

EMPIRICAL ESSAYS IN DIGITAL MARKETING

By

Michael Wai-Ming Wu

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Business Administration – Marketing – Doctor of Philosophy

2024

## **ABSTRACT**

The heightened usage of social media and other digital platforms in the modern age has generated a large number of interesting and impactful questions for business academics and practitioners. Consequently, marketing scholars have sought to utilize the tremendous amounts of data available in the digital domain to better understand the behaviors of various stakeholders within these arenas. As such, I present two essays aimed at addressing questions housed in the digital marketing space. The first essay considers how the introduction of providing the ability to display mid-roll advertisements (MRAs) on a live streaming platform affects viewing consumption. Utilizing a difference-in-differences estimation approach, the results surprisingly indicate that the introduction of MRAs increases audience consumption. This study also provides evidence of the mechanisms, suggesting that streamers make strategic adjustments to increase viewership after the introduction of MRAs to collect ad revenue. The second essay explores how dark humor and slang usage in social media text impacts virality by examining data from a social and digital media site. By employing a number of tools such as content analysis and topic modelling (via Latent Dirichlet Allocation) in conjunction with a fixed effects approach, this essay provides evidence that social media posts containing dark humor are more viral than those without it, despite the controversial nature of dark humor. Conceptually, we posit that knowledge of the specific context is required to value dark humor. This relationship is further compared against informative versus persuasive posts, finding that the effects of dark humor are generally stronger for posts which are persuasive. This essay simultaneously explores internet slang, suggesting that the use of internet slang is also positively related to virality. The findings from both essays should be of interest to both researchers and practitioners.

Copyright by  
MICHAEL WAI-MING WU  
2024

This dissertation is dedicated to my parents.  
Thank you for your never-ending belief in me.

## **ACKNOWLEDGEMENTS**

I am extremely appreciative to all the individuals who have helped me throughout my time at Michigan State University. In particular, I am very grateful to all the members of my dissertation committee: Dr. Sung Ham, Dr. Suman Basuroy, Dr. Forrest Morgeson, Dr. Eduardo Nakasone, Dr. Thomas Jeitschko, and Dr. Germán Pupato. These individuals have been an incredible set of mentors and teachers, providing me with the knowledge and support to flourish as a researcher. I am also thankful to all the other professors at Michigan State who have provided me with encouragement over the years.

Thank you to the terrific administrative and support staff at Michigan State, who were extremely accommodating throughout my doctoral program. I would also like to thank my peers for all the enjoyable memories and discussions shared. Lastly, I am immensely thankful for the support of my friends and loved ones, who have always supported me during this long journey.

## TABLE OF CONTENTS

ESSAY ONE: LIVE STREAMING CONSUMPTION AND STREAMER ABILITY TO PLAY MID-ROLL ADVERTISEMENTS.....	1
REFERENCES .....	45
APPENDIX A. Data Aggregation Procedure .....	52
APPENDIX B. Robustness Check with “Original” Dataset and Discussion of External Events.....	53
APPENDIX C. Falsification Tests Ruling Out Minor Sub-Policies.....	57
APPENDIX D. Alternative Aggregation Robustness Check .....	61
APPENDIX E. Historical Live Streaming Platform Survey Details .....	62
APPENDIX F. Data Visualizations and Robustness Checks of Baseline DiD Effects ....	70
APPENDIX G. Further Details of Dynamic Event Study Analysis .....	72
APPENDIX H. Further Robustness Checks Including Propensity Score Matching, Synthetic Difference-in-Differences, and Attrition Checks.....	74
APPENDIX I. Details of Primary Falsification Tests for COVID-19 .....	81
APPENDIX J. Further Details and Tests of DiD Mechanism Investigation .....	85
APPENDIX K. Further Details of Formal Mediation Analysis .....	87
APPENDIX L. Heterogenous Effects by Initial Success.....	88
APPENDIX M. Heterogenous Effects by Solo versus Social Content (versus Both).....	91
APPENDIX N. External Verification Survey Details .....	93
 ESSAY TWO: TOO GLOOMY OR TOO FUNNY? THE IMPACT OF DARK HUMOR AND SLANG ON SOCIAL MEDIA VIRALITY .....	 97
REFERENCES .....	122

# **ESSAY ONE: LIVE STREAMING CONSUMPTION AND STREAMER ABILITY TO PLAY MID-ROLL ADVERTISEMENTS**

## **ABSTRACT**

There has been little exploration on how audience content consumption may change in response to advertising permissions on live streaming platforms. Ads are utilized by brands to generate revenue through ad exposure, but is this benefit thwarted by the reduction of audience consumption of content? Using a dataset containing over 12 million observations in the live streaming space and a difference-in-differences estimation approach, we study the effects of a policy intervention by a live streaming platform which provided (some) streamers the ability to display mid-roll advertisements (MRAs). Although the ad avoidance literature infers that ad-supported content should be viewed unfavorably by audiences, our results indicate that providing the mere ability to introduce MRAs has a statistically significant positive effect on live streaming content consumption (average viewership and total hours watched). We suggest that a viable explanation for this response is through increases in broadcasting airtime, stream frequency and quality by streamers after the intervention, as these adjustments are drastically easier to adjust in a live streaming setting as compared to more “traditional” forms of media. We further explore the heterogeneity in these effects, in relation to initial success, streaming tenure, content activity, and how the effects evolve over time.

## INTRODUCTION

The popularity of live streaming, which is a setting where broadcasted (transmitted) streaming content is both generated and consumed simultaneously in real-time, has soared tremendously in recent years. For instance, *Forbes* reported that from 2019 to 2021, Amazon's live streaming platform Twitch experienced a 50 million boost in active users and further growth continues to be expected in the coming years (Hart 2021). Moreover, the global live streaming market, which includes platforms such as Twitch, Huya Live and YouTube (Live), is expected to be worth over four billion dollars by 2028 (Bloomberg 2022). To capitalize on this growing consumer market, large business entities such as Kentucky Fried Chicken and Electronic Arts have begun to promote or advertise themselves on live streaming services (Butler 2021).

Indeed, marketers have adopted digital marketing tools to effectively promote their brands to these sizable audiences who regularly consume live streaming content. One of the most publicized digital marketing tools that exists in the live streaming space are mid-roll advertisements (MRAs), where content is briefly paused midstream for a short commercial break, typically in a video format. Both "standard" and live streaming platforms such as YouTube, Facebook and Twitch have recently implemented or have experimented with MRAs as a method to enhance revenues for the platform (Perlberg 2017; Davey 2022; Grayson 2022). Generally, platforms introduce MRAs because brands that wish to advertise typically pay the host platform to have their ads displayed.

However, the question remains as to how advertising affects the content consumption of viewers. This is key as the purpose of MRAs is viewing exposure. Could potential reductions in consumption behaviors occur as a result of MRAs or could consumption somehow increase? This paper focuses on how consumption behaviors change after live streamers receive the ability



to display MRAs, as well as why these changes might occur. To explore these questions, we consider a one-time intervention shock on a live streaming platform in 2020 which provided the ability to display MRAs for a given sub-group of live streamers.<sup>1</sup>

Although brands typically benefit from increased sales or revenues as a consequence of advertising (e.g., Manchanda et al. 2006; Kim and KC 2020), other empirical evidence has also suggested that consumer audiences are generally avoidant towards advertisements (e.g., Wilbur 2008). Relatedly, in the digital space, ad blockers are frequently used by viewers to eliminate digital ads and streaming platforms typically require additional payment to a premium subscription tier to watch ad-free content. Thus, similar to the television space (e.g., Wilbur, Xu and Kempe 2013), one may posit that declines in viewership add to the cost of running ads (from the perspective of the brand or streaming platform). As such, streaming plans without ads are typically priced higher than those with ads, as observed in large platforms such as YouTube (YouTube Premium) or Twitch (Twitch Turbo) (Munson 2018; Miceli 2022).

Live streaming is an interesting setting to study how audiences respond to MRAs because (1) viewers are more active in this context as opposed to a more traditional media setting since they routinely interact with the streamer in real-time through the platform mechanisms and (2) streamers have the autonomy to change their stream quickly and inexpensively (such as their content streamed, the average duration they stream for, etc.). These features make the live streaming setting unique relative to “traditional” media contexts (such as television). Thus, the live stream experience is one that is comparable to live events (e.g., comedy shows, concerts, etc.), even though the service is consumed through a media device. Given the highly interactive nature between streamers and their audiences (in live streaming), ads could result in an even

---

<sup>1</sup> Live streamers stream on accounts called “channels”. Hence, “live streamers” and “channels” are synonymous with one another.

greater disruption for viewers especially since MRAs are typically not picture-in-picture. Hence, one might *a priori* anticipate that the effects of MRAs in a live streaming setting may induce even fiercer adverse consumption behaviors than in a traditional broadcast TV setting.

Meanwhile, MRAs concurrently provide a clear route for streamers to directly monetize their live streaming channels through revenue sharing by the platform. Traditionally, live streaming platforms have paid streamers in ad revenues through a cost per mile (CPM) scheme, meaning that ad revenues (including MRAs) for the streamer are dependent on the number of individuals who view each ad (Kinson 2020; Visuals By Impulse 2021; Parrish 2022). Thus, streamers have a monetary incentive to play MRAs to as many viewers as possible. This monetization and usage of MRAs could further alienate viewers and lead to immediate reductions in viewership and watch time due to ad avoidance. However, we propose one key reason for how displaying (or having the ability to display) MRAs may net an increase (rather than decrease) of streaming consumption, which is uniquely possible in the live streaming space.

Specifically, the incremental income from MRA ad revenues may motivate streamers to put greater effort towards adjusting their stream by extending streaming durations, streaming more frequently, or/and by increasing the quality of their stream content. These types of effort enhancements are made uniquely possible in a live streaming setting (as opposed to a more traditional medium) because of the sheer flexibility available to streamers to institute real-time changes to content quickly and without much cost. For example, streamers can decide how long they wish to stream for each session without any large, fixed costs and can even adjust their intended stream time during the broadcast itself. This flexibility does not exist in a traditional medium such as television, as television programs take months or years to film. Hence, the showrunners cannot easily add several minutes to each episode once it is filmed (or similarly,

cannot create another episode easily without time and cost). Likewise, filming is expensive and recording higher quality content is especially costly. Also, television networks often pre-set their broadcasting schedules and television showrunners do not have autonomous control of this schedule even if they were able to miraculously alter existing episodes quickly and cheaply.

In the live streaming space, we denote these analogous adjustments as “effort adjustments” or “streamer adjustments”.<sup>2</sup> Indeed, putting in more effort by (1) streaming for a longer duration per stream occurrence, (2) streaming more frequently, or (3) improving overall stream quality, should increase the amount of content available by a live streaming channel or improve existing content. These adjustments can increase existing average viewership or can allow current audience members to extend their viewing sessions since there is more content to consume. Consequently, we posit that it is conceivable that ad avoidance continues to exist but that the effects of streamer adjustments can overpower this negative effect, thereby creating a net positive outcome for the streamer and platform. Relatedly, Sridhar et al. (2011) study a comparable setting with platforms facing cross-market network effects. However, the live stream setting is differentiated as numerous individual streamers, rather than a singular platform, decide on how to implement adjustments. Ultimately, streamer adjustments are a feature unique to the live streaming setting which may potentially offset the negative effects of MRAs caused by ad avoidance and may even generate an overall positive effect on audience consumption. This is rather interesting, as this implies that the introduction of ads may alter the content itself.

To empirically explore the impact of introducing MRAs, we consider a 2020 policy implementation from a live streaming platform belonging to a large multinational technology

---

<sup>2</sup> It is possible to have “negative” streamer adjustments whereby the streamer produces less effort than before, but we discuss the conceptual implications of this later in the paper. This should not play a large role in our setting.

firm.<sup>3</sup> This policy shock (“the policy”), implemented several platform guidelines in the midst of our panel dataset. Critically, the major implementation was the introduction of the ability to display MRAs for partners, but not for non-partners (we discuss the differences between partnered channels and non-partnered ones later, but they are akin to receiving “verification” on social media or other digital platforms, thereby receiving additional benefits or abilities).<sup>4</sup> We also survey real live streamers who streamed on the platform to provide further institutional context. This is described later in the paper.

The policy generates an ideal quasi-experimental environment for a difference-in-differences (DiD) estimation, which can recover consistent estimates of a given treatment intervention conditional on the parallel trends assumption (Angrist and Pischke 2009; Goldfarb, Tucker, and Wang 2022). In particular, to examine the effects of introducing the ability to display MRAs (“MRA ability”), we focus on two measures of audience response behaviors: the average number of viewers and cumulative watch time per channel (we often refer to this as “total watch time”). We later conduct a variety of procedures to verify our results including staggered DiD analysis, propensity score matching (PSM), and conducting falsification tests to address alternative explanations.

Surprisingly, the findings from our DiD analysis suggest that overall, audiences responded favorably (rather than negatively) towards the MRA ability policy and consumed more live streaming content after the intervention. We show that introducing the option to use MRAs for streamers increased both the number of average viewers and the total watch time. Hence, platforms, brands, and streamers, which are all expected to profit from the additional ad

---

<sup>3</sup> We do not explicitly name the company to maintain anonymity.

<sup>4</sup> The intervention provides only the ability to display MRAs and does not force streamers to use them. However, we later discuss further why it is likely that most streamers utilize them (based on motivations by ad revenue).

revenues from MRAs, can extract gains without necessarily reducing consumer utility.

We then investigate the mechanisms of this relationship and find evidence showing that the airtime duration and the overall quality of the stream increased for the treated, partnered channels (on average) when compared to non-partnered channels. Concisely, we find evidence that partnered streamers adjusted to the intervention by streaming their channels for longer sessions and increasing their quality, which seemingly led to attracting more viewers and producing higher total watch times. Finally, we consider whether various degrees of heterogeneity exist with respect to this phenomenon. In particular, we explore (a) how the initial success of a streamer may influence the aggression of strategic adjustments implemented, (b) if established streamers implement MRAs differently than late-entry streamers, (c) whether the degree of social interaction the content activity entails matters, and (d) if the main consumption changes after introducing MRAs vary across time. A summary of the empirical analysis conducted is displayed in Table 1.

Table 1: Summary and Explanation of Empirical Tests

Method or test	Purpose of section	Research question
<i>Baseline DiD Estimates</i>	Establish baseline consumption behaviors	RQ1: Stream consumption due to MRA ability
<i>Callaway and Sant'Anna (2021)</i>	Address staggered intervention setting	RQ1
<i>Robustness Checks</i>	Ensure baseline consumption effects are robust	RQ1
<i>Threats to Validity</i>	Alleviate concerns about effects	RQ1
<i>Mechanism DiD</i>	Study mechanism effects via streamer adjustments	RQ2: Evidence of Streamer Adjustments
<i>Heterogenous Effects</i>	Explore differences in consumption effects and mechanisms	RQ3: Heterogeneity in Effects
<i>Verification Survey</i>	Validate streamers' usage of MRAs and adjustments	RQ1 & RQ2

Lastly, we conducted a verification survey on both university students and real live streamers. Their answers appear to corroborate our results and the mechanisms that we propose.

We now summarize the remainder of this paper. We first discuss the related literature and present our formal research questions. We then describe the empirical context of the platform and the data, and formally describe our DiD models which capture the relationship between MRA ability and live streaming viewing behaviors. We then present the results and then discuss the robustness and falsification checks. Next, we explore the proposed mechanisms and heterogeneous effects. Afterwards, we report and discuss the survey results. Finally, we present the managerial implications and suggestions for areas of future research.

## RELATED LITERATURE AND RESEARCH QUESTIONS

First, we provide a brief overview of live streaming and advertisements. We then amalgamate the above-mentioned ideas to discuss how MRAs and MRA ability can affect viewership in a digital and live streaming context. Lastly, we apply these concepts to discuss two ways in which MRAs may affect audience consumption, and formally state our empirically testable research questions.

### *Live Streaming Background and Overview*

The context of our research is based around live streaming, which can be classified as a type of user-generated content (UGC). UGC encompasses an expansive menu of activities including online reviews, digital blogs, social media, or any other digital consumer-created content (Fader and Winer 2012; Lamberton and Stephen 2016). During a live streaming session, a streamer will broadcast their content (typically, but not always, some form of entertainment) on a platform using their channel (we use “streamer” and “channel” interchangeably), while viewers can consume the content for no cost in real-time (Lin, Yao, and Chen 2021). As such, live streaming is a free service for viewers. Streaming sessions are often organized in advance to allow viewers to schedule time for watching the stream but can also occur spontaneously. The streamer can autonomously select the airtime duration for each session as well as the frequency. Examples of

content include playing video games, chatting with the viewing audience (via speaking or the chat box), performing with musical instruments, or cooking meals. Streamers can also sometimes display advertisements to generate revenue through processes such as ad exposure (via CPM) or through sponsorship support, but this is typically conditional on the regulations of the platform and their verification status (i.e., partnered, or non-partnered) (e.g., Grayson 2022).

Being “partnered” on a live streaming platform is an equivalent classification to social media verification and is typically associated with meeting a minimal threshold of quality (which can often be quite low). However, the exact selection of partnered channels by a platform is also rather fickle, as partnered accounts vary widely between observable traits such as the number of followers, average viewership, or the level of engagement. The benefits of being partnered can include increased visibility on the platform and advertising permissions. Importantly, we emphasize that it is typical for only partnered streamers to possess the ability or permissions to display ads, in part to protect the reputations of the platforms or brands themselves.

MRAs in a live streaming context are typically shown in a brief video format (which are typically 5 to 30 seconds in length). These video advertisements are relatively invasive, as the content typically comes to a temporary and sometimes abrupt halt (for the viewer) due to the overlay of the ad. Streamers can benefit financially from these ads but can also generate revenue from other sources such as donations, sponsorships or by growing a large audience, which in turn propels further donations and subscriptions (Grayson 2021; Lu et al. 2021; D’Anastasio 2022).

From the audience or viewer perspective, viewers can simply watch the content, chat through text in the public chatroom, or donate money to the streamer through two general alternative schemes: channel subscriptions (which provide benefits such as unique chat box badges or emoticons) and direct donations. Relatedly, a handful of studies have investigated why

viewing audiences consume live streaming content. Hilvert-Bruce et al. (2018) find that six socio-motivational factors (e.g., social interactions, sense of community, meeting new people, entertainment, etc.) encourage live streaming engagement. Thus, social motivations remain key incentives with regards to initial viewership, prolonged engagement, and positive sentiment (Hu, Zhang, and Wang 2017; Lin, Yao, and Chen 2021). These arguments are echoed by practitioners, with *The Wall Street Journal* suggesting that live streaming interactions in real-time fill a social void not found in prerecorded media (Needleman 2020). Though not the focus of this paper, we note that strong social interactions can increase purchasing in the digital space (Park et al. 2018).

#### *Advertisements and Consumption Behaviors in Digital and Live Streaming*

For brands, positive brand outcomes are commonly observed as intended with regards to general marketing initiatives through tactics such as ads (e.g., Fossen and Schweidel 2019; Kim and KC 2020). Relatedly, the digital marketing literature has found that ads can induce either positive or negative brand engagement and purchase likelihoods. The overall effect is conditional on factors including the content itself (e.g., degrees of humor or emotional engagement), the follower size, or the mood of the recipient (e.g., Berger and Milkman 2012; Lee, Hosanagar, and Nair 2018; Haenlein et al. 2020). Thus, positive outcomes for the brand are possible.

Our empirical setting is based on the introduction of the mere ability to display MRAs rather than the guarantee of usage. However, we propose (and later provide survey verification) that it is rather likely for MRA usage to be non-zero for streaming channels (once given the ability), due to the monetary incentives for streamers to run them, as well the frequent discussion of MRA usage by both brands and the media (Butler 2021; Grayson 2022). Moreover, revenues from ads on live streaming platforms can provide a stable and reliable income for streamers—something that live streamers themselves hope to acquire (Rubio-Licht 2022).



### *How Might Mid-Roll Advertisements in Live Streaming Discourage Content Consumption?*

We first discuss the rationale for how introducing MRA ability in a live streaming context could negatively impact content consumption behaviors (i.e., decreased viewership and lower total watch time). Succinctly, audiences may have an unfavorable disposition towards MRAs, causing them to leave the stream or reduce their watch time (provided that MRA usage itself is non-zero). This sentiment is rooted in the established ad avoidance literature, which suggests that large portions of consumers often wish to avoid viewing advertisements (e.g., Wilbur 2008; Wilbur, Xu, and Kempe 2013). On the viewer side, some reasons for this ad avoidant behavior include goal impediment and negative past ad experiences (Cho and Cheon 2004). Although many ad avoidance studies are conducted in the television space, a number of papers echo these sentiments to find similar conclusions in the digital and UGC domains, arguing that marketing-related phenomena may not necessarily translate into favorable consumer intentions and can sometimes paradoxically lower them (e.g., Cho and Cheon 2004; Doorn and Hoekstra 2013; Hennig-Thurau, Wiertz, and Feldhaus 2015; Chae, Bruno, and Feinberg 2019; Leung et al. 2022). Contrarily, one might argue that advertisements such as the Superbowl or other big event adverts are viewed as entertainment and could lead to increases in viewership (Whitler 2022). However, typical ads in live streaming are much more comparable to regularly broadcasted television advertisements, making them less appealing to audiences and are subsequently viewed as a general nuisance (Shevenock and Meyers 2021; D'Anastasio 2022; Miceli 2022).

### *How Might Mid-Roll Advertisements in Live Streaming Encourage Consumption?*

Although we have presented the case for a negative response to the introduction of MRA ability, we posit that unique factors related to the live streaming space can allow for an overall positive consumption response. The key reason why the consumption of content may increase in response

to the introduction of MRA ability is due to streamers increasing their effort through strategic streamer adjustments. We argue that this positive effect is possible because streamers are (1) able to direct their own content and can make enhancements rather quickly and without much cost, and (2) are motivated by the monetary benefits which come from displaying MRAs (ad revenues). In traditional media platforms such as television, making adjustments quickly (such as lengthening the time of an episode or playing more episodes per week than is planned) is much less feasible. Also, such changes are rather expensive (e.g., increasing the quality of an episode).

Ultimately, a live streamer's behavior can change easily through increasing their effort towards streaming activities. Said another way, higher audience consumption (more viewers or higher total hours watched) can potentially result in a larger number of people being exposed to each MRA played. This can increase revenue potential, as live streaming platforms have historically paid streamers via CPM (Kinson 2020; Visuals By Impulse 2021; Parrish 2022). To capitalize on MRA ability, streamers can enhance their previous streaming habits through (1) increasing the average airtime of each streaming session, (2) by streaming more frequently, or (3) by improving overall stream quality. This results in the opportunity to gain new viewers or to increase individual watch time for higher overall ad exposure (through increased viewer impressions per ad or by coupling additional ads in the extra time streamed). Broadly, we posit that streamers are incentivized to increase content availability or content quality (or both).

With respect to increasing airtime or stream frequency (content availability), the additional broadcasting time allows for audience members who crave more content the chance to watch for a longer duration and allows viewers who were unable to join previous broadcasting sessions (for reasons such as scheduling conflicts) more opportunities to tune in. Importantly, increases in airtime or frequency are likely to contain some combination of additional

entertainment content along with ads, as solely filling the increased broadcasting time with ads alone would likely decrease audience consumption due to ad avoidance and including only content without ads would not provide ad revenue. As such, in this extra broadcasting time, more content is likely to be accompanied by more ads, benefiting both the viewer (consuming more content than previously) and the streamer (more ad exposure). Twitch itself has researched how MRAs should accompany content and recommends their streamers run approximately one to three minutes of ads for every 15 minutes of additional content (Twitch 2023).

As for content quality improvements, previous studies have found that signals of quality can improve outcomes for a firm (e.g., Porter and Donthu 2008). Hence, it reasons that increases in streaming quality should also improve consumption outcomes. Improvements to quality can manifest in a number of ways including improved microphones, streaming new content, increasing engagement with the audience, or utilizing a higher definition camera. As a result, the potential success of a stream can increase with quality.

Ultimately, we posit that these potential increases in effort through streamer adjustments may be viewed positively by audiences (on average). Information cues which signal increased effort can also result in beneficial gains as potential consumers favorably respond to these indicators (e.g., Morales 2005; Porter and Donthu 2008). Altogether, these streamer enhancements create potential for higher revenues. Coupled with the notion that live streamers often want more reliable and stable incomes through playing ads (Rubio-Licht 2022), it reasons that if MRA ability was provided on a live streaming platform, the majority of streamers would feel incentivized by the opportunity to increase revenue and capitalize by (1) actually playing MRAs and (2) finding ways to maximize the number of people exposed to each ad.

Alternatively, what about the possibility of a streamer choosing to reduce content and fill

their stream with MRAs during the session to obtain ad revenue? Due to the reduced original content per stream and increased advertisements, consumers should respond negatively and reduce consumption to avoid these ads. Thus, this form of streamer adjustments returns to the argument rooted in ad avoidance, so we do not discuss this case in greater detail.

Lastly, we consider the potential heterogeneity in both the main effects and proposed mechanisms (with regards to introducing MRA playing ability for live streamers). We explore differences between streamers themselves by investigating which streamers make more aggressive adjustments based on pre-MRA ability success, whether there was divergence between new and established streamers, if there were differences in outcomes between streamers who stream social versus solo content, and whether there may be variability in the observed consumption changes across time. We next discuss these factors and how they may differentially affect viewing consumption below in further detail.

First, we posit that there is typically a range of heterogeneity in terms of streamer success, even among partnered channels. In particular, successful channels (with moderate or high levels of consumption) are more likely to have multiple avenues of streaming-related income. This can include external sponsorships from brands (providing payment) or higher amounts of personal donations. Thus, successful channels are less likely to be heavily reliant on MRA ad revenues due to the diversity of their streaming-related income streams. On the other hand, less successful streamers are likely to be rather incentivized to make streamer adjustments as it is less likely for them to have other avenues of streaming-related revenue. Consequently, we predict that channels with low success are more likely to put in the effort to aggressively implement streamer enhancements after being provided with MRA ability.

Second, the introduction of MRA ability may be viewed differently by established

streamers when compared to late-entry streamers. From one standpoint, established streamers may have been hoping for the ability to play MRAs for quite some time and are accordingly prepared to make strategic adjustments. From another perspective, established streamers may be habitual and resistant to adjusting their behaviors. As such, it is unclear whether established or late-entry streamers receive larger changes to consumption after MRA abilities.

Third, we consider the content played by streamers themselves. One way to categorize content is from the perspective of whether the content itself is based on a social activity or is a solo endeavor. For example, video games are an extremely common content activity in the live streaming space. Some video games are primarily independent endeavors (“singleplayer”), where the focus of the streamer’s social interactions are with the audience. Other types of video games are social endeavors where the streamer must also simultaneously socialize with other players (“multiplayer”). On the one hand, the attention given to the viewing audience may be lower in a multiplayer setting as the streamer divides their limited social resources into interacting with both other players as well as their audience. On the other hand, social content may add an enjoyable element of viewing these social interactions between players. Thus, it remains uncertain whether the effects of MRA ability may vary between solo and social content.

Fourth, we consider whether the audience consumption changes vary across time after streamers receive the ability to play MRAs. Specifically, we posit that initial streamer adjustments may have strong effects on viewer consumption. However, we suggest that potential effects may also eventually become weaker over time due to diminishing marginal returns.

#### *Mid-roll Advertisements in Live Streaming: Research Questions*

We posit that both primary forces (ad avoidance and streamer adjustments) are likely to exist simultaneously, but discovering the net overall effect is key for marketing academics and

practitioners alike, as platforms or brands may choose to not pursue the implementation of MRAs if audiences negatively respond to watching streams which contain them (due to reductions in ad exposure). Conversely, platforms and brands may wish to aggressively invest in MRAs if viewing behaviors are not negatively affected by the introduction of these ads (or if they can generate even larger audiences). The answer to whether the ability to use MRAs increases or decreases viewer consumption is an empirical one. In addition, it is critical to explore if our proposed mechanisms of strategic adjustments are empirically present. Finally, investigating the heterogeneity related to both the mechanisms and consumption effects may be helpful for firms. Therefore, we formally propose our primary research questions as follows:

*RQ1: How does the introduction of MRA ability affect live streaming consumption behaviors?*

*RQ2: Do live streamers make strategic adjustments in response to MRA abilities?*

*RQ3: What heterogeneity exists for (a) live streamers and how they implement strategic adjustments and (b) audience consumption changes?*

## DATA AND METHODS

### *Research Setting and Data Description*

The “full dataset”<sup>5</sup> used in this study comes from a live streaming platform in 2020 for 84 consecutive days, from February 18, 2020, to May 11, 2020 (though we sometimes refer to these dates by their “number” in the panel sequence such that day 1 is equivalent to February 18). This platform was one of the most popular live streaming platforms at the time and was known widely across the globe.<sup>6</sup> Similar to the vast majority of live streaming services, there was no cost to access and view live streaming content on this platform. Critically, this data contains daily

---

<sup>5</sup> This dataset comes from an agency that collects data from various live streaming services using public application programming interfaces (APIs).

<sup>6</sup> Information about the platform and data setting was reconstructed to the best of our ability from online news articles and a live streamer “historical” survey. The articles are not listed to protect the identity of the platform.

information on every channel (each unique streamer) and streaming occurrence on the platform over this period to create a panel dataset. Each channel is associated with a unique ID and contains both information about the channel itself as well as audience metrics. Channel-related variables include the name of the primary content, the primary language of the channel, the primary target audience, the airtime duration of each stream, and whether the channel was partnered or not. On the viewership side, the data contains the average number of viewers (AV) for each channel on that day, as well as the cumulative number of hours watched by the audience for that day, which act as our primary dependent variables (the measures of consumption behaviors).

The dataset also contains the value for peak viewers (PV), or the maximum number of viewers seen at one point for that given stream (as viewer numbers fluctuate throughout a stream). To operationalize a measure of quality (used for our mechanism analysis), we utilize this variable to create a measure based on viewer retention. Specifically, we calculate a form of “negative retention” (NR) by taking the difference between an observation’s PV and average viewership (AV), then dividing it by the level of PV. This acts as an alternative measure of (dis)quality as it captures how poor a channel was at retaining audience members throughout the stream, relative to the level of peak viewers in that stream.

On March 17 of our data, the platform enacted a major policy implementation. The major implementation was the introduction of partnered channels being provided with the ability to utilize MRAs—we also refer to the intervention as “MRA ability” or “MRA abilities”. This is the key intervening policy change examined in this study, as partnered channels were thereafter given the ability to display MRAs during their streams. Critically, all partnered channels received this ability, regardless of whether they intended to use them or not (though we assume

that average usage was likely positive).

Like most other live streaming platforms, the selection of which channels became partnered was left to the streaming platform itself even if the streamer themselves requested to be partnered. Archived news articles (unlisted to protect the platform identity) suggest that the selection for partnered channels was rather unpredictable after meeting a low minimal quota for number of followers and stream time. Indeed, in our data, the minimum number of viewers was 0 in the pre-intervention period across both partnered and non-partnered channels (likewise in the post-intervention period), implying that the low minimal requirements to be partnered did not imply a positive audience size and that selection was to some degree, unsystematic.

We took this dataset and aggregated it into 12 time periods. There are several reasons for this. First, aggregation of data for the purposes of our empirical estimation (DiD) provides a solution to address problems linked to serial correlation and incorrect grouped error terms (Bertrand, Duflo, and Mullainathan 2004; Gill, Sridhar, and Grewal 2017). Second, aggregation helps greatly with computational speed due to the large size of the full dataset. We first aggregated our data by each unique channel identifier for the 12 pre-determined time periods, each based on 7 days of observations. Periods five and beyond (March 17 and after) reflect the time period after the policy, and periods one to four indicate the time periods prior to the policy. We refer to this dataset as the “main dataset”.

Further details of the aggregation procedure are presented in Appendix A. Though the full dataset originally included extra days of data, we removed these days due to potential external events in order to keep maintain conservative empirical estimates (we address concerns about external events and further replicate the results with all the data in Appendix B). Hence, the main dataset contains 12 time periods of 7 days each. There are 4,716,929 unique channels



which streamed on the platform over this time, with a total of 12,511,578 channel-period observations over the 84 days. This is the primary dataset used in the main paper. Next, a small number of other minor sub-policies (primarily based on minor quality-of-life updates which are common in digital platforms) were additionally updated to the platform in addition to the main MRA policy, but we provide evidence in Appendix C that these minor adjustments were unlikely to significant effects on consumption behaviors. Finally, for robustness, we re-conduct our main analysis using an alternative aggregation method and present the results in Appendix D (verifying the main results).

Summary statistics of the main data can be found in Table 2. This table splits the data into the groups necessary for our DiD analysis: being partnered or not. Said another way, we present summary statistics of each group required for our analysis and this division is based on whether the channel ever received the treatment policy of being partnered. We further divide this data into the pre- and post-intervention periods (also required for our empirical analysis).

From this table, we can see that the typical partnered channel observation had higher viewership, total hours watched, airtime and NR over the non-partnered channels. Both partnered channels and non-partnered channels primarily targeted mature audiences of 18 and over with the primary channel language being English. We only include the top language and audience types instead of the comprehensive list due to spacing issues and because the vast majority of observations belong to these categories (there are 34 unique languages and 3 audience types in total). Lastly, the three categories listed were selected due to their simultaneous popularity across treatment groups. These content activities account for a substantial amount of platform content in each treatment group. Another reason we only display these three content categories is due to the vast number of unique categories (6,972).

Table 2: Summary Statistics Between Treatment and Non-Treatment Observations

Description		Non-Partnered		Partnered	
Treatment Time		Pre	Post	Pre	Post
N		3,026,488	9,473,080	3,993	8,017
Continuous Variables: Mean (Standard Error)					
$\ln(\text{Average Viewers})$	Avg. number of concurrent viewers	0.141 (0.342)	0.141 (0.333)	3.633 (1.048)	3.723 (1.039)
$\ln(\text{Hours Watched})$	Total Hours Consumed	0.149 (0.476)	0.141 (0.446)	4.891 (1.398)	5.057 (1.383)
$\ln(\text{Airtime (in minutes)})$	Duration the stream was live	3.421 (1.130)	3.402 (1.131)	5.356 (0.621)	5.432 (0.601)
$\ln(\text{Frequency})$	Occurrences streamed	0.356 (0.527)	0.389 (0.542)	1.426 (0.539)	1.517 (0.489)
$\ln(\text{Negative Retention})$	(PV-AV)/AV	0.534 (0.242)	0.715 (0.226)	0.416 (0.102)	0.860 (0.413)
Audience Type (% of total)	Target group of the stream				
Mature/18+		69.0%	67.2%	64.1%	67.6%
Primary Language (% of total)	Primary language of the channel				
English		86.1%	87.1%	88.1%	88.4%
Content (% of total)	Primary content displayed				
Top Activity 1		5.7%	8.6%	13.1%	19.7%
Top Activity 2		31.9%	33.8%	12.9%	8.9%
Top Activity 3		5.5%	3.3%	13.9%	7.6%

Note: Continuous variables are presented as average period values. There are some minor differences in sample size for categorical variables due to missing values. We report statistics for the natural log-transformed versions of most continuous variables, as they act as the main dependent variables in the paper. For simplicity, we occasionally refer to the log-transformed versions of the dependent variables using the “absolute” version, depending on the context.

*Historical Live Streaming Platform Survey (Further Data Context)*

We next provide further descriptive details related to the live streaming platform. Information of

the live streaming platform was evaluated from a number of archived news articles. From these articles, we ascertained details such as the date of the MRA ability intervention or the fact that the live streaming platform in our data was free to use for audience members, like many other competitors. However, to expand our understanding of the platform and the data setting, we asked real live streamers ( $N = 25$ ) who streamed on the focal platform at the time of the data, were partnered, and played MRAs, to complete an informational survey. Though not causal in nature, this survey provides a deeper general understanding of the historical or institutional context of our data. In particular, we asked these “historic streamers” a number of questions related to the platform and their behaviors (in what we call our “historical survey”).

We draw and verify several assumed conclusions based on this survey. First, streamers appeared to indeed generate MRA revenue from a CPM model. In particular, streamers received approximately two to three cents of revenue for each MRA view. Second, streamers generated income from a variety of sources including MRA-related ad revenue, channel subscriptions, external sponsorships, and donations by viewers. MRAs contributed about 6% of their total streaming related income. The other primary sources of streaming related income were subscriptions, external sponsorship support, and viewer donations (approximately 44%, 13%, and 29%, respectively). In the *Managerial Implications* section, we show that the income generated from MRAs was rather meaningful based on the information from this historical survey. Moreover, we leave the study of other income streams to future research, as the focus of this paper is focused on MRAs. Third, streamers did not appear to have control over the content contained in each MRA. Fourth, the vast majority of respondents indicated that the introduction of MRAs either increased or did not affect income generated from other income sources including subscriptions, external sponsorships, and tipping. This suggests that revenue

substitution by viewers was unlikely to exist. Lastly, most respondents thought generating MRA ad revenue was appealing and believed that the majority of other streamers on the platform would play MRAs (if possible). In short, these findings support the news articles and assumptions made in our data analysis. These survey details can be found in Appendix E.

### *Econometric Specification*

The focal policy presents itself in an ideal quasi-experimental setting to examine the impact of MRAs. In particular, we leverage a difference-in-differences (DiD) approach, which addresses a host of endogeneity issues and allows for the consistent identification of parameters conditional on the parallel trends assumption holding (Angrist and Pischke 2009; Goldfarb, Tucker, and Wang 2022). Estimating quasi-experimental settings using DiD is a common strategy in the marketing literature (e.g., Datta, Knox, and Bronnenberg 2018; Janakiraman, Lim, and Rishika 2018; Fisher, Gallino, and Xu 2019). DiD estimators possess the capability to capture the average treatment effect on the treated (ATT). Importantly, validity of the parallel trends assumption allows the biased selection or non-randomized assignment of the treatment as long as any selection bias is consistent across time (Roth et al. 2023). We assume that this was likely to be the case for this study, since the platform itself (not streamers or the channels themselves) ultimately decided which channels were to become partnered accounts and which did not in both the pre-treatment and post-treatment periods.<sup>7</sup> Moreover, the ATT in the context of this study refers to MRA ability only (as all partnered channels received this ability). This did not necessarily guarantee usage. Thus, we can alternatively think of this as the intent-to-treat (ITT) effect of MRA usage itself, although we maintain (and support via our two surveys) that it is rather likely that usage was non-zero due to the monetary incentives for streamers.

---

<sup>7</sup> Using the full dataset, we also find similar amounts of partnered observations in both the pre-treatment and post-treatment periods (proportion averages of approximately 0.3% and 0.2%, respectively).

We now apply the estimation method to the context of our study. The MRA ability main policy on March 17 (starting at period 5) was implemented for partnered channels only, thus making this segment the treatment group. In comparison, non-partnered channels did not receive the treatment and are thereby classified as the counterfactual control group. We specify our baseline DiD specification as follows:

$$(1) \text{ConsumptionBehavior}_{it} = \alpha + \beta T_i + \delta \text{Post}_t + \lambda(T_i \times \text{Post}_t) + \varepsilon_{it}.$$

In this specification,  $\text{ConsumptionBehavior}_{it}$  represents one of two outcome variables: the average viewership or the total (cumulative) hours watched for a given channel  $i$  at time  $t$ .  $T_i$  represents an indicator variable of whether the channel belongs to the treatment group of ever being partnered, and  $\text{Post}_t$  is another indicator variable which denotes the day the policy shock was introduced and all time periods afterwards (period 5 and beyond). Next,  $\varepsilon_{it}$  represents the error term while  $\alpha$  is the constant term. Most importantly, the parameter  $\lambda$  reflects the focal DiD term,  $T_i \times \text{Post}_t$ .

To account for endogenous concerns related to omitted variables, we include a set of covariates  $X_{it}$  in our DiD specification, which contains controls for the primary language of the streamer, the content streamed, and the target audience. These control variables are included to account for channel-level differences. One concern which would exist if we only included  $X_{it}$  in our next specification is that there could be potential issues of time-invariant heterogeneity as well as time-specific heterogeneity across channels and days. To address these issues, we exploit the panel nature of our data and include both channel and time fixed effects (FE). This FE approach is a common inclusion in DiD specifications with panel data (also called two-way fixed effects or TWFE). We further note that the main effects of  $T_i$  and  $\text{Post}_t$  are then removed from this specification, as they are captured by the incorporation of the two types of fixed effects (Goldfarb et al. 2022). Accordingly, our primary specification is:

$$(2) \text{ConsumptionBehavior}_{it} = \mu_i + \gamma_t + \lambda(T_i \times \text{Post}_t) + \tau X_{it} + \varepsilon_{it}.$$

In Equation 2<sup>8</sup>,  $\mu_i$  and  $\gamma_t$  capture channel and time FE, respectively. Parameter  $\lambda$  denotes the ATT of MRA ability. Robust standard errors are two-way clustered by the streamer and time.

## RESULTS

### *Support for Parallel Trends and Baseline Results of Focal Policy Shock*

One common way to provide support for the parallel trends assumption is by running the main DiD estimation with a placebo or “fake” treatment indicator and restricting the sample to the period prior to the actual intervention (e.g., Datta, Knox, and Bronnenberg 2018; Janakiraman, Lim, and Rishika 2018). This DiD placebo parameter is then inspected to see whether there is a statistically significant result. This determines whether there was a relative difference in the rate of change between the treatment groups between periods. Hence, a null result in the focal parameter between periods would suggest no difference and provides evidence to support the parallel trends assumption. Thus, we run a variant of the main estimation procedures by running three separate placebo regressions between each of the pre-treatment periods, in succession. We run separate placebo regressions between periods 1 and 2, 2 and 3, and 3 and 4. For each test, we set the latter period as the treatment period and the prior period as the non-treatment period. For each regression, we replace the actual treatment time ( $\text{Post}_t$ ) with a placebo.

We present the results in Table 3, using both the absolute values and the log-transformed versions of the dependent variables. As all of the focal parameters were found to be statistically insignificant, the placebo tests suggest that the rate of change over time with respect to both dependent variables was the same when comparing treated and untreated observations prior to the policy. Thus, this analysis supports the parallel trends assumption as both groups were

---

<sup>8</sup> We use the statistical package from Correia (2014) to estimate this equation, which also removes singleton observations.

trending similarly for at least 28 days prior to the policy. In Appendix F, visualization of these tests is provided through model-free graphical evidence of this main dataset, which visually appear to support the empirical placebo results.

Table 3: DiD Placebo Estimations

Dependent Variable	Focal $\lambda$ Parameter Estimates		
	P1 to P2	P2 to P3	P3 to P4
Average Viewers	-1.416 (3.549)	1.516 (1.764)	-5.511 (1.892)
$\ln(\text{Average Viewers})$	0.035 (0.021)	-0.010 (0.021)	-0.099 (0.018)
Hours Watched	-1.479 (39.520)	4.028 (15.508)	-4.260 (17.178)
$\ln(\text{Hours Watched})$	0.024 (0.032)	-0.005 (0.033)	-0.094 (0.029)
N	487,008	464,234	475,716

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Each cell represents the focal  $\lambda$  parameter of a separate TWFE DiD regression. “P” refers to period. Covariate controls and FE are included.

We next turn our attention to the baseline estimation of the effect of MRA ability on consumption behaviors, which is a formal test of Research Question 1. The results of the baseline DiD analysis are presented in Table 4. The main identified parameter of interest is  $\lambda$ , which represents the effect of MRA ability on the partnered accounts. The number of observations is lower than the main dataset due to the removal of singleton observations (Correia 2014) or missing covariate values—this also occurs throughout the remainder of the paper.

The DiD results support a positive effect of MRA ability on the partnered accounts by inducing increased levels of consumption behaviors with respect to both average viewers and hours watched. As we mentioned previously, this is a rather interesting result as ad avoidance predicts a negative effect, rather than a positive one. However, all main estimates of the focal parameter ( $\lambda$ ) were found to be positive and statistically significant. In particular, we find a 7.818 increase in average viewers for partnered channels after the policy introduction relative to non-partnered ones (Column 1  $\lambda = 7.818$ ,  $p < .05$ ). Moreover, we show that after policy A, that the

total hours watched significantly increased by 122.078 hours for partnered channels when compared to non-partnered channels (Column 3  $\lambda = 122.078$ ,  $p < .01$ ). In terms of percentage change, we find that partners had approximately 9.3% more average viewers (Column 2  $\lambda = 0.089$ ,  $p < .05$ ) and 16.3% higher total hours watched (Column 4  $\lambda = 0.151$ ,  $p < .01$ ) when compared to non-partnered channels after the shock.

Table 4: Baseline DiD Results of Consumption Behaviors

	(1)	(2)	(3)	(4)
	Average Viewers	$\ln(\text{Average Viewers})$	Hours Watched	$\ln(\text{Hours Watched})$
$T_i \times \text{Post}_t$	7.818** (3.255)	0.089** (0.030)	122.078*** (32.269)	0.151*** (0.032)
Constant	0.395*** (0.002)	0.144*** (0.000)	1.465*** (0.023)	0.158*** (0.000)
Observations	9,254,233	9,254,233	9,254,233	9,254,233
R-squared	0.884	0.732	0.906	0.795
Controls and FE	Yes	Yes	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . The transformation  $(e^\lambda - 1)$  was used for interpretations in the main text for log-transformed dependent variables. “Controls and Fixed Effects (FE)” in this paper refers to controlling for the three main content categories of target audience, primary language, and content activity as well as the inclusion of both streamer and period fixed effects.

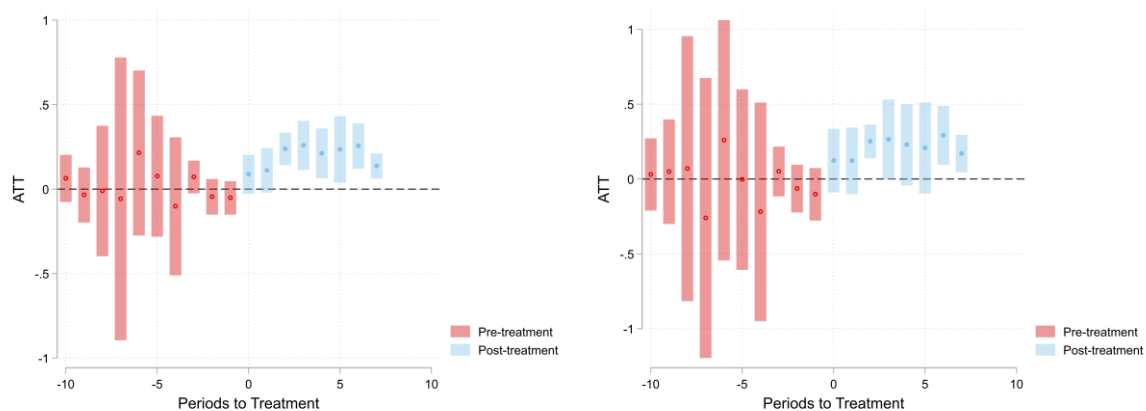
These results are large, positive effects which suggest substantial revenue and profit implications for both streamers, brands, and platforms. Finally, as a baseline robustness check, we also re-run these specifications by adding the following additional controls: (a) the one-period lagged dependent variable and (b) number of times the channel streamed in each period, recovering similar results (positive and statistically significant parameters). These are presented in Appendix F. Overall, we emphasize that these results reflect the overall net effect of the policy, after accounting for any competing forces between ad avoidance and streamer adjustments. These results are also a direct test of RQ1, which asks how the introduction of MRA ability affects live stream viewing behaviors.



## Addressing Staggered Intervention

Recent studies in the economics literature have introduced new methods to address staggered treatments in DiD estimations. Staggered treatments occur when treated units receive the intervention at different times, which may not generate the correct parameters under TWFE (Roth et al. 2023). Since some non-partnered channels may convert to partnered channels (as chosen by the platform) at different times after the main policy, our data setting can be viewed as one with staggered interventions (though this is a small portion of the data). Callaway and Sant’Anna (2021) (henceforth also known as CS) propose several similar estimators which generate dynamic DiD parameters that are relative to treatment time. Importantly, a generalized propensity score based on observable covariates is created based on each initial treatment period to predict the likelihood of being treated in a given time, thus suggesting that parameters are calculated based on comparable treated and non-treated observations (Callaway and Sant’Anna 2021; Cunningham 2021). We conduct this analysis and present the CS dynamic event study plots for both consumption behaviors in Figures 1a and 1b, respectively.

Figures 1a and 1b: CS Dynamic DiD Event Study Plots



Note: The left plot (1a) displays the dynamic event study plot for  $\ln(\text{average viewers})$  as the dependent variable, whereas the right one (Fig. 1b) represents the same plot for  $\ln(\text{hours watched})$ . The 95% confidence intervals are displayed.  $N = 2,507,176$  for both CS estimations.

The dynamic ATT's from the CS estimates are relative to the initial treatment period. From the figures, we can see that all pre-MRA ability ATT's are based around zero, providing continuing support of the parallel trends assumption (all pre-treatment ATTs were non-significant). After the intervention, we found statistically significant positive overall dynamic CS ATT's of the policy on both average viewers (ATT = 0.192,  $p < .01$ ) and total hours watched (ATT = 0.207,  $p < .01$ ). These results mirror our baseline analysis. Further details of this analysis are found in Appendix G. As the results appear to be approximately similar to those using TWFE, we utilize TWFE as our primary estimation method for the remainder of this paper.

#### *Further Robustness Checks*

To provide further support for the main results, we conduct a number of further robustness checks. First, though the CS estimators do also utilize propensity scores, we conduct a more standard propensity score matching (PSM) and weighting DiD analysis, similar to Ananthakrishnan, Proserpio, and Sharma (2023). Second, we conduct a synthetic difference-in-differences analysis (Arkhangelsky et al. 2021), which accounts for the dynamic nature of our DiD setting (like CS does), but additionally ensures the parallel trends through a re-weighting process. Third, we conduct an attrition test using only streamers who streamed in every period. These robustness checks are found in Appendix H, which all continue to support a positive and significant effect of the intervention on the treated partnered channels, relative to the control.

#### *Addressing Threats to Validity: Coronavirus (COVID-19) Pandemic*

One concern with our analysis is that the dataset overlaps in time with the first major wave of the COVID-19 pandemic which began in March 2020. On March 11, 2020, the World Health Organization (WHO) declared an official pandemic due to COVID-19, which was viewed as the onset of lockdowns and the commencement of households staying at home (Thebault, Meko, and

Alcantara 2021). Thus, one may wonder whether this initial lockdown, and not the focal MRA ability intervention, increased live streaming viewing behaviors for partnered streamers, as the implementation of the lockdowns was close to the MRA ability policy (March 17).

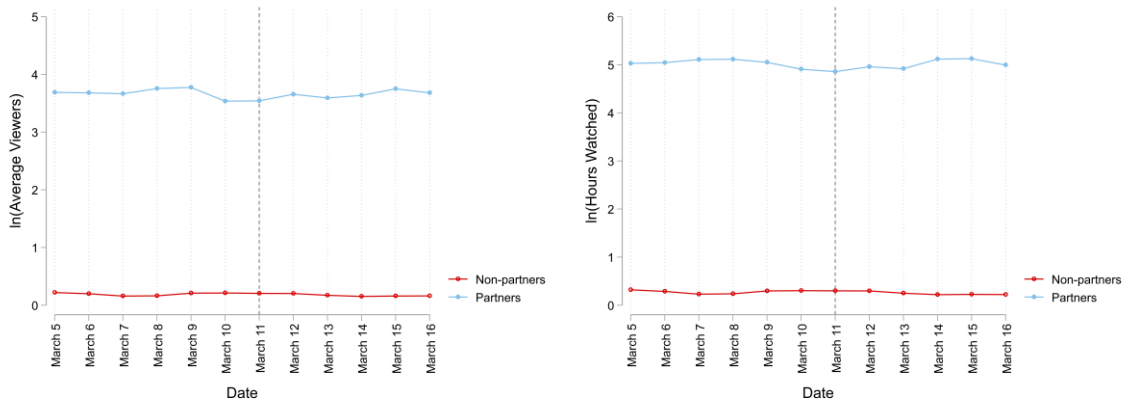
We do not believe that COVID-19 is a primary driver of our results for a number of reasons. First, differences between covariates are accounted for through the PSM and CS estimates. For example, COVID-19 lockdowns could have generated larger audience sizes, and these new viewers may have preferred the content played on partnered streamers over non-partnered ones. Our PSM and CS estimates respond to this threat, as after the matching procedure(s), analysis was conducted on observations with more similar levels of characteristics and observable covariates. Second, our inclusions of FE already account for cross-sectional and time invariant effects. Third, the pre-treatment trends between groups are rather parallel (e.g., Figures 1a and 1b), meaning that other exogenous shocks prior to COVID-19 did not affect the treatment groups differentially. Hence, there is less reason to believe that COVID-19 might also differentially affect partnered channels when compared to non-partnered ones.

However, to address the COVID-19 concern more directly, we conducted two empirical tests to provide evidence that the pandemic is an unlikely driving factor of our results. In the first test, we measure the short-term impact of the lockdowns to further show that there is no visible differential effect on consumption behaviors between groups. In particular, since there is a six-day window to measure the isolated effect of the lockdowns prior to the MRA ability policy, we run our DiD analysis and use March 11 (the lockdown day) as the date of the “intervention” to isolate the effect of the initial pandemic lockdown itself. Specifically, we take the full dataset and consider the timeframe from March 5 to March 16 and run a DiD on this subsample with March 11 as the treatment date (all days are prior to the MRA ability intervention). A null effect

of the main parameter would suggest that the lockdowns (in the context of this time frame) did not affect the level of average viewers or total hours watched for partnered channels only, thereby implying that the lockdowns did not affect the two groups differentially.

We display the plots of this analysis by displaying the log-transformed dependent variables over time in Figures 2a and 2b. We provide the formal regression results in Table A15 of Appendix I. No main parameters are found to be statistically significant. Figures created using the standard outcome variables also did not display any notable changes after March 11. These results provide support for the notion that at least in the short-term, the lockdowns did not generate a positive effect on consumption behaviors for partnered channels only.

Figures 2a and 2b: Short-Term Effects of COVID-19 Lockdowns on Consumption Behaviors



For further assurance that the lockdowns did not differentially affect partnered channels, we run a second empirical test. In particular, we run a specification controlling for potential between-group differential effects of the pandemic lockdowns. Specifically, we leverage data from the WHO, who collected data on daily COVID-19 cases from most countries (World Health Organization 2020). Using the primary language of the streamer to assume audience location, we include the interaction between new daily COVID-19 deaths and the focal term in our DiD analysis to control for differential effects of the pandemic. This analysis is presented in

Table A16 of Appendix I, and the results continue to indicate a positive and statistically significant focal parameter with regards to both average viewers (Column 1  $\lambda = 0.047$ ,  $p < .05$ ) and total hours watched (Column 3  $\lambda = 0.087$ ,  $p < .01$ ) despite controlling for the interaction, suggesting that the COVID-19 lockdowns are an unlikely driving force in our data setting.

Altogether, both tests indicate that the lockdowns were not a prominent factor in our analysis.

#### *Addressing Threats to Validity: Testing for Potential Violation of SUTVA*

##### *(Interference/Spillovers)*

Next, we address a second threat to validity. Interference (e.g., Johari et al. 2022) could arise due to an overwhelming shift in viewer allurements from the untreated to the treated group (spillover). Said another way, a viewership spillover between treatment groups would violate the Stable Unit Treatment Value Assumption (SUTVA), which is required for DiD analysis. However, we believe this is unlikely to be a focal driver of our results for several reasons. First, PSM considers smaller or moderate streams in order to match partnered observations with non-partnered ones (for example, the maximum observation of average viewers was 12,051.6 in the main dataset but only 9925.3 in the PSM sample). The largest streamers are expected to be the driver of interference since they are the most likely ones siphoning viewers from non-partnered channels, but our robustness checks with the matching procedures continued to support a positive effect. Second, characteristics between the treatment groups are more comparable after PSM and CS, meaning there are fewer reasons to prefer one channel over another. Third, streaming audiences who wish to watch several live streams simultaneously can use multiple browser tabs or use websites which allow viewers to watch several streamers on the same screen—and platforms themselves (such as Twitch) offer their own features so audiences can watch multiple streamers concurrently (Robinson 2022a). Lastly, non-partnered channels do not observe a drastic drop in

consumption in the during the post-periods as the interference argument suggests (e.g., Table 2).

For empirical assurance, we run two tests to demonstrate that interference (spillover) is unlikely to be a focal driver of our main results. In particular, we re-run our DiD analysis but ensure that the primary target language used for the treated group observations is different than the control group observations. After analysis using treatment groups with different languages, a focal DiD parameter that is still positive and significant would suggest that the effect is due to the MRA ability intervention in isolation and not due to the spillover of non-partnered viewers switching to partnered viewers, because audience members are far less likely to switch to another stream (on the aggregate) that speaks a different language than them. For example, an individual who typically watches a French speaking non-partnered streamer would be rather unlikely to switch to a Spanish speaking partnered stream due to the policy shock.

To that end, we run a series of DiD regressions using a variant of the full dataset (to ensure a large sample size) containing only observations where French or Spanish was the primary language of the stream. We run two series of tests for robustness. The first test (T1) uses French-oriented observations for the partnered treated group, whereas the control group is comprised solely of Spanish speaking streams. The second test (T2) replicates the T1 findings by using French targeted observations as the control group and Spanish streams as the treatment group. We selected these two languages due to the large prevalence of streamer observations using these languages in our data to maintain a large sample size, and to avoid issues with language “contamination” (for example, English is a language that is an official language across many countries). The findings are presented in Table 5, where we find positive and statistically significant effects of the intervention increasing viewing consumption of the treated group relative to the control, thus mirroring the main results. Consequently, we believe that interference

or spillover concerns are minimal and do not negate our main results.

Table 5: Spillover DiD Robustness Test

	(1)	(2)	(3)	(4)
	T1: $\ln(\text{Average Viewers})$	T1: $\ln(\text{Hours Watched})$	T2: $\ln(\text{Average Viewers})$	T2: $\ln(\text{Hours Watched})$
$T_i \times \text{Post}_t$	0.391** (0.191)	0.699** (0.334)	0.339*** (0.125)	0.535*** (0.187)
Constant	0.123*** (0.000)	0.171*** (0.000)	0.150*** (0.000)	0.218*** (0.000)
Observations	660,683	660,683	402,682	402,682
R-squared	0.667	0.690	0.717	0.749
Controls and FE	Yes	Yes	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . The short form “T” refers to “Test”. T1 uses French observations as the treated group and Spanish as the control, whereas T2 uses Spanish observations for partnered channels with French as the control.

### *Mechanism Investigation*

Given that we empirically obtain a robust positive effect of MRA ability on consumption behaviors, we argue that ad avoidance is not the focal mechanism behind the observed effects. Instead, we consider whether any meaningful streamer adjustments were made which could suggest a net increase in consumption behaviors. Stream airtime, stream frequency and quality are all proposed to increase as a consequence of streamers adjusting their streaming habits in the hopes of generating higher ad revenue, as it is very common for live streaming platforms to pay streamers through CPM. Ultimately, adjustments are possible due to the unique nature of live streaming, as streamers have great flexibility in making strategic adjustments (content availability and content quality). We reasonably assume a CPM based revenue model for streamers in our study (verified with our historical survey), as near the time of our data (2020), other large platforms, including Twitch, Facebook Gaming and YouTube also ran CPM based schemes (Kinson 2020; Visuals By Impulse 2021). Online news articles also suggest a CPM payment scheme for MRAs in our focal platform. As we previously argued that partnered

streamers should make streamer adjustments (irrespective of what non-partners do), we consider whether partnered streamers increased their airtime, stream frequency or quality after the intervention. For operationalizations, the mechanism variables used were log-transformed values of airtime, stream frequency (the number of times a channel streamed in a given period), and NR (which was our measure of quality).

We conduct a formal DiD analysis on the main dataset to explore whether partnered channels made strategic adjustments relative to the control group. The results are presented in Table 6. From this analysis, we find a positive and statistically significant effect of the intervention increasing airtime (Column 1  $\lambda = 0.037$ ,  $p < .01$ ) and reducing NR (Column 3  $\lambda = -0.022$ ,  $p < .01$ ). We do not find statistical evidence of streamers increasing frequency, although the focal parameter is still positive. Regarding NR, a negative focal parameter would thus suggest that streamers reduced their negative retention and hence, increased their quality. Thus, we provide evidence of strategic adjustments by partnered channels. Appendix J contains further details of this analysis, including support for the parallel trends assumption and additional robustness checks.

Table 6: DiD Focal Mechanism Results

	(1)	(2)	(3)
	$\ln(\text{Airtime})$	$\ln(\text{Frequency})$	$\ln(\text{NR})$
$T_i \times \text{Post}_t$	0.037*** (0.010)	0.025 (0.027)	-0.022*** (0.005)
Constant	3.473*** (0.000)	0.391*** (0.000)	0.536*** (0.000)
Observations	9,254,233	9,254,233	4,294,544
R-squared	0.544	0.507	0.464
Controls and FE	Yes	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . The sample size is smaller in the NR analysis due to occasions when average viewers equal zero, which created empty values.

In addition, we briefly consider ad avoidance from an empirical standpoint, which would



be consistent with the time watched per viewer decreasing. Although we do not have the number of unique viewers to calculate this variable precisely, we construct an approximating measure by taking the total hours watched and dividing by the number of average viewers. We run our DiD analysis using our constructed time watched per viewer measure as the dependent variable. We find that the focal parameter from this analysis is positive and significant ( $\lambda = 0.176$ ,  $p < .01$ ), suggesting that ad avoidance does not play a large role in our data setting and that channel surfing concerns are minimal, and if anything, viewers consume more content on average.

Taken together, our results suggest that partnered streamers adjust their streaming behaviors when MRA ability is introduced through quality as well as increased airtime, and that these changes lead to increases in consumption. Relatedly, our findings suggest that the positive influence of streamer adjustments are stronger than potential negative effects of ad avoidance, thus creating an overall net positive impact on consumption.

For completeness, we conduct a formal mediation analysis using the approach introduced by Baron and Kenny (1986). The Sobel test statistics of this mediation test for each of the three separate strategic adjustments (for both consumption variables) are presented in Table 7. Comparable with the TWFE DiD estimates, we find evidence of receiving MRA abilities positively increasing consumption behaviors primarily through the NR (quality) variable. We find mixed findings using airtime as a mediator and find no evidence using frequency. In short, this provides continuing support of quality improvements as the primary mechanism of the focal policy on consumption. Additional details of this analysis are found in Appendix K.

Table 7: Sobel Test Statistic for Formal Mediation Analysis

Dependent Variable	Mechanism Variable (Strategic Adjustment)		
	$\ln(\text{Airtime})$	$\ln(\text{Frequency})$	$\ln(\text{NR})$
$\ln(\text{Average Viewers})$	-3.75623469***	0.95609240	4.77288669***
$\ln(\text{Hours Watched})$	3.77230047***	0.9569710	4.65226003***

Note: \*\*\*  $p < 0.1$ . Columns 1 and 2  $N = 9,254,233$ . Column 3  $N = 4,288,343$ .

Ultimately, these results make sense for streamers who wish to make a living from live streaming (or at least to see how far they can get). Indeed, live streamers such as Valkyrae or Pokimane have spent tremendous amounts of time and effort into creating new content when they were smaller streamers, by streaming for long durations, and by improving their quality in order to grow their streams (Majumdar 2022; Morgan and Gray 2022). Therefore, many streamers are likely to increase their effort in response to revenue incentives.

#### *Heterogeneity Analysis: Differences by Initial Streamer Success*

Previously, we conceptually suggested that less successful streamers were more likely to be incentivized by MRA ad revenue than successful ones (who likely have other sources of streaming-related income) and would therefore input more effort into making streamer enhancements. We next classify the data into “low” (low success), “moderate” (moderate success) and “high” (high success) categories in relation to their viewership performance in the pre-treatment periods. Using the main dataset, we conduct a TWFE DiD analysis for each sub-group and report the DiD estimates of each regression. The results are presented in Table 8.

From Panel A, we note that each group adjusted their strategies differently. First, low performing streamers made aggressive improvements to airtime and quality. Second, moderate streamers made more modest adjustments to airtime and quality, but increased frequency as well. Finally, high streamers only made improvements to frequency and quality. Ultimately, all three mechanisms appeared to be viable strategies, but were utilized to various degrees.

The results of these adjustments are reflected by the sub-group changes in consumption (Panel B). It appears that despite differences in their implementations of strategic adjustments, all three groups appeared to receive increases with regards to audience consumption. However, low streamers seem to have generated the largest increases in consumption, but also made the

most aggressive adjustments. Further details of this analysis are found in Appendix L.

Table 8: DiD Parameter Estimates by Initial Viewership

Panel A: Mechanism Variables	Focal Parameter Estimates		
	Low	Moderate	High
$\ln(\text{Airtime})$	0.264** (0.100)	0.086*** (0.012)	0.038 (0.065)
$\ln(\text{Frequency})$	0.121 (0.089)	0.090*** (0.024)	0.207** (0.081)
$\ln(\text{NR})$	-0.065** (0.021)	-0.023*** (0.003)	-0.023** (0.009)
Panel B: Consumption Variables			
$\ln(\text{Average Viewers})$	0.488*** (0.148)	0.146*** (0.019)	0.334** (0.126)
$\ln(\text{Hours Watched})$	0.748*** (0.216)	0.227*** (0.024)	0.391*** (0.106)
N	6,112,488	28,203	3,112,101

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Controls and FE are included in estimations. Each cell represents the  $\lambda$  parameter and the corresponding standard error from separate estimations. For the NR cells, N = 2,940,493 for the “Low” group, N = 28,144 for the “Moderate” group, and N = 1,318,387 for the “High” group.

#### *Heterogeneity Analysis: Established vs. Late-Entry Streamers*

Next, we explore how established streamers may have implemented different strategies differently than newer (late-entry) streamers. To do so, we consider the time of the first appearance of a streaming channel. If the streaming channel had an observation in periods one or two, it suggests that the streamer was quite active for some time prior to the introduction of the focal policy. In contrast, we consider late-entry streamers who were only active prior to the MRA ability policy for a short duration prior or those who started after the intervention (periods 3 and after). We run our TWFE DiD analysis based on these sub-samples. We present the results in Table 9. The results suggest that established streamers are the group that generally make strategic adjustments. Consequently, these are the streamers that receive the bulk of the increases to viewing consumption.

Table 9: DiD Parameter Estimates by Established vs. Late-Entry Streamers

Panel A: Mechanism Variables	Focal Parameter Estimates	
	Established	Late-Entry
$\ln(\text{Airtime})$	0.034** (0.012)	0.174 (0.154)
$\ln(\text{Frequency})$	0.034 (0.026)	0.058 (0.092)
$\ln(\text{NR})$	-0.023*** (0.004)	-0.019 (0.018)
Panel B: Consumption Variables		
$\ln(\text{Average Viewers})$	0.089** (0.030)	0.049 (0.154)
$\ln(\text{Hours Watched})$	0.149*** (0.030)	0.132 (0.221)
N	4,536,036	4,717,089

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Controls and FE are included in estimations. Each cell represents the  $\lambda$  parameter and the corresponding standard error from each estimation. For  $\ln(\text{NR})$ ,  $N = 2,271,273$  for the “Established” estimation and  $N = 2,016,029$  for “Late-Entry”.

*Heterogeneity Analysis: Solo vs. Social Content*

We next investigate whether the sociability of the content activities a streamer played affected the impact of the focal MRA ability policy. To that end, we classify observations by content activities. We classify content activities into one of three general categories: solo (independent activities), social (multi-person activities), or both. Then, we run our DiD analysis on each subgroup to see if the effect holds for the set of observations in each category. Interestingly, our analysis suggests that positive increases to audience consumption after the focal policy on the treated group were found primarily in content activities that can prompt some degree of social interaction outside the audience. This analysis is detailed further in Appendix M.

*Heterogeneity Analysis: Treatment Effect Across Time*

One final consideration we make is whether the MRA ability intervention maintained persistent effects across time. Fortunately, the CS analysis conducted previously calculates the dynamic ATT’s across time for our data relative to the first treatment period. Hence, we can use the CS

analysis to observe how the treatment effects varied across time. From Table A8 in Appendix G, we can see that the increases in viewership do not occur immediately (in the first treatment period). After one or two periods after the initial treatment, the ATT's become positive suggesting that the consumption effects of the policy begin to work only after some time has passed. The positive and statistically significant effects then generally increase, peak, and eventually begin to decrease in magnitude (7 weeks after initial treatment). This pattern is also visually reflected in Figures 1a and 1b. Hence, it appears that the effects of the MRA policy begin to waver to some degree after approximately 7 weeks after the initial treatment week, suggesting that streamer adjustments may eventually be met with diminishing marginal returns.

#### VERIFICATION SURVEY

Lastly, we conducted a “verification survey” to corroborate our core findings, as we wished to verify that streamers were indeed incentivized by ad revenue, would make some sort of strategic adjustment, and would play some positive number of MRAs after being provided the ability to play them. This survey adds to our secondary data research because we were unable to observe whether streamers actually played MRAs when given the option to do so (though the historical survey already suggests that at least a reasonable proportion did play positive levels of MRAs). In the survey, we asked respondents to imagine a scenario where they were live streamers who previously could not play ads and were suddenly granted MRA abilities. We then explained that they could make ad revenue for every individual who views a displayed MRA and then asked about intended behaviors regarding this scenario. We ran a pre-test on students drawn from a large U.S. university (N=127) and then sent the main survey to real live streamers (N=85) who had recently been streaming on a live streaming platform. A summary of the results are shown in Table 10. We find that 77.65% of live streamers surveyed indicated that generating ad revenue

would be important to them and 70.59% would adjust their behavior to mitigate ad avoidance. These results are consistent with our secondary data findings and show that the ad incentives should entice streamers to play MRAs and make adjustments. Also, we find that most streamers (51.76%) would increase either frequency, airtime, quality or effort as we proposed.

Finally, the verification survey also included a section where we asked respondents to write any additional thoughts on MRA usage (if they had any). This allowed us to recover bodies of text (roughly) detailing additional strategies on how individuals might wish to implement MRAs. To explore this information, we turned to the machine learning approach of latent Dirichlet allocation (LDA) which clusters bodies of text into topics based on the notion that similar topics are likely to use the same words (Berger et al. 2020). For simplicity, we searched for two topics. Examples of top words found in the first topic were: “content”, “comedic”, and “enjoy”. Top words in the second topic included: “bathroom”, “water”, “breaks”, and “food”.

Table 10: Summary of Verification Survey Responses

Main Questions	Student Responses	Real Streamer Responses
How many MRAs would you play in an hour?	97.64% said 1 or more	77.65% said 1 or more
Would generating MRA ad revenue be appealing?	88.19% said “Yes”	77.65% said “Yes”
Are you likely to adjust to mitigate ad avoidance?	88.98% said “Yes”	70.59% said “Yes”
What type of adjustments would you make?	96.06% would increase at least one of airtime, frequency, quality, or effort.	51.76% would increase at least one of airtime, frequency, quality, or effort.
Would most live streamers play MRAs?	91.34% said “Yes”	74.12% said “Yes”
What other information you can tell us about MRAs?	Topic 1 top words: “content”, “comedic”, “enjoy”, “quality”. Topic 2 top words: “bathroom”, “restroom”, “water”, “breaks”, “food”, “natural”.	

Note: The questions are abridged in this table. Full details are presented in Appendix N.

Based on these words, we note that the first topic is related to the content of the ad itself, where streamers would prefer to play ads which are humorous or enjoyable for the audience. The

second topic focuses on timing, as streamers wish to play ads when no true content is occurring; for example, during a restroom or snack break. These associations indicate that streamers wish to strategically plan how to use MRAs in their stream. This could be one way for streamers to implement ads while minimizing backlash to these sponsored promotions. Though only exploratory, these results provide additional strategies on how streamers might play MRAs (if allowed to). Details of this survey and the LDA procedure are provided in Appendix N.

## DISCUSSION

This research uses a panel dataset of a live streaming platform in 2020 and a difference-in-differences estimation approach to study a policy intervention that introduced the ability to play MRAs. We find positive and statistically significant effects of MRA ability on live streaming consumption behaviors, which is rather surprising because ad avoidance suggests that ads are typically not desired (Research Question 1). We then provide support of increases in airtime, the number of streaming sessions (to some degree) and quality for partnered channels after the shock, underpinning a key mechanism behind the increase in consumption behaviors (Research Question 2). Conceptually, this can be explained by partnered streamers putting forth greater effort by increasing their airtime, frequency or quality to garner MRA revenues (to increase ad exposure). Finally, we explored the heterogeneity within these findings, providing evidence that (a) low success streamers made more aggressive strategic adjustments than more successful streamers, (b) established streamers made more aggressive adjustments than those who entered later, (c) streamers who play content that contains some degree of social interaction with other players appeared to generate the largest consumption increases, and (d) that the MRA policy effects first take a small amount of time to increase, peak, then appear to slowly weaken over time (Research Question 3). We then confirm the main findings with our verification survey.

### *Managerial Implications*

Managers and practitioners should find this study highly relevant. Though some brands already advertise on live streaming platforms, little research has been conducted about whether these are effective strategies or if brands can maintain advertising exposures across time without threatening a reduction in the audience base. We find evidence that introducing MRA ability resulted in increased live streaming consumption. This should encourage firms to invest resources into live streaming MRAs, as audiences are predicted to only consume more content from the channels which display these types of ads, thereby resulting in even more potential ad exposure. Also, we provide analysis on how streamers make adjustments to their behaviors and strategies in response to MRA ability permissions. This can help streaming platforms strategize their ad policies and can help identify streamers and brands they should partner with.

We next explore the financial impact of the MRA ability policy by leveraging information from the historical survey. Though the MRA revenue contribution (about 6% of total streaming income) for the average partnered streamer may seem “small” in comparison to other sources, streamers in the historical survey streamed on average about 4.6 hours per day and 5 days per week, while playing approximately 2 MRAs per hour. The survey respondents also had an average viewership of approximately 126 individuals, closely mirroring the average from our secondary data. Using the values from this survey (\$0.02 per MRA view), this results in roughly \$115.92 of income per week generated from playing MRAs ( $0.02 \text{ dollars} \times 2 \text{ MRAs per hour} \times 126 \text{ viewers} \times 4.6 \text{ hours per day} \times 5 \text{ days per week}$ ). From another lens, the average number of days streamed on the platform by our respondents was roughly 574.43 days. Assuming that these live streamers could play MRAs for this entire duration, this would result in MRA revenue generating around \$13,363.95 over the course of their streaming career on the platform (if they



were allowed to play MRAs the entire duration), which is a rather substantial amount (0.02 dollars  $\times$  2 MRAs per hour  $\times$  4.6 hours per day  $\times$  126 viewers  $\times$  576.43 total streaming days).

Next, we explore the average additional income generated after the introduction of the MRA ability policy. Our secondary data results in Table 4 suggest that after the policy intervention, partnered channels received approximately 8 additional viewers for the average streaming session. Combining this change with the information from the historical survey, this implies roughly that partnered streamers generated an additional \$7.36 per week on average than one may have expected (0.02 dollars  $\times$  2 MRAs per hour  $\times$  8 viewers  $\times$  4.6 hours per day  $\times$  5 days per week). This is rather non-trivial, as streamers may have run MRAs for many weeks and may also have lived in locations where a single U.S. dollar is worth a substantial amount. These calculations suggest that the income generated from playing MRAs is rather valuable.

We next consider the ad revenue received by the live streaming platform itself. Indeed, platforms keep some of the income given by brands by permitting those brands to push their MRAs on the platform (revenue sharing with the streamer). In percentage terms, we assume a 55-45 split (in favor of streamers), which is based on the current ad revenue split by another live streaming platform (Robinson 2022b). As the historical survey indicates that streamers received roughly 2 cents for each viewer watched an MRA, we assume that approximately 1.6 cents went to the platform for every MRA view ((0.02 dollars / 0.55 split for streamers)  $\times$  0.45 split for platform). From our main dataset, 1,141 unique live streaming partnered channels were found to have streamed on the platform after the MRA intervention. To keep calculations simple, we assume that each of these 1,141 streamers streamed in a similar pattern to the individuals in the historical survey. Hence, MRAs generated approximately an additional \$6,718.21 per week for the platform (0.016 dollars  $\times$  2 MRAs per hour  $\times$  8 viewers  $\times$  4.6 hours per day  $\times$  5 days per

week  $\times$  1,141 streamers). In our main dataset alone, we observe 8 weeks of post-intervention time. This would imply an additional \$53,745.66 of revenue for our focal live streaming platform generated from MRA revenue for this duration. This is a rather substantial amount of revenue generated for the platform, demonstrating the financial benefits of implementing MRAs.

#### *Limitations and Areas of Future Research*

There are several limitations and consequently, there are also potential areas of future research.

First, our secondary data does not include streamers' usage of MRAs. We only observe that streamers have the ability to use MRAs and because of the benefits, we assume that MRAs are implemented (although we provide a degree of support for this through our two surveys).

Relatedly, we also do not comment on how nuances such as targeting or time of day may play a role (e.g., Kanuri, Chen, and Sridhar 2018). Second, we do not focus on how MRAs affect brand consumption, but future researchers may wish to explore this topic. Third, this study addresses how the introduction of MRA abilities may affect consumption behaviors but does not comment on how varying levels of MRA advertising intensity may have differential effects. Fourth, it would also be worthwhile to compare the effects of MRAs in live streaming to that of regular streams or even traditional tv broadcasts. We leave these limitations as future research.

## REFERENCES

- Abadie, Alberto (2005), “Semiparametric Difference-in-Differences Estimators,” *The Review of Economic Studies*, 72 (1), 1–19.
- Abadie, Alberto, Alexis Diamond and Jens Hainmueller (2010), “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program,” *Journal of the American Statistical Association*, 105 (490), 493–505.
- Ananthkrishnan, Uttara, Davide Proserpio, and Siddhartha Sharma (2023), “I Hear You: Does Quality Improve with Customer Voice?” *Marketing Science*, 42 (6), 1143–1161.
- Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager (2021), “Synthetic Difference-in-Differences,” *American Economic Review*, 111 (12), 4088–4118.
- Angrist, Joshua D. and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton, NJ: Princeton University Press.
- Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager (2021), “Synthetic Difference-in-Differences,” *American Economic Review*, 111 (12), 4088–4118.
- Baron, Reuben M. and David A. Kenny (1986), “The Moderator- Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations,” *Journal of Personality and Social Psychology*, 51 (6), 1173–82.
- Berger, Jonah and Katherine L. Milkman (2012), “What Makes Online Content Viral?” *Journal of Marketing Research*, 49 (2), 192–205.
- Berger, Jonah, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel (2020), “Uniting the Tribes: Using Text for Marketing Insight,” *Journal of Marketing*, 84 (1), 1–25.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), “How Much Should We Trust Difference-in-Difference Estimates?” *Quarterly Journal of Economics*, 119 (1), 249–75.
- Bloomberg (2022), “Live Streaming Market Worth \$4.26 Billion by 2028 - Market Size, Share, Forecasts, & Trends Analysis Report with COVID-19 Impact,” *Bloomberg Business* (May 5), <https://www.bloomberg.com/press-releases/2022-05-05/live-streaming-market-worth-4-26-billion-by-2028-market-size-share-forecasts-trends-analysis-report-with-covid-19-impact>.
- Butler, Ricky Ray (2021), “Going Live Online: The State Of Live Streaming And The Opportunities For Brands,” *Forbes* (February 4), <https://www.forbes.com/sites/forbesagencycouncil/2021/02/04/going-live-online-the-state-of-live-streaming-and-the-opportunities-for-brands/>.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021), "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225, 200–230.

Chae, Inyoung, Hernán A. Bruno, and Fred M. Feinberg (2019), "Wearout or Weariness? Measuring Potential Negative Consequences of Online Ad Volume and Placement on Website Visits," *Journal of Marketing Research*, 56 (1), 57–75.

Cho, Chang-Hoan and Hongsik John Cheon (2004), "Why Do People Avoid Advertising on the Internet?" *Journal of Advertising*, 33 (4), 89–97.

Correia, Sergio (2014), "REGHDFE: Stata Module to Perform Linear or Instrumental-variable Regression Absorbing Any Number of High-dimensional Fixed Effects," *Boston College Department of Economics Statistical Software Components*, <https://ideas.repec.org/c/boc/bocode/s457874.html>.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin (2020), "Fast Poisson Estimation with High-Dimensional Fixed Effects," *The Stata Journal*, 20 (1), 95–115.

Cunningham, Scott (2021), "Callaway and Sant'anna DD estimator: A Story of Differential Timing and Heterogeneity," *Scott's Substack* (March 8), <https://causalinf.substack.com/p/callaway-and-santanna-dd-estimator>.

Datta, Hannes, George Knox, and Bart J. Bronnenberg (2018), "Changing Their Tune: How Consumers' Adoption of Online Streaming Affects Music Consumption and Discovery," *Marketing Science*, 37 (1), 5–21.

Davey, Lizzie (2022), "YouTube Ads for Beginners: How to Successfully Advertise on YouTube in 2022," *Shopify Blog* (April 6), <https://www.shopify.com/blog/youtube-ads>.

Doorn, Jenny van and Janny C. Hoekstra (2013), "Customization of Online Advertising: The Role of Intrusiveness," *Marketing Letters*, 24, 339–351.

D'Anastasio, Cecilia (2022), "Amazon's Twitch Seeks to Revamp Creator Pay With Focus on Profit," *Bloomberg* (April 27), <https://www.bloomberg.com/news/articles/2022-04-27/amazon-s-twitch-seeks-to-revamp-creator-pay-with-focus-on-profit>.

Fader, Peter S. and Russell S. Winer (2012), "Introduction to the Special Issue on the Emergence and Impact of User-Generated Content," *Marketing Science*, 31 (3), 369–371.

Fisher, Marshall L., Santiago Gallino, and Joseph Jiaqi Xu (2019), "The Value of Rapid Delivery in Omnichannel Retailing," *Journal of Marketing Research*, 56 (5), 732–748.

Fossen, Beth L. and David A. Schweidel (2019), "Social TV, Advertising, and Sales: Are Social Shows Good for Advertisers?" *Marketing Science*, 38 (2), 274–295.

Geyser, Werner (2022), "42 Useful Twitch Stats for Influencer Marketing Managers," *Influencer*

*Marketing Hub* (May 3), <https://influencermarketinghub.com/twitch-stats/>.

Gill, Manpreet, Shrihari Sridhar, and Rajdeep Grewal (2017), “Return on Engagement Initiatives: A Study of a Business-to-Business Mobile App,” *Journal of Marketing*, 81 (4), 45–66.

Golder, Peter N., Marnik G. Dekimpe, Jake T. An, Harald J. van Heerde, Darren S.U. Kim, and Joseph W. Alba (2023), “Learning from Data: An Empirics-First Approach to Relevant Knowledge Generation,” *Journal of Marketing*, 87 (3), 319–336.

Goldfarb, Avi, Catherine Tucker, and Yanwen Wang (2022), “Conducting Research in Marketing with Quasi-Experiments,” *Journal of Marketing*, 86 (3) 1–20.

Grayson, Nathan (2021), “The Twitch Hack Revealed Much More Than Streamer Salaries. Here Are 4 New Takeaways,” *The Washington Post* (October 8), <https://www.washingtonpost.com/video-games/2021/10/08/twitch-hack-leak-minimum-wage-pay-hasan/>.

Grayson, Nathan (2022), “Twitch Ad Update Offers Some Streamers Big Money, Others Pocket Change,” *The Washington Post* (June 20), <https://www.washingtonpost.com/video-games/2022/06/20/twitch-ad-incentive-money-payout-55-percent/>.

Haenlein, Michael, Ertan Anadol, Tyler Farnsworth, Harry Hugo, Jess Hunichen, and Diana Welte (2020), “Navigating the New Era of Influencer Marketing: How to Be Successful on Instagram, TikTok, & Co.,” *California Management Review*, 63 (1), 5–25.

Hart, Robert (2021), “Video And Live Streaming Apps Are Fueling A New Social Media Boom,” *Forbes* (September 6), <https://www.forbes.com/sites/roberthart/2021/09/06/video-and-live-streaming-apps-are-fueling-a-new-social-media-boom/?sh=6c3c29657781>.

Hennig-Thurau, Thorsten, Caroline Wiertz, and Fabian Feldhaus (2014), “Does Twitter Matter? The Impact of Microblogging Word of Mouth on Consumers’ Adoption of New Movies,” *Journal of the Academy of Marketing Science*, 43, 375-394.

Hilvert-Bruce, Zorah, James T. Neill, Max Sjöblom, and Juho Hamari (2018), “Social Motivations of Live-Streaming Viewer Engagement on Twitch,” *Computers in Human Behavior*, 84, 58–67.

Hirano, Keisuke and Guido W. Imbens (2001), “Estimation of Causal Effects Using Propensity Score Weighting: An Application to Data on Right Heart Catheterization,” *Health Services and Outcomes Research Methodology*, 2, 259–278.

Hu, Mu, Mingli Zhang, and Yu Wang (2017), “Why Do Audiences Choose to Keep Watching on Live Video Streaming Platforms? An Explanation of Dual Identification Framework,” *Computers in Human Behavior*, 75, 594–606.

Janakiraman, Ramkumar, Joon Ho Lim, and Rishika Rishika (2018), “The Effect of a Data Breach Announcement on Customer Behavior: Evidence from a Multichannel Retailer,” *Journal of Marketing*, 82 (2), 85–105.

Johari, Ramesh, Hannah Li, Inessa Liskovich, and Gabriel Y. Weintraub (2022), “Experimental Design in Two-Sided Platforms: An Analysis of Bias,” *Management Science*, 68 (10), 7069–7089.

Kanuri, Vamsi K., Yixing Chen, and Shrihari (Hari) Sridhar (2018), “Scheduling Content on Social Media: Theory, Evidence, and Application,” *Journal of Marketing*, 82 (6), 89–108.

Kim, Tongil TI and Diwas KC (2020), “Can Viagra Advertising Make More Babies? Direct-to-Consumer Advertising on Public Health Outcomes,” *Journal of Marketing Research*, 57 (4), 599–616.

Kinson, Anthony (2020), “Streaming Platform Comparison: Partner vs Affiliate,” *Gamesight* (July 30), <https://blog.gamesight.io/partner-vs-affiliate/>.

Lamberton, Cait and Andrew T. Stephen (2016), “A Thematic Exploration of Digital, Social Media, and Mobile Marketing: Research Evolution from 2000 to 2015 and an Agenda for Future Inquiry,” *Journal of Marketing*, 80 (6), 146–172.

Lee, Dokyun, Kartik Hosanagar, and Harikesh S. Nair (2018), “Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook,” *Management Science*, 64 (11), 5105–5131.

Legrand, Tim (2020), “From Insularity to Exteriority: How the Anglosphere is Shaping Global Governance,” Centre for International Policy Studies (October 1), <https://www.cips-cepi.ca/2020/10/01/from-insularity-to-exteriority-how-the-anglosphere-is-shaping-global-governance/>.

Leung, Fine, Flora Gu, Yiwei Li, Jonathan Z. Zhang, and Robert Palmatier (2022), “Influencer marketing effectiveness,” *Journal of Marketing*, 86 (6), 93–115.

Leuven, Edwin and Barbara Sianesi (2003), "PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing," *Boston College Department of Economics Statistical Software Components*, <https://ideas.repec.org/c/boc/bocode/s432001.html>.

Lin, Yan, Dai Yao, and Xingyu Chen (2021), “Happiness Begets Money: Emotion and Engagement in Live Streaming,” *Journal of Marketing Research*, 58 (3), 417–438.

Lu, Shijie, Dai Yao, Xingyu Chen, and Rajdeep Grewal (2021), “Do Larger Audiences Generate Greater Revenues Under Pay What You Want? Evidence from a Live Streaming Platform,” *Marketing Science*, 40 (5), 964–984.

Majumdar, Ripan (2022), “YouTube Stalwart Valkyrae Reveals How 15,000 Fans Got Her to

Start Streaming While Managing 9-Hour Work Shifts,” *Essentially Sports* (July 16), <https://www.essentiallysports.com/esports-news-youtube-stalwart-valkyrae-reveals-how-15000-fans-got-her-to-start-streaming-while-managing-9-hour-work-shifts/>.

Manchanda, Puneet, Jean-Pierre Dubé, Khim Yong Goh, and Pradeep K. Chintagunta (2006), “The Effect of Banner Advertising on Internet Purchasing,” *Journal of Marketing Research*, 43 (1), 98–108.

Miceli, Max (2021), “What is Twitch Turbo?” *Dot Esports* (October 13), <https://dotesports.com/streaming/news/what-is-twitch-turbo>.

Morales, Andrea C. (2005), “Giving Firms an “E” for Effort: Consumer Responses to High-Effort Firms,” *Journal of Consumer Research*, 31 (4), 806–812.

Morgan, Brandon and Gabran Gray (2022), “The Bizarre Story Behind Pokimane’s Rise to Fame,” *SVG* (May 13), <https://www.svg.com/369838/the-bizarre-story-behind-pokimanes-rise-to-fame/>.

Munson, Ben (2018), “YouTube Premium Arrives to Take on Hulu and Spotify,” *Fierce Video* (June 19), <https://www.fiercevideo.com/video/youtube-premium-arrives-to-take-hulu-and-spotify>.

Narang, Unnati and Venkatesh Shankar (2019), “Mobile App Introduction and Online and Offline Purchases and Product Returns,” *Marketing Science*, 38 (5), 756–772.

Needleman, Sarah E. (2020), “Everyone Is a Live-Streamer in Covid-19 Era,” *The Wall Street Journal* (August 9), <https://www.wsj.com/articles/everyone-is-a-live-streamer-in-covid-19-era-11596965400>.

Pailańir Daniel, and Damian Clarke (2022), “SDID: Stata module to perform synthetic difference-in-differences estimation, inference, and visualization,” *Boston College Department of Economics Statistical Software Components*, <https://ideas.repec.org/c/boc/bocode/s459058.html>.

Park, Eunho, Rishika Rishika, Ramkumar Janakiraman, Mark B. Houston, and Byungjoon Yoo (2018), “Social Dollars in Online Communities: The Effect of Product, User, and Network Characteristics,” *Journal of Marketing*, 82, 93–114.

Parrish, Ash (2022), “Twitch expands ad programs to pay streamers more money,” *The Verge* (June 14), <https://www.theverge.com/2022/6/14/23168185/twitch-ad-incentive-program-payouts-increase-1-billion-streamer-revenue>.

Pattabhiramaiah, Adithya, S. Sriram, and Puneet Manchanda (2019), “Paywalls: Monetizing Online Content,” *Journal of Marketing*, 83 (2), 19–36.

Perlberg, Steven (2017), “Facebook to Test Mid-Roll Video Ads,” *The Wall Street Journal*

(January 9), <https://www.wsj.com/articles/facebook-to-test-mid-roll-video-ads-1483996638>.

Porter, Constance Elise and Naveen Donthu (2008), “Cultivating Trust and Harvesting Value in Virtual Communities,” *Management Science*, 54 (1), 113–128.

Preacher, Kristopher J. and Geoffrey J. Leonardelli (2001), “Calculation for the Sobel test: An interactive calculation tool for mediation tests,” *QuantPsy: Kristopher J. Preacher*, <https://quantpsy.org/sobel/sobel.htm>.

Rios-Avila, Fernando, Pedro H. C. Sant'Anna, and Brantly Callaway (2021), “CSDID: Stata Module for the Estimation of Difference-in-Difference Models with Multiple Time Periods,” *Boston College Department of Economics Statistical Software Components*, <https://ideas.repec.org/c/boc/bocode/s458976.html>.

Robinson, Mika (2022a), “How to Watch Multiple Twitch Streams at Once,” *Stream Labs* (May 25), <https://streamlabs.com/content-hub/post/how-to-watch-multiple-twitch-streams-at-once>.

Robinson, Mika (2022b), “Twitch Ad Revenue: How Much Do Ads Pay?,” *Stream Labs* (November 10), <https://streamlabs.com/content-hub/post/how-much-do-twitch-ads-pay>.

Rosenbaum, Paul R. and Donald B. Rubin (1983), “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70 (1), 41–55.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe (2023), “What’s trending in difference-in-differences? A synthesis of the recent econometrics literature,” *Journal of Econometrics*, 235 (2), 2218–2244.

Rubio-Licht, Nat (2022), “Twitch wants streamers to make a steady paycheck,” *Protocol* (February 23), <https://www.protocol.com/bulletins/twitch-ads-incentive-program>.

Schwarz, Carlo (2018), “ldagibbs: A command for topic modeling in Stata using latent Dirichlet allocation,” *The Stata Journal*, 18 (1), 101–117.

Shevenock, Sarah and Alyssa Meyers (2021), “Consumers Think Streaming Ads Are Repetitive and Invasive. The Industry Says It’s Fixing It,” *Morning Consult* (October 18), <https://morningconsult.com/2021/10/18/ad-tech-streaming-services-poll/>.

Shutler, Ali (2022), “Twitch to scrap host mode because it apparently “limits a streamer’s growth potential”,” *NME Gaming* (September 7), <https://www.nme.com/news/gaming-news/twitch-to-scrap-host-mode-because-it-apparently-limits-a-streamers-growth-potential-3305969>.

Sridhar, Shrihari, Murali K. Mantrala, Prasad A. Naik, and Esther Thorson (2011), “Dynamic Marketing Budgeting for Platform Firms: Theory, Evidence and Application,” *Journal of Marketing Research*, 48 (6), 929–943.

Stephen, Bijan (2020), “The lockdown live-streaming numbers are out, and they’re huge,” *The*



Verge (May 13), <https://www.theverge.com/2020/5/13/21257227/coronavirus-streamers-arsenal-gg-twitch-youtube-livestream-numbers>.

Streamer Startup (2022), “How to Get Sponsored on Twitch: 7 Things You Need,” *Streamer Startup*, <https://www.streamerstartup.com/how-to-get-sponsored-on-twitch/>.

Thebault, Reis, Tim Meko, and Junne Alcantara (2021), “Sorrow and Stamina, Defiance and Despair. It’s Been a Year,” *The Washington Post* (March 11), <https://www.washingtonpost.com/nation/interactive/2021/coronavirus-timeline/>.

Twitch (2023), “Running Ads,” *Twitch* (February 25), <https://www.twitch.tv/creatorcamp/en/paths/monetize-your-content/running-ads/>.

Visuals By Impulse (2021), “How Do Facebook Streamers Make Money?” *Visuals By Impulse* (October 21), <https://visualsbyimpulse.com/how-do-facebook-streamers-make-money/>.

Whitler, Kimberly A. (2022), “Super Bowl Ads Provide Hollywood-Sized Entertainment,” *Forbes* (February 13), <https://www.forbes.com/sites/kimberlywhitler/2022/02/13/super-bowl-ads-provide-hollywood-sized-entertainment/?sh=2c9cfec5716>.

World Health Organization (2020), “WHO Coronavirus (COVID-19) Dashboard,” *World Health Organization*, <https://covid19.who.int/data>.

Wilbur, Kenneth C., (2008), “A Two-Sided, Empirical Model of Television Advertising and Viewing Markets,” *Marketing Science*, 27 (3), 356–378.

Wilbur, Kenneth C., Linli Xu, and David Kempe (2013), “Correcting Audience Externalities in Television Advertising,” *Marketing Science*, 32 (6), 892–912.

Wooldridge, Jeffrey (2010), *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. Cambridge, MA: MIT Press.

World Health Organization (2020), “WHO Coronavirus (COVID-19) Dashboard,” *World Health Organization*, <https://covid19.who.int/data>.

## APPENDIX

### *APPENDIX A. Data Aggregation Procedure*

For the aggregation procedure to create the “main dataset” in the main paper, we took the full dataset and assigned each day to one of twelve equally split time periods. Specifically, created four periods for the pre-treatment period (February 18 to March 16) and the remaining eight periods belonging to the post-treatment period (March 17 to May 11). By doing so, all periods contained seven days each. To that end, periods 5 to 12 denote the period after the introduction of MRAs. The specific dates included in each period are displayed in Table A1. This aggregation method allows continuous variables to contain the average value for that specific period. For categorical variables, we took the mode within each aggregate period for each channel ID and assigned it as the value for that period. Observations with multiple modes were dropped. Finally, a total of 12,511,578 observations remained in this dataset.

Table A1: Detailed Dates for Data Periods

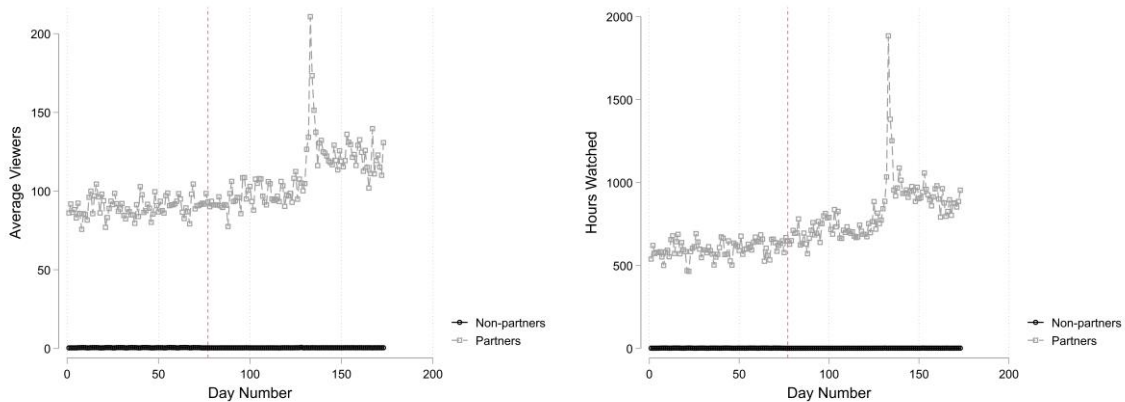
<u>Period</u>	<u>Dates</u>
1	February 18 to February 24
2	February 25 to March 2
3	March 3 to March 9
4	March 10 to March 16
5	March 17 to March 23
6	March 24 to March 30
7	March 31 to April 6
8	April 7 to April 13
9	April 14 to April 20
10	April 21 to April 27
11	April 28 to May 4
12	May 5 to May 11

Note: Periods 1 to 4 denote the pre-intervention periods, and periods 5 to 12 represent the post-intervention periods.

*APPENDIX B. Robustness Check with “Original” Dataset and Discussion of External Events*

Though the focal paper uses data from February 18, 2020, to May 11, 2020, the “original dataset” we obtained actually contains 173 consecutive days of data (starting on January 1, 2020), but due to potential externalities, removed the additional days for both the main dataset and the full dataset. In this section (Appendix B), we refer to dates in the original dataset by number in the sequence (for example, January 1 would correspond with Day 1). We then replicate the main findings using the original dataset. For simplicity, we report only the non-log transformed dependent variables with this analysis. Before presenting the formal results, we provide model-free evidence of the absolute versions of the focal consumption behaviors and how they vary across time, with respect to each treatment group. These graphs are displayed in Figures A1a and A1b. As in the main paper, the graphs appear to indicate a positive effect of the treatment.

Figures A1a and A1b: Graphical Visualization of Consumption Behaviors with Original Dataset



Note: The grey points represent the average value for partnered channel observations, whereas the black points represent the values for non-partnered ones (located very close to the x-axis).

We note that there is a positive outlier(s) or spike around days 133-135 (which is not contained in the main dataset or full dataset). After investigating this time period in the data, we notice that this large spike is driven primarily by a number of unusually large streams providing

content for exogenous events unrelated to the treatment intervention. The most notable of these events was a charity stream on day 133, whose purpose was raising awareness for a disease called Fibromyalgia, in addition to advocating for mental health. This particular stream generated the highest average viewers and the most hours watched within this three-day period by a substantial margin. Indeed, this stream generated more than 50,000 total hours watched on this day over the second most popular stream in terms of this time frame (a total hours watched of 123,841 hours). Similarly, the next largest stream during this time in terms of average viewers hosted a stream related to a video game contest for a cash prize. Hence, these were coincidental exogenous events in this time span.

Another observation is that the time frame in this original dataset corresponds to mid-May. Consequently, many universities around the globe (including ones in the U.S.) often have the start of their summer breaks during this period of time. Given that university students are likely a large portion of a stream's viewing audience, a sizable portion of the increase may be attributed to having leisure time over the summer break. Hence, we first reiterate that the positive spike is not explicitly related to the introduction of MRA ability. Although we believe that this spike would not have harmed the validity of the main findings, we removed the data from day 133 and onwards in the main paper to ensure that our data provided conservative estimates. We also removed data from before February 18 in our main dataset to help with computational speed (and due to the fact that we still had a reasonable timeframe to test for parallel trends).

For the formal analysis with the original dataset, we present simple placebo test results in columns 1 and 2 of Table A2, using the half-way point prior to the policy (day 38) as the marker for the placebo treatment. Similar to the main paper, we find no evidence of a difference between the partnered and non-partnered channels with respect to average viewers (Column 1  $\lambda = -1.396$ ,

$p > .10$ ) and total hours watched (Column 2  $\lambda = -6.067$ ,  $p > .10$ ) in this pre-intervention time period. Thus, continuing support for the parallel trends assumption is provided. We report the focal estimates of the intervention in columns 3 and 4 of Table A2. As in the main paper, we find positive and statistically significant effects of the intervention on the treated group.

For completeness, we briefly next turn our attention to comment on an external policy by the platform. Although the analysis in the original dataset uses data from 173 consecutive days, the “actual original dataset (AOD)” additionally contained days 174 to 205 but were removed prior to analysis due to a major platform announcement on day 174 (henceforth referred to as the “shutdown policy”). This announcement denoted the planned closure of the streaming platform on the 204<sup>th</sup> day.<sup>9</sup>

Table A2: Placebo Estimations and Baseline Results with Original Dataset

	(1) Placebo: Avg. Viewers	(2) Placebo: Hours Watched	(3) Average Viewers	(4) Hours Watched
$T_i \times Post_t$	-1.396 (2.653)	-6.067 (38.245)	20.494*** (4.781)	225.076*** (50.987)
Constant	0.892*** (0.006)	4.950*** (0.081)	0.751*** (0.009)	3.933*** (0.091)
Observations	13,082,898	13,082,898	40,420,923	40,420,923
R-squared	0.888	0.797	0.687	0.724
Controls and FE	Yes	Yes	Yes	Yes

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

Consequently, the data for this robustness check includes these values and therefore contains all possible observations contained in the AOD. Hence, observations from days 174 to 205 are included in this robustness analysis. A very small number of observations (116) were removed prior to analysis if they were non-partnered but belonged to a previously partnered

<sup>9</sup> The platform persisted for part of the 205<sup>th</sup> day as well.

channel. We then re-run our primary two-way fixed effects (TWFE) difference-in-differences (DiD) analysis using this dataset to verify that our baseline results hold regardless of the shutdown policy announcement. These results are presented in Table A3. The results remain robust as they are analogous to the original analysis, with positive and statistically significant estimates of both consumption behaviors after the MRA shocks on partnered channels (Column 1  $\lambda = 17.487$ ,  $p < .01$ ; Column 2  $\lambda = 201.888$ ,  $p < .01$ ). From these findings, we note that it is feasible to suggest that the findings in the main paper are rather conservative, giving further support to the validity of our main paper estimates.

Table A3: Difference-in-Differences Results with AOD

	(1)	(2)
	Average Viewers	Hours Watched
$T_i \times Post_t$	17.487*** (4.583)	201.888*** (48.859)
Constant	0.707*** (0.008)	3.739*** (0.085)
Observations	45,410,108	45,410,108
R-squared	0.682	0.691
Controls and Fixed Effects	Yes	Yes

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

### *APPENDIX C. Falsification Tests Ruling Out Minor Sub-Policies*

On the day of the focal policy, there were several other minor sub-policies added to the platform in addition to the primary MRA ability policy. The most prominent sub-policy change was the ability to “auto-host” for non-partnered (non-treated) channels (partnered channels already possessed this ability by March 17 of the data), which allowed a streamer to redirect a viewer to another stream when the focal channel transitioned offline. There were several other minor changes incorporated on this day related to other “quality-of-life” (QOL), which included changes to the homepage layout, higher pixel chat box emoticons and other minor improvements related to watching streams from a gaming device. We argue that the auto-hosting and QOL changes do not affect the findings from the paper for several reasons. First, minor QOL implementations occur very regularly across all live streaming platforms throughout time and the parallel trends assumption holds prior to the focal treatment shock. This implies that similar minor QOL changes administered earlier had no notable differential effect on viewing behaviors (between partnered and non-partnered channels). Second, later in this section, we find empirical evidence of a null effect of auto-hosting by using an earlier platform policy implementing auto-hosting for partnered channels prior to March 17. Since we find no statistical impact of auto-hosting on consumption behaviors for partnered channels, we conclude that any policy effects were essentially driven by providing MRA abilities. Third, the auto-hosting policy only affected non-partnered channels and should only contribute to underestimate our findings due to the comparative nature of the DiD analysis in the main paper which focuses on treatment groups. Finally, other live streaming platforms have recently removed auto-hosting features after noticing that expectations between the auto-hosted channel and new viewers were misaligned and thus, non-beneficial (Shutler 2022), providing further evidence that auto-hosting does not

drastically impact consumption.

For our empirical falsification test, we wish to provide evidence that auto-hosting was not a driving force behind the increases in audience consumption. Critically, we posit that this is not likely to be a concern since only non-partners were given the auto-hosting ability and partnered accounts already had this feature. Since we do not visually notice any substantial changes in non-partnered channels with respect to our dependent variables (Table 2), we do not believe this is likely to be a problem. Regardless, we address the issue more formally with this falsification test. To conduct the test, we use an earlier policy implemented by the platform on February 10. Importantly, the main implementation in the February 10 policy was the introduction of the auto-hosting feature for partnered accounts only (whereas non-partnered channels did not yet have this feature).

Consequently, if the empirical analysis suggests that there were no effects related to auto-hosting (on the day February 10 policy for partnered only channels), and if the March 17 main policy contains both auto-hosting introductions (for non-partnered channels) as well as the addition of MRAs, then we can conclude that our empirical estimations from the main policy were caused solely due to the introduction of the mid-roll advertisements. We critically note that prior to estimation, our checks for the parallel trends in the main paper occur after February 10. This suggests there were no notable differential effects of auto-hosting affecting the rate of change in consumption behaviors between treatment groups prior to the MRA intervention. Regardless, we formally test for an effect by running our TWFE DiD estimations on the original dataset using February 10 as the time of treatment and restricting the sample to between February 3 and February 16 to isolate the effect of auto-hosting on partnered channels (providing us with 7 days prior to treatment and 7 days after). The results of this analysis are presented in



Table A4.

Table A4: Falsification Tests Ruling Out Auto-Hosting with Full Dataset				
	(1)	(2)	(3)	(4)
	Average Viewers	$\ln(\text{Average Viewers})$	Hours Watched	$\ln(\text{Hours Watched})$
$T_i \times \text{Post}_t$	-1.349 (2.119)	0.007 (0.014)	-5.162 (20.225)	0.001 (0.026)
Constant	1.096*** (0.007)	0.232*** (0.000)	6.210*** (0.063)	0.377*** (0.000)
Observations	1,881,247	1,881,247	1,881,247	1,881,247
R-squared	0.931	0.845	0.836	0.824
Controls and FE	Yes	Yes	Yes	Yes

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

We find no statistical evidence of auto-hosting features having any influence on either of the dependent variables in Table A5. Said another way, the null coefficients on the difference-in-difference parameters suggests that auto-hosting features do not have any notable effect on consumption behaviors. These results therefore reinforce the validity of our main findings, as we can conclude that the main policy effects were driven by introducing MRA abilities and not the introduction of auto-hosting features.

We next briefly discuss why auto-hosting may not have made a positive impact on audience consumption. Critically, one focal reason is likely based on discrepancies between audience expectations and delivered auto-hosted channel content. Audience members often have sets of favorite streams they prefer (as well as preferences for content or quality), and if the auto-hosted stream does not provide the same utility or meet the same expectations as the original channel, the audience member will be left unhappy. If this mismatch of expectations were to occur once, why should they wait to see who is auto-hosted the next time? Alternatively, if the auto-hosted channel is of high quality and is entertaining, why would they need help from another streamer to redirect viewership? To illustrate this point using an example from another

platform, Twitch has recently removed auto-hosting features, arguing that this creates a mismatch between a viewer's expectations and the perceived entertainment value of the streamer providing the content (Shutler 2022). Consequently, we remain relatively unsurprised that auto-hosting does not have an impact on consumption in our own data setting.

*APPENDIX D. Alternative Aggregation Robustness Check*

Next, we refer back to the construction of the aggregated dataset, where we initially aggregated data from the full dataset to the main dataset using mean values. However, we also generated the sum values of each consumption variable in each period (in contrast to the mean values used in the main paper). Thus, using the aggregated sum value accounts for having a fluctuating number of days streamed in a given period. After using the alternative aggregation method by taking the sum of each continuous variable across each time period rather than the average, we conduct our standard two-way fixed effects difference-in-differences (TWFE DiD) analysis and find comparable results with positive and statistically significant focal parameters. The results are presented in Table A5.

Table A5: Sample Analysis with Summed Dependent Variables

	(1) Average Viewers	(2) <i>ln</i> (Average Viewers)	(3) Hours Watched	(4) <i>ln</i> (Hours Watched)
$T_i \times Post_t$	70.940*** (19.286)	0.137*** (0.026)	896.156*** (217.954)	0.192*** (0.031)
Constant	1.379*** (0.013)	0.224*** (0.000)	7.411*** (0.160)	0.268*** (0.000)
Observations	9,254,233	9,254,233	9,254,233	9,254,233
R-squared	0.916	0.756	0.899	0.783
Controls and FE	Yes	Yes	Yes	Yes

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

### *APPENDIX E. Historical Live Streaming Platform Survey Details*

In this section, we describe the motivations and details of our first survey, the historical live streaming platform survey (HLSPS). To begin, the live streaming platform at the time of this research was defunct, as discussed in Appendix B. As a consequence, much of our understanding of the context of the secondary data and platform was based on a number of archived news articles. These articles provided a large amount of information, such as the date of the focal policy, but lacked some further details related to both MRAs and the platform in general. As such, we aimed to supplement our knowledge of the platform by conducting a survey on live streamers who actually streamed on the secondary data live streaming platform at the time when it was still functioning. Our secondary dataset (the original dataset) contained the digital “usernames” (also known as “handles”) of each live streamer. Fortunately, online personas (especially in live streaming) very often use the same usernames across digital platforms. These handles are often fairly unique, as they may contain special characters, numbers, and/or are generally uncommon. We then used search engines to find individuals who were partnered streamers on the platform. Ultimately, we found 25 usable respondents who were willing to complete the historical informational survey in exchange for a gift card payment (valued at \$40.00 USD). These respondents verified that they were partnered streamers on the focal live streaming platform, and additionally played MRAs. We note that the answers of the survey may be somewhat imprecise, as the survey was conducted in 2023, but the questions were based on streaming experiences much earlier in time in 2020. Participants had to be at least 18 years of age to participate in the survey.

The full survey and responses are found in Table A6. We make several (non-causal) conclusions based on this survey (which are also leveraged in the main text). First, streamers

appeared to receive MRA revenue based on a CPM payment scheme (as supported by archived news articles). Streamers made approximately two cents for every MRA view (Q5). Second, MRA revenue contributes approximately 6% of a streamer’s streaming-related income (Q7). While this amount may initially appear to be somewhat small, as seen in the *Managerial Implications* portion of the main paper, MRA revenue still contributes a substantial amount of income for both the streamer and the platform. We leave the exploration of these alternative income sources for future research. Third, streamers were unable to select which types of MRAs were played (Q12). Fourth, the utilization of MRAs did not appear to decrease other income streams such as subscriptions, external sponsorships, and tipping. Indeed, the vast majority of respondents indicated that their income from other streaming sources either increased or were maintained after the introduction of MRA abilities (Q8). Fifth, the majority of respondents indicated that they were interested in making income from MRAs (Q9) and thought that other streamers would also be interested in making revenue from MRAs (Q14). Information utilized in the *Managerial Implications* portion of the paper was gathered from this survey. Also, the average number of MRAs played in an hour was approximately 2 (1.88 was the average of Q3 answers). On average, streamers streamed for about 4.6 hours per day (Q19) and 5 days per week (Q20). The average streamer was on the platform for approximately 576.43 days (Q15).

Table A6: Historical Live Streaming Platform Survey

Survey Questions	Live Streamer Response Options (if Applicable)	Live Streamer Responses
Q1. Did you ever play video advertisements (ad-breaks or mid-roll advertisements) when live streaming on the live streaming platform <i>the platform</i> ?	“Yes” or “No”. If “No”, survey ends, and respondent is not considered for this survey.	All included responses (25) said “Yes”

Note: To protect the identity of our platform, we replaced the actual name of the platform with “*the platform*”. Minor details such as the introduction of the survey and Institutional Review Board (IRB) agreement information were removed from this table for conciseness.

Table A6 (cont'd)

Q2. Please consider your experiences as a former "partnered" live streamer on the platform "*the platform*", then provide your answers to the following questions. As well, recall that on March 17, 2020, *the platform* provided partnered streamers the ability to play mid-stream video ads, which are also known as ad-breaks or mid-roll advertisements (MRAs). Recall that your answers are anonymous, and will not be connected to your streamer handle or ID.

Q3. For every hour you streamed on *the platform*, around how many MRAs would you play (assuming that the average MRA length is 15 seconds)?

Q4. In which ways did playing MRAs help you make ad revenue on *the platform*?

Q5. About how much money did you make per viewer for each MRA on *the platform*? Please write a numeric response in dollars. For example, you would write "0.005" if you received half a cent per view (for a single MRA), or "0.01" if you received one cent per view.

"0", "1", "2", "3" or "4 or more"

"Revenue per viewer for each MRA played", "Flat fee (for each ad played or for a bundle of ads you had to play)", "Uncertain (or cannot recall)", "No revenue (no explicit payment)". Note that respondents could select more than 1 answer.

Numeric written answers were provided here

12 said "1", 7 said "2", 3 said "3" and 3 said "4 or more"

14 said "Revenue per viewer for each MRA played", 1 said "Flat fee (for each ad played or for a bundle of ads you had to play)", 11 said "Uncertain (or cannot recall)" and 1 said "No revenue (no explicit payment)"

15 respondents provided a numeric answer (the remaining ones said they could not recall). The average of written numeric answers was 0.023.

Table A6 (cont'd)

Q6. What were the ways in which you made streaming-related revenue on *the platform*? You may select more than one answer for this question.

“MRA-related ad revenue”, “Channel subscription payments”, or “External sponsorships (not from platform)”, “Other ad-related revenue”, “Streaming contract (provided by *the platform*)”, “Tipping by viewers (donations)”, “Other”, “No revenue (no explicit payment)”. Note that respondents could select more than 1 answer.

19 said “MRA-related ad revenue”, 24 said “Channel subscription payments”, 19 said “External sponsorships (not from platform)”, 10 said “Other ad-related revenue”, 3 said “Streaming contract (provided by *the platform*)”, 24 said “Tipping by viewers (donations)”, 10 said “Other”, 0 said “No revenue (no explicit payment)”

Q7. List the percentage-wise breakdown of the contribution to your streaming-related revenue on *the platform* based on each of the items you checked off in the previous question. Write your answers numerically out of 100. For example, if you checked off “MRA-related ad revenue” and “Channel subscriptions”, and you made 50% of income on both, you would write “50” beside “MRA-related ad revenue”, “50” beside “Channel subscriptions”, and “0” for the remaining items.

Written numeric answers for each of the items in Q6

Average values across respondents (approximate):  
 5.75% from “MRA-related ad revenue”,  
 44.16% from “Channel subscription payments”,  
 12.67% from “External sponsorships (not from platform)”,  
 29.17% from “Tipping by viewers (donations)”,  
 5.42% from “Other”,  
 and 0% from “No revenue (no explicit payment)”

Q8. How did playing MRAs on *the platform* affect your other streams of revenue in terms of increasing, decreasing or staying the same?

“Increase”, “Stayed the same”, “Decreased”, or “I did not make revenue from this source”

a. Channel subscription revenue

8 said “Increased”, 15 said “Stayed the same”, 1 said “Decreased”, and 1 said “I did not make revenue from this source”

Table A6 (cont'd)

<p>b. External sponsorships revenue (not from platform)</p>		<p>6 said “Increased”, 14 said “Stayed the same”, 1 said “Decreased”, and 4 said “I did not make revenue from this source”</p>
<p>c. Other ad-related revenue</p>		<p>1 said “Increased”, 17 said “Stayed the same”, 0 said “Decreased”, and 7 said “I did not make revenue from this source”</p>
<p>d. Streaming contract revenue (provided by <i>the platform</i>)</p>		<p>1 said “Increased”, 12 said “Stayed the same”, 0 said “Decreased”, and 12 said “I did not make revenue from this source”</p>
<p>e. Tipping revenue by viewers (donations)</p>		<p>1 said “Increased”, 23 said “Stayed the same”, 1 said “Decreased”, and 0 said “I did not make revenue from this source”</p>
<p>Q9. Was generating MRA ad revenue appealing to you overall?</p>	<p>“Yes” or “No”</p>	<p>14 said “Yes”, and 11 said “No”</p>
<p>Q10. On March 17, 2020, <i>the platform</i> provided the ability to play MRAs for partnered channels. Roughly how long did it take for you to start playing mid-roll advertisements (MRAs)?</p>	<p>“Immediately”, “After a few days”, “After a few weeks”, or “More than a few weeks”</p>	<p>12 said “Immediately”, 11 said “After a few days”, 1 said “After a few weeks”, and 1 said “More than a few weeks”</p>
<p>Q11. Recall that you can generate ad revenue by displaying MRAs (we assume). Please answer the following questions based on how you may have adjusted your behavior in the following ways, after you were given the ability to play MRAs.</p>	<p>“Increase”, “Stayed the same”, or “Decreased”</p>	
<p>a. The length or duration of each streaming session</p>		<p>4 said “Increased”, 20 said “Stayed the same”, and 1 said “Decreased”</p>



Table A6 (cont'd)

<p>b. The number of times you streamed per week</p>		<p>5 said “Increased”, 20 said “Stayed the same”, and 0 said “Decreased”</p>
<p>c. The production quality of your stream (e.g., purchasing better equipment for higher camera or microphone quality, bringing on guests, etc.)</p>		<p>10 said “Increased”, 14 said “Stayed the same”, and 1 said “Decreased”</p>
<p>d. The effort you put into each stream (e.g., thought, preparation, planning, etc.)</p>		<p>7 said “Increased”, 18 said “Stayed the same”, and 0 said “Decreased”</p>
<p>Q12. Did you have any flexibility or control over what type of MRA was played or the content that the MRA contained?</p>	<p>“Yes” or “No”</p>	<p>3 said “Yes”, and 22 said “No”</p>
<p>Q13. How often did the MRA content align with your actual streaming content? For example, high alignment would be when an MRA for a video game was played when you were playing video games whereas low alignment would be when an MRA for cooking would play while you were playing video games.</p>	<p>“Never”, “Sometimes”, “About half the time”, “Most of the time”, or “Always”</p>	<p>5 said “Never”, 11 said “Sometimes”, 3 said “About half the time”, 5 said “Most of the time”, and 1 said “Always”</p>
<p>Q14. Do you think the majority of live streamers on <i>the platform</i> would have played MRAs on their streams if given the chance to?</p>	<p>“Yes” or “No”</p>	<p>21 said “Yes”, and 4 said “No”</p>
<p>Q15. Approximately how many total days did you stream for on <i>the platform</i>? Please put a numeric response (for example, you would write “91” if you thought you streamed for 91 total days).</p>	<p>Written numeric answers</p>	<p>Average values across respondents (approximate): 576.43 days</p>
<p>Q16. What is an estimation of your average viewership per stream while on <i>the platform</i>?</p>	<p>“0 to 100 viewers”, “101 to 200 viewers”, “201 to 300 viewers”, “301 to 400 viewers”, or “401 viewers or more”</p>	<p>12 said “0 to 100 viewers”, 10 said “101 to 200 viewers”, 1 said “201 to 300 viewers”, 0 said “301 to 400 viewers”, and 2 said “401 viewers or more”</p>

Table A6 (cont'd)

<p>Q17. How many followers did you have on <i>the platform</i>?</p>	<p>“0 to 1,000 followers”, “1,001 to 5,000 followers”, “5,001 to 10,000 followers”, “10,001 to 50,000 followers”, or “50,001 followers or more”</p>	<p>0 said “0 to 1,000 followers”, 3 said “5,001 to 10,000 followers”, 19 said “10,001 to 50,000 followers”, and 1 said “50,001 followers or more”</p>
<p>Q18. How many viewer subscriptions did your channel have on <i>the platform</i>?</p>	<p>“0 to 1,000 subscriptions”, “1,001 to 5,000 subscriptions”, “5,001 to 10,000 subscriptions”, “10,001 to 50,000 subscriptions”, or “50,001 subscriptions or more”</p>	<p>22 said “0 to 1,000 subscriptions”, and 3 said “5,001 to 10,000 subscriptions”. All other choices had 0 respondents.</p>
<p>Q19. What was the average duration of one of your typical streaming sessions in hours on <i>the platform</i>?</p>	<p>“1 hour or less”, “2 hours”, “3 hours”, “4 hours” or “5 hours or more”</p>	<p>2 said “3 hours”, 6 said “4 hours” and 17 said “5 hours or more”. All other choices had 0 respondents.</p>
<p>Q20. What is the average number of days you streamed per week on <i>the platform</i>?</p>	<p>“1 day”, “2 days”, “3 days”, “4 days”, “5 days”, “6 days”, or “7 days”</p>	<p>2 said “3 days”, 4 said “4 days”, 11 said “5 days”, 6 said “6 days” and 2 said “7 days”. All other choices had 0 respondents.</p>
<p>Q21. How knowledgeable are you about MRAs in general?</p>	<p>“Not knowledgeable at all” to “Very knowledgeable” (5 point scale)</p>	<p>5 said “2”, 9 said “3”, 7 said “4” and 4 said “5”</p>
<p>Q22. Is there any other information you can tell us about how you may have played (or not played) MRAs while on <i>the platform</i>? Put “0” if there is nothing you wish to say.</p>	<p>Written responses</p>	<p>Only 4 written responses were given and were evaluated as rather trivial in terms of new information or were redundant, as the topics were generally captured in the main HLSPS survey. For example, one respondent commented on a “hot key” button they used to play ads. Another respondent</p>

Table A6 (cont'd)

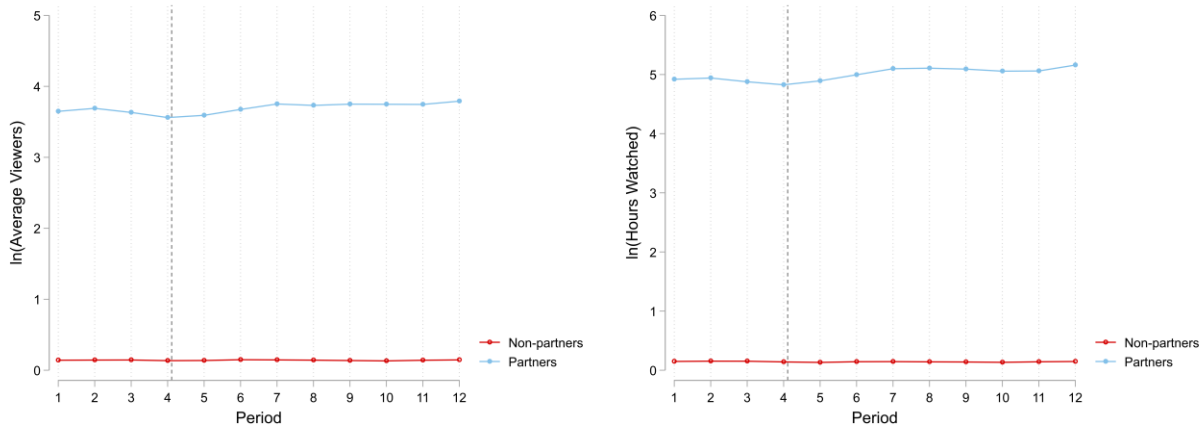
		discusses how they wish they would have known which MRA would be played (similar to Q12). These responses are available upon request.
Q23. What is your age in years?	“18 to 25”, “26 to 30”, “31 to 40”, or “41 and older”	3 said “18 to 25”, 7 said “26 to 30”, 10 said “31 to 40”, and 5 said “41 and older”
Q24. What is your gender?	“Male”, “Female”, “Other”, or “Prefer not to say”	16 said “Male”, 8 said “Female”, 1 said “Other”

---

*APPENDIX F. Data Visualizations and Robustness Checks of Baseline DiD Effects*

We present graphical figures as well as initial robustness checks of our main baseline effects of the MRA ability intervention as seen in Table 4 of the main paper. We provide the graphical representations of our main dataset in Figures A2a and A2b. In particular, we plot the average level of consumption behaviors across time for both treatment groups (blue is partnered and red is non-partnered). We note that rate of change differences between treatment groups from periods 1 to 4 are reflected in Table 3 of the main paper.

Figures A2a and A2b: Graphical Visualizations of Main Dataset



Note: Figure A2a is displayed on the left, while A2b is shown on the right. The blue dots represent the average values for partnered channels, whereas the red points represent the non-partnered ones. We put the treatment line immediately after period 4 only for visualization.

Relatedly, we see that in the period prior to the main intervention policy, the rate of change for both partners and non-partners appears relatively flat prior to the policy with respect to both consumption behaviors. This is graphical support for the parallel trends assumption, as both groups were trending rather comparably in the pre-intervention timeframe. After the intervention, we observe increased levels in both outcome variables for partnered channels but not for non-partnered ones as demonstrated by the positive slopes for partnered channels. This is graphical evidence suggesting that there may be an increase of consumption behaviors for

partnered channels.

Next, we conduct a robustness check of our main results. In particular, we re-run the two-way fixed effects DiD analysis by adding two additional controls: (1) the lagged dependent variable (DV) and (2) the number of observations found in each period for the streamer (the frequency). The results are presented in Table A7, where we continue to find positive and statistically significant effects of the focal DiD parameter, despite the addition of these additional control variables, thus supporting the conclusions found in the main paper.

Table A7: Baseline Results with Lagged DV's and Frequency as Additional Controls

	(1) Average Viewers	(2) ln(Average Viewers)	(3) Hours Watched	(4) ln(Hours Watched)
$T_i \times \text{Post}_t$	5.691** (2.204)	0.092** (0.035)	94.180*** (20.676)	0.149*** (0.034)
Frequency	0.015*** (0.002)	0.006*** (0.000)	0.225*** (0.027)	0.033*** (0.001)
Lagged DV	0.249** (0.098)	-0.043 (0.027)	0.223* (0.121)	0.038 (0.027)
Constant	0.465*** (0.059)	0.169*** (0.005)	1.970*** (0.363)	0.162*** (0.008)
Observations	3,493,422	3,493,422	3,493,422	3,493,422
R-squared	0.880	0.843	0.917	0.867
Controls and FE	Yes	Yes	Yes	Yes

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. DV = Dependent variable.

### *APPENDIX G. Further Details of Dynamic Event Study Analysis*

We next further discuss the usage of the main dataset for the staggered intervention analysis using Callaway and Sant’Anna (2021), which we also refer to as “CS”. To begin, observations which switched from non-partnered to partnered half-way through a period were coded as being “partnered” for the entirety of the period. This coding was required to utilize the CS estimator. We utilized the method proposed by CS with inverse probability weighting (Abadie 2005) to analyze our aggregated dataset.<sup>10</sup>

Using CS, we report the results of the dynamic event study estimates for both of the log-transformed focal dependent variables, whereby the ATT’s are calculated based on the notion that different treated units may receive the treatment at varying times (also known as treatment cohorts). For example, treated units at period 5 (which are partnered at this time) receive the ability to display MRAs for eight periods whereas treated units at period 11 are only able to display MRAs for two periods. For computational ease (due to the vast number of covariates), we restricted our analysis to only observations belonging to the most common language (English) and most frequent target audience (18+/Mature). We further considered the content between the two treatment groups, finding eight overlapped activities in the top 15 content activities for both treatment groups. We additionally filtered the data to only include these eight content activities (the three activities found in Table 2 are included in this set). A final total of 2,507,176 observations were used. The results of this dynamic event study aggregation analysis are presented in Table A8.

Each ATT cohort is presented relative to the initial treatment. These tables are also visualized in Figures 1a and 1b of the main paper. We observe that all of the pre-treatment period

---

<sup>10</sup> We use the package by Rios-Avila et al. (2021) to conduct the CS estimations and to construct the plots.

ATT's are not statistically different than zero. This provides support for the parallel trends assumption. After the intervention, we find that the vast majority of treatment cohorts had a positive and statistically significant ATT. As well, with regards to both variables, we find a positive and statistically significant overall post-treatment ATT. This provides further support for the main effects of a positive increase of consumption behaviors for treated channels after the intervention.

Table A8: Dynamic Event Study Estimates for Average Viewers and Total Hours Watched

	<i>ln</i> (Average Viewers)			<i>ln</i> (Total Hours Watched)		
	ATT	95% CI		ATT	95% CI	
Pre-Treatment Average	0.012	-0.087	0.111	-0.019	-0.163	0.125
Post-Treatment Average	0.192***	0.106	0.277	0.207***	0.066	0.348
<b>Periods Until Treatment</b>						
10	0.063	-0.076	0.203	0.031	-0.211	0.272
9	-0.035	-0.198	0.128	0.049	-0.301	0.398
8	-0.011	-0.397	0.375	0.069	-0.816	0.954
7	-0.058	-0.895	0.779	-0.261	-1.196	0.675
6	0.214	-0.274	0.702	0.259	-0.543	1.061
5	0.076	-0.281	0.434	-0.004	-0.607	0.599
4	-0.102	-0.510	0.306	-0.219	-0.949	0.511
3	0.071	-0.026	0.168	0.050	-0.116	0.216
2	-0.046	-0.151	0.060	-0.064	-0.223	0.096
1	-0.052	-0.151	0.047	-0.102	-0.278	0.074
<b>Times After Treatment</b>						
0	0.088	-0.028	0.203	0.123	-0.089	0.335
1	0.111*	-0.021	0.242	0.121	-0.101	0.344
2	0.238***	0.142	0.335	0.252***	0.139	0.364
3	0.259***	0.113	0.405	0.263*	-0.004	0.531
4	0.212***	0.064	0.359	0.229*	-0.043	0.500
5	0.234***	0.037	0.432	0.207	-0.097	0.511
6	0.255***	0.120	0.389	0.292***	0.095	0.488
7	0.136***	0.062	0.211	0.170**	0.045	0.295

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. We note that these estimates are based on a generalized propensity score (part of the CS estimator), making observables between treatment groups rather comparable for this analysis. CI = Confidence Interval. N = 2,507,176.

*APPENDIX H. Further Robustness Checks Including Propensity Score Matching, Synthetic Difference-in-Differences, and Attrition Checks*

One of the primary concerns with the dataset is that the non-treated and treated groups are compositionally different prior to the intervention. Though the previous CS analysis already utilizes a propensity score to allow for closer comparison groups, for redundancy, we supplement our DiD estimations with a more standard propensity score matching (PSM) and weighting procedure. Matching methods are used in tandem with weighted DiD estimations to address concerns that there may be potential discrepancies in observable characteristics between the treated and untreated groups (e.g., Datta, Knox, and Bronnenberg 2018; Janakiraman, Lim, and Rishika 2018; Fisher, Gallino, and Xu 2019; Narang and Shankar 2019; Ananthakrishnan, Proserpio, and Sharma 2023). PSM procedures estimate the propensity score, or the likelihood of being treated, through a logit or probit model by using the covariates that are observed in the data (Rosenbaum and Rubin 1983).

Ananthakrishnan, Proserpio, and Sharma (2023) conduct a PSM TWFE analysis in a dynamic DiD setting to reinforce the findings from their research, and we conduct our analysis in a mirrored fashion. We made additional adjustments to assist in the matching procedure as to ensure that the treatment groups were even more comparable (primarily by limiting some of the covariate categories to the “top” or most frequent occurrences). In particular, we took the main dataset, and kept only observations that were English targeted and Mature/18+ (as the majority of the data is found in these categories for language and target audience, as seen in Table 2). As such, the primary matching variable was the content activity. Like Ananthakrishnan, Proserpio, and Sharma (2023), we note that standard TWFE PSM in a staggered DiD setting is rather challenging, and additionally match on the content covariate based on the first appearance of a



streamer. We observe that after the PSM procedure, the treatment groups are more comparable than in Table 2 of the main paper. The covariate balance table after the PSM procedure can be found in Table A9.

Table A9: Difference in Covariates After PSM Procedure

Matching Variables	Non-partnered	Partnered	<i>P-value</i>
Audience Type (Mature/18+ only)	1.000	1.000	1.000
Language (English only)	1.000	1.000	1.000
Content Activities (Index 1 to 5155)	2022.7	2025.2	0.975

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Mean values of each covariate for each treatment group are presented. The  $p$ -value of the  $t$ -test to determine whether there is a difference between both groups is reported.

We used Leuven and Sianesi (2003) to conduct the PSM procedure. We next present the placebo tests for the parallel trends assumption using this PSM sample in Table A10 in same fashion as in Table 3 of the main paper. None of the placebo DiD parameters are seen to be statistically significant, indicating support for the parallel trends assumption. We also present the visualization of the data for observations with a PSM weight in Figures A3a and A3b, which supports rather comparable slopes between the two groups between each period prior to the intervention (and providing further support for the parallel trends assumption). The formal results of the PSM TWFE analysis are presented in Table A11.

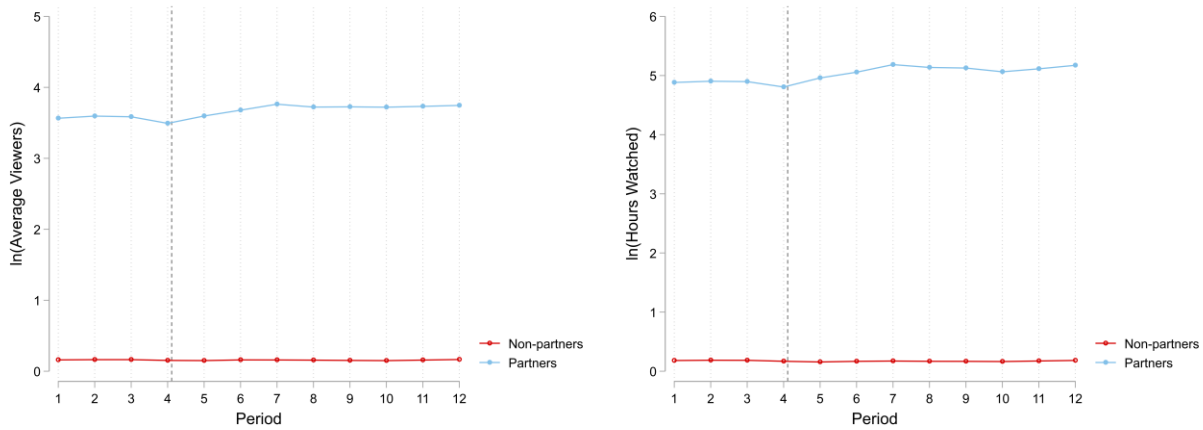
We continue to obtain a statistically significant positive focal effect across all specifications (Column 1  $\lambda = 0.122$ ,  $p < .01$ ; Column 2  $\lambda = 0.171$ ,  $p < .01$ ). These numbers are comparable to the original TWFE findings, as they reflect roughly a 13.0% ( $e^{0.122} - 1$ ) increase in average viewers and a 18.6% ( $e^{0.171} - 1$ ) increase in total hours watched for treated channels relative to non-partnered ones after the intervention. This provides some evidence that the original or “vanilla” TWFE results may actually underestimate the true effects, as the estimates found from the main dataset in Table 4 were lower in magnitude.

Table A10: PSM DiD Placebo Estimations

Dependent Variable	Focal $\lambda$ Parameter Estimates		
	P1 to P2	P2 to P3	P3 to P4
$\ln(\text{Average Viewers})$	0.026 (0.038)	-0.053 (0.045)	-0.072 (0.037)
$\ln(\text{Hours Watched})$	0.014 (0.062)	-0.044 (0.077)	-0.112 (0.064)
N	380,686	357,062	367,226

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Each cell represents the focal  $\lambda$  parameter of a separate TWFE DiD regression. The shortform P refers to “period”. Controls and FE are included.

Figures A3a and W3b: Graphical Visualizations of PSM Dataset



Note: Figure W3a is displayed on the left, while W3b is shown on the right. The blue points and lines refer to the average levels of consumption for partnered channels, whereas the red points and lines denote the consumption for non-partnered channels.

Table A11: Difference-in-Differences with PSM and Weighting

	(1) $\ln(\text{Average Viewers})$	(3) $\ln(\text{Hours Watched})$
$T_i \times \text{Post}_t$	0.122*** (0.028)	0.171*** (0.033)
Constant	2.822*** (0.013)	3.867*** (0.014)
Observations	5,983,920	5,983,920
R-squared	0.966	0.955
Controls and FE	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . The transformation  $(e^\lambda - 1)$  was used for interpretations.

Another robustness test we implement is based on synthetic DiD (SDID) (Arkhangelsky

et al. 2021). The “classic” synthetic control (SC) method is a data-focused technique based on a weighted combination of control observations to better construct a counterfactual group with respect to outcome variables in the pre-treatment period (Abadie, Diamond, and Hainmueller 2010; Pattabhiramaiah, Sriram, and Manchanda 2019). Critically, SDID combines benefits of both DiD and SC by putting more weight on control units which more closely resemble treated units in pre-periods as well more weight on the pre-time periods themselves which are similar to treated periods—while also accounting for scenarios with staggered treatments (similar to CS), and concurrently ensuring the time trend is parallel through the reweighting procedure (Arkhangelsky et al. 2021). Though we find support for the parallel trends assumption in our TWFE DiD analysis as well as our staggered DiD estimations, this SDID analysis simply provides further evidence while ensuring this key assumption.

Prior to conducting the SDID estimations, we first filtered the data for computational ease. Similar to the CS analysis, we took the main dataset and kept only observations which belonged to the top language and audience (English and 18+/Mature). We were able to keep observations from all content activities without running into any computational issues. Said another way, we filtered the dataset to only include top covariates (in terms of frequency) to ensure characteristics were relatively similar between groups. As a result of restricting the covariates through this filtering procedure, we were able to (1) ensure a similar comparison pool between covariates and (2) assist in the computational speed of SDID while still keeping the most common parts of the data, as SDID estimations created a large computational burden. Finally, SDID estimations work primarily with balanced panel datasets (Abadie, Diamond, and Hainmueller 2010; Pattabhiramaiah, Sriram, and Manchanda 2019; Pailañir and Clarke 2022) and as a result, we removed observations (and channels) which caused any imbalances to form a

balanced panel. Our SDID analysis was estimated using the clustered placebo method (Arkhangelsky et al. 2021). All SDID results in this study were generated using PailaÑir and Clarke (2022).

We now present the results of the SDID analysis. Dependent variables were log-transformed. This analysis resulted in a total of 81,240 observations. We report the average ATTs for both consumption variables in Table A12. We also report the differences in the treatment and control outcomes for each treatment cohort for average viewers in Table A13a and for total hours watched in Table A13b (the data filtering and SDID procedure removed observations treated at periods 9 and 10). For visualization, figures for the first treated cohort (period 5) are presented in Figures A4a and A4b. We note that the blue area reflects the lambda weight in Arkhangelsky et al. (2021). Similar to before, we find an overall positive and statistically significant effect of the policy on average viewers (Avg. ATT = 0.110,  $p < .01$ ) and total hours watched (Avg. ATT = 0.068,  $p < .01$ ). This corresponds to a 11.63% ( $e^{0.110} - 1$ ) increase in average viewers and an 7.04% ( $e^{0.068} - 1$ ) increase in total hours watched for partnered channels relative to non-partnered channels after the shock.

The final robustness check we conduct is based on potential attrition. In particular, we conduct our TWFE DiD analysis only on streamers who were found to have been active in every single period of time (no attrition). There are two main reasons for this test. First, this test removes concerns of new streamers who may have entered the platform as a result of the COVID-19 pandemic. Hence, this analysis is focused on established streamers only and not ones that entered as a result of the pandemic. As a result, this test accounts for (to some degree) some of the potential influences caused by the pandemic (though we present further evidence against the pandemic driving the results in the main paper and in Appendix I). Second, later in Appendix

L, we use this attrition sample to combat concerns that our heterogenous analysis (found in the main paper) may be confounded with newly emerging streamers versus unsuccessful established streamers. We conduct the analysis and report the DiD estimates on this “no attrition” sample in Table A14, finding similar positive and statistically significant effects as in Table 4 of the main paper.

Table A12: SDID ATT Averages

	<i>ln</i> (Average Viewers)		<i>ln</i> (Hours Watched)	
	ATT	SE	ATT	SE
Post-Intervention Average	0.110***	0.015	0.068***	0.0013

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. SE = Standard error.

Table A13a: SDID Outcome Differences Between Treatment Groups for *ln*(Average Viewers)  
Differences Between Outcomes For Each Treatment Cohort

Period	Cohort 5	Cohort 6	Cohort 7	Cohort 8	Cohort 11	Cohort 12
1	3.279	2.673	2.810	2.862	2.417	2.521
2	3.280	2.673	2.814	2.867	2.424	2.528
3	3.279	2.673	2.808	2.866	2.423	2.524
4	3.277	2.678	2.811	2.863	2.421	2.525
5	3.357	2.677	2.810	2.869	2.421	2.513
6	3.397	2.918	2.812	2.867	2.417	2.526
7	3.480	3.231	2.900	2.865	2.425	2.529
8	3.374	2.850	2.879	3.112	2.424	2.522
9	3.358	2.937	2.896	3.057	2.402	2.528
10	3.353	3.110	3.175	3.197	1.516	2.523
11	3.338	3.105	2.992	3.154	2.530	2.543
12	3.281	3.278	2.897	3.119	2.494	2.594

Note: Periods 1 to 4 denote the pre-intervention periods, and periods 5 to 12 represent the post-intervention periods. Treatment cohorts 9 and 10 are not present due to the data filtering procedure as well as the SDID procedure.

Table A13b: SDID Outcome Differences Between Treatment Groups for *ln*(Hours Watched)  
Differences Between Outcomes For Each Treatment Cohort

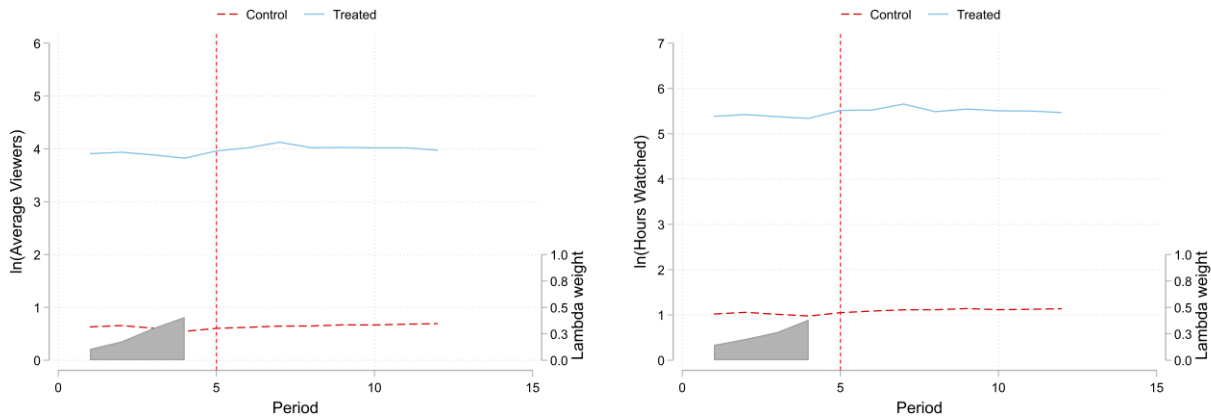
Period	Cohort 5	Cohort 6	Cohort 7	Cohort 8	Cohort 11	Cohort 12
1	4.365	4.259	3.448	3.715	2.934	2.909

Note: Periods 1 to 4 denote the pre-intervention periods, and periods 5 to 12 represent the post-intervention periods. Treatment cohorts 9 and 10 are not present due to the data filtering procedure and SDID procedure.

Table A13b (cont'd)

2	4.366	4.259	3.451	3.725	2.953	2.917
3	4.365	4.259	3.430	3.722	2.962	2.919
4	4.363	4.260	3.440	3.719	2.959	2.916
5	4.464	4.261	3.445	3.729	2.929	2.903
6	4.438	4.478	3.448	3.722	2.958	2.920
7	4.542	4.780	3.291	3.714	2.951	2.930
8	4.372	4.424	3.333	3.986	2.952	2.907
9	4.405	4.418	3.275	3.963	2.962	2.923
10	4.391	4.683	3.559	3.855	0.411	2.912
11	4.376	4.678	3.412	4.000	3.048	2.583
12	4.330	4.903	3.263	4.109	3.260	3.038

Figures A4a and A4b: SDID Analysis for Focal Dependent Variables



Note: Periods 5 and capture the post-treatment periods. The control group is shown with the red dotted line, whereas the treatment group is shown with the blue solid line. The graphs display observations for the first treatment cohort (period 5).

Table A14: Robustness Check of Baseline DiD Estimates with “No Attrition” Sample

	(1)	(2)
	<i>ln(Average Viewers)</i>	<i>ln(Hours Watched)</i>
$T_i \times Post_t$	0.099*** (0.025)	0.144*** (0.024)
Constant	0.522*** (0.000)	0.913*** (0.000)
Observations	349,517	349,517
R-squared	0.945	0.927
Controls and FE	Yes	Yes

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

*APPENDIX I. Details of Primary Falsification Tests for COVID-19*

In this section, we provide further details of our empirical tests which support the view that the COVID-19 pandemic lockdowns did not have a large impact on our empirical MRA ability setting. In the first test, we refer back to Figures 2a and 2b of the main paper, which display our data six days prior and six days after the initial lockdowns. We note that from these figures, there does not appear to be an increase in viewing consumption for partnered channels (and not non-partnered channels). We now present the more formal DiD regressions of this test by using this timeframe and setting the date of the initial WHO lockdowns as the treatment. Null focal parameters would suggest that after the lockdowns, the two groups did not have differential changes to their consumption (which we would expect if COVID-19 was differentially increasing the viewership or consumption of partnered channels more than non-partnered ones). Said another way, a null focal parameter would suggest that COVID-19 lockdowns did not affect partnered channels differently than non-partnered channels, and hence, the main DiD estimations are reasonably unaffected (as DiD estimations are based on the treatment group comparisons). We present the results of this initial lockdown DiD test in Table A15, where we find no statistically significant parameters, suggesting that the lockdowns did not differentially affect the treatment groups in the short term.

Table A15: Results for Isolated Effect of Initial COVID-19 Lockdowns

	(1)	(2)
	<i>ln</i> (Average Viewers)	<i>ln</i> (Hours Watched)
$T_i \times Post_t$	-0.038 (0.034)	-0.037 (0.038)
Constant	0.225*** (0.000)	0.362*** (0.000)
Observations	1,584,789	1,584,789
R-squared	0.852	0.830
Controls and FE	Yes	Yes

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

We next detail the second check to control for potential differential effects of the COVID-19 lockdowns. In particular, we aim to include a control for how the lockdowns may have differentially affected treatment groups based on the three-way interaction through multiplying the focal interaction of  $T_i \times Post_t$  with the daily new deaths caused by COVID-19 (higher deaths were the driving trigger for lockdowns). To do so, we leverage data from the World Health Organization (2020), who provided daily updates on new COVID-19 deaths for most countries across the globe. This captures some essence of the aggression of the lockdowns. We were able to then match this dataset to the live streaming dataset, as the full dataset also included daily observations from the streaming platform (and thus, this analysis is conducted at the daily level). For consistency, we kept only the same days as found in the main dataset for this analysis. As the COVID-19 cases were provided by country, we sought to find a way to isolate observations in the full dataset to properly match the pandemic data to the live streaming data. However, the live streaming full dataset did not include data segregated by countries.

Instead, we turn to the primary language category and use this covariate as a proxy for country. As most of our data comes from English speaking streamers (as seen in Table 2), we focus our attention for this test on the “Anglosphere”, which includes the five main English-speaking countries in the globe: the United States, Canada, Australia, New Zealand, and the United Kingdom (Legrand 2020). Hence, we assume that most of the viewers who watch English speaking streams are from these Anglosphere countries. To run the test, we take the full dataset and remove all observations which do not indicate English as the primary language. We then use data from the World Health Organization (2020), which contains the values of daily Coronavirus deaths and kept only the data from these Anglosphere countries. We take the average number of new deaths across these five countries (caused by COVID-19). Lastly, we combined the two



datasets.

As for the empirical specification itself, we first take the original focal DiD interaction and multiply it by a variable  $COVID_t$  which represents the average number of new COVID-19 deaths on that particular day. For notational simplicity, we make a slight adjustment to equation 2 and present a generalized form of the DiD equation by combining the previous interaction of  $T_i \times Post_t$  into a single indicator  $TreatInd_{it}$  (these are identical in essence). Said another way, regarding this new variable, an observation receives the treatment indicator of 1 if the MRA policy has been implemented and the observation is also from a partnered channel. Next, we multiply  $TreatInd_{it}$  by  $COVID_t$  and include it in the DiD regression. By adding this new three-way interaction, we control for differential effects of the COVID-19 lockdowns between treatment groups. We formally present the equation used as equation 3:

$$(3) \text{ ConsumptionBehavior}_{it} = \mu_i + \gamma_t + \lambda(TreatInd_{it}) + \beta(TreatInd_{it} \times COVID_t) + \tau X_{it} + \varepsilon_{it}.$$

The results of this test are presented in Table A16. The results continue to show that the primary treatment indicator of the MRA policy is positive and statistically significant. Though we believe that differential effects of COVID-19 on the treatment groups are rather unlikely to occur to a large degree and is nevertheless addressed to some degree by our matching procedure(s) as well as fixed effects, this auxiliary DiD analysis continues to produce a positive and statistically significant focal parameter. As such, we continue to conclude that COVID-19 is not a focal driver of our main findings. We also include a robustness check by re-running this analysis but additionally include the one day lagged dependent variable as an additional control, finding similar results. Finally, we note that the baseline inclusion of  $COVID_t$  in equation 3 is wiped out by fixed effects.

Table A16: Test of Main Effect Controlling for Differential Impact of COVID-19

	(1)	(2)	(3)	(4)
	$\ln(\text{AV})$	$\ln(\text{AV})$	$\ln(\text{HW})$	$\ln(\text{HW})$
TreatInd <sub>it</sub>	0.047** (0.019)	0.052*** (0.017)	0.087*** (0.023)	0.086*** (0.019)
TreatInd <sub>it</sub> × COVID <sub>t</sub>	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
Lagged DV		0.183*** (0.004)		0.175*** (0.003)
Constant	0.191*** (0.000)	0.236*** (0.001)	0.278*** (0.000)	0.440*** (0.002)
Observations	17,500,720	5,733,071	17,500,720	5,733,071
R-squared	0.761	0.853	0.761	0.809
Controls and Fixed Effects	Yes	Yes	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Data used is from the full dataset and the sample selected (on the streamer side) is specifically from channels with English as the primary language. AV = Average Viewers, and HW = Hours Watched.

We attribute the lack of COVID-19 impact (from both COVID-19 tests) to the live streaming competition space, as the platform had to compete with Twitch’s massive market share dominance. During the initial months of the 2020 lockdown from March to April (which overlaps with a large portion of our dataset), Twitch gained the vast majority of “lockdown viewers” by growing around 50%, whereas other high-end competitors such as YouTube (referring to the live streaming sections of the platform) grew only 14% over these two months (Stephen 2020). The data from the platform used in this study was from neither Twitch nor YouTube but was much more comparable in terms of popularity to YouTube at the time of the data. It can further be inferred that the platform used in this study was the least popular outlet relative to the other major competitors (we also find news articles supporting this, but do not name them to protect the platform). Thus, we conclude that the lockdowns did not play a notable role in increasing consumption in our data setting, as most “lockdown viewers” primarily resorted to Twitch for their live streaming needs.

*APPENDIX J. Further Details and Tests of DiD Mechanism Investigation*

In this section, we discuss details related to the investigation of the mechanisms proposed in the main paper. First, we present the placebo tests with respect to all three of these mechanism variables to test for the parallel trends assumption. We follow the same placebo procedure as in the main paper of the consumption variables. These results are presented in Table A17, which show no statistical difference in slopes between the two groups in the pre-treatment period for all three mechanism variables. This provides support for the parallel trends assumption.

Table A17: Placebo Tests for Mechanism Variables

Dependent Variable	Focal $\lambda$ Parameter Estimates		
	P1 to P2	P2 to P3	P3 to P4
$\ln(\text{Airtime})$	0.039 (0.020)	0.062 (0.022)	0.021 (0.021)
$\ln(\text{Freq})$	-0.030 (0.021)	0.053 (0.020)	0.042 (0.019)
$\ln(\text{NR})$	0.003 (0.004)	0.003 (0.004)	-0.011 (0.004)
N (for Airtime and Freq)	487,008	464,234	475,716

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Controls and FE are included in all estimations. P = period, Freq = Frequency. For Negative Retention (NR), N = 207,292 for P1 to P2, N = 197,436 for P2 to P3 and N = 198,308 for P3 to P4.

Next, we provide a robustness check of the mechanism results found in Table 6 of the main paper. In particular, similar to the analysis in Table A7, in Table A18, we present the results of our mechanism DiD analysis which includes the lagged DV and the number of observations in each period (Frequency) as additional controls. We make several observations. First, we note that quality (negative retention) is still strongly negative and statistically significant, reinforcing the findings from Table 6 of the main paper. Second, the focal parameter on airtime continues to be significant and positive. Finally, frequency appears to decrease slightly when including these additional controls, suggesting that streamers may slightly decrease their frequency after the intervention. Ultimately, airtime and quality improvements appear to be

the focal ways in which streamers make positive strategic adjustments.

Table A18: Robustness DiD Check for Mechanism Variables

	(1)	(2)	(3)
	<i>ln</i> (Airtime)	<i>ln</i> (Freq)	<i>ln</i> (NR)
$T_i \times Post_t$	0.033** (0.013)	-0.032** (0.012)	-0.026*** (0.006)
Frequency	0.121*** (0.001)	0.388*** (0.001)	0.005*** (0.000)
Lagged DV	-0.118*** (0.018)	-0.024*** (0.001)	-0.179*** (0.017)
Constant	3.898*** (0.070)	-0.269*** (0.001)	0.616*** (0.009)
Observations	3,493,422	3,493,422	1,368,457
R-squared	0.626	0.957	0.533
Controls and FE	Yes	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . DV = Dependent Variable.

*APPENDIX K. Further Details of Formal Mediation Analysis*

In the main paper, we conduct a formal mediation analysis (in Table 7) to support the general findings uncovered from the mechanism analysis (Table 6). In particular, we conduct this analysis by finding the indirect effect proposed by Baron and Kenny (1986), first finding the “Path A” and “Path B” estimates and standard errors, for each mediator and additionally for both consumption dependent variables using our TWFE DiD regressions. We then calculated the Sobel test statistic to test for indirect mediation.<sup>11</sup>

---

<sup>11</sup> For computational ease, we made the calculations using Preacher and Leonardelli (2001).

### *APPENDIX L. Heterogenous Effects by Initial Success*

We next discuss the details of our heterogenous effects analysis where we grouped streamers into low, moderate and high groups based on their success in the pre-treatment period. To do so, we first calculate the average viewership for each streamer across the pre-treatment timeframe. Next, we note that the average stream across the main dataset had an average viewer count of 0.4, with a standard deviation of 12.3. Somewhat comparably, other average live streaming platforms have been found to have an audience size of approximately 27.7 average viewers (Geyser 2022).

Generally, viewer success in the live streaming space is often measured in absolute terms. Indeed, external sponsorships often consider the absolute level of viewers prior to approaching a streamer as a sponsor (Streamer Startup 2022). In broad accordance with these rough measures, we segment the data by defining low streamers to have 10 average viewers or less, moderate streamers to have 11 average viewers to 200 average viewers, and high streamers to have more than 200 viewers, across the pre-treatment periods. Finally, we ran a separate TWFE DiD regression on each sub-sample for our consumption and mechanism variables.

To summarize this analysis, all three types of streamers appeared to make some degree of strategic adjustments (Table 8, Panel A). When making adjustments, low streamers appeared to make very aggressive quality improvements (as seen by the statistically significant absolute parameter size relative to the other two groups) as well as fairly large increases to their airtime. Moderate streamers did make adjustments to all three strategic variables but did not seem to adjust airtime or quality as aggressively as low performing streamers. Finally, high performing streamers appeared to make adjustments primarily through increasing frequency and quality.

In response to these adjustments, all three groups appeared to have statistically significant parameters indicating increases in consumption (Table 8, Panel B). In particular, low performing

streamers appeared to have the largest gains in consumption (indicated by the parameter magnitudes). We posit this may be due to the very aggressive quality increases made by low performing channels.

We run one additional robustness check related to this heterogenous analysis by initial success. In particular, one concern is that this heterogenous analysis is that our “low” success category may be capturing both new streamers who have great potential with streamers who are well-experienced but rather unsuccessful. To disentangle this effect, we re-run this same heterogenous analysis on the “no attrition” sample discussed in Appendix H. By running the analysis on this no attrition data (where every user streamed in every period), we can ensure that “low” performing streamers are indeed rather established and simply “unsuccessful” as they have streamed for every possible period prior to the intervention (28 days in total) with a low viewership, as opposed to a new streamer who may have started streaming less than a week before the intervention. This analysis is presented in Table A19.

Table A19: DiD Parameter Estimates by Initial Viewership on No Attrition Sample

Panel A: Mechanism Variables	Focal Parameter Estimates		
	Low	Moderate	High
<i>ln</i> (Airtime)	0.265*** (0.079)	0.052*** (0.013)	-0.011 (0.087)
<i>ln</i> (Frequency)	0.208* (0.100)	0.053** (0.019)	0.156 (0.091)
<i>ln</i> (NR)	-0.089*** (0.022)	-0.017*** (0.003)	-0.015 (0.012)
Panel B: Consumption Variables			
<i>ln</i> (Average Viewers)	0.424** (0.160)	0.080*** (0.020)	0.278* (0.149)
<i>ln</i> (Hours Watched)	0.813*** (0.237)	0.131*** (0.024)	0.266* (0.134)
N	329,546	19,245	514

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Controls and FE are included in estimations. Each cell represents the  $\lambda$  parameter and the corresponding standard error from separate estimations. For the *ln*(NR) cells, N = 252,226 for the “Low” group, N = 19,245 for the “Moderate” group, and N = 514 for the “High” group.

We note that the results are rather similar to the results in Table 8, with low performing streamers (who are explicitly not new streamers) making the most aggressive adjustments and also receiving the largest increases to their viewership. The other results are still rather similar, with moderate streamers utilizing all three types of streamer adjustments. In this (small) sample, high performing streamers did not appear to make any adjustments, and we find marginal evidence that they received increases to consumption (which can be partially attributed to the smaller sample size). Irrespectively, these results continue to suggest that it is the low performing streamers who make the most aggressive strategic adjustments, and these are the streamers who received the greatest increases to consumption.



*APPENDIX M. Heterogenous Effects by Solo versus Social Content (versus Both)*

We turn our attention to the heterogenous analysis conducted based on content activity. To classify content activities as solo or social content, we first identify that the vast majority of the content activities are video games. From the perspective of video games, solo activities can be classified as “singleplayer” and social games as “multiplayer”. To categorize these video games, we turn to “PCGamingWiki” (<https://www.pcgamingwiki.com/wiki/Home>), a website which contains information of a vast number of video games across time, including singleplayer (S), multiplayer (M) or if the game has components of both (B). As there was a substantial amount of content activities to code in the main dataset (6,972), we instead focused on coding the content activities that partnered streamers played. The reason for this is because the number of observations of partnered channels was substantially lower than the non-partnered ones (e.g., Table 2), and is therefore the limiting factor for an adequate sample size when running the DiD analysis. Hence, we first filtered the data and found that there were 479 unique content activities that partnered channels played in our main dataset.

We then used PCGamingWiki to classify the vast majority of content categories as S, M or B. However, several of the video games were unable to be found and were simply verified through a general internet search. Additionally, several of the content categories would not be classified as video games. For example, “IRL” is a common content activity across live streaming platforms where the live streamer films themselves in the physical world (doing any type of activity such as exercising, cooking, etc.). Another type of content that would not be considered a video game was “creative”, where streamers could broadcast themselves working on activities such as drawing or painting. Ultimately, these types of activities were classified as B. Finally, two categories could not be matched and were removed from the analysis. We present

the results of this analysis in Table A20.

Table A20: DiD Consumption Estimates by Content Activity Type

	(1)	(2)	(3)	(4)	(5)	(6)
	S:	S:	M:	M:	B:	B:
	$\ln(AV)$	$\ln(HW)$	$\ln(AV)$	$\ln(HW)$	$\ln(AV)$	$\ln(HW)$
$T_i \times Post_t$	0.038 (0.074)	0.067 (0.100)	0.062* (0.033)	0.102** (0.040)	0.136*** (0.037)	0.208*** (0.037)
Constant	0.240*** (0.000)	0.379*** (0.000)	0.117*** (0.000)	0.132*** (0.000)	0.192*** (0.000)	0.215*** (0.000)
Observations	171,156	171,156	4,828,543	4,828,543	2,453,642	2,453,642
R-squared	0.794	0.802	0.775	0.836	0.757	0.809
Controls and Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . S = Solo, M = Multiplayer or Social, B = Both, AV = Average Viewers, HW = Hours Watched.

From these results, we note that streamers who played solo content did not appear to experience increases in their viewership consumption after the ability to play MRAs (as indicated by the null DiD parameters in columns 1 and 2). Moreover, we find some evidence of consumption increases (after the intervention) through hours watched for those who streamed social (or multiplayer) content (Column 3  $\lambda = 0.062$ ,  $p < .10$ ; Column 4  $\lambda = 0.102$ ,  $p < .05$ ). Finally, we find positive and statistically significant DiD parameters for both consumption behaviors for streamers who played content activities that could accommodate both solo and social activities (Column 5  $\lambda = 0.136$ ,  $p < .01$ ; Column 6  $\lambda = 0.208$ ,  $p < .01$ ). Ultimately, we find the strongest evidence of the MRA ability policy improving consumption for streamers who are focused on content activities that possess some degree of social interaction with other players or other streamers.

## *APPENDIX N. External Verification Survey Details*

In this section, we provide further details of our external survey. The purpose of this survey is to verify the findings of the main study. In particular, we wanted to confirm that individuals are motivated enough by ad revenue such that they will intend to make streaming adjustments. Moreover, we hoped to verify that usage of ads would be positive after being provided the ability to display MRAs. We first ran a pre-test on students from a large U.S. university (usable N=127), made minor adjustments (with the removal of one question), and then sent the survey out to a number of live streamers (usable N=85) via email. Surveys which were unfinished, or incomplete were not included in the final analysis. All participants were at least 18 years of age. Students were compensated with course credit and live streamers were compensated with gift cards with a monetary value of approximately \$40.00 USD. We present the final survey sent to the real live streamers in Table A21 (we remove minor details such as the introduction to the survey, indicators to click to the next page, and IRB related information). A total of 64 respondents provided responses for Q14, the written answer.

For the machine learning latent Dirichlet allocation (LDA) procedure, many words overlapped between topics (such as “MRA”) and were thus ignored. We pooled answers across both the student survey and the final real streamer survey prior to conducting the LDA topic modelling. Schwarz (2018) was used to run this analysis, which is based on a Gibbs Sampling algorithm where we use 1000 iterations for the burn-in period and 50 iterations between individual samples. Prior to running the LDA algorithm, we removed stop words based on a 2014 Google Code project (<https://code.google.com/archive/p/stop-words/>). Finally, the “top words” discussed in the main paper are within the top 75 words within each topic in terms of frequency. Other words were disregarded due to their common usage between topics (for

example, the words “ads” and “stream” were frequent words between both topics but did not provide any unique information).

Finally, we note that the question about specific adjustments in the summary survey in the main paper is based on question 6 of the full survey. To get the reported numbers in the main table, we simply created a new variable indicating if a respondent answered “Increase” to at least one of the four options in question 6. This is because streamers have a variety of options in which to make strategic adjustments and may not necessarily wish to make all adjustments.

Table A21: Full Survey Questions and Responses for Real Live Streamers

Survey	Live Streamer Response Options (if Applicable)	Live Streamer Responses
<p>Q1. Please consider the following situation, then answer the questions that follow. Imagine that you are a live streamer who streams entertainment content on a live streaming platform. Previously, you were unable to play mid-stream video ads, which are also known as mid-roll advertisements (MRAs). One day, the platform provides you with the ability to display MRAs. You can now make ad revenue by playing MRAs while streaming. The revenue you receive increases based on how many people see the ad. You are not forced to play these MRAs, but can choose to do so if you wish.</p>		
<p>Q2. How likely would you be to play MRAs on your stream?</p>	<p>“Extremely unlikely” to “Extremely likely” (5 point scale)</p>	<p>14 reported “1”, 19 reported “2”, 8 reported “3”, 31 reported “4”, 13 reported “5”</p>
<p>Q3. Suppose you stream for an hour. Within that hour, how many MRAs would you play (assuming that the average MRA length is 15 seconds)?</p>	<p>“0”, “1”, “2”, “3” or “4 or more”</p>	<p>19 reported “0”, 19 reported “1”, 25 reported “2”, 5 reported “3”, 17 reported “4”</p>

Table A21 (cont'd)

<p>Q4. Would generating MRA ad revenue be appealing to you?</p>	<p>“Yes” or “No”</p>	<p>19 reported “No”, 66 reported “Yes”</p>
<p>Q5. Viewers often dislike or avoid ads. Because of this, are you more likely to adjust your streaming behavior in some way to mitigate this "ad avoidance"?</p>	<p>“Yes” or “No”</p>	<p>25 reported “No”, 60 reported “Yes”</p>
<p>Q6. Recall that you can potentially generate ad revenue by displaying mid-roll advertisements (MRAs). Please answer the following questions based on how you might adjust your behavior in the following ways, after you are given the ability to play MRAs.</p>	<p>“Decrease”, “Stay the same”, or “Increase”</p>	
<p>a. The length or duration of each streaming session</p>		<p>5 reported “Decrease”, 56 reported “Stay the same”, 24 reported “Increase”</p>
<p>b. The number of times you stream per week</p>		<p>2 reported “Decrease”, 68 reported “Stay the same”, 15 reported “Increase”</p>
<p>c. The production quality of your stream (e.g., purchasing better equipment for higher camera or microphone quality, bringing on guests, etc.)</p>		<p>1 reported “Decrease”, 50 reported “Stay the same”, 34 reported “Increase”</p>
<p>d. The effort you put into each stream</p>		<p>3 reported “Decrease”, 54 reported “Stay the same”, 28 reported “Increase”</p>
<p>Q7. Do you think the majority of live streamers would play MRAs on their streams if given the chance to?</p>	<p>“Yes” or “No”</p>	<p>22 reported “No”, 63 reported “Yes”</p>
<p>Q8. This next set of questions is about your real live streaming experience. Recall that your answers are anonymous, and will not be connected to your streamer handle or ID.</p>		

Table A21 (cont'd)

Q9. What is an estimation for your average viewership per stream?	5 categories (“0 to 100 viewers”, “101 to 200 viewers”, “201 to 300 viewers”, “301 to 400 viewers” or “401 viewers or more”)	63 said category 1, 13 said category 2, 3 said category 3, 4 said category 4, 2 said category 5
Q10. How many followers do you have?	5 categories (“0 to 1,000 followers”, “1,001 to 5,000 followers”, “5,001 to 10,000 followers”, “10,001 to 50,000 followers” or “50,001 followers or more”)	3 said category 1, 30 said category 2, 18 said category 3, 26 said category 4, 8 said category 5
Q11. What is the average duration of one of your typical streaming sessions in hours?	5 categories (“1 hour or less”, “2 hours”, “3 hours”, “4 hours” or “5 hours or more”)	0 said category 1, 1 said category 2, 15 said category 3, 30 said category 4, 39 said category 5
Q12. What is the average number of days you stream per week?	7 categories (“1 day”, “2 days”, “3 days”, “4 days”, “5 days”, “6 days”, or “7 days”)	1 said 1 day, 3 said 2 days, 22 said 3 days, 20 said 4 days, 16 said 5 days, 18 said 6 days, 5 said 7 days
Q13. How knowledgeable are you about MRAs in general?	“Not knowledgeable at all” to “Very knowledgeable” (5 point scale)	4 said “1”, 13 said “2”, 25 said “3”, 22 said “4”, 21 said “5”
Q14. Is there any other information you can tell us about how you might play (or not play) MRAs? Put "0" if there is nothing you wish to say.		41 said “0”, 44 gave a written answer
Q15. What is your age in years?	“19 or younger”, “20”, “21”, “22” or “23 and older”	5 said “19 or younger”, 4 said “20”, 4 said “21”, 4 said “22”, 68 said “23 and older”
Q16. What is your gender?	“Male”, “Female”, “Other”, Prefer not to say”	43 said “Male” and 42 said “Female”

## **ESSAY TWO: TOO GLOOMY OR TOO FUNNY? THE IMPACT OF DARK HUMOR AND SLANG ON SOCIAL MEDIA VIRALITY**

### **ABSTRACT**

The impact of peculiar (but common) text elements such as dark humor and slang on the virality of social media communications has been left relatively uninvestigated, despite the prevalence and usage of these tools by social media users and firms in marketing communications. In this study, we leverage data from the information and social media site “Reddit” to explore both of these text-related characteristics. In particular, we investigate how the usage of a sub-form of humor we identify as “dark humor” can impact the virality of social media posts. We also simultaneously consider the effects of internet slang on virality. To empirically explore these topics, we leverage a dataset from a sub-forum of Reddit (a “subreddit”) called “/r/wallstreetbets”, which focuses on stocks and trading. Using text-based content analysis tools and a fixed effects estimation approach, we provide evidence to suggest that dark humor can aid in generating virality, despite potential negative connotations associated with this style of humor. We further find that slang also positively influences virality despite its common usage. Moreover, using machine learning applications to generate topic models, we find that the positive effects of dark humor persist for posts which are persuasive rather than those which are informative. We determine that slang is beneficial for posts which are neither informative nor persuasive. Finally, we additionally provide evidence that dark humor improves virality for posts created after normal working hours, whereas slang is beneficial for posts created during working hours.

## INTRODUCTION

The attributes encoded in written text for social media communications have the potential to influence or persuade a countless number of individuals. Indeed, marketing scholars have recently explored how the power of text characteristics such as humor, sentiment, and emotion may affect consumer-related outcomes such as virality<sup>12</sup>, purchase intentions or sales. This study aims to contribute to this literature by examining two common but relatively unexplored characteristics of social media text: dark humor and internet-specific slang.

Humor has been widely documented as a rather positive marketing tool for firms looking to reach consumer audiences (Isaza 2022). Indeed, marketing scholars have also identified encouraging benefits of humor including improving attitudes towards advertisements as well as generating higher purchase intentions (Eisend 2009). These findings are not simply confined to traditional media outlets, as marketing communications in the digital domain can also receive benefits from using of humor (e.g., Borah et al. 2020).

Another equally critical characteristic of social media content is sentiment. A sizable number of electronic word-of-mouth (eWOM) studies have considered sentiment in social media communications. One might initially posit that negative eWOM as compared to positive eWOM may generate higher virality, as it could be viewed as more “interesting”. However, it is also feasible that negative content may simply be viewed unfavorably. Indeed, Berger and Milkman (2012) provide evidence that online content which is positively slanted can also be more viral.

When considering these two social media characteristics of humor and sentiment, we observe a potential discrepancy. Specifically, what is the virality of a social media communication when it is humorous but concurrently possesses negative sentiment? We identify

---

<sup>12</sup> Like many business sources, we define “virality” as the feasibility of content to reach or spread to individuals digitally (Barron 2018).



this case as being representative of dark humor (which has also been labeled as black humor or gallows humor in various literatures). Negative sentiment can be viewed as being aligned with psychological descriptions of dark humor, which has been described to utilize negative characteristics rooted in irony, satire, sarcasm, and cynicism—often for the purpose of ridicule or self-deprecation (Dionigi, Duradoni, and Vagnoli 2022). Dark humor is widely used in social media, as observed by news outlets and the general public (e.g., Berg 2020). However, it remains unclear whether business practitioners should utilize dark humor as a marketing tool in their own communications. On one hand, humor in general is broadly viewed to produce positive brand outcomes related to virality or consumption (e.g., Eisend 2009). On the other hand, content containing negative sentiment is less likely to become viral (e.g., Berger and Milkman 2012). Indeed, the use of dark humor has been noted as being capable of making the receiver of such a communication offended or uncomfortable (Bashforth 2021).

Ultimately, it remains uncertain as to how dark humor will be received by general audiences in a digital setting. Could the unique domain of social media allow for a positive response to dark humor? This study aims to address this question. This topic should be of relevance to both academics and practitioners, as both groups are gaining interest in how text-related characteristics can affect viewer perceptions, particularly due to the increasing amount of text-related media in the digital age (Berger et al. 2020).

From a conceptual standpoint, we posit that the easy accessibility of information from the internet allows social media users to more easily understand the context in which dark humor is utilized, generating a higher likelihood to appreciate the content. Said another way, we argue that viewers of social media posts must typically have knowledge of the context in order for dark humor to be positively received. Importantly, dark humor is often riddled with context-specific

references. As a baseline example, a meme (which can be defined as a wide-spread piece of digital content which is typically humorous or amusing) referencing the COVID-19 pandemic became shared across the internet—this meme also contained an element of dark humor: “A year from now, you’ll all be laughing about this virus...Not all of you, obviously.” (Williams 2020). The context required the recipient of the joke to have a reasonable understanding of the genuine threat that Coronavirus brought to the global population. In another example from the digital media platform Reddit, the use of the phrase “Apes together strong” which insultingly refers to other Redditors (users of Reddit) as “apes” became synonymous with trading or dealing with particular stocks on the subreddit “/r/wallstreetbets”. This piece of content essentially suggests that all Redditors should buy or hold their stocks in solidarity regardless of volatility and concurrently acts as a self-deprecating snub, suggesting a lack of intelligence of such individuals who comply with these actions (somewhat ironically). Relatedly, these Redditors having been observed calling each other, including themselves “...an army of ‘dumb’ day traders...” and “...degenerative apes protecting their own species” (Vincent 2021). Hence, although being called an “ape” is very likely to be taken as an insult in many situations, /r/wallstreetbets users are more likely to be familiar with the context and may even find it positively humorous. Consequently, individuals who are familiar with the context in which dark humor is being utilized are more likely to appreciate it, generating a potentially positive effect on virality.

Generally, humor itself has been noted as being effective or ineffective conditional on the situational context. For instance, Scott, Klein, and Bryant (1990) find that humorous advertising is effective when promoting social events but is rather ineffective for business events. As we are interested in digital communications which utilize dark humor, we focus on two types: informational and persuasive. Similar to how a firm’s general communications can be

informative or persuasive (e.g., Shankar and Kushwaha 2021), we posit that the content of social media posts may also be classified as informational or persuasive. Indeed, Eisenbeiss, Hartmann, and Hornuf (2023) consider informative versus persuasive posts on Facebook and Twitter posts regarding crowdfunding participation and define informative posts as posts which simply provide information whereas persuasive posts attempt to influence decision-making choices—and further find that persuasive posts do have the potential to increase investments over informative ones, conditional on a number of other factors. We apply this categorization to the context of dark humor usage in digital media, suggesting that social media communications which contain dark humor may be perceived differently based on whether the content itself is informational or persuasive.

Adjacent to dark humor, internet slang is another common linguistic trait found in social media content. An example of such a phrase is “LOL” which represents “laugh out loud”. However, few studies have examined the effects of this attribute on social media virality. Liu et al. (2019) conduct one of the few studies exploring this construct but find somewhat mixed results as internet slang was found to improve consumer attention but negatively impacted brand and product evaluation. In their Web Appendix, Li, Chan, and Kim (2019) also provide some initial evidence that internet slang may be perceived to be an inappropriate tool for business settings (due to reductions in competence perceptions), but also suggest that more empirical research is necessary. We aim to build on these studies by further examining the impact of internet slang in a social media setting.

To empirically investigate these questions, we leverage data from the social media and information website Reddit. The primary dataset comes from the subreddit /r/wallstreetbets, where Redditors discuss stocks and other financially related topics. We utilize a number of text-

related content analysis tools, including a machine learning Latent Dirichlet Allocation (LDA) algorithm (which extracts topics from the text) to uncover text-related characteristics of the raw data (Berger et al. 2020). Finally, we estimate parameters of a regression model utilizing these text-related characteristics using a fixed effects approach. The results suggest that dark humor does indeed increase virality as manifested by higher aggregated user scores. However, dark humor does not appear to increase the discussion of the thread (which we consider another term for a social media post), as we are unable to find an effect of dark humor on the number of comments. Moreover, we find that the effect of dark humor on virality is prominent in posts which are persuasive rather than informative and for those that are posted after standard working hours as opposed to during work hours. Finally, we find that internet slang appears to generally increase user score as well.

The remainder of this paper is structured in the following manner. First, we further discuss the conceptual background of the paper and propose formal research questions. Second, we describe the empirical setting, detail the analysis, and present the results. Finally, we discuss the implications of these findings and suggest areas of future research.

## RELATED LITERATURE AND RESEARCH QUESTIONS

### *Social Media and Virality*

As a result of the increasing usage of social media by various marketing stakeholders, including consumers and firms, marketing scholars have sought to understand the nuances of this form of communication. Indeed, the benefits that can be provided by optimal social media strategies are rather limitless. Such factors explored by business scholars include the information encoded in the content itself, the text used in a communication or the participants in the social exchange (e.g., the creators of the social media content or the recipients). Ultimately, understanding these

characteristics is important for firms as there are downstream implications through key outcomes such as consumer sentiment, brand awareness or purchase intentions (e.g., Colicev et al. 2018; Fossen and Schweidel 2019; Rust et al. 2021).

This study broadly focuses on two critical aspects found in social media: text characteristics and virality. Indeed, words and language are key pieces of many social media communications. Critically, one of the key aspirations of social media content is to become viral, meaning consumers share the content and depending on the context, may have increased purchase intent after viewing the content (Akpınar and Berger 2017). Hence, we emphasize the applications of this study, as text in social media communications have been increasingly observed as a source linked to virality (Berger et al. 2020; Rosario, de Valck and Sotgiu 2020).

#### *Dark Humor and Virality*

Humor has been widely established as generally having a positive effect on improving marketing communications (Scott, Klein, and Bryant 1990; Eisend 2009). Applications of humor in the social media sphere are no different, which also find that social media communications which contain humor are more likely to be favorably perceived (e.g., Tucker 2015; Lee, Hosanagar, and Nair 2018). However, to the best of our knowledge, dark humor is one particular sub-type of humor which has not been explored in a digital marketing context. The majority of studies in the marketing and business literatures have typically assessed humor as a singular construct and have not considered whether dark humor in particular may differentially impact recipient perceptions or actions when compared to more typical forms of humor.

Dark humor is differentiated from most types of “vanilla” or standard humor as it often possesses some type of negative spirit. Dionigi, Duradoni, and Vagnoli (2022) explore these differences and explain how dark humor contains negative characteristics such as irony, satire,

sarcasm, and cynicism, and are also often rooted in some type of ridicule whereas more standard “light” humor encapsulates cooperation, benevolence, positive emotion and cognitive capabilities. Most studies in the marketing domain exploring humor have assumed or utilized light humor. For example, Scott, Klein, and Bryant (1990) experimentally manipulate humor with a picnic flier containing text suggesting that food will be provided by an experienced and sophisticated chef but is also accompanied with the image of a cartoon chef flipping a burger. Ultimately, is the general positive effect of humor on consumer perceptions or actions driven solely by light humor or should dark humor be expected to also drive positive outcomes?

We assert that the answer to this question is rather unclear. On one hand, the “dark” aspect of dark humor has the potential to be rather off-putting. Dionigi, Duradoni, and Vagnoli (2022) find that dark humor is predicted by psychopathy and Machiavellian traits. In a similar vein, Allen, Ash, and Anderson (2022) find that individuals who adhere to moral disengagement and schadenfreude are more likely to find comedy in unpleasant topics such as media violence. Applied to the social media space, one may surmise that dark humor may not assist in virality due to the distastefulness of negative content. On the other hand, we have previously noted that communications possessing humor are likely to generate favorable outcomes (e.g., Eisend 2009).

Critically, we posit that contextual knowledge is key to finding dark humor enjoyable. More specifically, recipients of dark humor require a level of understanding—why and what a piece of content is communicating, for a positive response to occur (Chmielewski 2018). For example, Schnurr and Rowe (2008) examine workplace emails to find that employing subversive emails which contain humor referencing institutional norms can help in guiding new work-related policies. These messages are not likely to have ever been sent, nor understood or valued, if the emails were sent to randomly selected university students who would have no knowledge

of the workplace, instead of the co-workers. Indeed, Willinger et al. (2017) find that comprehension of dark humor is related to cognitive processing and higher levels of education, implying that a degree of proficiency or comprehension is required for valuing dark humor as compared to “light” humor. Extrapolating this thought, a large proportion of dark humor is contextual and requires the recipient to understand the “dark” aspect of the joke or quip and to simultaneously understand the humorous part of the communication rather than perceiving it as offensive.

We apply this conceptual notion to the social media space. Importantly, we posit that the ease and fast accessibility of the internet allows individuals on social media to search about the context of a particular humorous joke when it is not understood. Thus, as many dark humor references require context, dark humor in the social media space when applied towards a random individual is more likely to be understood or valued when compared to dark humor used in real life towards a stranger who may then be unable to grasp the backdrop of the joke (as the usage of social media already implies existing familiarity with the internet).

Ultimately, there are mixed reasons for how dark humor may affect social media virality. Dark humor is often viewed as offensive or crude, but the accessibility of the internet can allow recipients to understand the context of the communication, thereby also allowing for positive responses. Hence, we formally propose our first research question:

*RQ1: How does the use of dark humor in social media affect virality?*

### *Internet Slang and Virality*

Internet slang (or simply “slang” in the context of this study) is another trait of social media text which is frequently used. Phrases or abbreviations such as “LOL” or “TL;DR” (which stands for “too long; didn’t read”) are commonly used across most social media platforms. However,

despite the prevalent usage of slang, only a small number of studies have examined its usage in the social media space. For example, Li, Chan, and Kim (2019) briefly explore this topic (and emotions simultaneously) from the perspective of warmth and competence—finding that slang is rather inappropriate in business settings. Lee, Hosanagar, and Nair (2018) also conduct one of the few other studies regarding this topic, discovering that content which contains too much slang may be perceived as being too common or dull, thus lowering social media engagement.<sup>13</sup> However, both studies consider slang as a rather secondary variable or combine it with another social media factor, confounding the isolated effects of slang.

We extend these notions to consider the effect of slang on virality in a social media setting, as the directionality is also rather unclear. To that end, we similarly consider this relationship through a similar lens of warmth versus competency (Judd et al. 2005; Aaker, Vohs, and Mogilner 2010). Specifically, we posit that social media posts which utilize slang may be alternatively viewed as warm or incompetent. On the one hand, viewers of a social media post may enjoy the casual and thus friendly nature of the text, improving recipient perceptions and increasing the likelihood of it being shared. On the other hand, frequent usage of slang in a piece of media may erode perceptions of intelligence, competence, or experience, thus disillusioning recipients of a post due to the lack of formality and sophistication in their language. Hence, we formally propose:

*RQ2: How does the use of internet slang affect social media virality?*

*Persuasive vs. Informative*

We next turn our attention to explore the boundary conditions of RQ1 and RQ2. In this section, we consider whether the above focal relationships will hold for informative versus persuasive

---

<sup>13</sup> Lee, Hosanagar, and Nair (2018) do not disentangle or distinguish emoticons from text slang.



posts. A number of economists and business academics have explored differences between informative and persuasive content, with many empirical results suggesting that persuasive marketing communications have the potential to be more effective in inducing positive outcomes when compared to informative ones, even if both strategies are rather serviceable (e.g., Narayanan, Manchanda, and Chintagunta 2005; Eisenbeiss, Hartmann, and Hornuf 2023). Lee, Hosanagar, and Nair (2018) do find that social media engagement is typically lower for informative posts, unless there are brand personality attributes in conjunction with this content.

When applying these findings to a social media context, we posit that text effects which have positive impacts on virality should be stronger for persuasive posts. We reason this is the case due to the “social” component of social media. Specifically, individuals commonly use social media to interact with other individuals and to hear their opinions. In contrast, more traditional informational sources such as textbooks, or academic experts are typically viewed as being more reliable for information but are less likely to be received well in a “social” context where the primary goal is entertainment or relaxation (via social interactions). However, we acknowledge that the controversial nature of dark humor and the unsophisticated nature of slang cloud this particular proposition. Hence, we formally reveal our next research question:

*RQ3: How do dark humor and slang usage in social media affect virality when considering an informative versus persuasive piece of content?*

*After Typical Working Hours (vs. During Work Hours)*

Finally, we turn our attention to assessing whether dark humor and slang may have differential effects on virality when comparing the posting time relative to the standard working hours. More specifically, individuals are shown to be fatigued after the stresses of the workplace (Zohar 1999). Echoing this thought, Park et al. (2020) show that working higher hours per week is

highly associated with increases in stress, depression, and suicidal thoughts. Relatedly, dark humor is often used as a coping mechanism for stress or negative emotion. Indeed, studies have demonstrated that employees do employ dark humor to reduce stress or negative emotions in a number of fields including emergency services (e.g., Rowe and Regehr 2009) and in biological laboratories (Dueñas, Kirkness, and Finn 2020).

Hence, we posit that social media posts which were created after typical working hours may have higher natural or authentic alignment with dark humor usage, as the poster may have been coping with a full day of workplace stress. As such, potential viewers of such a post may also find the content to be more authentic. Indeed, authenticity has been seen to be linked to positive outcomes in the marketing literature (e.g., Becker, Wiegand, and Reinartz 2019). Similarly, slang usage in social media posts created after working hours may have been both created and viewed by individuals who were fatigued by using formal language during working hours, and thus, the usage of verbal shortcuts may be viewed as more authentic. We also do not distinguish between weekends and weekdays, as “work hours” on weekends are often filled with daily chores, which are also known to generate negative emotion and stress (McIntyre, Korn, and Matsuo 2008). Regardless, the divisive nature of both dark humor and slang may confound this proposition. Consequently, we propose the final research question:

*RQ4: Do the effects of dark humor and slang in social media text affect virality differently for content created during work hours as opposed to after work hours?*

## DATA AND METHODS

### *Data Background*

This paper utilizes one key dataset to examine our research questions. Importantly, we employ data from the digital media site Reddit. Reddit is an extremely popular social media and news

platform utilized widely across the globe. In 2021, more than 50 million daily active users contributed to over 50 billion page views (Tepper and Curry 2021). The primary data for this study comes from the subreddit /r/wallstreetbets.

A social media platform such as Reddit is a particularly interesting setting to study dark humor and slang for two reasons. First, the vast availability of data on this platform allows for more options to provide empirical evidence supporting theoretical claims, as well as a reasonable sample size for statistical analysis. Second, dark humor is utilized heavily in the social media space (e.g., Berg 2020), and Reddit is no exception. Indeed, the marketing research firm Ipsos has found that the Millennial and Gen Z demographics (who are the most prominent users of social media) prefer dark humor more than previous generations (Chessey and Ranly 2022).

#### *Data Description and Empirical Method*

The primary dataset is based on thread posts from the subreddit /r/wallstreetbets. This panel dataset of 831,165 unique posts from 418,745 users across four years was obtained from a data hosting site called Kaggle, where individuals can obtain data from application programming interfaces (APIs) and post them for public use (Fontes 2021). The original dataset technically contains 1,118,875 observations, but “corrupted” observations (e.g., empty cells, observations spread across several rows, etc.) were removed from the final dataset. For simplicity, we kept data only from 2018 to 2021. Each observation (or post) contains the title, score, unique id, number of comments (at the time of the data scraping), and the timestamp of the post. For simplicity, we set the temporal aspect of the panel to be at the monthly level, resulting in 38 consecutive months of data (starting in 2018, and ending in February of 2021). We discuss these variables in more depth later in this section.

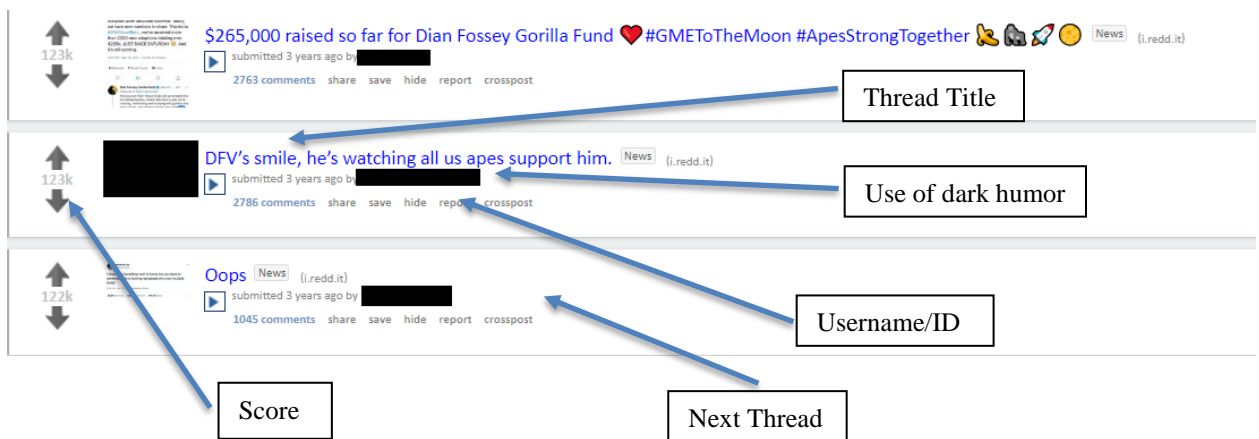
The process in which a Reddit user (or anyone who accesses the site) retrieves

information on a particular subreddit is two-fold. First, users must arrive at the subreddit itself. Subreddits can be found through a number of avenues, including search engine queries or the front page of Reddit (which provides a list of the most popular recent posts across all subreddits). Second, after arriving at a subreddit, the user can view the top “threads” or posts. The titles of these threads are shown in a sequential order based on parameters set by the user. One the most popular ways to view threads is by sorting the page by overall rating across time. Relatedly, on the subreddit, users can view the “score” rating or the number of upvotes minus downvotes of the thread. Threads do not visibly allow downvoting past a score of zero (or do not allow downvoting altogether). The viewer can read the title of each post instantaneously but can also click on an “arrow” icon which can display the body of main text (if there is any). Alternatively, the viewer can click on the thread and be redirected to a page dedicated to the post, displaying the thread title, any additional body of text or images (which are optional) and user response comments. Interestingly, many threads also put the core idea of the thread in the title itself but do not write a body of text. A thread must have a title to be posted, but not necessarily a body. The body text for observations in our dataset was not scraped, and thus, was unavailable for our research. A visual example from /r/wallstreetbets is displayed in Figure 1 (usernames and images are covered to protect identities).

Critically, we utilize the text in the thread title to generate variables related to our analysis. The most important of these is the construction of our novel variable, dark humor. To operationalize this variable, we first construct a general measure of the degree of humor stemming from each post. To do so, we identify a number of humorous “cultural” features unique to /r/wallstreetbets (henceforth referred to as WSB) during 2021. Indeed, the creation and usage of particular memes (widespread digital content that is humorous) specific to WSB surged

in 2021. In particular, a number of text memes (e.g., “apes together strong”, “buy high sell low”, “to the moon”) as well as a number of emojis used in a humorous “meme” fashion (e.g., rocket ships “🚀” or diamonds “💎”) were seen to have been originated from or were heavily used on WSB in 2021 (Hartwig 2021). Many of the active Redditors on this subreddit were aware of these specific humorous memes, and often partook in utilizing them as well. Hence, most Redditors on this subreddit were familiar with these specific memes or jokes. Individuals who did not understand these references were able to find the context of these memes by conducting a brief query from any major search engine. For emphasis, many of these specific meme references were considered humorous or comical. In order to generate a measure of humor, we used content-related text tools to count the number of occurrences these text or emoji memes appeared in a thread title. Hartwig (2021) was used to find the list of WSB-specific memes. This generated a variable containing some semblance of the degree of humor contained in the thread title.

Figure 1: Example Posts (Threads) From /r/wallstreetbets Subreddit



Next, we identify whether the post was negatively charged in terms of sentiment (consistent with the “dark” portion of dark humor). Analyzing the overall sentiment of a piece of text is common in the digital space, where computer programs count the number of positive or negative words based on a dictionary to indicate the general level of sentiment (Berger and

Milkman 2012; Berger et al. 2020). We use the “sentimentr” package to run this sentiment analysis (Rinker 2021) on each title post, which utilizes this general methodology by counting the number of positive and negative words, accounting for “valence shifters” (and other nuances) and then ultimately providing an overall numerical score of the sentiment. Negatively charged posts have scores which are less than zero, whereas positively charged ones have scores which are greater than zero. Indeed, negative words such as “abandon” contribute to a negative score based on these dictionaries and are also simultaneously related to the themes of dark humor (and thus likely to be connected to some form of irony, satire, sarcasm, or cynicism) (Rinker 2021; Dionigi, Duradoni, and Vagnoli 2022).

Finally, we amalgamate the variables constructed from humor and sentiment to generate our novel measure of dark humor. Critically, we define dark humor posts as ones containing a sizable amount of negative sentiment but are also simultaneously humorous. To do so, we calculated the mean levels of humor and sentiment across all observations. Then, we created an indicator variable equaling 1 if the observation was greater or equal to the mean value in humor as well as lower or equal to the mean value of sentiment. Ultimately, this dummy variable acts as an indicator of whether a post was classified as one which contains dark humor (one which is high in humor and low in sentiment).

As for internet slang, phrases such as “JK” (just kidding) are commonly used in social media posts. Hence, we searched for a list of internet slang terms which are commonly used and counted the amount of slang phrases used (BSC Team 2020). These are phrases or words that are common in social media posts and have become part of the typical vocabulary in the digital space. To construct the variable, we take the natural log of this count (plus one) and use this variable (which we call  $Slang_{it}$ ) as the focal operationalization of slang.

The primary dependent variable we use in this analysis is the score, which is calculated by the number of upvotes (or individuals who “liked” the post) minus downvotes (the individuals who “disliked” the post). This measure is quite critical, as higher scored posts are more likely to be read and shared by other Redditors. It is important to note that this variable is not allowed to be negative (as per the administrators of Reddit), and so, we take the natural log of this variable to avoid skewness (adding 1 to each value beforehand). Combining all of these variables, the baseline estimation regression (equation 1) can be defined as:

$$(1) \text{Score}_{it} = \beta_0 + \beta_1 \text{DarkHumor}_{it} + \beta_2 \text{Slang}_{it} + \varepsilon_{it},$$

where  $\text{Score}_{it}$  is our measure of virality. More specifically, it is the log-transformed (plus 1) score of a given post  $i$  on month  $t$ .  $\text{DarkHumor}_{it}$  is a dummy variable indicating whether the post is considered one with dark humor and  $\text{Slang}_{it}$  is the natural log of the total count of slang (plus 1).

Although we could solely estimate this model using ordinary least squares (OLS), we next consider endogenous threats to the consistency of the parameter estimates. To begin, we consider whether omitted variables may be related to dark humor or slang, which in turn could be driving changes in the score. To address this to a certain degree, we include a set of control variables. First, we include a control for whether the post was given a label for being “not safe for work” (NSFW). Reddit posts that are indicated as being NSFW require that the viewer acknowledge that they are 18 or older to view the content contained within. Second, we include a control variable accounting for the number of characters of the thread (post length). Third, we include the absolute degree of sentiment as a control. We also control for the hour in the day a post was made using a set of control variables  $X_{it}$ . In addition to these control variables, we exploit the panel nature of our dataset. In particular, we use a fixed effects (FE) approach and include FE for each individual user and for each time period, controlling for time-invariant and

user-invariant heterogeneity (Angrist and Pischke 2009). Hence, we present our primary estimation equation (equation 2), based on FE and where the variables  $\gamma_i$  and  $\tau_t$  capture user and monthly fixed effects, respectively:

$$(2) \text{Score}_{it} = \gamma_i + \tau_t + \beta_1 \text{DarkHumor}_{it} + \beta_2 \text{Slang}_{it} + X'_{it} \delta + \varepsilon_{it}.$$

Next, to explore how the effects of dark humor and slang affect persuasive or informative posts, we turn to a machine learning application to generate topic models (LDA). Marketing scholars are increasingly utilizing LDA algorithms as a tool to generate topics (decided upon by researcher judgement or loose statistical tests) based on the similarities of words found in the text itself (Berger et al. 2020). Succinctly, LDA assumes that similar topics make use of similar words and groups words which co-occur together. Using this, the LDA algorithm assigns the probability that a piece of text belongs to a certain topic. We utilize Schwarz (2018) to run these topic modelling algorithms. Following other studies in the marketing domain, we take an unsupervised approach (e.g., Hollenbeck 2018) and removed all stop words (words used universally across all topics, such as “each” or “as”) based on a 2014 Google Code project (<https://code.google.com/archive/p/stop-words/>). We decided upon generating two topics for simplicity. This generated the likelihood that a post belonged to either the first or second topic (Topic 1 or Topic 2). Threads with equal probabilities of being in a given topic were assigned as being “neutral”. The most frequent words belonging to each topic were also generated from this process.

After running the LDA algorithms to determine the likelihood that a post belonged to either of the two topics generated, we sorted through the most frequently used words of each topic to determine the content contained in each. Examples of Topic 1’s top words (within the top 30) were “i’m”, “selling”, “today”, “hold”, “tomorrow”, “bought”, “let’s” and “shares”. On



the other hand, examples of some of the most common words from Title Topic 2 included “funds”, “market”, “gamestop”, “covid”, “price”, “earnings”, “daily” and “robinhood”. Consequently, as a result of the words suggesting action (e.g., selling, hold) and a personal slant (e.g., i’m, let’s), we identify Topic 1 as possessing words which suggested the declaration or request of tangible action, and hence, we categorize this as the persuasion topic. Topic 2 contained words which were more descriptive and was therefore labelled to be the informative topic. Finally, we decided a sub-sample analysis would be an appropriate way to assess the differences between informative, persuasive, and neutral posts after exploring the main effects of dark humor and slang.

We additionally generated a dummy variable to further classify observations to explore whether dark humor and slang effects differ based on whether the post was made after standard working hours or not. The creation time of each post was used to create this indicator. We denote all times before 6 a.m. and after 6 p.m. to be defined as “after work” and use this definition to construct the indicator variable. We assume locations of users do not impact our analysis, since post times were based on Greenwich Mean Time (GMT) but simultaneously acknowledge this as a limitation of our research related to RQ4. Similar to persuasive vs. informative topics, we analyze this classification using a sub-sample analysis.

## RESULTS

### *Main Findings*

We now formally estimate our data setting using equation 2. We present the main results in Table 1. The results suggest a positive and statistically significant effect of dark humor on score ( $\beta_1 = 0.023$ ,  $p < .05$ ). Similarly, slang appears to have a statistically significant positive effect on score ( $\beta_2 = 0.042$ ,  $p < .05$ ). These results are a direct test of RQ1 and RQ2, as the results of this

analysis suggest that social media posts which contain dark humor can improve virality, while the use of more internet slang in the post can also increase virality. We note that these conclusions are reached by the addition of our set of control variables and including fixed effects.

Table 1: Baseline FE Estimates

Dependent Variable: Score <sub>it</sub>	(1) FE
DarkHumor <sub>it</sub>	0.023** (0.009)
Slang <sub>it</sub>	0.042** (0.017)
Sentiment <sub>it</sub>	0.017* (0.010)
CharLength <sub>it</sub>	0.000*** (0.000)
NSFW <sub>it</sub>	0.022 (0.034)
Constant	0.956*** (0.005)
Observations	541,754
R-squared	0.374
Time of Day Control and FE	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Cluster-robust standard errors were clustered by ID and month. For robustness, we re-run this specification while including a control for the topic (persuasive vs. informative) and find very similar results with  $\beta_1 = 0.023$  ( $p < .05$ ) and  $\beta_2 = 0.041$  ( $p < .05$ ). As well, we run another specification including daily fixed effects and recover similar results ( $\beta_1 = 0.020$ ,  $p < .05$ ;  $\beta_2 = 0.052$ ,  $p < .05$ ).

We next turn our attention to RQ3. To do so, we take a sub-sample approach. In short, we estimate the sub-sample of observations which are classified as informative, persuasive, or neutral. The results of this sub-sample analysis are presented in Table 2, where column 1 displays the FE results for the informative posts, column 2 presents the same for the persuasive posts, and column 3 provides the results for threads which were neutral. Critically, we can see

that the parameters on dark humor are positive only in column 2 for persuasive posts (column 2  $\beta_1 = 0.037$ ,  $p < .05$ ). The results do not suggest an effect of dark humor on virality for the informative or neutral sub-sample. Relatedly, slang appears to increase virality for neutral posts (column 3  $\beta_2 = 0.176$ ,  $p < .10$ ) but not for informative or persuasive ones. Ultimately, we suggest that these results support the notion that the positive impact of dark humor persists for persuasive posts while the positive impact of slang improves virality for posts which are more neutral.

Table 2: Informative vs. Persuasion vs. Neutral Posts

Dependent Variable: Score <sub>it</sub>	(1) Informative	(2) Persuasive	(3) Neutral
DarkHumor <sub>it</sub>	-0.000 (0.013)	0.037** (0.017)	-0.000 (0.015)
Slang <sub>it</sub>	0.027 (0.023)	-0.003 (0.022)	0.176* (0.088)
Sentiment <sub>it</sub>	0.009 (0.015)	0.027** (0.011)	0.039 (0.023)
CharLength <sub>it</sub>	0.001*** (0.000)	0.000 (0.000)	0.001* (0.000)
NSFW <sub>it</sub>	0.042 (0.051)	0.040 (0.040)	-0.015 (0.084)
Constant	1.033*** (0.007)	0.946*** (0.006)	0.929*** (0.008)
Observations	193,870	182,554	52,854
R-squared	0.409	0.426	0.474
Time of Day Control and FE	Yes	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Cluster-robust standard errors are clustered by ID and month.

Finally, we turn our attention to the effect of dark humor and slang on posts written after typical working hours (versus during work hours) to explore RQ4. We run our FE regression on the sample of posts created during work hours, then again on the sample for posts created after typical working hours. The results are presented in Table 3, with column 1 displaying the FE results for posts created during work hours while column 2 displays the same analysis for posts

created after typical working hours. The results suggest that dark humor usage improves virality for posts created post-working hours (column 2  $\beta_1 = 0.023$ ,  $p < .01$ ), but we find no evidence of dark humor improving the score for posts created during work hours. Interestingly, slang was seen to improve virality for posts created during work hours (column 1  $\beta_2 = 0.047$ ,  $p < .05$ ).

Table 3: During vs. After Work Posts

Dependent Variable: Score <sub>it</sub>	(1) During Work	(2) After Work
DarkHumor <sub>it</sub>	0.015 (0.012)	0.023*** (0.008)
Slang <sub>it</sub>	0.047** (0.023)	0.030 (0.023)
Sentiment <sub>it</sub>	0.032* (0.016)	0.018* (0.010)
CharLength <sub>it</sub>	0.000*** (0.000)	0.000*** (0.000)
NSFW <sub>it</sub>	0.052 (0.057)	0.049 (0.036)
Constant	0.963*** (0.007)	0.978*** (0.005)
Observations	210,652	260,793
R-squared	0.428	0.398
Time of Day Control and FE	Yes	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Cluster-robust standard errors are clustered by ID and month.

### *Impacts on Conversation Generation*

Thus far, we have provided general evidence of dark humor and slang usage in social media posts increasing virality. This is because as the score value increases, posts are more likely to be visible for new users exploring the subreddit (for instance, categorization by “top” pushes posts with the highest score to the top of the subreddit page) and are also more likely to be shared by other Redditors. However, we additionally consider whether dark humor and slang could also increase discussion or further dialog on each post. To do so, we utilize our baseline FE approach

but replace the dependent variable with  $Comments_{it}$ , which is the log-transformed number of comments (plus one).<sup>14</sup> We report the results in Table 4. Interestingly, the results suggest that both dark humor and slang are statistically insignificant and thus, no evidence could be found suggesting that dark humor or slang generate further discussion on a post.

Table 4: FE Estimates for Comments as the Dependent Variable

	(1)
Dependent Variable:	FE
$Comments_{it}$	
$DarkHumor_{it}$	-0.013 (0.018)
$Slang_{it}$	0.029 (0.021)
$Sentiment_{it}$	-0.075*** (0.020)
$CharLength_{it}$	0.000 (0.000)
$NSFW_{it}$	-0.152*** (0.046)
Constant	1.103*** (0.013)
Observations	541,754
R-squared	0.410
Time of Day Control and FE	Yes

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Cluster-robust standard errors are clustered by ID and month.

## DISCUSSION

This research uses data from the social media and information site Reddit to study how dark humor and slang in social media text communications affect virality. The results suggest that in general, both dark humor and slang positively influence social media virality. Conceptually, we posit that dark humor increases virality when the context is understood by all parties involved in

<sup>14</sup> The number of comments may have varied over time, but we use the number of comments at the time of data extraction.

the interaction, as is likely with Redditors browsing WSB. For slang, we posit that the theoretical mechanism driving the positive impact is due to perceptions of slang usage being perceived as warm and casual. Next, we find that the effect of dark humor improving virality holds for persuasive posts but not for informative ones. Slang was seen to improve virality for posts which were neither persuasive nor informative (which we labelled as “neutral”). Finally, we find that the positive effect of dark humor persists for posts which were created after typical working hours. Interestingly, we do not find evidence of slang improving virality for posts created after work hours, but instead find the positive impact on virality for posts made during work hours.

Managers should find these results interesting as social media is increasingly being utilized by various stakeholders (including firms, governments, and consumers) to communicate with one another. Dark humor is commonly found in social media communications, but the impact of dark humor on virality has not been explored in the marketing domain (to the best of our knowledge). Likewise, internet slang has also not been heavily studied despite its prevalent usage. Moreover, understanding the context for when to use these text-related factors, such as considering informative versus persuasive content, may prove to be critical in crafting social media strategies. Practitioners can utilize these findings to craft communications in a manner that can improve communications, as virality or engagement are key to generating positive outcomes which are more tangible (e.g., Akpınar and Berger 2017). In general, this study finds results indicating that the usage of dark humor in slang in social media may assist in generating virality.

Like many papers, this study contains several limitations and areas for future research. First, the mechanisms of this study are only theoretically argued and not empirically explored. Expansions of this research can explore these mechanisms. Second, the empirical approach utilizes a fixed effects approach, but future studies may wish to consider additional causal

inference methods which may help alleviate any remaining endogenous concerns. Third, no financial outcomes are directly linked to dark humor in this study. Finally, the WSB dataset used in this paper is obtained from a public non-institutional data source, which may be biased.

## REFERENCES

- Aaker, Jennifer, Kathleen D. Vohs, and Cassie Mogilner (2010), “Nonprofits are Seen as Warm and For-Profits as Competent: Firm Stereotypes Matter,” *Journal of Consumer Research*, 37 (2), 224–37.
- Akpınar, Ezgi and Jonah Berger (2017), “Valuable Virality,” *Journal of Marketing Research*, 54 (2), 318–330.
- Allen, Johnie J., Sabrina M. Ash, and Craig A. Anderson (2022), “Who Finds Media Violence Funny? Testing the Effects of Media Violence Exposure and Dark Personality Traits,” *Psychology of Popular Media*, 11 (1), 35–46.
- Althoff Tim, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky (2014), “How to Ask for a Favor: A Case Study on the Success of Altruistic Requests,” Proceedings of the International Conference on Weblogs and Social Media (ICWSM), <https://snap.stanford.edu/data/web-RedditPizzaRequests.html>.
- Angrist, Joshua D. and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton, NJ: Princeton University Press.
- Aschwanden, Christie (2016), “We Asked 8,500 Internet Commenters Why They Do What They Do,” FiveThirtyEight (November 28), <https://fivethirtyeight.com/features/we-asked-8500-internet-commenters-why-they-do-what-they-do/>.
- Barron, Sophia Bernazzani (2018), “5 Research-Backed Methods of Achieving Product Virality (and 6 Stories of Companies That Did It),” *HubSpot* (August 6), <https://blog.hubspot.com/service/virality-definition>.
- Bashforth, Emily (2021), “Why do we use dark humour as a coping mechanism?” *Patient* (December 12), <https://patient.info/news-and-features/why-do-we-use-dark-humour-as-a-coping-mechanism>.
- Becker, Maren, Nico Wiegand, and Werner J. Reinartz (2019), “Does It Pay to Be Real? Understanding Authenticity in TV Advertising,” *Journal of Marketing*, 83 (1), 24–50.
- Berg, Matt (2020), “On social media, teens find dark (and sometimes silly) humor amid coronavirus news,” *The Boston Globe* (March 13), <https://www.bostonglobe.com/2020/03/13/lifestyle/social-media-teens-find-dark-sometimes-silly-humor-amid-coronavirus-news/>.
- Berger, Jonah and Katherine L. Milkman (2012), “What Makes Online Content Viral?” *Journal of Marketing Research*, 49 (2), 192–205.
- Berger, Jonah, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel (2020), “Uniting the Tribes: Using Text for Marketing Insight,” *Journal of*



*Marketing*, 84 (1), 1–25.

Bhattacharya, Abhi Neil A. Morgan, and Lopo L. Rego (2021), “Customer Satisfaction and Firm Profits in Monopolies: A Study of Utilities,” *Journal of Marketing Research*, 58 (1), 202–222.

Borah, Abhishek, Sourindra Banerjee, Yu-Ting Lin, Apurv Jain, and Andreas B. Eisingerich (2020), “Improvised Marketing Interventions in Social Media,” *Journal of Marketing*, (84) 2, 69–91.

BSC Team (2020), “All the English Internet Slang You Need to Know in 2020,” *British Study Centres*, <https://www.british-study.com/en/blog/internet-slang>.

Chessey, Kelsey and Phillip Ranly (2022), “Laughter is the best medicine,” *Ipsos* (April 26), <https://www.ipsos.com/en-us/knowledge/new-services/Laughter-is-the-best-medicine>.

Chmielewski, Michael (2018), “Content isn't king, CONTEXT IS KING!” LinkedIn (May 2), <https://www.linkedin.com/pulse/content-isnt-king-context-mike-chmielewski/>.

Colicev, Anatoli, Ashwin Malshe, Koen Pauwels, and Peter O’Connor (2018), “Improving Consumer Mindset Metrics and Shareholder Value Through Social Media: The Different Roles of Owned and Earned Media,” *Journal of Marketing*, 82, 37–56.

Correia, Sergio (2014), “REGHDFE: Stata Module to Perform Linear or Instrumental-variable Regression Absorbing Any Number of High-dimensional Fixed Effects,” *Boston College Department of Economics Statistical Software Components*, <https://ideas.repec.org/c/boc/bocode/s457874.html>.

Dueñas, Angelique N., Karen Kirkness, and Gabrielle M. Finn (2020), “Uncovering Hidden Curricula: Use of Dark Humor in Anatomy Labs and its Implications for Basic Sciences Education,” *Medical Science Educator*, 30, 345–354.

Dionigi, Alberto, Mirko Duradoni, and Laura Vagnoli (2022), “Humor and the dark triad: Relationships among narcissism, Machiavellianism, psychopathy and comic styles,” *Personality and Individual Differences*, 197, 111766.

Eisenbeiss, Maik, Sven A. Hartmann, and Lars Hornuf (2023), “Social media marketing for equity crowdfunding: Which posts trigger investment decisions?” *Finance Research Letters*, 52, 103370.

Eisend, Martin (2009), “A meta-analysis of humor in advertising,” *Journal of the Academy of Marketing Science*, (37), 191–203.

Fossen, Beth L. and David A. Schweidel (2019), “Measuring the Impact of Product Placement with Brand-Related Social Media Conversations and Website Traffic,” *Marketing Science*, 38 (3), 481–99.

Goldfarb, Avi, Catherine Tucker and Yanwen Wang (2022), “Conducting Research in Marketing with Quasi-Experiments,” *Journal of Marketing*, 86 (3) 1–20.

Hartwig, Justin (2021), “WallStreetBets Slang and Memes,” *Investopedia* (February 2021), <https://www.investopedia.com/wallstreetbets-slang-and-memes-5111311>.

Hollenbeck, Brett (2018), “Online Reputation Mechanisms and the Decreasing Value of Chain Affiliation,” *Journal of Marketing Research*, 55 (5), 636–654.

Isaza, Juan (2022), “Being Funny Pays Off: Let's Bring Humor Back To Advertising,” *Forbes* (December 22), <https://www.forbes.com/sites/forbesbusinesscouncil/2022/12/22/being-funny-pays-off-lets-bring-humor-back-to-advertising/?sh=6347da6a1261>.

Judd, Charles M., Laurie James-Hawkins, Vincent Yzerbyt, and Yoshihisa Kashima (2005), “Fundamental Dimensions of Social Judgment: Understanding the Relations between Judgments of Competence and Warmth,” *Journal of Personality and Social Psychology*, 89, 899–913.

Kaplan, Jessica (2023), “These Are the Most Uniquely Popular Texting Acronyms in Every State,” *Reader's Digest* (March 27), <https://www.rd.com/article/most-popular-texting-acronyms-every-state/#:~:text=The%20most%20common%20acronym%20in,uses%20per%20every%20100%2C000%20tweets..>

Lee, Dokyun, Kartik Hosanagar, and Harikesh S. Nair (2018), “Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook,” *Management Science*, 64 (11), 5105–5131.

Li, Xueni (Shirley), Kimmy Wa Chan, and Sara Kim (2019), “Service with Emoticons: How Customers Interpret Employee Use of Emoticons in Online Service Encounters Get access Arrow,” *Journal of Consumer Research*, 45 (5), 973–987.

Liu, Shixiong, Dan-Yang Gui, Yafei Zuo, and Yu Dai (2019), “Good Slang or Bad Slang? Embedding Internet Slang in Persuasive Advertising,” *Frontiers in Psychology*, 10 (1251), 1–12.

McIntyre, Kevin P., James H. Korn, and Hisako Matsuo (2008), “Sweating the small stuff: how different types of hassles result in the experience of stress,” *Stress and Health*, 24 (5), 383–392.

Narayanan, Sridhar, Puneet Manchanda, and Pradeep K. Chintagunta (2005), “Temporal Differences in the Role of Marketing Communication in New Product Categories,” *Journal of Marketing Research*, 42 (3), 278–290.

Park, Sungjin, Hyungdon Kook, Hongdeok Seok, Jae Hyoung Lee, Daeun Lim, Dong-Hyuk Cho, and Suk-Kyu Oh (2020), “The negative impact of long working hours on mental health in young Korean workers,” *PLoS ONE*, 15(8): e0236931.

Fontes, Raphael (2021), “Reddit - r/wallstreetbets,” *Kaggle* (February 16),

<https://www.kaggle.com/datasets/gpreda/reddit-wallstreetsbets-posts>.

Rinker, Tyler (2021), “sentimentr: Calculate Text Polarity Sentiment,” The Comprehensive R Archive Network (Oct 12), <https://cran.r-project.org/web/packages/sentimentr/index.html>.

Rosario, Ana Babić, Kristine de Valck and Francesca Sotgiu (2020), “Conceptualizing the electronic word-of-mouth process: What we know and need to know about eWOM creation, exposure, and evaluation,” *Journal of the Academy of Marketing Science*, 48, 422–448.

Rowe, Alison and Ceryl Regehr (2009), “Whatever Gets You Through Today: An Examination of Cynical Humor Among Emergency Service Professionals,” *Journal of Loss and Trauma*, 15, 448–464.

Rust, Roland T., William Rand, Ming-Hui Huang, Andrew T. Stephen, Gillian Brooks, and Timur Chabuk (2021), “Real-Time Brand Reputation Tracking Using Social Media,” *Journal of Marketing*, 85 (4), 21–43.

Schnurr, Stephanie and Charley Rowe (2008), “The “Dark Side” of Humour. An Analysis of Subversive Humour in Workplace Emails,” *Lodz Papers in Pragmatics*, 4 (1), 109–130

Schwarz, Carlo (2018), “ldagibbs: A command for topic modeling in Stata using latent Dirichlet allocation,” *The Stata Journal*, 18 (1), 101–117.

Scott, Cliff, David M. Klein, and Jennings Bryant (1990), “Consumer Response to Humor in Advertising: A Series of Field Studies Using Behavioral Observation,” *Journal of Consumer Research*, 16 (4), 498–501.

Shankar, Venkatesh and Tarun Kushwaha (2021), “Omnichannel marketing: Are cross-channel effects symmetric?” *International Journal of Research in Marketing*, 38, 290–310.

Tepper, Taylor and Benjamin Curry (2021), “Reddit IPO: What You Need To Know,” *Forbes* (December 16), <https://www.forbes.com/advisor/investing/reddit-ipo/>.

Tellis, Gerard J., Deborah J. MacInnis, Seshadri Tirunillai, and Yanwei Zhang (2019), “What Drives Virality (Sharing) of Online Digital Content? The Critical Role of Information, Emotion, and Brand Prominence,” *Journal of Marketing*, 83 (4), 1–20.

Tucker, Catherine E. (2015), “The Reach and Persuasiveness of Viral Video Ads,” *Marketing Science*, 34 (2), 281–296.

Vincent, James (2021), “WallStreetBets donates more than \$350,000 to gorilla charity to prove ‘apes together strong’,” The Verge (March 18), [https://www.theverge.com/2021/3/18/22337670/wallstreetbets-donate-gamestop-profits-gorillas-fund-apes-together-strong?fbclid=IwAR0p\\_yJFnTCHiwiwoR5EsRGNg1nybMaZY2lc6-vJWwpWcoCvD0siTnaPsS0](https://www.theverge.com/2021/3/18/22337670/wallstreetbets-donate-gamestop-profits-gorillas-fund-apes-together-strong?fbclid=IwAR0p_yJFnTCHiwiwoR5EsRGNg1nybMaZY2lc6-vJWwpWcoCvD0siTnaPsS0).

Williams, Alex (2020), “It’s OK to Find Humor in Some of This,” *The New York Times* (April 2022), <https://www.nytimes.com/2020/04/22/style/coronavirus-humor.html>.

Willinger, Ulrike, Andreas Hergovich, Michaela Schmoeger, Matthias Deckert, Susanne Stoettner, Iris Bunda, Andrea Witting, Melanie Seidler, Reinhilde Moser, Stefanie Kacena, David Jaeckle, Benjamin Loader, Christian Mueller, and Eduard Auff (2017), “Cognitive and emotional demands of black humour processing: the role of intelligence, aggressiveness and mood,” *Cognitive Processing*, 18, 159–167.

Zohar, Dov (1999), “When things go wrong: The effect of daily work hassles on effort, exertion and negative mood,” *Journal of Occupational & Organizational Psychology*, 72 (3), 265–283.