relatively homogeneous text structure (Yang et al., 2019). Analysis of social media comments such as those on YouTube and other platforms has been more challenging due to shorter text, context dependence, and a vast number of potential topics.

In the studies of social media, labeling comments has also been a difficult task. For some widely used and popular tasks such as hate speech and misinformation recognition, high accuracy has been achieved through the collective efforts of many researchers, the creation of topic-specific dictionaries, and the manual work of social media moderators (Bruns and Stieglitz, 2012). For less studied topics, many researchers rely on manual labels for comments (Yang et al., 2019). Since we use over 1.6 million user comments in our analysis, manual labeling is not feasible. Therefore, we use a keyword search methodology to label comments (MacAvaney et al., 2019). We select at random a subset of 3,000 comments and manually label them for the presence of each mechanism. Next, we split the labeled comments into train (2,000 comments) and test (1,000 comments) sets. We use the train set to identify keywords for automated labeling of the rest of the comments. In Table 1, we show examples of phrases indicating the presence of the mechanisms. We list the full set of phrases used for labeling in Web Appendix B.

Table 4.1: Comment Labeling Examples

Categories	Definitions	Example Comments	Keyphrase Examples
Habit	Indicates regularity of exercising	I cannot express enough how grateful I am for your classes, particularly in these times! I look forward to every class every morning :-) Thank you so much, lots of love and gratitude!	"daily routine", "exercise every day", "stick to"
Community Mentions	support from other followers or the coach	Hi Tim! I also will love to keep up with the challenge for the quarantine it brings a sense of community and its great to practice with you. I really appreciate your work and sharing it with us. Namaste!	"practice together", "support", "community"
Progress	Indicates improvement in physical performance	As always I enjoyed your class, my body has become so much stronger, thank you so much!!!!	"was able to do", "improvement", "feel stronger"
Praise	Compliment to the workout video	This was such a wonderful Monday morning practice <3 wow. Thank you!	"great workout", "really enjoyed the workout", "love this video"

After our labeling procedure, we get 70,643 comments labeled as “habit” (4.5% of the total comments), 52,692 comments labeled as “community” (3.3%), 38,211 comments labeled as “progress” (2.4%,) and 221,624 comments labeled as “praise” (14%.) We estimate the performance of the labeling procedure by calculating common metrics, such as precision, recall, and F_1 score, on the test set. We report these metrics in Table 4.2. We see that the F_1 score for Community, Progress, and Praise is around 70-75%. Habit comments are harder to predict, potentially due to the variety of potential phrases used for indicating habit, as well as evidence of lack of habit (negation) or plans to develop an exercise habit rather than stating current regular exercise. These performance magnitudes are common for sentence classification tasks (Timoshenko and Hauser, 2019; Song et al., 2021).

Table 4.2: Labeling Performance

Label	Precision (%)	Recall (%)	F_1 (%)
Habit	54.2	64	58.7
Community	69.7	75.7	73
Progress	60.9	80.6	69.4
Praise	80.5	72	76

4.2.4 Summary Statistics

After collecting information on the daily channel performance and user comments, we construct the following three datasets. First, we utilize the information about daily channel-wide views, engagements, new followers, and views per follower to estimate the effect of fitness challenges on channel performance during and after the challenges. We present the summary statistics in Panel A of Table 3.

Next, we study the behavioral changes of the channel viewers during and after the challenges. To do so, we construct a channel-day-level dataset on the number of viewers’ comments on each of the proposed mechanisms of interest: habit formation, communal support,

and feeling of competence. Panel B of Table 3 shows the summary statistics for the comment label counts.

Finally, we additionally examine the effect of challenges on individual users' activities by examining users' commenting behaviors. Specifically, we utilize the fact that YouTube makes it possible to track individual users via their unique ids and create a panel dataset with the variable of interest $Comment_{ict}$, which equals 1 if user i left a comment on channel c on date t , and 0 otherwise. Summary statistics for this dataset are presented in Panel C of Table 3.

4.3 Empirical Strategy and Results

In this section, we present our estimates for the effect of fitness challenges on the aggregate daily performance of the channel, the analysis of mechanisms driving the effects, and the commenting activity of individual users. In addition, we perform a series of robustness checks for our findings.

4.3.1 Identification Strategy

We aim to estimate the effect of the fitness challenges on different performance metrics at the aggregate channel level and individual YouTube user level. For this, we define the binary treatment variable $Challenge_{ct}$ as follows. It equals 1 if channel c on day t hosted a challenge, and 0 otherwise.

In addition to determining the treatment effect of fitness challenges, we are interested in examining whether the treatment effect lasts after the challenges are over. It is important to ascertain the long-term effects of challenges from the perspectives of the channels, because channels, which have invested time and resources into creating fitness challenges, need to understand the long-term implications of these challenges. Furthermore, from the perspective of the users, we find it imperative to understand whether they continue to exercise regularly even after the challenges are over. We define $PostChallenge_{ct}$ to be equal to 1 if channel c offered a fitness challenge recently, as of point of time t , and 0 otherwise. More specifically,

Table 4.3: Summary Statistics

Panel A: Daily Aggregate Channel Performance					
Variable	Obs	Mean	Std. Dev.	Min	Max
Daily Views	21,743	66,520.61	163,327.6	0	2,159,543
Daily Engagements	21,743	1,189.127	2,594.805	0	57,793
New Followers	21,743	613.573	2,660.818	-7176	100,000
Views per Follower	21,743	.147	.469	0	32.6

Panel B: User Comment Labels					
Variable	Obs	Mean	Std. Dev.	Min	Max
Total Comments	21,743	53.842	138.896	0	5359
Habit Comments	21,743	2.375	8.07	0	368
Community Comments	21,743	1.929	13.495	0	815
Progress Comments	21,743	1.21	4.538	0	186
Praise Comments	21,743	7.659	16.224	0	487

Panel C: Individual Users					
Variable	Obs	Mean	Std. Dev.	Min	Max
Comment	108,562,833	.007	.083	0	1

for each challenge, we mark the following n days as Post Challenge after the end of the challenge, where n equals the length in days of that specific challenge. For example, if a channel hosted a 30-day fitness challenge between April 1 and April 30, 2021, then the variable *PostChallenge* will be equal to 1 for the next 30 days after the end of the challenge; that is, May 1-May 30, 2021 will be labeled as Post Challenge days on this channel. In order to estimate the treatment effect of challenges, we apply several identification strategies in

our empirical analysis.

First, we employ the popular difference-in-differences (DiD) framework (Fredriksson and Oliveira, 2019; Lechner et al., 2011). This framework allows for estimating the causal effects of non-random treatments (in our case, fitness challenges) under several assumptions. We compare the performance of the channels that hosted fitness challenges (treatment group) at a given time with the ones that did not (control group). The DiD method is a reliable estimate in cases where the parallel trends assumption is reasonable, given that no other processes happened during the treatment time that could affect the outcomes. In order to check the validity of this assumption, we perform the formal parallel trends check in the periods prior to the treatment and confirm similar behavior of the treated and control channels prior to the fitness challenges. This analysis is presented in more detail in Section 4.6.

Second, we apply a novel Counterfactual Estimates method (Liu et al., 2021; Xu, 2017) that alleviates some of the concerns traditionally connected with DiD estimations. This method, closely related to generalized synthetic controls Xu (2017), uses re-weighting of the control group observations in order to approximate the dynamics of the treated units and predict counterfactual values for the treated units in absence of treatment. This method relaxes the DiD assumptions by allowing for unobserved time trends for each observational unit. With this, the estimation does not have to rely on the parallel trends assumption. We employ this method and demonstrate that the estimates are similar to the ones from the DiD method.

4.3.2 *Effect of Challenges on Channel Performance*

First, we apply the DiD framework described above to the analysis of aggregate-level channel performance during and after the challenges. For this, we estimate the effect of the fitness challenges and the presence of carryover effects using the following regression equation:

$$y_{ct} = \beta_0 + \beta_1 \text{Challenge}_{ct} + \beta_2 \text{PostChallenge}_{ct} + \alpha_c + \gamma_t + \epsilon_{ct} \quad (4.1)$$

where y_{ct} is a performance variable such as views, engagements, new followers, or views per follower of channel c on day t . As defined above, $Challenge_{ct}$ and $PostChallenge_{ct}$ are dummy variables for the treatment and carryover effects of the fitness challenges. We also control for channel- and year-month-, and day-of-week-fixed effects. Additionally, we add channel-specific linear and quadratic time trends because many of the channels in the sample increased their follower count during the three-year observation span.¹² The main coefficients of our interest are coefficients β_1 and β_2 , where β_1 and β_2 capture the effect of fitness challenges during and after the challenges, respectively.

We estimate Model 4.1 and present the results in Table 4.4. We find that during a fitness challenge the total daily views for the hosting channel go up by 22.8%. We find a similar positive effect for all the other performance variables of interest. Namely, the daily engagements increase by 42.9%, the number of new followers goes up by 51.2%, and the views per follower go up by 3.2%.

These estimates demonstrate that challenges can improve a channel’s standing from several important aspects. First, daily views of the channel increase, which would likewise increase the channel’s profits coming from the viewing activity.¹³ Increased engagements may indicate an improved interest in the channel during the fitness challenge. We also see that the channel experiences accelerated growth during the challenge which is beneficial for the subsequent channel development. Finally, the increased number of views per follower indicates that, on average, each user watches videos uploaded by the treated channel more often, which is an important outcome for policymakers and consumers, as it leads to a more frequent physical activity per user. Importantly, the estimates of β_2 show that the effect persists after the challenge is over, and its magnitude is comparable to the challenge effect

¹²Previous research studying the life cycle of creators have noticed that popular social media accounts have a common performance pattern with growth phase followed by a plateau and then decline (Garimella and West, 2021). By including the linear and quadratic time trends, we control for the natural productivity and audience interest trends.

¹³YouTube creators have several ways of making money through YouTube, including views coming from accounts with Premium membership and displayed ads <https://support.google.com/youtube/answer/72857?hl=en>.

Table 4.4: Effect of Challenges on Aggregate Daily Performance

	(1)	(2)	(3)	(4)
	LogViews	LogEngagements	LogNewFoll	LogViewsPerFollower
Challenge	0.228*** (0.0797)	0.429*** (0.114)	0.512** (0.199)	0.0321*** (0.00694)
PostChallenge	0.143** (0.0652)	0.186** (0.0757)	0.680** (0.302)	0.0260*** (0.00714)
Channel Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Observations	21,742	21,742	21,723	21,742
R ²	0.943	0.935	0.456	0.714

Reported standard errors are clustered at the channel level.

itself. The presence of carryover effects is reassuring because it suggests longer-term effects of the fitness challenges.

4.3.3 Counterfactual Estimates

Next, we apply a Counterfactual Estimates (CE) (Liu et al., 2021) method that allows for the treatment effect estimations under weaker assumptions than DiD. Namely, the CE method extends the DiD framework and allows for unobservable time-varying confounders with parametric restrictions on these unobservable components. Accounting for unobservable components eliminates some estimation concerns common for the DiD method, such as the presence of pre-trends and other unobserved channel-level differences.

The CE method produces estimates by modeling counterfactuals for the treated units at the treatment time periods, i.e., $Y_{ct}(0)$ for all the c, t s.t. $D_{ct} = 1$, where $Y_{ct}(0)$ is the outcome for unit c at time t in the absence of treatment. In particular, CE models counterfactuals in the following general form:

$$Y_{ct}(0) = X_{ct}\beta + h(U_{ct}) + \epsilon_{ct},$$

where $Y_{ct}(0)$ is the untreated potential outcome for channel c at time t , X_{ct} are observed covariates, U_{ct} are unobservable attributes, and h is a known parametric function. Liu et al. (2021) suggest using the Interactive Fixed Effects model that consists of one or several terms $\lambda_c f_t$ that allow for unit-specific effect changes over time, where λ_c are channel-specific and f_t are the time-specific components. The number of these terms in the final model is selected by cross-validation. Then, the optimal model is estimated.¹⁴

Table 4.5 reports the estimates from the CE estimation. We find that the estimated effects from the CE model are very similar to those from DiD (21.8% vs. 22.8% for the increase in daily views, 44.4% vs. 42.9% for the increase in daily engagements, 63.1% vs. 51.2% for the increase in the new follower count, and 3.3% vs. 3.2% for the increase in views per follower.)

¹⁴For CE estimation, we use the Stata package *fect* released by Liu et al. (2021).

Table 4.5: Effect of Challenges on Daily Performance - Counterfactual Estimates

	(1)	(2)	(3)	(4)
	LogViews	LogEngagements	LogNewFoll	ViewsPerFollower
Challenge	0.218*** (0.074)	0.444*** (0.117)	0.631** (0.248)	0.033*** (0.0085)
PostChallenge	0.132*** (0.0468)	0.192*** (0.0708)	0.577** (0.258)	0.023*** (0.0044)
Channel Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Observations	21,742	21,742	21,723	21,742

Bootstrapped standard errors are reported at the individual business level ($n = 200$).

In Sections 4.2 and 4.3, we demonstrated that fitness challenges have a positive effect on the aggregate daily-level YouTube channel performance. We find that the results of these estimations are positive and significant for both the standard DiD and the more robust counterfactual estimates model. In Section 4.6, we perform additional analyses for the validity of these estimates, such as robustness and parallel trends checks. Now, we turn to analyze the effect of the challenges on the audiences' reactions and activity.

4.3.4 Comment Content Analysis

In the previous sections, we find that fitness challenges help fitness channels improve their performance. Now, we examine the underlying mechanisms driving the challenge effectiveness by utilizing the content of the user comments. For businesses, it is important to monitor online social chatter because it helps explain the sentiment towards a product (Zhang and Moe, 2021) and specific ways that users interact with the product (Timoshenko and Hauser,

2019). From the methodological point of view, trends in UGC topics have been used to understand the underlying reasons for user behavioral changes (Deng et al., 2022). As previously discussed, we identify three mechanisms that potentially explain the challenges’ effect on the users’ exercise frequency. We label the comments according to whether they mention the three mechanisms of our interest.

We test whether the relative frequency of user comments that mention these mechanisms is higher during and after challenges. For that, we count the number of comments labeled as “habit,” “community,” or “progress” for each day and every channel and test whether the number of these comments is higher during and after challenges. We also compare the frequency of these comments during and after the challenges using one of the common types of comments under the workout videos—praise for the workout video. We estimate Model 1 to test this assumption. Since the number of engagements increases during and after fitness challenges, an increase in the number of comments mentioning the mechanisms might also go up. To remove this concern, we also control for the total number of comments that a given channel received on a given day. We estimate the model and present the results in Table 4.6.

In column 1, we find that the relative frequency of the habit-themed comments goes up by 31.5% during the fitness challenges and by 11.6% after the challenges. Since fitness challenges are the only intervention during this period, we conclude that challenges are the main driving factor for a change in the comment composition. People report that they are exercising more regularly, and this self-reported metric agrees with our previous finding of increased views per follower. We find similar changes in the relative frequency of the comments mentioning community support and physical progress (for them, the effects during the challenges are 42.5% and 35.7%, and the effects after the challenges are 13.6% and 12.5%, respectively.)

These findings show that the mechanisms of habit formation, communal support, and feeling of competence, previously found to be working in other settings, also promote more active exercise and engagement in the context of YouTube workouts. Namely, Charness and Gneezy (2009) and Acland and Levy (2015) find that in offline settings, a monetary nudge

Table 4.6: Mechanisms

	(1)	(2)	(3)	(4)
	LogCommentsHabit	LogCommentsCommunity	LogCommentsProgress	LogCommentsPraise
Challenge	0.315** (0.129)	0.425* (0.210)	0.357* (0.183)	0.0542 (0.0616)
PostChallenge	0.116* (0.0654)	0.136* (0.0730)	0.125** (0.0574)	-0.0583 (0.0395)
LogTotalComments	0.328*** (0.0487)	0.275*** (0.0590)	0.208*** (0.0393)	0.638*** (0.0383)
Channel Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Observations	21,742	21,742	21,742	21,742
R ²	0.765	0.749	0.680	0.869

Reported standard errors are clustered at the channel level.

to exercise regularly for some time helps promote exercise habit formation. We find that the non-monetary nudge to exercise in online settings can create similar effects. Next, although online platforms have been found to be limited in the level of interpersonal interaction, we find that the creators may influence the level of support and community they provide to their followers. Finally, column 3 demonstrates that the users also mention progress in their physical performance more during and after the challenges. This finding is in accordance with previous studies of weight-loss communities that demonstrate a positive effect of challenges (Bojd et al., 2022) and extends them to wider online contexts.

In contrast to the results in columns 1-3, column 4 demonstrates that the portion of the comments praising the videos does not change during and after fitness challenges. This finding validates the estimates in the other columns. Namely, we see that the overall sentiment towards the videos posted on the channel remains unchanged. This means that fitness challenges have attributes that differentiate them from the rest of the content, and this difference

is not related to the video quality characteristics.

4.3.5 Effect of Challenges on Individual User Activity

Next, we turn to the individual user activity analysis by incorporating the comment data. Since each comment contains the unique commenter’s ID, we can track the individual user dynamics over time. We create a panel data structure at the user-day-channel level with the variable of interest *Comment* equal to 1 if a given user left a comment on a given channel on a given day, and 0 if not. We aim to estimate whether the user’s propensity to comment increases during and after challenges. For this, we estimate the following equation:

$$\begin{aligned} Comment_{ict} = & \beta_0 + \beta_1 Challenge_{ct} + \beta_2 PostChallenge_{ct} + \\ & + \beta_3 Tenure_{ict} + \alpha_i + \gamma_t + \epsilon_{ict} \end{aligned} \quad (4.2)$$

Here, $Comment_{ict}$ is the dummy for whether user i left one or several comments on channel c at time t . $Challenge_{ct}$ and $PostChallenge_{ct}$ have the same definition as in the previous models and indicate whether channel c had a challenge or recently ended a challenge at time t . We control for user- and time-fixed effects α_i and γ_t . We also control for the time since the first comment left by the user $Tenure_{ict}$ (Brzozowski, 2009). The intuition behind including this variable is the decreasing activity of users over time since first commenting on the channel because of dropout. Therefore, the inclusion of this variable improves the fit of the model.

In the customer relationship management literature, it is common to estimate user behaviors using cohorts, i.e., compare the behavior of customers who joined around the same time (Ascarza and Hardie, 2013; Schweidel et al., 2008). We use all the commenters who left their first comment in a given month as a cohort. This approach allows us to compare the effects of challenges among similar groups of customers and ensure that the number of observations per user in the estimation dataset is the same (the observations for each customer start at the same point of time.) For the analyses below, we use three different cohorts: the users who left their first comment in September 2019, in March 2020, and in September 2020.

These three cohorts represent interesting groups of users: the ones who first commented on a yoga channel during the start of the Covid-19 pandemic, as well as the ones who joined 6 months prior to or after that date.

Since the dependent variable in Model 4.2 only takes values 0 and 1, we run a fixed-effects logistic regression. We present these results in Table 4.7. We find that at the individual level fitness challenges also increase the activity of users. Namely, the coefficients of interest are all positive and significant for all three cohorts. In order to understand the magnitude of the effect, we calculate the average marginal effects of the variables $Challenge_{ct}$ and $PostChallenge_{ct}$. We present the results in Table B.3 of the Appendix. According to these estimations, a commenter is more likely to leave a comment during the challenge by around 7% and after the challenge by 2%. As previous studies found (Zheng et al., 2015), higher user engagement leads to higher loyalty development, which creates the base for more active consumption of workout videos.

The aggregate channel analysis, along with the analysis of individual video performance and individual user activity, demonstrates that fitness challenges increase channel performance and user activity.

4.3.6 Robustness Checks

In this section, we demonstrate the validity of the DiD approach in our empirical setting and perform robustness checks to eliminate potential concerns with our findings. First, we discuss the DiD validity and present the results of the parallel trends test. Next, we address the concerns of potential outliers driving the effect and remove potential over-performers in terms of the channels or time periods and re-estimate the main effects model.

Difference-in-Differences Method Validity

The validity of the DiD framework lies in the parallel trends assumptions. We check that this assumption holds for the aggregate performance dataset by estimating the following regression equation:

Table 4.7: Effect of Yoga Challenges on Individual Commenting Behavior

	(1)	(2)	(3)
	Comment	Comment	Comment
Challenge	1.336*** (0.0582)	0.9721*** (0.06597)	1.645*** (0.0984)
PostChallenge	0.5278*** (0.0813)	0.2528*** (0.0696)	0.4613*** (0.145)
Tenure	-0.2952*** (0.0438)	-0.341*** (0.0212)	-0.1786*** (0.108)
Individual FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	1,065,525	2,512,346	250,893
Cohort	Sept, 2019	March, 2020	Sept, 2020

Individual-level bootstrapped standard errors are reported in parentheses.

$$\text{LogViews}_{ct} = \sum_{j=-15}^{15} \beta_j \text{Challenge}_{ct}(t = k + j) + \lambda_c + \gamma_t + \epsilon_{ct}. \quad (4.3)$$

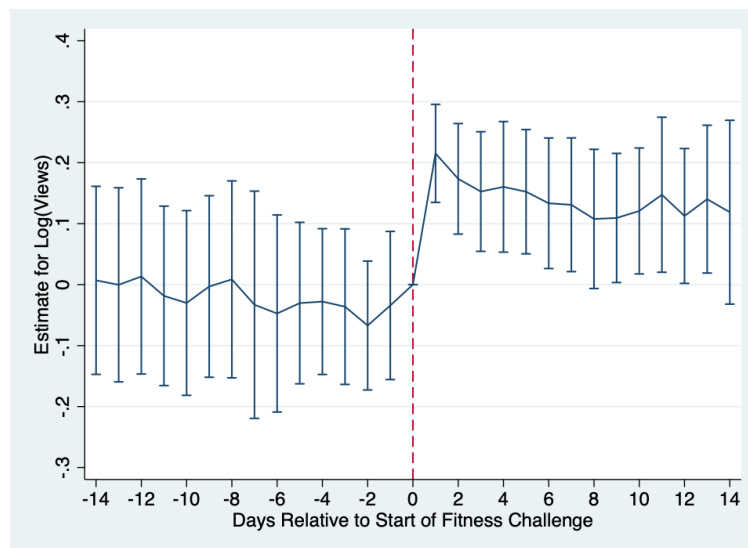
Here, we use a popular approach of estimating the lags and leads of the treatment effect (Autor, 2003) for checking whether treated units demonstrated different trends in outcomes before the start of the challenge. Namely, we estimate the treatment coefficients for 15 periods before and after the beginning of fitness challenges. In the equation, each coefficient β_j ($j = -15, -14, \dots, 14, 15$) represents the treatment coefficient for the treated units j days before or after the start of the challenge. It is expected that the lead coefficients (i.e., the coefficients for the time points before the start of the fitness challenges) are not statistically different from zero. We present the plot with estimates for the lead and lag dummies below. All the point estimates to the left of the dashed line present the channel's performance before the start of the challenge, compared to the control group, and all the observations to the right are point estimates for the challenge effect after the start of the challenge. The plot demonstrates that the channel hosting fitness challenge does not show a statistically different behavior before the start of the challenge (the estimates are all insignificant), but had a clear increase in performance after the start of the challenge, since all the point estimates are positive and much higher in magnitude, compared to the pre-challenge trends, and all but one are statistically significant.

Removing Outliers

One of the potential concerns might be the heterogeneity of the audience reaction to challenges for different YouTube channels. For example, Yoga with Adriene, which is the largest channel in our sample, with over 10 million followers, is famous for its yearly yoga challenges.¹⁵ Given these differences from the competitors, Yoga with Adriene channel may elicit a more positive audience response to the challenges compared to other channels in our sample. In order to test this, we rerun Model 1 but excluded all the observations for Yoga

¹⁵<https://www.vox.com/the-goods/2020/5/6/21241805/yoga-with-adriene-youtube-workout-at-home>

Figure 4.2: Parallel Trends Check: Lagged Coefficients Estimates



Note. The plot shows the number of views in the treatment group compared to the control group. The horizontal axis denotes the number of days before the start of the challenge. 95% confidence intervals are provided for the estimated coefficients. The red dashed line represents the start of the challenge.

with Adriene channel. We present the estimates in Table 4.8. The estimates are still similar in magnitude and statistically significant. Therefore, the positive effect is not only driven by the market leader, and smaller channels also reap the benefits of hosting yoga challenges.

Table 4.8: Robustness Check: Effect of Challenges on Daily Performance

	(1)	(2)	(3)	(4)
	LogViews	LogEngagements	LogNewFoll	LogViewsPerFollower
Challenge	0.170** (0.0662)	0.343*** (0.0960)	0.413* (0.218)	0.0295*** (0.00840)
PostChallenge	0.130* (0.0715)	0.169* (0.0820)	0.687* (0.355)	0.0271*** (0.00857)
Channel Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Observations	20,478	20,478	20,461	20,478
R ²	0.921	0.918	0.441	0.710

4.4 Discussion and Conclusion

Understanding ways to help people exercise more regularly has been a challenging task for researchers and practitioners alike. In this study, we examine whether the release of fitness challenge programs by fitness creators can be a helpful strategy to help followers exercise more consistently. We use the context of yoga channels posting follow-along workouts on YouTube. We find that fitness challenges are helpful in increasing both a creator’s performance and a follower’s exercise activity. We find that fitness challenges significantly increase multiple metrics for channel performance, such as daily views, engagements, new followers, and views per follower. We utilize the comment data in order to understand the potential drivers

of these phenomena. Namely, we find evidence that the relative frequency of comments mentioning habit formation, community support, and physical shape improvement increases, compared to other comment topics. Notably, we find that the effect of the challenges persists, and the hosting channels enjoy increased performance after the challenge is over. Finally, we find evidence that individual user engagement is higher during and after the fitness challenges. Combined with the improved views per follower metric, this finding suggests that fitness challenges benefit the audience at the individual level, providing the potential for increased exercising frequency.

Our study contributes to the understanding of exercise incentives that are of value to both researchers and practitioners. First, we contribute to several streams of the literature. We extend previous studies that found that monetary incentives help people start exercising and promote exercise habit formation (Charness and Gneezy, 2009; Acland and Levy, 2015). Unlike these studies, we propose that online content strategies are effective in habit formation for exercising, in the absence of external monetary stimuli. Besides, we contribute to the studies of online platforms and their potential to improve health-related outcomes. Namely, we explore the role of YouTube in promoting exercise activity, which has been understudied. Additionally, we explore the potential driving mechanisms that drive the effectiveness of fitness challenges. Third, we contribute to the emerging stream of literature that analyzes what online video content attributes facilitate audience and performance growth (Park et al., 2023; Krijestorac et al., 2019). Unlike most studies in this domain, ours focuses on a channel-wide strategy that content creators may apply, as opposed to single-video features.

From the application perspective, our findings have valuable practical implications for creators, viewers, and policymakers. Creators may find the strategy of investing in a fitness challenge risky, since it usually includes the release of multiple workouts in a short period of time and the audience might not commit to following the suggestions for a long time because it might not meet their personal needs and not work with their schedules. We find that, on the contrary, fitness challenges significantly increase important performance metrics and lead to increased interest in a channel. In addition, we provide evidence that these effects endure

after the challenge is over. This demonstrates long-term positive performance that serves as additional motivation to invest into this type of content. In addition to content creators, content creator sponsors may find it beneficial to help creators with hosting the challenges. These collaborations would assist content creation and help sponsors reap the long-term benefits as well. The findings are also intriguing for healthcare practitioners because the information may help them design more sustainable and effective wellness plans for their patients. Similarly, the findings are of interest to policymakers who may help stimulate content creators to provide more exercise programs such as fitness challenges. Finally, our findings are important for the users themselves since they provide practical suggestions for more sustainable exercise routines.

Our study highlights a novel physical exercise market. We believe there are several directions for future research. In this chapter, we use the context of free workouts posted on YouTube. Paid fitness apps and websites have been gaining popularity recently and they, too, contain workout programs along with standalone workouts. It would be interesting to study the effect of the fitness programs compared to the standalone workouts in the context of paid fitness content. Given that the customers paying for the workout could be more motivated than the consumers of the free content and the potential differences in the content of the paid and free content, different results might be found. Studying the outcomes on other platforms might also be interesting to understand the potential heterogeneous effects of fitness programs on users. One of the challenges of the YouTube platform is the fact that individual viewing behavior is not available to researchers and content creators; other platforms might provide ground for studying individual behaviors in more detail.

Additionally, it might be interesting to track the market dynamics among fitness content creators and the long-term effects of challenges. Namely, we noticed that the strategy of hosting a fitness challenge has become more common. In this case, the effect of fitness challenges on individual channel performance might decrease over time. Furthermore, users might find the challenges less interesting and follow the fitness program less strictly. Finally, here, we only focus on synchronous fitness challenges. Some other types of fitness challenges

that we found on YouTube and other platforms include posting daily exercise calendars, posting the whole program at the same time so that the users start on their own time, and short exercise challenges that are meant to be added to regular exercise rather than serving as a full exercise schedule. Given the rich potential of the online fitness industry, we believe that these directions will provide further understanding of the components of sustainable fitness electronic systems.

Chapter 5

CONCLUDING REMARKS

This dissertation aimed to understand the impact of recently developing phenomena in the context of social media on socially beneficial online outcomes. We selected two novel artifacts and studied the contexts where their effect are prominent.

Chapter 3 delved into the influence of User-Generated Content (UGC) on the performance of minority-owned businesses, particularly during periods of heightened societal interest. While UGC has been extensively researched, few studies have specifically examined its impact on underperforming products and its potential to narrow the earning gap between businesses. Our investigation, therefore, fills this research gap by scrutinizing the role of online reviews in the context of political consumerism. The results reveal that reviews mentioning Black ownership significantly increase foot traffic to Black-owned restaurants, particularly during heightened consumer interest in supporting Black causes. Furthermore, our findings highlight the nuanced impact of reviews in cities with lower segregation and racial isolation, emphasizing the role of online content in fostering political consumerism and promoting interracial exposure.

In Chapter 4, our attention shifts to the burgeoning market of video content creation on social media platforms, specifically within the genre of physical fitness. With the increasing time spent online by users, understanding the impact of content creation and posting strategies becomes crucial. Focusing on the fitness genre, we explore the effectiveness of fitness challenges as a content creation and engagement strategy. The research, encompassing 21 popular yoga channels on YouTube and analyzing 37 fitness challenges, reveals that these challenges significantly enhance channel performance and user engagement. The mechanisms behind this effectiveness include habit formation, community support, and improved physical

performance, offering valuable insights for fitness content creators, consumers, sponsors, and policymakers.

In summary, this dissertation contributes to the evolving literature on social media by unraveling the intricate dynamics of User-Generated Content and content creation strategies, shedding light on their substantial influence on offline outcomes. The findings expand our understanding of these phenomena and also have practical implications for businesses, consumers, and policymakers navigating the dynamic landscape of contemporary digital interactions.

BIBLIOGRAPHY

- Acland, D. and Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science*, 61(1):146–160.
- Adjerid, I., Loewenstein, G., Purta, R., and Striegel, A. (2022). Gain-loss incentives and physical activity: the role of choice and wearable health tools. *Management Science*, 68(4):2642–2667.
- Agarwal, A., Hosanagar, K., and Smith, M. D. (2011). Location, location, location: An analysis of profitability of position in online advertising markets. *Journal of marketing research*, 48(6):1057–1073.
- Agarwal, S. and Sen, A. (2022). Antiracist curriculum and digital platforms: Evidence from black lives matter. *Management Science*.
- Ananthakrishnan, U. M., Basavaraj, N., Karmegam, S. R., Sen, A., and Smith, M. D. (2023). Book bans in american libraries: Impact of politics on inclusive content consumption. *Available at SSRN*.
- Aneja, A., Luca, M., and Reshef, O. (2023). Black ownership matters: Does revealing race increase demand for minority-owned businesses? Technical report, National Bureau of Economic Research.
- Archak, N., Ghose, A., and Ipeirotis, P. G. (2011). Deriving the pricing power of product features by mining consumer reviews. *Management Science*, 57(8):1485–1509.
- Ascarza, E. and Hardie, B. G. (2013). A joint model of usage and churn in contractual settings. *Marketing Science*, 32(4):570–590.

- Athey, S., Blei, D., Donnelly, R., Ruiz, F., and Schmidt, T. (2018). Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. In *AEA Papers and Proceedings*, volume 108, pages 64–67.
- Athey, S., Ferguson, B., Gentzkow, M., and Schmidt, T. (2021). Estimating experienced racial segregation in us cities using large-scale gps data. *Proceedings of the National Academy of Sciences*, 118(46):e2026160118.
- Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics*, 21(1):1–42.
- Babar, Y., Adeli, A., and Burtch, G. (2022). The effects of online social identity signals on retailer demand. *Available at SSRN 4292001*.
- Bates, T. and Robb, A. (2013). Greater access to capital is needed to unleash the local economic development potential of minority-owned businesses. *Economic Development Quarterly*, 27(3):250–259.
- Baumeister, R. (2018). *Self-regulation and self-control: Selected works of Roy F. Baumeister*. Routledge.
- Becker, G. S. and Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy*, 96(4):675–700.
- Beshears, J., Lee, H. N., Milkman, K. L., Mislavsky, R., and Wisdom, J. (2021). Creating exercise habits using incentives: The trade-off between flexibility and routinization. *Management Science*, 67(7):4139–4171.
- Bischoff, K. and Reardon, S. F. (2014). Residential segregation by income, 1970–2009. *Diversity and disparities: America enters a new century*, 43.
- Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). cem: Coarsened exact matching in stata. *The Stata Journal*, 9(4):524–546.

- Bojd, B., Song, X., Tan, Y., and Yan, X. (2022). Gamified challenges in online weight-loss communities. *Information Systems Research*.
- Bruns, A. and Stieglitz, S. (2012). Quantitative approaches to comparing communication patterns on twitter. *Journal of Technology in Human Services*, 30(3-4):160–185.
- Brynjolfsson, E. and Smith, M. D. (2000). Frictionless commerce? a comparison of internet and conventional retailers. *Management science*, 46(4):563–585.
- Brzozowski, M. J. (2009). Watercooler: exploring an organization through enterprise social media. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, pages 219–228.
- Buchanan, L., Bui, Q., and Patel, J. K. (2020). Black lives matter may be the largest movement in u.s. history. *The New York Times*.
- Butler, B. S. (2001). Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information Systems Research*, 12(4):346–362.
- Butler, J. S. (2012). *Entrepreneurship and self-help among black Americans: A reconsideration of race and economics*. Suny Press.
- Carrera, M., Royer, H., Stehr, M., and Sydnor, J. (2020). The structure of health incentives: Evidence from a field experiment. *Management Science*, 66(5):1890–1908.
- Charness, G. and Gneezy, U. (2009). Incentives to exercise. *Econometrica*, 77(3):909–931.
- Chatterjee, P. (2001). Online reviews: Do consumers use them? *Advances in Consumer Research*, 28:129–133.
- Chernev, A., Böckenholt, U., and Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25(2):333–358.

- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022a). Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921):108–121.
- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022b). Social capital ii: determinants of economic connectedness. *Nature*, 608(7921):122–134.
- Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354.
- Chiou, L. and Tucker, C. (2020). Social distancing, internet access and inequality. Technical report, National Bureau of Economic Research.
- Copeland, L. and Boulianne, S. (2022). Political consumerism: A meta-analysis. *International Political Science Review*, 43(1):3–18.
- Cutler, D. M. and Glaeser, E. L. (1997). Are ghettos good or bad? *The Quarterly Journal of Economics*, 112(3):827–872.
- Davis, D. R., Dingel, J. I., Monras, J., and Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, 127(4):1684–1738.
- De Zúñiga, H. G., Copeland, L., and Bimber, B. (2014). Political consumerism: Civic engagement and the social media connection. *New media & society*, 16(3):488–506.
- Deci, E. L. and Ryan, R. M. (1985). Cognitive evaluation theory. In *Intrinsic motivation and self-determination in human behavior*, pages 43–85. Springer.
- Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management science*, 49(10):1407–1424.

- Deng, Y., Zheng, J., Khern-am nuai, W., and Kannan, K. (2022). More than the quantity: the value of editorial reviews for a user-generated content platform. *Management Science*, 68(9):6865–6888.
- Droogendyk, L., Wright, S. C., Lubensky, M., and Louis, W. R. (2016). Acting in solidarity: Cross-group contact between disadvantaged group members and advantaged group allies. *Journal of Social Issues*, 72(2):315–334.
- Dyrlund, A. K. and Wininger, S. R. (2006). An evaluation of barrier efficacy and cognitive evaluation theory as predictors of exercise attendance. *Journal of Applied Biobehavioral Research*, 11(3-4):133–146.
- Erkan, I. and Evans, C. (2016). The influence of ewom in social media on consumers' purchase intentions: An extended approach to information adoption. *Computers in human behavior*, 61:47–55.
- Fairlie, R. W. and Robb, A. M. (2007). Why are black-owned businesses less successful than white-owned businesses? the role of families, inheritances, and business human capital. *Journal of Labor Economics*, 25(2):289–323.
- Farrell, D., Wheat, C., and Grandet, C. (2019). Place matters: Small business financial health in urban communities. *Available at SSRN 3462771*.
- Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information systems research*, 19(3):291–313.
- Fredriksson, A. and Oliveira, G. M. d. (2019). Impact evaluation using difference-in-differences. *RAUSP Management Journal*, 54:519–532.
- Fulk, J., Schmitz, J., Steinfield, C. W., et al. (1990). A social influence model of technology use. *Organizations and Communication Technology*, 117:140.

- García, R. E. and Darity Jr, W. A. (2021). Self-reporting race in small business loans: A game-theoretic analysis of evidence from ppp loans in durham, nc. In *AEA Papers and Proceedings*.
- Garimella, K. and West, R. (2021). Evolution of retweet rates in twitter user careers: Analysis and model. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 15, pages 1064–1068.
- Gentzkow, M. and Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, 126(4):1799–1839.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., and Vermeersch, C. M. (2016). *Impact evaluation in practice*. World Bank Publications.
- Ghose, A., Ipeirotis, P. G., and Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, 60(7):1632–1654.
- Ghose, A. and Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management science*, 55(10):1605–1622.
- Giugni, M., Lorenzini, J., Giugni, M., and Lorenzini, J. (2017). Social networks and social capital. *Jobless Citizens: Political Engagement of the Young Unemployed*, pages 63–77.
- Hirschman, E. C. (1980). Innovativeness, novelty seeking, and consumer creativity. *Journal of Consumer Research*, 7(3):283–295.
- Hong, P. M. and Peoples, C. D. (2021). The ties that mobilize us: Networks, intergroup contact, and participation in the black lives matter movement. *Analyses of Social Issues and Public Policy*, 21(1):541–556.
- Horowitz, J. M. (2021). Support for black lives matter declined after george floyd protests, but has remained unchanged since. *Pew Research Center*.

- Hou, Y. and Poliquin, C. W. (2023). The effects of ceo activism: Partisan consumer behavior and its duration. *Strategic Management Journal*, 44(3):672–703.
- Howell, S. T., Kuchler, T., Snitkof, D., Stroebel, J., and Wong, J. (2021). Racial disparities in access to small business credit: Evidence from the paycheck protection program. Technical report, National Bureau of Economic Research.
- Jacobsen, E. and Dulrud, A. (2007). Will consumers save the world? the framing of political consumerism. *Journal of Agricultural and Environmental Ethics*, 20:469–482.
- Jurgens, D., McCorriston, J., and Ruths, D. (2015). An analysis of exercising behavior in online populations. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 9, pages 630–633.
- Kim, M. (2022). How can I be as attractive as a Fitness YouTuber in the era of COVID-19? The impact of digital attributes on flow experience, satisfaction, and behavioral intention. *Journal of Retailing and Consumer Services*, pages 102778–102778.
- Krijestorac, H., Garg, R., and Saar-Tsechansky, M. (2019). Personality-based content engineering for rich digital media. *Available at SSRN 3366561*.
- Lechner, M. et al. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics*, 4(3):165–224.
- Levi, M. and Linton, A. (2003). Fair trade: a cup at a time? *Politics & Society*, 31(3):407–432.
- Li, Z. and Wang, G. (2020). The role of on-demand delivery platforms in restaurants. *Available at SSRN 3813891*.
- Liaukonytė, J., Tuchman, A., and Zhu, X. (2023). Frontiers: spilling the beans on political consumerism: do social media boycotts and buycotts translate to real sales impact? *Marketing Science*, 42(1):11–25.

- Liu, D., Santhanam, R., and Webster, J. (2017). Toward meaningful engagement: a framework for design and research of gamified information systems. *MIS Quarterly*, 41(4).
- Liu, L., Wang, Y., and Xu, Y. (2021). A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. *arXiv preprint arXiv:2107.00856*.
- Liu, L., Wang, Y., and Xu, Y. (2022). A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. *American Journal of Political Science*.
- Liu, X., Zhang, B., Susarla, A., and Padman, R. (2019). Go to youtube and call me in the morning: Use of social media for chronic conditions. *Forthcoming, MIS Quarterly*.
- Luo, H. and Zhang, L. (2021). Scandal, social movement, and change: Evidence from #metoo in hollywood. *Management Science*, pages 1–19.
- MacAvaney, S., Yao, H.-R., Yang, E., Russell, K., Goharian, N., and Frieder, O. (2019). Hate speech detection: Challenges and solutions. *PloS One*, 14(8):e0221152.
- Massenkoff, M. and Wilmers, N. (2023). Rubbing shoulders: Class segregation in daily activities. *Available at SSRN 4516850*.
- Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social forces*, 67(2):281–315.
- Mayya, R., Ye, S., Viswanathan, S., and Agarwal, R. (2020). Who forgoes screening in online markets and why? evidence from airbnb. *MIS Quarterly*.
- McKinsey (2022). The rise of the inclusive consumer. *McKinsey & Company*.
- Micheletti, M., Stolle, D., et al. (2004). *Politics, products, and markets: Exploring political consumerism past and present*. transaction publishers.

- Milkman, K. L., Minson, J. A., and Volpp, K. G. (2014). Holding the hunger games hostage at the gym: An evaluation of temptation bundling. *Management Science*, 60(2):283–299.
- Mundt, M., Ross, K., and Burnett, C. M. (2018). Scaling social movements through social media: The case of black lives matter. *Social Media+ Society*, 4(4):2056305118807911.
- Neilson, L. A. (2010). Boycott or buycott? understanding political consumerism. *Journal of Consumer Behaviour*, 9(3):214–227.
- Neilson, L. A. and Paxton, P. (2010). Social capital and political consumerism: A multilevel analysis. *Social Problems*, 57(1):5–24.
- Netzer, O., Feldman, R., Goldenberg, J., and Fresko, M. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science*, 31(3):521–543.
- Park, K., Lee, S., Doosti, S., and Tan, Y. (2023). Provision of helpful review videos: Effects of video characteristics on perceived helpfulness. *Production and Operations Management*.
- Pattabhiramaiah, A., Overby, E., and Xu, L. (2021). Spillovers from online engagement: How a newspaper subscriber’s activation of digital paywall access affects her retention and subscription revenue. *Management Science*.
- Perry, A., Rothwell, J., and Harshbarger, D. (2020). Five-star reviews, one-star profits: The devaluation of businesses in black communities. *Brookings Institution*.
- Phillips, N. E., Levy, B. L., Sampson, R. J., Small, M. L., and Wang, R. Q. (2021). The social integration of american cities: Network measures of connectedness based on everyday mobility across neighborhoods. *Sociological Methods & Research*, 50(3):1110–1149.
- Piercy, K. L., Troiano, R. P., Ballard, R. M., Carlson, S. A., Fulton, J. E., Galuska, D. A., George, S. M., and Olson, R. D. (2018). The physical activity guidelines for americans. *Jama*, 320(19):2020–2028.

- Rackow, P., Scholz, U., and Hornung, R. (2015). Received social support and exercising: An intervention study to test the enabling hypothesis. *British Journal of Health Psychology*, 20(4):763–776.
- Richard, M., Christina, M. F., Deborah, L. S., Rubio, N., and Kennon, M. S. (1997). Intrinsic motivation and exercise adherence. *International Journal of Sport Psychology*, 28(4):335–354.
- Royer, H., Stehr, M., and Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics*, 7(3):51–84.
- Schoenmueller, V., Netzer, O., and Stahl, F. (2023). Frontiers: Polarized america: From political polarization to preference polarization. *Marketing Science*, 42(1):48–60.
- Schweidel, D. A., Fader, P. S., and Bradlow, E. T. (2008). Understanding service retention within and across cohorts using limited information. *Journal of Marketing*, 72(1):82–94.
- Sevilla, J., Lu, J., and Kahn, B. E. (2019). Variety seeking, satiation, and maximizing enjoyment over time. *Journal of Consumer Psychology*, 29(1):89–103.
- Shi, L., Mayya, R., and Ye, S. (2023). Location divide in digital platforms? evidence from a natural experiment. *Evidence from a Natural Experiment (April 18, 2023)*.
- Sokolova, K. and Perez, C. (2021). You follow fitness influencers on YouTube. But do you actually exercise? How parasocial relationships, and watching fitness influencers, relate to intentions to exercise. *Journal of Retailing and Consumer Services*, 58:102276.
- Song, X., Petrak, J., Jiang, Y., Singh, I., Maynard, D., and Bontcheva, K. (2021). Classification aware neural topic model for COVID-19 disinformation categorisation. *PloS One*, 16(2):e0247086.

- Stolle, D., Hooghe, M., and Micheletti, M. (2005). Politics in the supermarket: Political consumerism as a form of political participation. *International political science review*, 26(3):245–269.
- Stolle, D. and Micheletti, M. (2013). *Political consumerism: Global responsibility in action*. Cambridge University Press.
- Strohacker, K., Galarraga, O., and Williams, D. M. (2014). The impact of incentives on exercise behavior: a systematic review of randomized controlled trials. *Annals of Behavioral Medicine*, 48(1):92–99.
- Timoshenko, A. and Hauser, J. R. (2019). Identifying customer needs from user-generated content. *Marketing Science*, 38(1):1–20.
- Uetake, K. and Yang, N. (2020). Inspiration from the “biggest loser”: Social interactions in a weight loss program. *Marketing Science*, 39(3):487–499.
- Verma, S. and Yadav, N. (2021). Past, present, and future of electronic word of mouth (ewom). *Journal of Interactive Marketing*, 53:111–128.
- Waldfogel, J. (2007). *The tyranny of the market: Why you can't always get what you want*. Harvard University Press.
- Williams, J. P. (2021). A year after george floyd’s killing, white allyship fades. *US News*.
- Wilson, P. M., Mack, D. E., and Grattan, K. P. (2008). Understanding motivation for exercise: a self-determination theory perspective. *Canadian Psychology/Psychologie Canadienne*, 49(3):250.
- Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1):57–76.
- Yan, L. and Tan, Y. (2014). Feeling blue? Go online: an empirical study of social support among patients. *Information Systems Research*, 25(4):690–709.

- Yang, M., Ren, Y., and Adomavicius, G. (2019). Understanding user-generated content and customer engagement on Facebook business pages. *Information Systems Research*, 30(3):839–855.
- Zervas, G., Proserpio, D., and Byers, J. W. (2021). A first look at online reputation on airbnb, where every stay is above average. *Marketing Letters*, 32:1–16.
- Zhang, K. and Moe, W. (2021). Measuring brand favorability using large-scale social media data. *Information Systems Research*, 32(4):1128–1139.
- Zhang, S., Mehta, N., Singh, P. V., and Srinivasan, K. (2021). Frontiers: Can an artificial intelligence algorithm mitigate racial economic inequality? an analysis in the context of airbnb. *Marketing Science*, 40(5):813–820.
- Zhang, X., Wang, J., Kwan, M.-P., and Chai, Y. (2019). Reside nearby, behave apart? activity-space-based segregation among residents of various types of housing in beijing, china. *Cities*, 88:166–180.
- Zhao, K., Lu, Y., Hu, Y., and Hong, Y. (2022). Direct and indirect spillovers from content providers’ switching: Evidence from online livestreaming. *Information Systems Research*.
- Zheng, X., Cheung, C. M., Lee, M. K., and Liang, L. (2015). Building brand loyalty through user engagement in online brand communities in social networking sites. *Information Technology & People*, 28(1):90–106.

APPENDIX A

1 CEM Table

Table A.1: Effect of Yelp Reviews Robustness Checks - CEM

	(1)
	LogVisits
AfterBlackOwnedReview=1	0.109*** (0.0330)
AfterBlackOwnedReview=1 × BlackInterest=1	0.0415*** (0.0138)
Observations	18880
R ²	0.870

Note:

The dependent variable is the log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.* p<0.10, ** p<0.05, *** p<0.01

2 Generalized Synthetic Controls

We apply a counterfactual estimate (CE) method (Liu et al., 2022) that allows weaker treatment effect estimations under assumptions than DiD. In particular, the CE method extends the DiD framework and allows for unobservable time-varying confounders with parametric restrictions on these unobservable components. Accounting for unobservable components

eliminates some estimation concerns common to the DiD method, such as pre-trends and other unobserved channel-level differences.

The CE method produces estimates by modeling counterfactuals for treated units in treatment periods, i.e., $Y_{ct}(0)$ for all c, t s.t. $D_{ct} = 1$, where $Y_{ct}(0)$ is the outcome for unit c at time t in the absence of treatment. In particular, counterfactuals from CE models are in the following general form:

$$Y_{ct}(0) = X_{ct}\beta + h(U_{ct}) + \epsilon_{ct},$$

where $Y_{ct}(0)$ is the untreated potential outcome for channel c at time t , X_{ct} are observed covariates, U_{ct} are unobservable attributes, and h is a known parametric function. Liu et al. (2021) suggest using the Interactive Fixed Effects model that consists of one or several terms $\lambda_c f_t$ that allow for unit-specific effect changes over time, where λ_c are channel-specific, and f_t are time-specific components. The number of these terms in the final model is selected by cross-validation, and the optimal model is then estimated.¹ We present our results in Table A.2. Our estimate from the CE model for the effect of online reviews on a Black-Owned restaurant receiving a Black-Owned review is 6.1%.

Table A.2: Effect of Yelp Reviews Robustness Checks - CE

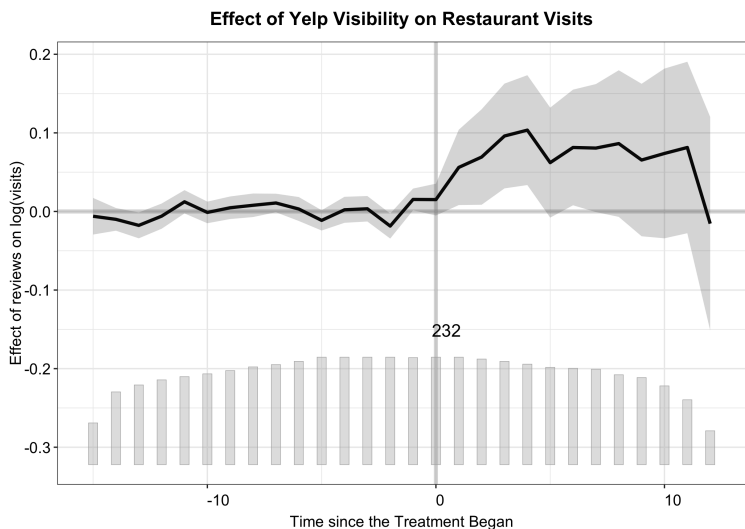
(1)	
LogVisits	
AfterBlackOwnedReview=1	0.0611**
	(0.0307)
Observations	18880

In addition to the ATT estimate, the CE framework is convenient for understanding the dynamics of the treatment effects and checking for pre-trends. Specifically, the estimation produces a visualization that shows ATT_t over time, where t is the number of periods that

¹For CE estimation, we use the Stata package *fect* released by Liu et al. (2021).

have passed since the start of the treatment. We present this plot in Figure A.1. One of the goals of this visualization is to verify a good fit. In other words, the predicted treatment effects in the pre-treatment periods should be close to zero, as in a typical linear-time model. We observe that the treatment effects are small in magnitude and insignificant in the pre-treatment periods and are high above the zero line and significant after the treatment begins.

Figure A.1: Dynamic Treatment Effects



3 Additional Robustness Checks

4 Time Spent at a Restaurant

In this subsection, we verify that the outcome variable, defined as a logarithm of total foot traffic, is a reasonable approximation of sales. We utilize an alternative approach to rule out short visits unrelated to actual sales, similar to (Li and Wang, 2020). In particular, we use an alternative metric SafeGraph provides that breaks down the visits by the times spent at the location. For this analysis, we only include visitors who spent more extended periods (between 10 minutes and 2 hours) at the venue. We define the new outcome variable *LogEatInVisits* as the logarithm of the number of more extended visits. We present our

results in Table A.3.

Table A.3: Effect of Yelp Reviews on Dine-In Visits

	(1)	(2)
	LogEatInVisits	LogEatInVisits
AfterBlackOwnedReview=1	0.167*** (0.0379)	0.160*** (0.0374)
AfterBlackOwnedReview=1 × BlackInterest=1		0.0302* (0.0161)
Observations	17646	17646
R ²	0.849	0.849

Note:

The dependent variable is the log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.*
p<0.10, ** p<0.05, *** p<0.01

As the table demonstrates, our results still hold and are similar to the baseline results, where we include all foot traffic. This result demonstrates that the overall foot traffic is a good approximation for longer visits.

5 Removing restaurants that were treated during times of heightened national interest in Black causes

In this analysis, we remove restaurants that obtain their first Black-owned reviews during increased national interest in Black causes. Times of increased national interest refer to February of 2019, 2020, and 2021 (February is celebrated as Black History Month in the US) and June and July of 2020 (months corresponding to nationwide BLM protests in 2020). Only 21 restaurants obtained their first review mentioning Black-ownership in this month. We remove these restaurants from our analysis and present the results below. The results

are consistent with the main analysis.

Table A.4: Effect of Yelp Reviews on Black-Owned Restaurant Visits After Removing Restaurants That Received Their First Black-Owned Review During Months Of High Black Interest

	(1)	(2)
	LogVisits	LogVisits
AfterBlackOwnedReview	0.108** (0.0434)	0.0998** (0.0431)
AfterBlackOwnedReview \times BlackInterest		0.0451*** (0.0141)
Observations	16952	16952
R ²	0.865	0.865

Note:

The dependent variable is the log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.*
p<0.10, ** p<0.05, *** p<0.01

6 Visual Cues of Black Ownership

One of the potential factors that may impact our estimates of the impact of online reviews revealing Black ownership on consumption is offline visual cues. For example, a restaurant can add posters outside the venue that inform potential customers about minority ownership.

To address this concern, we manually checked more than five thousand photos on the Yelp page of the 738 restaurants in our sample. The photos include visuals of the exteriors and interiors and checked if the restaurants had any posters or signs identifying them as Black-owned. Such visual cues are rare, with only 17 restaurants having such displays. We re-run our baseline analysis excluding these restaurants from the sample and present our

results in Table A.5. Our results remain consistent even after removing restaurants with posters and signs indicating Black ownership.

Table A.5: Effect of Yelp Reviews on Black-Owned Restaurant Visits After Removing Restaurants with Offline Labels

	(1)	(2)
	LogVisits	LogVisits
AfterBlackOwnedReview	0.121*** (0.0352)	0.112*** (0.0350)
AfterBlackOwnedReview \times BlackInterest		0.0471*** (0.0131)
Observations	18928	18928
R ²	0.865	0.865

Note:

The dependent variable is the log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.*
p<0.10, ** p<0.05, *** p<0.01

Placebo Test

The Placebo Test is a standard additional method for verifying the parallel trends assumption. It compares the performance of the control group with the performance of the placebo group, which is usually a subsample of control units with randomly assigned "treatment" periods (Gertler et al., 2016). This test checks whether the placebo group deviates from the control group's performance. The two groups should demonstrate similar performance trends over time since none had an actual treatment.

For this test, we select only the restaurants that were never visible as Black-owned on Yelp. Next, we randomly choose half of them in the "treatment group." The restaurants from

the treatment group also get assigned a random month at the start of the treatment. Next, we run the basic specification on this sample. We present our results in table A.6. We find that none of the regression coefficients attributed to treatment are significant, demonstrating the test's validity.

Table A.6: Effect of Yelp Reviews on Black-Owned Restaurant Visits - Placebo Test

	(1)	(2)
	LogVisits	LogVisits
FalseTreatment	-0.0273 (0.0465)	-0.0301 (0.0461)
FalseTreatment \times BlackInterest		0.0132 (0.0266)
Business FE	Yes	Yes
Year-Month FE	Yes	Yes
Observations	8,700	8,700
R ²	0.851	0.851

Note:

The dependent variable is the log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.*
 $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7 Alternate Specification: Conditional on receiving a Black-owned review

Our results may be affected if there are systematic differences between Black-owned restaurants that receive at least one review mentioning Black ownership in our study period and those that never receive such reviews. Therefore, we exclude all Black-owned restaurants that never receive any reviews mentioning Black ownership in our sample and thus are never

featured on the Yelp search when a consumer searches for “Black-owned” restaurant ²

Thus, this analysis only includes the impact of reviews mentioning Black-ownership on foot traffic conditional on receiving at least one review mentioning minority-ownership. We use the main specification in equation 4.1 and present the results in and present our results in Table A.7. Our results are consistent with our main results.

Table A.7: Effect of Yelp Reviews on Black-Owned Restaurant Visits

	(1)	(2)
	LogVisits	LogVisits
AfterBlackOwnedReview=1	0.140*** (0.0383)	0.131*** (0.0385)
AfterBlackOwnedReview=1 × BlackInterest=1		0.0520** (0.0203)
Observations	10462	10462
R ²	0.864	0.864

Note:

The dependent variable is the log of visits. The unit of observation is restaurant-month. Cluster-robust standard errors are reported at the restaurant level. All the regressions include restaurant level, year-month and city-year-month (time trends at the city level) fixed effects.* p<0.10, ** p<0.05, *** p<0.01

8 Calculating Exposure

Massenkoff and Wilmers (2023) use the same data from SafeGraph that indicates the number of users from a census block group. We closely follow (Massenkoff and Wilmers, 2023) to calculate the probability that a restaurant patron of one group would encounter another patron of any other pre-defined group. We can calculate the exposure metrics for any group.

²As mentioned earlier, we do not include restaurants that have voluntarily adopted badges.

In our analysis, we focus on racial groups. We aggregate the visits from each census block group to a census tract (CT) level. Further, we categorize each census tract as White or non-White based on the dominant racial group. First, we calculate the exposure of members of each group G in each CT to each establishment k without counting the focal group. The metric is calculated using the equation below:

$$FirmExposure_{iGk} = \frac{Visitors_{k(G-i)}}{\sum_{m \in I: m \neq i} Visitors_{k(m)}} \quad (1)$$

We calculate the $FirmExposure_{iGk}$ metrics for each group G in each census tract i for each restaurant without including the focal CT i . Then, we average the $FirmExposure_{iGk}$ metrics weighted by total visits from each CT i . In other words, the exposure of the non-White group (NW) to the White group (W) has two components. The numerator is the product of Firm exposure of members of W CTs at a restaurant k and the total number of visitors to the restaurant k from CT i . We sum over all W CTs indexed by i . Then, we sum the total visits by residents of W CTs. More specifically, we define $Exposure_k(NW, W)$ as

$$Exposure_k(NW, W) = \frac{\sum_{i \in W} FirmExposure_{iGk} \times Visitors_{k(i)}}{\sum_{i \in W} Visitors_{k(i)}} \quad (2)$$

The above equation gives us the mean exposure of White CT residents to non-White CT residents, weighted by total visits to a restaurant k . The procedure is exactly similar to the one followed in Massenkoff and Wilmers (2023) except that we do not sum over k . $Exposure_A(NW, W)$ is 0.65, which indicates that a visitor from a non-White CT at restaurant A is likely to be exposed to 65% of White residents. Given that there are only two groups, the average exposure is 50% when no segregation exists.

Thus, we calculate the exposure for residents from non-White CTs to White CTs for each of the 738 restaurants in our sample over all the months in our study period. We use this as a dependent variable to study how reviews mentioning Black ownership change the exposure experienced by different racial groups at different restaurants.

9 *EatOkra*

We leverage EatOkra data to identify Black-owned restaurants, which might not have been identified as such on Yelp. Other studies in this literature address the issue of ground truth by performing image analytics on business owners’ images (Zhang et al., 2021). However, Yelp restaurants rarely have photos of business owners, making this issue even more difficult. Fortunately, we obtain this information from the website “EatOkra”³. Starting in 2016, EatOkra initially used various online resources to list Black-owned restaurants.

EatOkra is the only resource to identify Black-owned restaurants in a central location. We contacted the chiefs of Black Business Bureaus in all the ten cities in our data through our university’s Consulting and Business Development Center, which had a Black Future Co-op Fund and Black-Led Business program. We even interviewed one of the Black business leaders in Seattle. All our contacts replied that such a consolidated list does not exist, and even governments that contract out to minority-business owners don’t have access to such a list and rely on user-generated content in the form of lists that were available on forums and social media.

We looked into how EatOkra obtained these lists, and according to an interview with Salon⁴, “Over the next six months, Edwards built the website whenever he wasn’t working at his full-time job at a construction project management software company. The couple input all the data themselves—some 350 mostly New York-based restaurants they’d unearthed like so many diners before: through word of mouth, articles, and informally shared Google Doc lists. They debuted the app in May 2016, calling it EatOkra in a nod to their Southern heritage and the ingredient’s literal and symbolic role as a binding agent and connector of Black Americans. ” They also started having partnerships with other organizations like U.S. Black Chambers (USBC) which offers certification for Black businesses⁵. In summary,

³<https://www.eatokra.com/>

⁴<https://www.salon.com/2022/09/13/eatokra-aims-to-be-the-next-for-black-owned-businesses/>

⁵see https://certification.byblack.us/?_gl=1*12rxe79*_ga*MTI5NzQyODEyMS4xNjYxNzg1MzAy*_

EatOkra is a combination of user-generated lists, with manual verification and certifications from USBC.

As of September 2022, “there are more than 11,000 Black-owned restaurants nationwide are listed on EatOkra.” and they were awarded one of 15 annual App Store Awards in December 2021.

Table B.1: Comment Labeling Keywords

Categories	Keyphrases
Habit	habit, regularly, daily routine, ve been following you, ve been a big fan, for over 2 years, do yoga every day, ve been doing your classes, doing your exercises every, ve been doing actively yoga for about a year, ve been doing yoga for about a year, I do it every morning, year after year, ve been practicing for a long, yoga every day, yoga everyday, your yoga videos since the beginning, couple of years, love starting my day with your, love starting my day with yoga, ve done more than a hundred hours of your classes, every single day, your yoga workouts for 4 years, your yoga workouts for 2 years,your yoga workouts for 3 years, for a few years, for a few months, videos everyday, videos every day, videos for a while, stay consistent, consistently, videos for quite a while, for the past several years, for several years, for years, ve been working out with you, ve been doing yoga with you, stick to, videos almost every day, videos every day, practice every day, exercise every day, ve been following, ve been doing, every morning, for a while, daily, every single day, consecutive, been following, been watching, with you for a long time, ve been practicing yoga, each day, bedtime routine, morning routine, practicing every day
Community	community, millions of us, millions of people, practicing together, practice together, practice together, comment section, everyone, shout out, hugs to anyone, everybody, communication, support, practicing with, fwfg family, with other people, facebook group, journey together, around the world, all the people, with more than, fellow, who else, yogi friends, with my friend, hundreds of people, see you all

Progress no longer have back pain, both feet off, progress, was able to do, feel stronger, could complete, finally did, come so far, first successful, crow, split, was totally able, recover from, incredible result, hand stand, handstand, head stand, headstand, lost weight, more flexible, seeing the results, see the results, feeling stronger, stronger, am stronger, transforming my body, closer to the ground, improved, improvement, to the next level

Praise good yoga flow, amazing class, amazing sequence, amazing practice, amazing workout, loved this video, favorite practice, favorite routine, beautiful practice, perfect practice, love this, awesome as usual, had a great time, just what i needed, perfect start to, perfect way to start, fantastic class, great video, amazing after this routine, love this, relaxing practice, lovely practice, it was amazing, awesome stretch, awesome practice, awesome class, great practice, great class, great flow, great session, loved this practice, the best practice, favorite class, favorite flow, favorite practice, loved it, perfect class, perfect practice, perfect flow, perfect sequence, this is what i needed, loved this, great sequence, this was exactly what i needed, very good practice, very good video, exactly what i needed, love it, my favourite practice, awesome workout, beautiful flow, beautiful class, nice sequence, nice practice, this was perfect, nice workout, nice session, nice flow, this was my favourite, this is my favourite, wonderful practice, that was amazing, awesome flow, wonderful flow, my favorite yoga, my favorite practice, this was my favorite, this was lovely, this was great

Table B.2: List of Yoga Channels Used in Analysis

Channel Name	Number of Followers, thousands	Number of Videos	Number of Challenges
Arianna Elizabeth	143	278	2
Boho Beautiful Yoga	2,310	436	0
Breathe and Flow	413	385	3
BrettLarkinYoga	478	773	0
Candace Cabrera	307	532	0
Cat Meffan	263	400	1
DoYogaWithMe	132	503	0
Five Parks Yoga w/ Erin Sampson	240	348	0
Jessica Richburg	252	113	0
Move With Nicole	623	121	1
SarahBethYoga	1,470	373	0
Travis Eliot	268	256	0
YOGATX	342	677	1
Yoga For Cure Videos	269	310	9
Yoga With Adriene	10,700	625	4
Yoga With Bird	267	574	4
Yoga With Tim	244	360	6
Yoga with Allie Van Fossen	154	282	0
Yoga with Kassandra	1,840	660	3
YogiApproved	279	242	0
yoginimelbourne	193	224	1

The numbers are reported as of January, 2022. Channel names are reported according to their spelling on youtube.com. Only the challenges posted within the study timeframe are included into counts.

Table B.3: Effect of Yoga Challenges on Individual Users - Probability Margins

	(1)	(2)	(3)
	Comment	Comment	Comment
Challenge	0.0696*** (0.0139)	0.0579*** (0.00868)	0.0725*** (0.00707)
PostChallenge	0.0275*** (0.00672)	0.0151*** (0.00449)	0.0177*** (0.0394)
Individual FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	1,065,525	2,512,346	250,893
Cohort	Sept, 2019	March, 2020	Sept, 2020

Individual-level standard errors are reported in parentheses.