Essays on User Behavior in Online Platforms

Kyung Min Park

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Yong Tan, Chair

Ming Fan

Stephanie Lee

Program Authorized to Offer Degree:

Business Administration

University of Washington

Abstract

Essays on User Behavior in Online Platforms

Kyung Min Park

Chair of the Supervisory Committee:
Michael G. Foster Endowed Professor Yong Tan
Information Systems and Operations Management

In this dissertation, I investigate user behavior in two different domains: online shopping and online learning. The first essay examines the provision of helpful product video reviews. As the provision of online review videos grows and consumers increasingly rely on them for their purchase decisions, understanding factors that contribute to the perceived helpfulness of video reviews becomes critical for video review management. I find that video viewers perceive online product review videos that are more visually stimulating, more positive, and less subjective as more helpful. In addition, review videos are perceived as more helpful when reviewers reveal their faces in their reviews and when reviewers speak with faster speech rates and lower voice pitch. The second essay studies the dynamics of students' online social engagement in the context of online learning and how the recent COVID-19 pandemic affected it. I propose a Hidden Markov

Model (HMM) with students' self-efficacy as hidden states to investigate the dynamics of elementary school students' online social behavior. The result suggests that the amount of proactive social activities from teachers, peers, and parents have positive effects on students' social engagement in any given self-efficacy state. The analysis on state transition suggests that the social influences from teachers, parents, and peers on students' self-efficacy are positive for students with high self-efficacy. When students have lower self-efficacy, the peer influence on their self-efficacy is negative. However, the positive influences from teachers and parents can mitigate such negative peer influences. The results also imply that students' self-observation of their past behavioral patterns positively affects their self-efficacy. Finally, the results suggest that the relative importance of parental influence and students' personal influence increased during the pandemic. The reduced face-to-face schooling can explain such change during the pandemic.

TABLE OF CONTENTS

Page

List of Figures …………………………………………………………………………………… iv

List of Tables ……………………………………………………………………………………….. v

Chapter 1. Introduction ……………………………………………………………………….... 1

Chapter 2. Literature review ………………………………………………………………….... 4

2.1. Provision of helpful review videos ……………………………………………….. 4

2.2. Social interaction in online learning ………………………………………………… 6

Chapter 3. Provision of helpful review videos: Effect of video characteristics on perceived helpfulness …………………………………………………………………………………….. 8

3.1. Introduction ……………………………………………………………………………… 8

3.2. Theoretical Framework ……………………………………………………………….. 13

3.2.1. Video Characteristics and Perceived Helpfulness of Video Reviews ……. 13

3.2.2. Voice Feature and Perceived Helpfulness of Video Reviews …………… 15

3.2.3. Review Content and Perceived Helpfulness of Video Reviews ………….. 17

3.3. Data and Empirical Framework ………………………………………………….. 20

3.3.1 Product Review Video Data ……………………………………………….. 20

3.3.2. Key Variables Construction ……………………………………………… 21

3.4. Empirical Model ……………………………………………………………………… 31

3.4.1. Multinomial Propensity Score Weighting ……………………………….. 32

3.4.2. Tobit Model ……………………………………………………………… 33

3.5. Estimated Results …………………………………………………………………... 35

3.5.1. Main Results ……………………………………………………………... 35

3.5.2. Robustness Checks 36

3.6. Controlled Experiment 42

3.6.1. Procedure and Measures 42

3.6.2. Results and Discussion 43

3.7. Conclusions 46

3.7.1. Managerial and Practical Implications 47

3.7.2. Future Research and Limitations 47

Chapter 4. Student Social Engagement in Online Learning: An Empirical Study 49

4.1. Introduction 49

4.2. Theoretical Framework 53

4.2.1. Student self-efficacy 53

4.2.2. Short-term Social Influence and Student Personal Characteristics 55

4.3. Empirical Model 56

4.3.1. Hidden Markov Model (HMM) for Student Online Social Engagement ... 56

4.3.2. State Transition Matrix 58

4.3.3. State-Dependent Outcome 59

4.3.4. Estimation 60

4.4. Data 61

4.4.1. Research Context 61

4.4.2. Variable Construction 62

4.5. Estimation Results 67

4.5.1. State-dependent outcome 70

4.5.2. State Transition 72

4.6. Conclusions ………………………………………………………………………... 77

Chapter 5. Concluding Remarks …………………………………………………………... 79

Bibliography ………………………………………………………………………………. 81

Appendix A: Data Collection Process ……………………………………………………... 93

Appendix B: Visual Variability Construction ……………………………………………... 95

Appendix C: Distribution of Helpfulness ………………………………………………….. 98

Appendix D: Topic Modeling with Latent Dirichlet Allocation (LDA) Algorithm …………….. 99

Appendix E: Propensity Score Weighting Validation ……………………………………..102

Appendix F: Robustness Checks using Alternative Samples ………………………………….. 105

Appendix G: Time Discounting Factor in Exponential Smoothing ……………………………. 106

LIST OF FIGURES

Page

Figure 3.1. Examples of frame distance 23

Figure 4.1. Conceptual Framework 55

Figure 4.2. Hidden Markov Model of Students' Online Social Behavior 57

Figure B.1. Visual Variability Construction 96

Figure B.2. Examples of Frame Distance 97

Figure C.1. Distribution of Helpfulness 98

Figure D.1. Topic Coherence Scores 100

Figure E.1. Summary of Balance Metrics 104

LIST OF TABLES

Page

Table 3.1. Hypotheses and Variables …………………………………………………………… 19

Table 3.2: Variable Description …………………………………………………………………… 28

Table 3.3: Descriptive Statistics …………………………………………………………………… 29

Table 3.4: Number of Review Videos in Each Product Category ……………………………… 29

Table 3.5: Effects of Review Characteristics on Perceived Helpfulness ……………………….. 38

Table 3.6. Robustness Check – Alternative Measurement of Explanatory Variables ………….. 39

Table 3.7. Robustness Check – Alternative Measurement of Perceived Helpfulness ………….. 41

Table 3.8. Differences in Review Helpfulness Scores ………………………………………… 45

Table 4.1. Variable Description …………………………………………………………………… 65

Table 4.2. Descriptive Statistics …………………………………………………………………… 66

Table 4.3. Selection of the Number of States …………………………………………………… 68

Table 4.4. Estimates for 2-state HMM ………………………………………………………….. 69

Table 4.5. Intrinsic state transition matrix ……………………………………………………… 72

Table 4.6. Changes in state transition matrix – Social feedback ………………………………. 75

Table 4.7. Changes in state transition matrix – Self-observation ……………………………… 76

Table D.1. 19 Topics and Their Top 10 Relevant Keywords ………………………………...… 101

Table F.1. Robustness Check – Estimation Result Employing Different Samples …………….. 105

Table G.1. Evaluation of different discount factor ……………………………..………...….. 106

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Yong Tan. I am grateful for his unconditional support, invaluable advice, enthusiasm, and patience during my PhD study. Without his assistance and dedicated involvement in every step throughout the past five years, I would have never made this far of doctoral study. I will never forget the moments when he supported me and believed in my capabilities when even I was not sure about myself. He was beyond an academic mentor for me. I extremely feel fortunate to have him as my advisor.

My thanks also go to the other committee members: Stephanie Lee, Ming Fan, and Quan Wen. Their insightful comments and constructive feedback have guided me to improve this dissertation. Especially, I was very lucky to have a chance to work with Stephanie Lee. Her immense knowledge and passion for research have encouraged me all the time of my PhD study. I truly feel thankful for her continuous support. My appreciation also goes to Ming Fan, as his enthusiasm and far-sighted suggestions have certainly inspired me. I am also grateful to Quan Wen for providing helpful comments and teaching logical thinking.

My special thanks also go to Shawna Reimers, Jaime Banaag, Beau Kirkeby, and other staff at the Foster School of Business for their consistent support and caring during my study. I thank so much to my peers in the Mackenzie Hall who have supported me throughout this journey: Zhijin Zhou, Tongxin Zhou, Kyung Sun Rhee, Zixuan Meng, Vipul Aggarwal, Keehae Moon, and Seo Eun Chang. I will always remember the discussions and joyful times I had with you.

Finally, I sincerely thank my friends back in Korea who have sent unlimited emotional supports across the sea: Hyewon Lee, Yejin Ham, Jiyoung Shin, Bumsoo Kim, Gwang Kim, and Minkyu (Mark) Kim. Thanks for staying with me during hard times and celebrating my achievements.

DEDICATION

I dedicate my dissertation to my beloved family.

To my parents, Youngshin Park and Dongsook Kim, who have supported me with endless love. Throughout my study, you have always been there to offer words of encouragement. Without your support, I would have never been completed this journey.

To my husband, Alex Amato, who has been my best friend and love of my life. I will never finish appreciating you for everything you do every day for me.

To my sister and brother-in-law, Kyungjin Park and Hyunwoo Sohn, who have been cheering me and sharing the happiness of my achievements.

Chapter 1. INTRODUCTION

Thanks to the widespread of the Internet and the improvement in relevant technologies, various offline activities have been transformed into online activities. Online settings have distinctive characteristics (e.g., limited social presence) compared to their corresponding offline settings. Therefore, online users may have different decision-making strategies or express unique behavioral patterns. Accordingly, it is crucial to understand online-specific user behavior to design effective business platforms. In this dissertation, I investigate user behavior in two different domains: online shopping and online learning.

The first essay examines the provision of helpful product review videos. With the rapid growth and popularity of YouTube, an increasing number of consumers rely on online product review videos to obtain product-related information. Visual and sound cues in review videos can convey product experiences in rich and immersive ways and have a powerful impact on consumers' perceptions and intentions to purchase the product. As the provision of online review videos grows and consumers increasingly rely on them for their purchase decisions, understanding factors that contribute to the perceived helpfulness of video reviews becomes critical for video review management. The first essay examines how various visual characteristics —visual stimulation level and face presence—, voice characteristics of reviewers—speech rate, pitch, and speech dynamics—, and review content characteristics—positivity and subjectivity— are associated with the perceived helpfulness. I collect detailed data on 13,840 electronic product review videos posted on YouTube and employ video content analysis, speech recognition, and natural language processing techniques to extract visual, speech, and sentimental characteristics. I find that video viewers perceive online product review videos that are more visually stimulating, more positive, and less subjective as more helpful. In addition, review videos are perceived as more helpful when reviewers reveal their faces in their reviews and when reviewers speak with faster speech rates and lower voice pitch.

The second essay studies the dynamics of students' online social engagement and how the recent COVID-19 pandemic affected it. I analyze elementary school students' behavior data in the context of an online learning platform. This study relies on Bandura's social cognitive theory as a theoretical ground to understand student social behavior (Bandura, 1977). Building upon social cognitive theory, I posit three significant factors that affect students' social behavior in online learning communities: (1) students' self-efficacy, (2) social influence from others, and (3) students' self-observation of their behavior patterns. Self-efficacy plays a distinctive role in the causal structure of Social Cognitive Theory by affecting individuals' motivation of performing behaviors. However, self-efficacy is often unobservable. To address this issue, I propose a Hidden Markov Model (HMM) to investigate the dynamics of students' online social engagement and to what extent the external shock from the recent COVID-19 pandemic brought in the dynamics. The result suggests that the amount of proactive social activities from teachers, peers, and parents have positive effects on the amount of students' social activities in a given self-efficacy state. The analysis on state transition suggests that the social influences from peers, teachers, and parents on students' self-efficacy are positive for students with higher self-efficacy. For students with lower self-efficacy, the social influence from peers is negative, which can be potentially explained by peer dynamics. However, the social influences from teachers and parents remain positive for the students with lower self-efficacy. This result suggests that the positive impact from adults can mediate the negative impact from peers and highlights the importance of the social role of adults in online learning environments. The result also indicates that students' self-observation of their past social interactions positively affects their self-efficacy. During the pandemic, the relative effect of parents' social proactive activities on students' social engagement increases compared to that of teachers and peers. Also, the relative effect of students' self-observation on students' self-efficacy increases during the pandemic, whereas that of other users' social feedback decreases during the pandemic. The reduced face-to-face schooling can explain such change during the pandemic.

Chapter 2. LITERATURE REVIEW

2.1. Provision of Helpful Review Videos

The previous literature on text reviews highlights the importance of helpful product reviews in multiple aspects. Helpful and informative reviews can help companies attract new consumers and increase their brand and product awareness (e.g., Kim and Choi 2012, Berger et al. 2010). For example, helpful reviews can reduce perceived risks associated with making purchase decisions in online shopping. Consumers' perceived risks when making purchase decisions are closely related to consumers' perception of retailer credibility, and perceived risks can be mediated by consumer review credibility (Kim and Choi 2012). In addition, Berger et al. (2010) show that even negative reviews can increase product purchase likelihood and sales when the product is relatively unknown by increasing product awareness. Furthermore, from the perspective of online platforms, the presence of more helpful and credible reviews can help online platforms grow by increasing consumers' loyalty to platforms (Chevalier and Mayzlin 2006, Connors et al. 2011, Spool 2009). Relatedly, this essay also contributes to a stream of literature that examines factors that affect the perceived helpfulness of online text reviews. The perceived helpfulness of a review is defined as the perceived value of information diagnosticity of a review (Mudambi and Schuff 2010, Yin et al. 2014). The prior literature, which exclusively focuses on online text reviews, identifies different factors, including product-related features (Mudambi and Schuff 2010), reviewer-related features (Wang et al. 2019), and reviews' textual characteristics (Ghose and Ipeirotis 2011, Yin et al. 2014), that affect text reviews' perceived helpfulness. Despite the presence of the prior literature that examines the perceived helpfulness of text reviews, the literature on the perceived helpfulness of review videos is absent so far.

Online review videos, with more realistic and dynamic visual and sound cues, when compared with traditional text reviews, can convey product experiences in richer ways and influence consumers' information acquisition, perceptions about the product, and intentions to purchase the product. The media richness theory from the information systems literature shows that richer media that contains more communication cues, such as visual and vocal information, is more effective in delivering messages (Daft and Lengel 1986; Dennis and Kinney 1998). Relatedly, the multimedia learning theory also shows that information presented by utilizing multiple sensorimotor channels, such as including human speech, environmental sounds, images, and videos, can enhance learning performance and experience, and ultimately help form mental representations of objects (Mayer 1997; Mayer and Moreno 1999; Carney and Levin 2002). The prior literature shows that videos can attract consumers' attention by presenting stimuli for different processing channels and enhance cognition (Hong et al. 2004; Jiang and Benbasat 2007). Xu et al. (2015) conducts a laboratory experiment to show that video-based online reviews are perceived as more credible, helpful, and persuasive than text online reviews. Although the prior literature supports the potential merits of video reviews, the literature on how to design and provide helpful video reviews is absent. Therefore, I extend the literature to examine how various video content characteristics and voice characteristics of reviewers are associated with the perceived helpfulness so that appropriate guidelines can be developed for consumers, businesses, review platforms, and product reviewers.

2.2. Social interaction in online learning

Online learning communities usually contain three major components - instruction, technology, and social interaction (Carabajal et al., 2003; Tu and Corry, 2002). Instruction contains learning materials, contents, and resources. Technology offers a platform and communication tools that realize online teaching and learning. The last component, social interaction, enables students to maintain some degree of sense of belonging and mutual social-emotional bonds that leads to good relationships with other students and teachers. The social dimension of online learning is not directly related to learning tasks; however, it is a crucial component of effective online learning. In nature, online learning environments often have limited social interactions compared to traditional face-to-face classroom learning. This limitation makes students experience the feeling of isolation in online learning, which is a great challenge for educators (Palloff and Pratt, 1999). Isolation can affect students' attitudes towards online learning and hinder students from achieving successful learning in online environments, and insufficient interactions with other students and instructors may lead to the feeling of isolation (McInnerney and Roberts, 2004). Indeed, in Muilenburg and Berge's (2005) survey-based factor analysis study, they found that a lack of social interaction was the single most critical hindrance to students in online learning. Sit et al. (2005) also revealed that the inadequate opportunity for human interaction is a significant barrier in online learning and makes students feel negative towards their experience in online learning. To overcome the limited social interactions in online learning, Harrisim et al. (1995) suggested that providing a space for informal discourse, such as a virtual cafe, is crucial to designing an effective online educational platform. They further explained that "computer-mediated communication is capable of supporting socio-emotional communication as well as task-oriented communication; in fact, without personal communication, the group will not be nurtured". Therefore, fostering students' social interaction engagement is essential to increase students' learning effectiveness. In the second essay, I study the dynamics of students' online social engagement and how the recent COVID-19 pandemic affected it.

Chapter 3. PROVISION OF HELPFUL REVIEW VIDEOS: EFFETS OF VIDEO CHARACTERISTICS ON PERCEIVED HELPFULNESS

3.1. Introduction

With the pervasiveness and popularity of social media and online reviews, the provision, operation, and management of online consumer reviews have become increasingly important (Chen and Xie 2008, Gu and Ye 2014, Sun and Xu 2018). Online reviews, which became a new word-of-mouth (WOM) channel, can help companies attract consumers, increase their brand and product awareness, improve consumer loyalty, and increase sales, and, as a result, online reviews are recognized as a valuable asset to business organizations (e.g., Chevalier and Mayzlin 2006, Kim and Choi 2012, Berger et al. 2010). Notably, with the recent rapid growth and popularity of YouTube—the world's largest video sharing platform—an increasing number of consumers rely on social media and online review videos to obtain product-related information. In fact, more than 50 percent of consumers today utilize online videos before they make any purchase decisions, and over 50,000 years' worth of product review YouTube videos have been watched on mobile devices between July 2015 and June 2017.[1] Compared with traditional text reviews, online review videos, with visual and sound cues, can convey product experiences in richer and more immersive ways and have a powerful impact on consumer's perceptions and intentions to purchase the product; the media richness theory and the multimedia learning theory show that employing multiple media cues makes messages more clear, salient, and attention-grabbing (Daft and Lengel 1986; Mayer

[1] Think with Google, Product Review Video Watch Time Statistics, https://www.thinkwithgoogle.com/data/product-review-video-watch-time-statistics/

1997; Dennis and Kinney 1998; Xu et al. 2015). Although online review videos have become a new major review channel in need of management, the literature on review videos has been very limited so far, and I aim to address the gap in the literature.

As the provision of online review videos grows and consumers increasingly rely on them for their purchase decisions, understanding factors that contribute to the perceived helpfulness of video reviews becomes critical for online video review management. There has been a dramatic increase in the number of review videos that are posted on YouTube,[2] and the availability of too many reviews and excessive information can generate an information overload problem (Jones et al. 2004), as consumers may only need a small number of helpful reviews when making purchase decisions. Helpful reviews help consumers make better decisions and maximize consumer's satisfaction (Kohli et al. 2004, Qazi et al. 2016). From the perspectives of companies and online platforms, understanding which product review videos are more helpful and promoting them improves consumers' satisfaction, engagement, and loyalty (Chevalier and Mayzlin 2006, Connors et al. 2011). In addition, identifying factors that affect perceived helpfulness can help product reviewers create more helpful review videos, and, consequently, generate more video views, which can directly increase reviewers' earnings from online platforms. For example, YouTubers can make $3 to $5 for every 1,000 video views.[3] As the first study to examine the perceived helpfulness of product review videos, this study examines online product review videos and identifies key video features that affect the perceived helpfulness of review videos to facilitate the management and operation of online review videos.

---

[2] Think with Google, Why Overwhelmed Shoppers are Turning to Online Video for Help, https://www.thinkwithgoogle.com/data-collections/shopping-research-data/

[3] Intuit Turbo, How Much Do YouTubers Make & How to Become a Youtuber, https://turbo.intuit.com/blog/relationships/how-much-do-youtubers-make-5035/#2

I examine review videos from multiple dimensions and analyze 1) visual characteristics of review videos, 2) reviewers' vocal characteristics of review videos, and 3) the actual content and messages that reviewers include in their review videos. More specifically, I first develop and examine hypotheses on the association between visual characteristics —visual stimulation level and face presence—and perceived helpfulness. First, I build upon the literature on visual attention and cognitive information processing to posit that consumers will perceive more visually stimulating videos as more helpful. Second, I rely on prior social presence literature to hypothesize that consumers will perceive video reviews as more helpful when reviewers reveal their faces in their review videos. In addition, I explore the influence of reviewers' voice characteristics on perceived helpfulness. As the literature on phonetics and communications finds that voice characteristics affect the credibility and persuasiveness of speech, I also examine voice characteristics—speech rate, pitch, and pitch variability—of helpful review videos. Finally, I explore the effect of review content characteristics—positivity and subjectivity—and perceived helpfulness. I build upon prior psychology literature on the need for cognition (NC) theory to posit that consumers may perceive positive review videos as more helpful. Lastly, I hypothesize that more subjective reviews would be perceived as less helpful because subjective messages can hurt the credibility and persuasiveness of messages.

To examine the hypotheses on video features of helpful review videos, I collect detailed data on 13,840 electronic product review videos that are published between February 2016 and January 2019. Electronic products are one of the most common product types for review videos, and 64 percent of electronics shoppers watch electronic review videos on YouTube one week prior to actual purchases (Jarboe 2015). In addition to collecting data on basic video and reviewer information, I utilize various techniques to collect and process the data. I conduct video frame analysis and shot-level label detection to extract visual characteristics of review videos. Also, to analyze the vocal characteristics of reviewers, I extract fundamental frequency contours of voice data, which characterizes the speaking style of a person. Furthermore, I employ speech recognition techniques to generate video transcriptions and analyze the content of reviewers' messages. Then, I utilize natural language processing techniques to extract sentimental features from speech transcriptions.

I employ a reviewer fixed effects model to empirically examine the extent to which visual, voice, and sentiment characteristics affect the perceived helpfulness of product review videos. The model includes a wide range of control variables on video and review characteristics. To control for other review content, I use the Latent Dirichlet Allocation (LDA) model (Blei et al. 2003), a machine-learning algorithm model that reduces information dimensionality and discovers latent topics of review videos. I additionally utilize a propensity score weighting method to correct for observable characteristics that may have affected reviewers' choices of which product type to review. Furthermore, I employ a Tobit model as an additional empirical strategy to account for the bounded nature of the review helpfulness score. Consistent with the hypothesis, I find that the increase in visual stimulation and the reviewer face presence increases the perceived helpfulness of reviews. Furthermore, faster speech rate and lower voice pitch are associated with higher perceived helpfulness of reviews. In online review videos, more positive and objective information is perceived as more helpful. The results are robust to various robustness checks, including employing alternative measurements of key explanatory variables and perceived helpfulness.

I additionally conduct two controlled experiments to isolate the causal effects of review characteristics on the perceived helpfulness. In the controlled experiments, participants watched and evaluated two review videos. I manipulated vocal characteristics at two levels (lower-pitched voice vs. higher-pitched voice) in the first experiment, and I manipulated visual characteristics at two levels (brighter video vs. darker video) in the second experiment. The results show that the decrease in voice pitch increase in the video brightness improves the perceived helpfulness when all other review video characteristics remain constant, which is consistent with the previous findings.

The rest of this chapter is organized as follows. In the next section, I discuss the related literature, provide theoretical background, and develop the hypothesis. In Section 3.3, I provide more details regarding the research setting and describe the data in detail. In Section 3.4, I describe the empirical models. Section 3.5 presents the estimation results and various robustness checks to ensure that the results are robust. Section 3.6 conducts controlled experiments and presents the results. Section 3.7 concludes and discusses implications for future work.

3.2. Theoretical Framework

In the remainder of this section, I develop hypotheses on the association between video content characteristics—reviews' information accessibility, positivity, and subjectivity—and voice characteristics of reviewers—speech rate, pitch, and speech dynamics—and perceived helpfulness.

3.2.1. Video Characteristics and Perceived Helpfulness of Video Reviews

Visual Stimulation

Higher visual stimulation in video can improve perceived review helpfulness by improving viewers' attention span. Attention is the first step of the psychological process of information receiving and processing (Chang et al. 2019). Miller (1987) proposed a hierarchical model of cognitive processing and style. The model describes cognitive processing in multiple stages (perception, memory, and thought), where individuals can identify different cognitive styles. Specifically, the perception stage consists of pattern recognition and attention. Studies on attention have also found that attention plays an important role in enhancing visual memory (Inoue & Takeda, 2012) and learning effectiveness (Chang et al., 2019; Chen and Wu, 2015). A learning process needs sustained attention to achieve effective learning and memory (Chen and Wu, 2015). Therefore, retaining attention can affect perceived helpfulness of review by increasing the effectiveness of viewers' information processing. However, humans' attention span is limited, and voluntary attention is actually a continual returning of attention to its task when the attention wanders away (Broadbent, 1958). According to the visual attention literature, various type of visual stimuli (e.g., motion, luminance, objects) predicts visual attention and gaze time (for an extensive review of various visual attention models, see Borji and Itti (2012)). Therefore, a visually stimulating review video can help viewers to pay continuous attention to the review video and receive the information more efficiently. Therefore, I hypothesize that:

H1: *More visually stimulating videos are perceived as more helpful.*

Facial Presence of Content Creator

Content creators can expose and show their faces in review videos. Alternatively, creators can choose to create review videos without showing their faces. I build upon the literature on social presence (Fulk et al. 1990; Gefen and Straub 2004) to posit that featuring reviewers' faces in review videos can increase the perceived helpfulness of videos by increasing social presence in the review video. The social presence refers to "the degree to which the medium facilitates awareness of the other person and interpersonal relationships during the interaction" (Fulk et al. 1990). In face-to-face communications, nonverbal cues, such as facial expression, eye movement, and posture, contribute to the degree of social presence in communication (Short et al. 1976). Traditional electronic communication media, such as emails, text messages, and websites, has a much lower social presence when compared with face-to-face communications (Rice, 1993). However, showing a reviewer's face in review videos can increase the degree of social presence in the review video by conveying nonverbal communication cues such as facial expressions, eye contact, and posture.

The prior literature on social presence theory highlights the importance of social presence in understanding information and trusting the source of information. The prior literature on online retail context finds that increasing the social presence in their website (e.g., increasing a sense of human contact, personalness, and sociability) can have a positive impact on consumers' perception of the retailer's credibility and their purchase intentions (Gefen and Straub 2004; Lu et al. 2016). The literature on online learning shows that perceived social presence predicts perceived learning and satisfaction in online courses (Kizilcec et al. 2014; Lyons et al., 2012; Richardson and Swan, 2003). Despite the popularity of online video platforms, there has been limited literature on the effect of social presence on online videos. Social presence in a review video can improve the credibility of information and facilitate consumers' understanding of review videos, which, in turn, can improve the perceived helpfulness of the review video. Therefore, I hypothesize that showing content creators' faces in review videos increase the social presence and improves the perceived helpfulness of review videos.

H2: *Review videos that show reviewers' faces are perceived as more helpful.*

3.2.2. Voice Features and Perceived Helpfulness of Video Reviews

In video reviews, reviewers use their own voices to deliver review information. Therefore, distinct from text reviews, video reviews embody additional vocalic cues that are not present in text data. The literature on phonetics and communications finds that voice characteristics affect the credibility and persuasiveness of speech (see Burgoon et al. (1990) and Peterson et al. (1995) for a comprehensive review), which are important determinants of perceived review helpfulness. The prior literature finds that the persuasiveness of a speaker depends on the speaker's vocal cues, and experimental and empirical studies in various domains have shown that voice characteristics of a speaker, including speech rate, pitch, and variability, significantly affect the perceptions of the speaker's credibility, persuasiveness, and attractiveness (e.g., Apple et al. 1979, Hall 1980, Burgoon et al. 1990, Chebat et al. 2007). Therefore, I posit three hypotheses on the effects of voice characteristics—speech rate, pitch, and variability—on perceived review helpfulness.

First, speech rate is found to be positively correlated with higher speech credibility (Apple et al., 1979; Chebat et al., 2007; Peterson et al., 1995; Rosenberg and Hirschberg, 2009). For example, in the context of marketing, Peterson et al. (1995) show that a successful salesperson exhibits a higher speaking rate. A different stream of literature shows that lower-pitched voices increase credibility and effectiveness of speech. Apple et al. (1979) conducted experiments with male speakers' speech answering two interview questions and found that speakers with higher voice pitch were judged less emphatic and less truthful. In examining political speech, Tigue et al. (2012) find that candidates' speech pitch has a significant impact on voting behavior, and male politicians with lower-pitched voices have advantages in elections. Similarly, Laustsen et al. (2015) also find that men and women prefer to vote for candidates, both male and female, with lower-pitched voices. Furthermore, Wang et al. (2018) find that lower-pitched voices had greater persuasiveness in conversations than higher-pitched voices. Finally, greater pitch variability can improve speech perception and improve the persuasiveness and credibility of messages. Burgoon et al. (1990) find that greater pitch variability enhances the pleasantness of voice that fosters favorable judgment. Similarly, Rosenberg and Hirschberg (2009) find that variability in pitch and amplitude is greater in persuasive and attractive speech.

Despite the prevalence of online review videos, there have not yet been empirical studies that examine the effect of voice characteristics on perceived helpfulness, as review videos are a relatively newer form of product reviews. To understand factors that affect perceived helpfulness and effectiveness of review videos, it is important to understand the effects of not only speech content but also speech voice characteristics. Therefore, I build upon literature on communications and hypothesize that:

H3: *Speech rate in a review video has a positive effect on the perceived review helpfulness.*

H4: *Reviewer's voice pitch has a negative effect on the perceived review helpfulness.*

H5: *Review videos with a greater pitch variability are perceived more helpful than review videos* with a smaller pitch variability.

3.2.3. Review Content and Perceived Helpfulness of Video Reviews

Review Positivity.

Reviewers can include positive and negative information in their product review videos. By building on the prior psychology literature on the need for cognition (NC) theory, I posit that positivity bias may be present in YouTube review videos. The NC, first introduced and developed in Cacioppo and Petty (1982), refers to the extent to which individuals engage in active information processing and enjoy effortful cognitive endeavors. As YouTube is not a retail platform, to watch product review videos, consumers need to exert extra effort and search for videos. Therefore, consumers who watch product review videos on YouTube are likely to have a high level of NC because they made conscious efforts to search for and watch review videos. The NC theory finds that individuals with a high level of NC are less likely to change their original beliefs, and that they are more resistant to messages that counter their original beliefs about products (Haugtvedt and Petty 1992). Furthermore, Tormala and Petty (2002) find that when consumers resist persuasive messages that counter their initial beliefs, consumers become more convinced of their initial beliefs. At the same time, consumers who view product review videos on YouTube are likely to have a positive attitude towards the product. For example, 55 percent of consumers search for a product on Google, and then visit YouTube to learn more about the product they already have in mind (Haller 2019). In addition, consumers who watch product review videos are more likely to visit a store and purchase the same product that they saw in videos.[4]

---

[4] Think with Google, Online Video Shopping Decisions, https://www.thinkwithgoogle.com/data/online-video-shopping-decisions/

Based on the NC theory and observations that consumers who watch online review videos on YouTube are likely to 1) have a high level of NC and 2) have a positive initial attitude toward the product, I predict that YouTube viewers would disfavor product review videos with negative reviews that contradict their original positive perception of products. Consumers may instead perceive positive review videos as more helpful because positive reviews would strengthen their initial beliefs of the product. As a result, I hypothesize that:

H5: *Positive review videos are perceived as more helpful than negative review videos*.

Review Subjectivity.

Product reviews can contain subjective and objective information, and the literature on consumer skepticism finds that consumers are more skeptical of subjective claims than of objective claims (Ford et al. 1990), which is consistent with Nelson's (1970) findings that experience-based claims are less credible than objective claims in describing product attributes. Therefore, the level of subjectivity can affect the perceived helpfulness of reviews by influencing the credibility of messages, and subjective text reviews are perceived as less helpful when compared with objective text reviews (Ghose and Ipeirotis 2011). I hypothesize that more subjective video reviews would be perceived as less helpful. Subjective messages can hurt the credibility and persuasiveness of messages (Appelman and Sundar 2016, Ford et al. 1990), and additional visual and sound aids that are present on video reviews can help consumers more easily assess subjectivity of video reviews. I, therefore, hypothesize that:

H6: *Subjective review videos are perceived as less helpful than objective review videos*.

In Table 3.1, I summarize the six hypotheses on the effect of review characteristics on perceived review helpfulness.

Table 3.1. Hypotheses and Variables

Hypotheses	Variable
H1: More visually stimulating videos are perceived as more helpful.	*Visual Variability* Brightness
H2: Review videos that show reviewers' faces are perceived as more helpful.	*Face Presence*
H3: Speech rate in a review video has a positive effect on the perceived review helpfulness.	*Speech Rate*
H4: Reviewer's voice pitch has a negative effect on the perceived review helpfulness.	*Pitch*
H5: Review videos with a greater pitch variability are perceived more helpful than review videos with a smaller pitch variability.	*Pitch Variability*
H6: Positive review videos are perceived as more helpful than negative review videos.	*Positivity*
H7: Subjective review videos are perceived as less helpful than objective review videos.	*Subjectivity*

3.3. Data and Empirical Framework

3.3.1. Product Review Video Data

To examine the hypotheses, I examine electronic product review videos uploaded on YouTube. I focus on electronic product review videos because they are one of the most popular product types for review videos. I identify and collect data on product review videos uploaded on YouTube by the following steps. First, I select 40 keywords and combine each keyword with the word "review" to identify and search for electronic product review videos that were published between February 2016 and January 2019. Examples of search keywords include "iPad review" and "Samsung Galaxy review." For example, a product review video titled "9.7 Inch iPad Pro Review" published on March 25, 2016, which is identified from the search keyword "iPad review," reviews an electronic product, Apple iPad Pro (9.7-inch). After keyword-based searches, I exclude videos with less than 1,000 views to ensure that each video had enough viewers to determine the helpfulness of the review video. I exclude videos with less than ten total votes to ensure that like and dislike votes can represent the perceived review helpfulness of review videos.[5] As I focus on review videos that primarily use speech to deliver their core information, I exclude videos that convey their contents by primarily using written language or visual aids without using many verbal descriptions. To do so, I exclude videos with an average duration-to-word larger than one second per word. The threshold is determined based on the observation that the average conversation rate for English speakers in the U.S. is 150 words per minute. I additionally observe a discontinuity in duration-to-word-count values around the threshold. After the data collection process, the data includes a total of 13,840 electronic review videos uploaded between February 2016 and January

[5] I do a robustness check on different samples and show the result in Appendix F.

2019. Appendix A provides additional details regarding the data collection and cleaning process.

For all 13,840 electronic product review videos, I employ multiple data sources and techniques to obtain information on 1) visual characteristics, 2) speech characteristics, and 3) review content characteristics, 4) video and reviewer characteristics.

3.3.2. Key Variables Construction

In this section, I describe the key variables. Table 3.2 presents descriptions of the key variables in the model, and Table 3.3 provides the descriptive statistics of the key variables.

3.3.2.1. Video and Reviewer Characteristics

I download all electronic product review videos included in the sample and collect data on video characteristics, including video URL, video title, and video creator. Then, I collect additional data on videos by using *YouTube-dl*, which is a command-line program that allows us to efficiently collect other relevant video information without using YouTube API. Collecting data through YouTube API can be time- and resource-consuming because it imposes limits on query size. Using *YouTube-dl*, I obtain additional video information, such as video view count, like count, dislike count, comment count, video tags, and video category. I additionally collect information on reviewers (i.e., video creators who uploaded review videos), such as a reviewer's name, region, channel URL, number of subscribers, and genre.

3.3.2.2. Video Visual Characteristics

As the human vision can distinguish and process around 10-12 images-per-second, a stream of images that move at a high rate is perceived as motion (Read and Meyer 2000). Therefore, a video can be viewed as a collection of static images that are referred to as frames. A typical video is comprised of 30 to 60 frames per second (FPS), and each frame consists of many pixels, which are represented in the RGB (red, green, and blue) color space. Therefore, to analyze visual information, I follow approaches taken in the video mining literature (Li et al., 2019) and extract frames that are evenly distributed at 10-second intervals. I employ a total of 681,726 frames to extract videos' visual characteristics and measure visual stimulation level. Furthermore, I conduct a shot-level analysis to check whether reviewers reveal their faces in their review videos. A shot is continuous footage between two edits and a collection of frames that shares most of their visual and semantic contents. By using Google Video AI, I detect a total of 307,874 shot changes in the dataset and analyze the visual information in each shot.[6]

Applying frame analysis and shot-level video analysis, I construct three variables: Visual Variability, Brightness, and Face Presence. In Section 6.2, I show that the findings are robust to alternative measurements of speech voice characteristics.

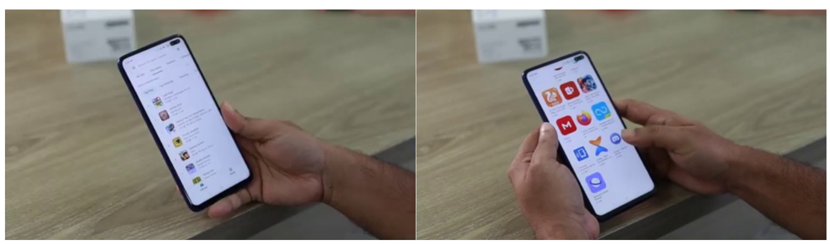
Visual Variability and Brightness.

To dynamic visual stimulation in a review video, I examine how rapidly the review video screen changes. Following Li et al. (2019), I construct the variable *Visual Variability*, which measures the dynamic changes in review videos by calculating the distance between nearby video frames. Figure 3.1 shows the example of the frame distance. In Example A, both image frames show smartphone browsing activities without significant changes in background and posture. However, the background, posture, and point of view drastically change in Example B. The frame distance is much higher in Example B (537660.40) than in Example A (19406.61). *Visual*

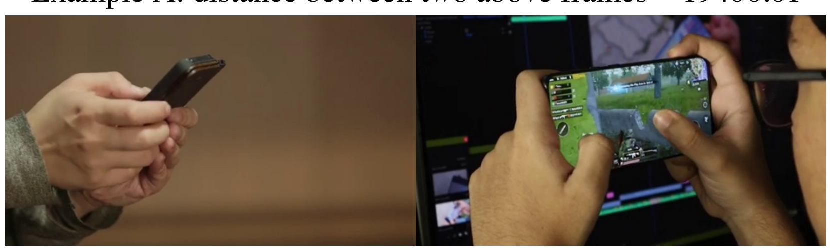
---

[6] Google Video AI uses pre-trained machine learning models to automatically analyze video content and detect video's shots and object (https://cloud.google.com/video-intelligence/docs).

Variability is the log value of the average frame distance between consecutive frames in each video. A higher value of *Visual Variability* indicates a higher visual stimulation. Appendix B provides additional details on the construction of the variable *Visual Variability*. Brightness is a widely used static visual measure in visual attention models (e.g., Itti et al., 1998; Rajashekar et al., 2008). In color appearance terminology, brightness refers to the attribute of visual sensation about the amount of light that an area appears to emit (Fairchild 2005). *Brightness* of a review video is the average of its frame brightness values, which are the average of pixel brightness across all pixels. In *RGB* color space, pixel brightness is the arithmetic mean of the values of red, green, and blue coordinates: $Brightness = (R + G + B)/3$.



Example A: distance between two above frames = 19406.61



Example B: distance between two above frames = 537660.40

Figure 3.1: Examples of frame distance

Face Presence

I analyze the face presence in video reviews by using Google Video AI's shot-level label detection, which identifies general objects in shots. Based on the label detection result, I consider a shot as a "face shot" (i.e., the shot contains a human face) if the shot is labeled with at least one of the labels under the "person" category with a confidence value higher than 0.75. The value of

Face Presence of a review video takes one if the video includes at least one face shot.

3.3.2.3. Speech Voice Characteristics

I additionally collect the data on reviewers' speech voice characteristics. I use FFmpeg, a software that helps process and convert multimedia content, to extract audio data from video files.[7] Then, I analyze the voice data and extract speech voice characteristics by generating fundamental frequency contours of voice data.[8] The fundamental frequency refers to the frequency at which the vocal cords vibrate during a speech and characterizes the speaking style of a person. For example, a voice with a lower fundamental frequency is lower-pitched. I analyze the fundamental frequency to measure reviewers' speech rate, pitch level, and pitch variability.

Speech Rate.

To compute the variable *Speech Rate*, which measures the speed of the reviewer's talk, I detect syllable boundaries and pauses in voice data. The variable *Speech Rate* is defined as the number of syllables pronounced per speech time in seconds: *Speech Rate=(Number of Syllables)/(Video Duration)*. A higher speech rate indicates that a reviewer speaks faster than a reviewer with a lower speech rate.

Pitch and Pitch Variability.

Following the literature on phonetics, the variable *Pitch* is defined as the median of the fundamental frequency of a reviewer's voice. The variable *Pitch Variability*, defined as the inter-quartile range of the voice fundamental frequency, measures how dynamic the reviewer's voice is

---

[7] FFmpeg, https://www.ffmpeg.org/

[8] I use a Python library, *my-voice-analysis*.

in terms of the voice pitch. A voice variability indicates that a reviewer talks in more dynamic ways by utilizing a wider range of his or her voice pitch. Although I separate speech sound from non-speech sound to create voice fundamental frequency contour, the voice fundamental frequency contour of reviewers' speech may potentially contain some noise. For example, some intermittent background noise may not be perfectly removed when the voice fundamental frequency contour is extracted from audio data. To address the potential concern, I use the median and inter-quartile range, which are outlier-robust distribution measurements, of the voice fundamental frequency to measure voice pitch and pitch variability.

3.3.2.4. Review Content Characteristics

In addition to basic video and reviewer information, it is important to understand the actual content of review videos and the messages that reviewers deliver. Therefore, I analyze reviewers' messages by utilizing speech recognition techniques and generating video transcription. When reviewers upload product review videos on YouTube, reviewers have an option to include their own video subtitles. For videos with subtitles, I download original subtitles that reviewers submitted. However, the proportion of videos with subtitles submitted by creators is less than 10%. For videos without submitted subtitles, I employ YouTube's speech recognition technique to auto-generate transcriptions of speech. To ensure the quality of the auto-generated transcription, I focus the analysis on videos created by English-speaking reviewers because YouTube's speech recognition has the best performance when detecting English. After I generate and download transcriptions for all videos, I clean the transcription data by converting the timestamp transcription files (in WebVTT format) into text files (in textfile format). I additionally clean transcription data by removing redundant parts and checking for errors. I introduce and define variables that capture review content characteristics: positivity and subjectivity.

Positivity and Subjectivity.

I extract sentimental and other speech features from speech transcription using natural language processing techniques.[9] The variable *Positivity*, measures the positive sentiment of a speech transcription where the *Positivity* of -1.0 indicates the most negative sentiment and 1.0 indicates the most positive sentiment. The variable *Subjectivity* measures how subjective or objective a review video is and takes a value between 0.0 and 1.0, where 0.0 indicates the most objective sentiment and 1.0 indicates the most subjective sentiment. The positivity and subjectivity of the entire speech are calculated as the average scores of all adjectives included in the speech.

3.3.2.5. Perceived Helpfulness

Perceived helpfulness is defined as the perceived value of information diagnosticity of a review (Mudambi and Schuff 2010, Yin et al. 2014). I define helpfulness of a review as the percentage of people who liked the review, which is consistent with the prior literature on text review helpfulness (Mudambi and Schuff 2010, Yin et al. 2014). On YouTube, video viewers can either like or dislike a video. The variable *Helpfulness*, which is the main outcome variable in the empirical model, is the number of likes divided by the total number of likes and dislikes received. Many online vendors, such as Amazon and eBay, provide a measurement of review helpfulness. For example, Amazon has a helpful button for each review to collect feedback from readers.[10] Similarly, I use the percentage of people who liked the review as a metric for review helpfulness.

---

[9] I implement Python's natural language processing package, *TextBlob*, to analyze sentimental features of audio transcriptions.

[10] The helpfulness information is shown under each review, such as "8 people found this helpful."

Khan (2017) shows that the information-seeking motive of viewers is a significant positive predictor for liking, disliking, and commenting on videos on YouTube. Considering that the primary purpose of review videos is providing product information, like and dislike votes of review videos well represent how viewers satisfy with review videos based on their information-seeking motive.

The variable *Helpfulness* is defined as the number of likes divided by the total number of like and dislike votes. The distribution of the variable *Helpfulness* is a skewed *J* shape distribution between 0 and 1, and such type of distribution is commonly observed in various rating platforms (Ho et al. 2017, McGlohon et al. 2010) (See Appendix C for details). I additionally examine alternative measurements of perceived helpfulness, including likes per view and positivity of comments, and show that results are robust to alternative measurements.

3.3.2.6. Product Types and Video Topics

Although I restrict the analysis to electronic product reviews, perceived helpfulness may differ across video topics and product types. To account for such differences across videos, I additionally 1) control for product types by manually labeling all videos' product categories and 2) control for video topics by employing natural language processing and machine learning algorithms.

***Review Product Types*.**

I additionally control for product categories in the main empirical model by manually labeling review videos based on product types. I label videos based on video titles and classify all review videos into the following six product categories: wearable devices, smartphones, tablets, laptops, other portable devices (e.g., portable speakers), and larger electronic devices (e.g., desktop computers). For example, the product review video titled "9.7 Inch iPad Pro Review" uploaded on

March 25, 2016 is categorized as a tablet review. Table 3.4 presents all six product category labels and their distributions.[11]

Table 3.2: Variable Description

Variables	Description
Helpfulness	The percentage of people who liked the product review video. This was derived by dividing the number of likes by the total like/dislike votes in response to the product review video.
Visual Variability	Measurement of dynamic visual stimulation level of a review video. This is derived by taking log of average distances between nearby frames in the review video.
Brightness	Measurement of the amount of light that a review video emits. This is derived by averaging frame brightness values for each review video.
Face Presence	The binary variable indicates reviewer's face presence in a review video. 1 indicates that there is a face presence in a review video; 0 otherwise.
Speech Rate	The ratio between the number of syllables spoken by a reviewer and the video duration (in second).
Voice Pitch	The median value of the voice fundamental frequency of a reviewer's talk.
Pitch Variability	The inter-quantile range of the voice fundamental frequency of a reviewer's talk.
Positivity	The positivity score of the review speech. Positivity is float which lies in the range of [-1, 1]. The Positivity of -1.0 indicates most negative sentiment and 1.0 indicates most positive sentiment.
Subjectivity	The subjectivity score of the review speech. Subjectivity is float which lies in the range of [0, 1] where 0 means a very objective statement and 1 means a very subjective statement.
Word count	The number of words in the review speech.
Total votes	The number of total like/dislike votes in response to the product review video.

[11] If a review video reviews multiple products in different categories, I label the review video with multiple categories. Out of all 13,840 review videos, the majority of videos, only 81 videos review two product categories, and only one video reviews three product categories.

Table 3.3: Descriptive Statistics

Numeric Variables	Min	Q1	Median	Mean	Q3	Max
Helpfulness	0	85.056	91.667	87.583	95.298	100
Visual Variability	4.527	10.453	10.746	10.665	10.964	11.800
Brightness	15.990	96.670	113.233	115.089	131.506	244.976
Face Presence	0	0	1	0.577	1	1
Speech Rate	0.080	3.264	3.814	3.682	4.220	6.482
Voice Pitch	86.3	118.0	131.3	137.475	149.6	264.1
Pitch Variability	7.0	26.0	36.0	39.989	48.0	233
Positivity	-0.206	0.142	0.179	0.180	0.217	0.557
Subjectivity	0.100	0.462	0.495	0.495	0.528	0.871
Word count	111	748	1,174	1,459.4	1,816	29,902
Total votes	10	54	147	901.6	502	156,466
Categorical Variable	Categories					
Year	2016:4,215	2017: 5,020	2018: 4,311	2019: 294		

Notes: The table reports descriptive statics of the main variables (N=13,840).

Table 3.4: Number of Review Videos in Each Product Category

Wearable Device	Smartphone	Tablet	Laptop	Other Portable Device	Larger Device
963	9,462	1,196	879	236	1,104

Video Topics via Machine Learning Algorithms.

Although all product review videos in the data review electronic products, videos may cover different content and it is important to control for such heterogeneity. Different review videos may focus on different product features (e.g., design, functionality, portability), and some product features may be more important for certain product types. For example, portability and weights may be important features for mobile electronics, such as smartphones and laptops, but maybe less important features for larger devices, such as desktop computers. To control for review content, I rely on natural language processing and machine learning algorithms. I use a machine-learning algorithm model, the Latent Dirichlet Allocation (LDA) model (Blei et al. 2003), to understand the information included in review videos and reduce review information dimensionality. The details of the LDA algorithm specifications can be found in Appendix D. The *Cv* coherence score, developed by Röder and Hinneburg (2015), is employed to find the optimal number of topics of the LDA algorithm. Additional details regarding the *Cv* coherence score is discussed are Appendix D. Using the *Cv* coherence score, 19 different topics are identified. Appendix Table D.1 shows the ten most related words in each topic.

3.4. Empirical Model

To empirically examine the extent to which video characteristics and voice characteristics affect the perceived helpfulness of product review videos, I estimate a reviewer fixed effects (FE) model of the following main equation:

$$
\begin{aligned}
Helpfulness_{it} = {} & \beta_0 + \beta_1 Visual\,Variability_{it} + \beta_2 Brightness_{it} + \beta_3 Face\,Presence_{it} \\
& + \beta_4 Speech\,Rate_{it} + \beta_5 Pitch_{it} + \beta_6 Pitch\,Variability_{it} + \beta_7 Positivity_{it} \\
& + \beta_4 Subjectivity_{it} + X_{it}\gamma + \alpha_i + \varepsilon_{it} \qquad (1)
\end{aligned}
$$

where $Helpfulness_{it}$ measures the share of like votes in percentage of the review uploaded by reviewer *i* on time *t*. I employ a linear reviewer fixed effects model to account for time-invariant unobserved reviewer characteristics. Reviewers' time-invariant unobservable characteristics, such as their overall reputation, experience level, and video techniques, can have heterogeneous effects on the perceived helpfulness of review videos. To address this issue, I exploit the unique panel nature of the dataset, where I observe multiple review videos uploaded by the same reviewers over time. With a reviewer fixed-effects model, I can examine how the perceived helpfulness of the same reviewers' review videos changes with visual, voice, and review content characteristics.

The model includes the three main variables that measure video visual characteristics of a review uploaded by reviewer *i* on time *t*. —$Visual\ Variability_{it}$, $Brightness_{it}$, *and* $Face\ Presence_{it}$, —and three variables that measure voice characteristics—$Speech\ Rate_{it}$, $Pitch_{it}$, and $Pitch\ Variability_{it}$. —and two variables that measure review content characteristics—$Positivity_{it}$, and $Subjectivity_{it}$. I additionally include a wide range of control variables, where X_{it} denotes a vector of video characteristic controls and reviewer characteristic controls of a review. First, I control for the quantity of information included in the review video by controlling for the number of words (in thousands) in a video speech. I additionally control for the total number of votes that a video received. To capture the seasonality and time trend in video activities, I include video publication year-month dummies to control for the year and month the video is uploaded. The model also includes product type controls to account for six different product types that I manually classified. To account for differences in video topics, the model includes video topic controls, and controls for 19 video topics that are identified using the LDA technique. α_i is the unobserved time-invariant individual effect of reviewer *i* (e.g., reviewer's filming skill level).

I employ the reviewer fixed-effects model as the main specification to exploit within-reviewer variations and to examine factors that determine the perceived helpfulness of video reviews. At the same time, the presence of possible omitted variables that affect both video characteristics and perceived helpfulness may lead to a potential endogeneity problem. Therefore, I take several empirical strategies to address the potential concern. First, I include a comprehensive set of control variables as described in previous sections. Second, I additionally utilize a propensity score weighting method to correct for observable characteristics that may have affected reviewers' choices of which product type to review. Third, as an additional empirical strategy, I employ a Tobit model as an additional specification to account for the bounded nature of the dependent variable.

3.4.1. Multinomial Propensity Score Weighting

I use multinomial propensity score weighting estimation to account for differences in observable characteristics that may affect reviewers' review product type choices (Rosenbaum and Rubin 1983). I first estimate a boosted logistic regression tree model to predict reviewers' choice of review product type based on observable reviewer characteristics and video publication month controls.[12,13] The model finds a set of parameters that maximize the likelihood of observing the review product types in the data. Then, I estimate each reviewer i's multinomial propensity score $\hat{p}_{ijt}$ for all product type $j \in$ {wearable device, smartphone, tablet, laptop, other portable devices, larger device} on time t, where $\hat{p}_{ijt}$ denotes the probability of choosing product type *j* conditional on observable reviewer characteristics and video publication month controls. The multinomial propensity score estimates are transformed into weights such that each online review in the reviewer fixed-effects model is weighted by $\omega_i = \frac{1}{\hat{p}_{ij^*t}}$, where j^* denotes the chosen review product type. The weights are utilized to generate a synthetic sample in which the distribution of observable reviewer characteristics and video publication time is balanced across review product types. Appendix E shows that the propensity score weighting method balances the covariates of different review product types.

3.4.2. Tobit Model

I employ a Tobit model as an additional specification to account for the bounded nature of the dependent variable as the variable *Helpfulness* is constrained to lie between 0 and 100 and is censored (Mudambi and Schuff 2010, Yin et al. 2014). The Tobit model additionally helps address a potential selection problem resulting from the fact that not all video viewers cast a vote (McDonald and Moffitt 1980). However, the Tobit model with fixed effects can be biased because of the incidental parameters problem (Neyman and Scott 1948, Lancaster 2000). For consistent and unbiased estimates, the non-linear nature of the Tobit model requires a sufficiently large

[12] The Toolkit for Weighting and Analysis of Nonequivalent Groups (*twang*) package is used for the estimation.
[13] I select a dominant review product type for the reviews with multiple review product types.

number of observations compared to the number of parameters in a model.[14] As a Tobit model with reviewer fixed effects would likely be biased, I estimate a Tobit model using both the original sample. The model additionally controls for reviewers' characteristics, including reviewers' gender and popularity, which is measured by reviewers' subscribers.

[14] The number of coefficients to estimate in the model increases linearly with the number of reviewers. In the data, there are 1,504 reviewers who created multiple reviews, and they created 11,195 review videos.

3.5. Estimated Results

3.5.1. Main Results

Table 3.5 presents the main results that examine the relationship between video review characteristics and perceived helpfulness of reviews. In the first column, I only include eight explanatory variables and do not include any control variables. Column 2 shows the result of the main model, reviewer fixed effects model with a full set of controls including video characteristics controls (the number of words in a video speech, the total number of votes that a video received, and video publication month), product type, and video topic. In column 3, I estimate a multinomial propensity score reviewer fixed-effects model.

In column 2, six key main explanatory variables —*Visual Variability*, *Brightness, Face Presence, Speech Rate*, *Pitch*, and *Positivity* — have statistically significant effects on perceived helpfulness that are consistent with the proposed hypothesis. In column 2, I find that an increase in *Variability* and *Brightness* increases the perceived helpfulness of reviews (significant at 1 percent and significant at 10 percent, respectably), which provides empirical evidence that viewers prefer review videos with higher visual stimulation. The positive estimate of coefficient on *Face Presence* (significant 10 percent) supports Hypothesis 2 that posits a positive effect of reviewer's face presence on perceived review helpfulness. In examining the relationship between voice characteristics and perceived helpfulness, I find that a faster speech rate and lower voice pitch are associated with higher perceived helpfulness (both significant at 1 percent), which provides empirical evidence consistent with Hypothesis 3 and Hypothesis 4, respectively. The estimate of coefficient on *Pitch Variability* is not statistically significant. However, the result on *Pitch Variability* is consistent across all models, and it suggests that viewers may perceive reviews with higher speech dynamicity as more helpful. The estimate of coefficient on the variable *Positivity* shows that when the positivity score increases by 0.1, it is associated with approximately 1 percentage point increase in the perceived helpfulness when holding other variables constant (column 2, significant at 1 percent). Therefore, different from text reviews, in online reviews, positive information is perceived as more helpful, which is consistent with the need for cognition (NC) theory. Compared with an average of 87.6 percentage points perceived helpfulness, an increase of 1 percentage point implies a 1.14 percent increase in latent perceived helpfulness. The estimated results on *Subjectivity* indicate that viewers prefer more objective video reviews, which is consistent with Hypothesis 4 despite losing statistical significance in column 2. In column 3, the multinomial propensity score weighting estimation consistently shows that six of the key explanatory variables have statistically significant effects on perceived helpfulness after accounting for differences in observable characteristics that may affect reviewers' review product type choices. In column 4, the Tobit model is estimated to account for the bounded nature of *Helpfulness*. The main results remain consistent and robust.

3.5.2. Robustness Checks

3.5.2.1. Alternative Measurements of Key Explanatory Variables

In this section, I show that the results are robust to alternative measurements of key explanatory variables: *Visual Variability*, *Brightness*, *Face Presence,* and *Pitch*. First, I define an alternative measurement of *Visual Variability*, which is based on the average shot length in seconds in review videos. A shorter average shot length indicates that a video changes shots more rapidly, and therefore, the video is more visually dynamic. For ease of comparison with the previous result with the original *Visual Variability*, I inverse the average shot length and take the log of it to construct its alternative variable. Therefore, like the original *Visual Variability*, alternative *Visual Variability* increases when a review video is more visually stimulating. Second, I define an alternative *Brightness* using the HSB (for Hue, Saturation, and Brightness) model instead of the RGB model. HSB model is an alternative way to represent images. To calculate the alternative *Brightness,* I convert each frame's RGB representation into HSB representation. Then, I calculate video brightness based on the HSB-brightness score, which is max(R, G, B) for each pixel. Third, I define an alternative *Face Presence* by re-defining a set of labels that defines "face shot". Previously, I defined "face shot" with all the labels under the "person" category. To construct the alternative *Face Presence*, I manually select a set of labels that can indicate a face presence in a shot (e.g., emotion, facial expression, and smile). Lastly, I construct an alternative *Pitch* variable. The original pitch measurements employ the median of the voice fundamental frequency to be more robust to any outliers. This time, I take the mean of the voice fundamental frequency to construct the alternative *Pitch* variable.

To examine the robustness of the results, I re-estimate the main model with alternative measurements by replacing each original variable at a time with its alternative variable while keeping other key variable measurements the same. Table 3.6 presents the robustness check estimation results, where bold texts indicate the estimated coefficients on alternative variables. The results consistently show that the main results are robust to alternative measurements of voice characteristics. In the first column, the original *Visual Variability* is replaced by its alternative variable, and the result shows that review videos with higher visual variability are perceived as more helpful (significant at 1 percent). Similarly, employing alternative brightness, face presence, and pitch measurements yield consistent and statistically significant results.

Table 3.5: Effects of Review Characteristics on Perceived Helpfulness

	(1) Reviewer FE w/o Controls	(2) Reviewer FE	(3) Propensity Score Weighted Reviewer FE	(4) Tobit
Visual Characteristics				
Visual Variability	1.887*** (0.313)	1.783*** (0.326)	1.696*** (0.328)	2.904*** (0.285)
Brightness	0.009*** (0.004)	0.007* (0.004)	0.008** (0.004)	0.006 (0.004)
Face	0.310 (0.231)	0.420* (0.238)	0.246 (0.237)	1.568*** (0.241)
Voice Characteristics				
Speech Rate	0.722*** (0.193)	0.725*** (0.198)	0.879*** (0.200)	2.320*** (0.159)
Pitch	-0.032** (0.007)	-0.036*** (0.008)	-0.034*** (0.008)	-0.027*** (0.006)
Pitch Variability	0.009 (0.008)	0.009 (0.008)	0.003 (0.008)	0.008 (0.007)
Review Content Characteristics				
Positivity	10.353*** (1.719)	9.878*** (1.755)	9.513*** (1.741)	12.142*** (1.902)
Subjectivity	-4.701** (1.949)	-3.098 (1.972)	-3.641* (1.964)	-6.364*** (2.151)
Control Variables	No	Yes	Yes	Yes
N	13,840	13,840	13,745	13,840

Notes: *p<0.1; **p<0.05; ***p<0.01. Standard errors are reported in parentheses under coefficients. The model additionally controls for word counts, total votes, product category, content-level topic, and publication month. In the Tobit model includes log of the number of subscribers and reviewer gender as well.

Table 3.6: Robustness Check – Alternative Measurement of Explanatory Variables

Replaced original variable:	(1) Visual Variability	(2) Brightness	(3) Face Presence	(4) Pitch
Visual Variability	**0.283***** **(0.137)**	1.900*** (0.324)	1.897*** (0.325)	1.897*** (0.325)
Brightness	0.008** (0.004)	**0.020*** **(0.012)**	0.007* (0.004)	0.007* (0.004)
Face Presence	0.445* (0.245)	0.497** (0.237)	**0.483**** **(0.243)**	0.491** (0.237)
Speech Rate	0.688*** (0.199)	0.718*** (0.197)	0.715*** (0.197)	0.671*** (0.197)
Pitch	-0.039*** (0.008)	-0.038** (0.008)	-0.038*** (0.008)	**-0.036***** **(0.009)**
Pitch Variability	0.008 (0.008)	0.007 (0.008)	0.007 (0.008)	0.011 (0.009)
Positivity	9.750*** (1.749)	9.636*** (1.745)	9.617*** (1.745)	9.698*** (1.746)
Subjectivity	-3.072 (1.965)	-3.432* (1.961)	-3.457* (1.962)	-3.493* (1.962)
Control Variables	Yes	Yes	Yes	Yes
N	13,840	13,840	13,840	13,840

Notes: *p<0.1; **p<0.05; ***p<0.01. The model additionally controls for word counts, total votes, product category, content-level topic, and publication month. Estimated coefficients on alternative variables are in bold.

3.5.2.2. Alternative Measurements of Perceived Helpfulness

In measuring the perceived helpfulness of reviews, the ratio of the number of like votes to total votes, which is employed in the main empirical specification, is almost equivalent to traditional measurements of perceived helpfulness used in prior literature on text reviews (Mudambi and Schuff 2010, Yin et al. 2014). At the same time, the interactive and social nature of the YouTube platform allows for other additional alternative measurements of perceived helpfulness as the platform allows researchers to observe video view counts and comments. In this section, I show that the findings are robust to employing alternative outcome measurements that capture review perceived helpfulness. I introduce two alternative measurements of perceived helpfulness: *Likes per View* and *Comment Positivity*.

First, the variable *Likes per View*, computed as the percentage value of the ratio of the number of like votes to the total number of views, measures the share of the audience who likes the video and helps account for the fact that not all viewers cast a vote for videos. Second, *Comment Positivity* captures an audience's sentiment toward videos. YouTube video viewers can actively engage with review videos by leaving comments on videos. Shao (2009) finds that comments can adequately capture audience engagement levels as it measures the extent to which users actively responds to posted content. *Comment Positivity* measures the sentiment of text comments where the positivity score of 0 indicates most negative sentiment and the positivity score of 100 indicates most positive sentiment.[15] To construct the variable *Comment Positivity*, I collect and analyze all 1,970,621 text comments left on all review videos in the data set. I exclude review videos with fewer than ten comments to ensure that the sentiment of comments can represent the perceived helpfulness of videos. The limited contextual information of short texts, such as single comments,

[15] I employ Python's natural language processing package, *TextBlob*, to analyze sentimental features.

makes the sentiment analysis of short texts challenging and less accurate compared to the sentiment analysis of longer texts (Dos Santos and Gatti 2014). Therefore, I pool all comments on each review video into one document and compute the sentiment score of the pooled comments.

Table 3.7 presents four different estimation results with two different reviewer fixed effects model where the outcome variable measures two measurements of perceived helpfulness. I estimate the same main empirical models, except that I control for the log of the number of views in place of the number of votes when the outcome variable measures *Likes per View*. Similarly, when the outcome variable is *Comment Positivity*, I control for the log of the number of comments in place of the number of votes.

Table 3.7: Robustness Check – Alternative Measurement of Perceived Helpfulness

	Like per View (%)		Comment Positivity (%)	
	Reviewer FE	Propensity Score Weighted FE	Reviewer FE	Propensity Score Weighted FE
Visual Variability	0.244*** (0.072)	0.206** (0.073)	0.635** (0.265)	0.782*** (0.266)
Brightness	0.001 (0.001)	0.002* (0.001)	-0.004 (0.003)	-0.004 (0.003)
Face Presence	0.087* (0.053)	0.080 (0.053)	0.337* (0.181)	0.293 (0.182)
Speech Rate	0.162*** (0.044)	0.156*** (0.044)	0.164 (0.153)	0.116 (0.155)
Pitch	-0.0002 (0.002)	-0.0001 (0.002)	-0.002 (0.006)	-0.006 (0.006)
Pitch Variability	0.001 (0.002)	0.001 (0.002)	0.010 (0.006)	0.013** (0.006)
Positivity	0.818*** (0.388)	0.767** (0.386)	8.782*** (1.346)	8.826*** (0.343)
Subjectivity	-0.482 (0.436)	-0.424 (0.436)	-1.778 (1.504)	-1.010 (1.508)
Control Variables	Yes	Yes	Yes	Yes
N	13,840	13,745	13,840	13,745

Notes: *p<0.1; **p<0.05; ***p<0.01. The model additionally controls for word counts, total votes, log of the number of subscribers, reviewer gender, product category, content-level topic, and publication month.

The results show that the main results are mostly robust to alternative measurements of perceived helpfulness.

3.6. Controlled Experiment

I conducted two controlled experiments to isolate the causal effects of review characteristics on the perceived helpfulness. In the controlled experiments, participants watched and evaluated two review videos. In the first experiment, I manipulated vocal characteristics at two levels (lower-pitched voice vs. higher-pitched voice), and in the second experiment, I manipulated visual characteristics at two levels (brighter video vs. darker video).

3.6.1. Procedure and Measures

The study used a 30-second review video that reviews the camera features of iPhone 13, which was one of the newest smartphone models from Apple at the time of the study. To remove the potential confounding effect of review content and other review video characteristics, I used the same review video and manipulated the vocal or visual characteristics. In the first experiment, in which the vocal characteristics are manipulated, I rely on the prior literature on vocal characteristics (e.g., Chattopadhyay et al. 2003; Oleszkiewicz et al. 2017) to experimentally lower or raise voice pitches using iMovie and create a lower-pitched and a higher-pitched video. Similarly, in the second experiment, the visual characteristics are manipulated using iMovie to create a brighter and a darker video.[16]

I recruited a total of 337 participants from Amazon Mechanical Turk (MTurk) and compensated them for their participation. In the cover story, participants were asked to imagine that they are looking for a new smartphone and seeking to learn about the camera features of smartphones before they make a purchase decision. 176 participants who participated in the first

[16] I lowered the pitch level of the original video to generate a lower-pitched video and raised the pitch level of the original video to generate a higher-pitched video. I ensured that there are no vocal or visual distortions in all review videos included in the experimental studies.

experiment watched one video with a lower-pitched voice and another video with a higher-pitched voice. 137 participants who participated in the second experiment watched one brighter video and one darker video. In both studies, two review videos were shown in random orders. After watching the review videos, participants were asked to score each review video based on their perception of review helpfulness using a 0–100 point scale. I measure review helpfulness using three items: helpful, useful, and informative (Sen and Lerman 2007, Lei et al. 2021). To conduct manipulation checks, I asked participants to report their perceptions of the manipulated review characteristics using a seven-point Likert scale.

3.6.2. Results and Discussion

I checked the manipulations of voice pitch and video brightness. Results reveal that the perceived voice pitch of the lower-pitched review video is significantly lower than that of the higher-pitched review video ($M_{lower\ pitch} = 3.81$ vs. $M_{higher\ pitch} = 5.43$, $p < 0.001$), and perceived video brightness of the brighter review video is significantly higher than that of the darker review video ($M_{darker} = 4.59$ vs. $M_{brighter} = 5.29$, $p < 0.001$). Therefore, the manipulation of both vocal and visual characteristics was successful.

Table 3.8 presents the results of the experiment. Panel A presents the perceived helpfulness score differences between the lower-pitched and higher-pitched review videos. The positive differences of the means indicate that participants scored the lower-pitched review video to be more helpful, more informative, and more useful (Panel A, column 1, all significant at 1 percent). Therefore, results indicate that an increase in voice pitch lowers the perceived helpfulness of review videos when all other review video characteristics remain constant. Similarly, Panel B presents the perceived helpfulness score differences between the brighter and darker review videos

(Panel B, column 1, all significant at 1 percent). The positive differences of the means indicate that the increase in the video brightness improves the perceived helpfulness when all other review video characteristics remain constant, which is consistent with the previous findings.

I additionally employ the controlled experiments to explore the potential mechanisms that underly the effects of vocal and visual characteristics on the perceived review helpfulness. In implementing the first experiment on reviewers' voice pitch, I asked participants to report their perceptions of other vocal-related characteristics—trustworthiness and ease of comprehension—in a seven-point Likert scale. I find that the perceived trustworthiness and ease of comprehension were significantly higher for the lower-pitched review video (trustworthiness: $M_{lower\ pitch} = 5.63$ vs. $M_{higher\ pitch} = 5.24$, p-value < 0.001; ease of comprehension: $M_{lower\ pitch} = 6.10$ vs. $M_{higher\ pitch} = 5.56$, p-value < 0.001). The results suggest that lower voice pitch has positive effects on perceived trustworthiness and ease of comprehension of a reviewer's speech, which in turn, can affect the perceived review helpfulness. The finding is consistent with the previous communications literature (e.g., Apple et al., 1979; Wang et al. 2018), which finds that lower-pitched voices had greater persuasiveness in conversations than higher-pitched voices. During the second experiment on video brightness, participants were asked to report their perceptions of the visual characteristics of review videos in the following two dimensions using a seven-point Likert scale: (1) the review video's visuals helped me to understand the reviewer's explanation, and (2) the review video's visuals matched the reviewer's explanation. The results indicate that the brighter video's scores are significantly higher than the darker video's scores (help understand the explanation: $M_{brighter} = 5.88$ vs. $M_{darker} = 5.19$, p-value <0.001;match the explanation: $M_{brighter} = 6.00$ vs. $M_{darker} = 5.30$, p-value< 0.001). Therefore, brighter visuals can increase the perceived helpfulness of review videos by delivering the review content more effectively and helping the video audience understand the content better.

Table 3.8: Differences in Review Helpfulness Scores

	(1) **Mean**	(2) **Std. Error**	(3) **t-value**	(4) ***p*-value**
Panel A: Voice Pitch: $M_{lower\ pitch} - M_{higher\ pitch}$				
Review was helpful	14.080	1.773	7.942	0.000
Review was informative	8.453	2.116	3.996	0.000
Review was useful	9.263	2.243	3.129	0.000
Panel B: Video Brightness: $M_{brighter} - M_{darker}$				
Review was helpful	8.591	1.618	5.308	0.000
Review was informative	6.506	1.787	3.640	0.000
Review was useful	6.784	1.930	3.515	0.000

Note. The Panel A reports the score difference between the lower-pitched review video and the higher-pitched review video. Each score difference was generated by deducting the score of the higher-pitched review video from the score of the lower-pitched review video. (N=176). The Panel B reports the score difference between the brighter video and the darker review video. Each score difference was generated by deducting the score of the darker review video from the score of the brighter review video. (N=137).

3.7. Conclusions

As online product review videos can deliver richer and more abundant product-related information than text reviews, online review videos have become a major new channel through which consumers obtain product information. As the first study in the literature to examine the perceived helpfulness of online product review videos, this study examines how various visual characteristics of reviewers—visual variability, brightness, and face presence—, voice characteristics of reviewers—speech rate, pitch, and speech dynamics—, and review content characteristics—information accessibility, positivity, and subjectivity—affect the perceived helpfulness. I find that video viewers perceive online product review videos that are more visually stimulating, more positive, and less subjective as more helpful. In addition, review videos are perceived as more helpful when reviewers reveal their faces in their reviews and when reviewers speak with faster speech rates and lower voice pitch.

This study contributes to the extant literature on online reviews in multiple dimensions. I build upon the literature on information systems, operations management, communications, marketing, and psychology to introduce and examine new metrics that capture unique features of online video reviews. Some findings are unique to online review videos and differ from findings from online text reviews. The finding that the audience perceives positive review videos as more helpful supports the presence of the positivity bias in video reviews. The finding contrasts with findings from previous literature that examines online text reviews that are present on online retail platforms, which documents negativity bias in which negative information is perceived as more helpful (Kuan et al. 2015, Yin et al. 2014). In addition, different from text reviews, reviewers use their own voices and visual aids to deliver review information, and review videos contain vocalic and visual cues that are not present in text reviews.

3.7.1. Managerial and Practical Implications

The findings have important implications for content generators (i.e., product reviewers), companies, online platforms, and policymakers. Content generators who wish to increase the perceived helpfulness of their review videos, and thereby increase their revenue from videos, can undertake more effective strategies in expressing their opinion and delivering their messages. Also, content generators can follow strategies to improve the perceived helpfulness of their reviews. Similarly, when companies market their products by creating their own review videos or sponsoring product reviewers, companies can provide guidelines for the reviewers. Furthermore, it would be in the interest of online video platforms to understand how different factors affect the audience's perceived helpfulness, so that the platforms can better understand their audience's needs and potentially encourage the publication of review videos that are regarded as more helpful. Finally, the findings can have important implications for policymakers as regulatory bodies, including the Federal Communications Commission (FCC) and Federal Trade Commission (FTC), newly design and set guidelines for content on social media to help protect businesses and consumers.

3.7.2. Future Research and Limitations

There are various ways this study's findings can be further extended by future research. First, as this study focuses on electronic review videos, future research can extend the study by examining review videos in other product categories and examining other video types. For example, online videos have become important channels not only for product reviews, but also for education, entertainment, and advertisement. Second, there has been a growing interest in using more advanced video technology, such as live streaming services, 360-degree, AR, and VR technologies. As video technologies continue to advance, future research can explore how future (and futuristic) video technologies can help reshape the online video review landscape. Furthermore, with the growth and ubiquity of unstructured multimedia data, future research on product review videos can employ advanced machine learning and computer vision techniques, such as face recognition (Zhao et al. 2003) and facial expression recognition (Fasel and Luettin 2003), to analyze non-verbal attributes of reviewers. Finally, as the prevalence and popularity of online review videos is a recent phenomenon, future research can examine the long-term effects of online review videos on brand awareness and perception.

Chapter 4. STUDENT SOCIAL ENGAGEMENT IN ONLINE LEARNING: AN EMPIRICAL STUDY

4.1. Introduction

In January 2020, the World Health Organization (WHO) declared the coronavirus disease 2019 (COVID -19) as a public health emergency worldwide. As the reported case of COVID-19 rapidly increases, WHO announced a pandemic in March 2020. While the infection rate of COVID-19 was different among countries, the COVID-19 pandemic caused school closures in 151 countries and affected 1.3 billion students worldwide in April 2020.[17] For example, in Korea, the country-wide school closure happened until late May 2020, then schools partially opened for the remaining semester periods. Students experienced a sudden shift away from the traditional face-to-face schooling during the pandemic. The demand for online learning rapidly increased since using online education tools became necessary for schools to deal with the pandemic-initiated school closures.

As a result, understanding the factors that contribute to online learning effectiveness is crucial to improving the learning experience in online settings. Prior literature in online education has shown that effective online learning is based on technological support, well-designed learning materials, an adequate amount of interactions between the teacher and students, and the creation of a sense of community (Sun and Chen 2015). Among those factors, this study focuses on understanding students' social interaction in online learning settings. Numerous studies in online

[17] UNESCO, "Education: From disruption to recovery", (https://en.unesco.org/covid19/educationresponse)

education have supported the strong correlation between social interaction and achieving success in online learning (e.g., Brindley et al. 2009; Cox and Cox 2008; Ke 2010; Yang et al. 2014).

In this study, I investigate the dynamics of students' online social engagement and how the COVID-19 pandemic affected it. As a theoretical ground to understand student social behavior, this study relies on Bandura's social cognitive theory (Bandura, 1977). Building upon social cognitive theory, I posit three major factors that affect students' social behavior in online learning communities: (1) students' self-efficacy, (2) social influence from others, and (3) students' self-observation of their past behavior patterns. Self-efficacy plays a distinctive role in the causal structure of Social Cognitive Theory by affecting individuals' motivation of performing behaviors. Self-efficacy is individuals' perceptions about their ability to exercise control over their behavior and environment. Bandura (2001) demonstrated that individuals with stronger self-efficacy have "greater aspirations, higher motivational investments in their undertakings, and stronger staying power and persistence in the face of impediments". In the context of online learning communities, students with stronger self-efficacy exert more effort to make social interactions and develop good relationships with others and continuously contribute to grow the social-emotional bonds in their group.

However, students' self-efficacy is often unobservable. To deal with the unobservable student self-efficacy, this study proposes a Hidden Markov Model (HMM) that models students' latent self-efficacy state as a Markov chain. HMM models the underlying unobservable stochastic process between different self-efficacy states that determines the sequence of observable outcomes, which is the level of student social engagement in online learning communities. The HMM is estimated based on the dataset collected from a leading online learning platform in Korea.

The results indicate that recent proactive social activities from all user groups (i.e., teachers, peers, and parents) have positive influences on students' social engagement. The result also suggests that the parental influence on the student's social engagement increases during the pandemic, but teachers and peers' social influences decreased during the pandemic. The increased perceived social distance towards students' classmates and teachers during the pandemic can be the potential explanation. The pandemic setting gives students fewer opportunities to have face-to-face interactions with teachers and peers. In contrast, the students' perceived social bond with parents may increase during the pandemic because of the increased time spent with their parents. This environmental change may account for the strengthened parental influence on students' online social engagement compared to that of teachers and peers.

The analysis on self-efficacy state transition shows that the amount of social feedback on the platform and students' self-observation on their past social engagement dynamically shape students' self-efficacy level. The analysis on state transition suggests that the social influences from peers, teachers, and parents on students' self-efficacy are positive for students with higher self-efficacy. For students with lower self-efficacy, the social influence from peers is negative, which can be potentially explained by peer dynamics. However, the social influence from teachers is positive for the students with lower self-efficacy. This result suggests that the positive impact from teachers can mediate the negative impact from peers and highlights the importance of the social role of teachers in online learning environments. The result also indicates that the cumulative number of social activities of a student has a positive effect on state transition regardless of the student's self-efficacy state. This implies that students' self-observation on their past engagement in social interactions can improve students' self-efficacy in building good relationships with others online. The overall social influence on state transitions becomes weaker during the pandemic. In contrast, the effect of self-observation on state transition is stronger during the pandemic. The reduced perceived group membership induced by a lack of face-to-face schooling can explain such change during the pandemic.

4.2. Theoretical Framework

I employ Bandura's social cognitive theory to model students' engagement behavior in social interaction in online settings. Social cognitive theory is a theoretical framework that provides a causal model to understand individuals' behavior in social environments (Bandura, 1977). Building upon social cognitive theory, I posit three significant factors that affect students' social behavior in online learning communities: (1) students' self-efficacy, (2) short-term impact of other's social activities, and (3) students' self-observation of their past behavior patterns.

4.2.1. Student self-efficacy

Self-efficacy plays a crucial role in the causal structure of Social Cognitive Theory by affecting individuals' motivation of performing behaviors. Bandura (1977) defined self-efficacy as an individual's belief in her ability to exercise control over her behavior and social environmental events. Individuals with stronger self-efficacy have "greater aspirations, higher motivational investments in their undertakings, and stronger staying power and persistence in the face of impediments" (Bandura 2001). For example, in online learning communities, students with stronger self-efficacy exert more effort to make social interactions and develop good relationships with others and continuously contribute to grow the social-emotional bonds in their group.

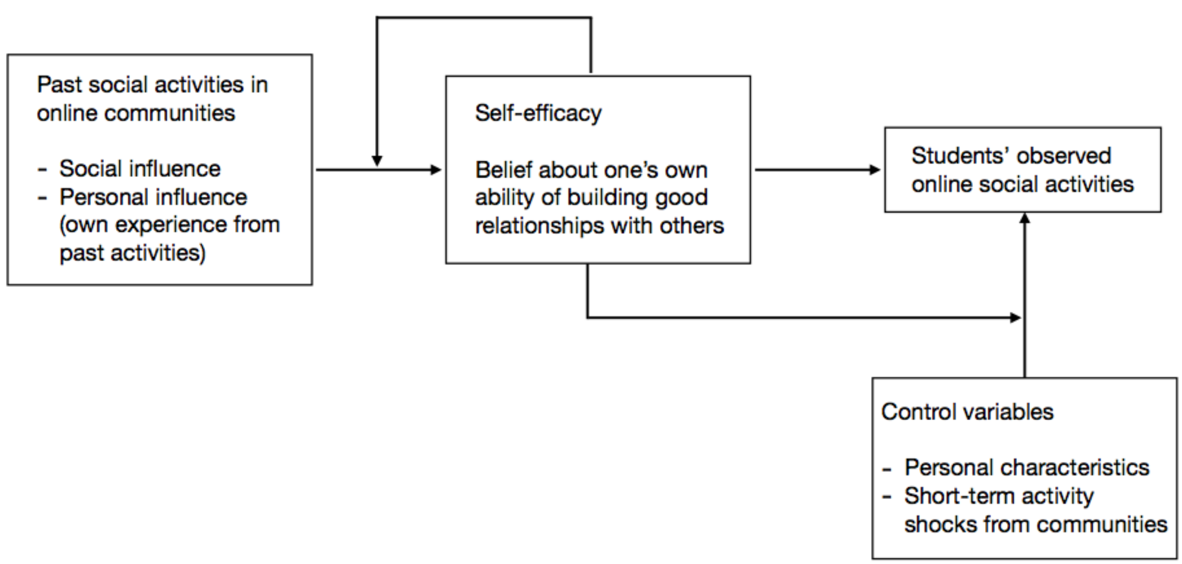
Various social factors can form a student's self-efficacy. Prior literature in student engagement has found several factors that determine students' engagement in social interaction. First, teacher and peer influence on students' social engagement in various age groups has been shown by many studies (Rimm-Kaufman et al. 2015; Martin and Rimm-Kaufman 2015; Taylor and Parsons 2011; Pianta et al. 2012; Kiefer et al. 2015). For example, Rimm-Kaufman et al. (2015) showed the importance of the social role of teachers. By examining the relationship between teacher-student interaction quality and 5th grade student engagement in mathematics classrooms, they found that students' social, cognitive, and emotional engagement were higher when teachers provided stronger emotional support. Kiefer et al. (2015) found that the support from teachers and peers is social and academic in nature and helps middle school students by improving their motivation, engagement, and sense of belonging in the school. Parents' support is another crucial factor that can influence students' self-efficacy in social interaction. For example, Ansong et al. (2017) studied students' emotional and behavioral engagement in Ghana and found that social support from classmates and parents has a strong association with student engagement. Compared to the traditional face-to-face classroom setting, the effect of various social factors on students' social engagement in an online learning setting is not well-explored. Peer and teacher influence on students' social engagement is relatively more explored than parental influence. For instance, Cho et al. (2010) demonstrated that students' monitoring for peer and teacher interactions is positively related to their perceived social presence and connectedness to the community. The recent COVID-19 pandemic and corresponding lockdowns have forced schools to increase their use of online learning tools. Considering that more schools have adopted hybrid-learning systems, which combine face-to-face learning and online learning, it is crucial to analyze a more extensive set of social factors and study how they affect students' social engagement in online environments.

In addition to social factors, students' past behavior patterns also play a substantial role in shaping their self-efficacy. According to Bandura (1997), individuals' past behavior affects self-efficacy belief since individuals' own past experience forms the judgment of their own capabilities. For example, students may have a higher self-efficacy after experiencing positive results (e.g., a new friendship, a meaningful discussion) after they made some social interactions with others. In contrast, a negative social experience (e.g., peer rejection, bullying) can discourage students' social engagement (Juvonen et al. 2012).

4.2.2. Short-term social influence and student personal characteristics

Social cognitive theory posits that individuals' behavior in social environments is inherently driven by their self-efficacy. While individuals' self-efficacy provides the baseline of their behavior in social settings, recent online social activities from others may also create a short-term impact on individuals' behavior (Zhou et al. 2021). For example, students can write a comment as a response to a recent post of their teacher. Other users' observable activities on online platforms may attract more attention from students and create an opportunity for students to make a new social interaction with others in their communities. Students' personal characteristics (e.g., age) may also influence their online social behavior. Therefore, I assume that other users' recent social activities exert short-term activity shocks on students' online social behavior. Figure 4.1 presents the graphical representation of the conceptual framework.

Figure 4.1: Conceptual Framework



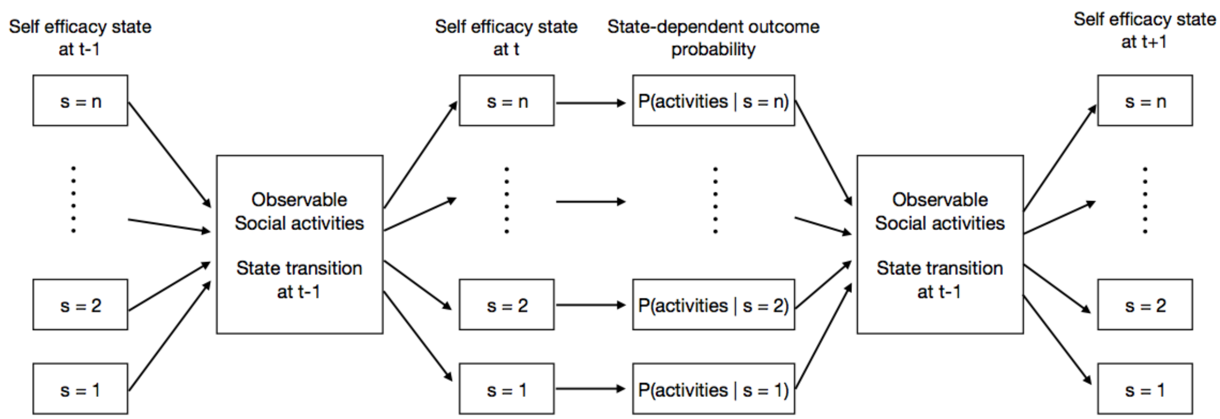
4.3. Empirical Model

4.3.1. Hidden Markov Model (HMM) for student online social engagement

This study builds on the Social Cognitive Theory, which posits that individuals' self-efficacy is an important determinant of their social behavior. However, students' self-efficacy level is not observable. To account for the unobservable self-efficacy level of students, I employ a Hidden Markov Model (HMM) that models students' latent self-efficacy state as a Markov chain. HMM models the underlying unobservable stochastic process that determines the sequence of observable outcomes. An HMM has a finite set of hidden states that account for the outcome generation process, and transitions between states follow a Markov process, in which any given time period's state depends on the state in previous periods. (Rabiner, 1989). HMMs have been widely applied to model online behavior in different contexts. For example, Singh et al. (2011) studied developers' learning dynamics in the context of open-source software projects; Yan and Tan (2014) investigated the effect of social support exchange on patients' mental health in the context of an online healthcare community; Zhou et al. (2021) modeled the influence of online social activities on individuals' weight-management behavior in online weight-loss communities.

In this study, the HMM models individual students' self-efficacy on the platform as a hidden state. Figure 4.2 demonstrates a graphical representation of the proposed framework. In every time period, a student is probabilistically in one self-efficacy state, in which 1 demonstrates the lowest self-efficacy state and n is the highest. The student's state in period t is based on the social influence from the platform (e.g., observable social support activities from other users in her classroom), her self-observation of her own past behavior, and her self-efficacy state in period t-1. After a state transition occurs at time t, her self-efficacy state and other variables (e.g., the student's relevant characteristics and recent proactive social activities on the platform) determine the outcome at time t. In the model, the outcome sequence Y corresponds to individual students' social engagement on the platform, which is the number of social activities made by students in each period.

Figure 4.2: Hidden Markov Model of Students' online social behavior



4.3.2. State Transition Matrix

The first model component of HMM is the state transition matrix for the underlying Markov process. I assume that there are discrete n hidden states that represent students' self-efficacy level from the lowest self-efficacy state 1 to the highest self-efficacy state n. The state transition matrix contains the transition probabilities from one state to another. Students can observe other users' social feedback activities (e.g., other students, parents, and the teacher) and their own social behavior patterns on the platform. Such observations can form and influence students' self-efficacy level to engage in social activities on platforms. A student can move to a higher self-efficacy state if the impact of the observation is higher than a certain threshold. However, if the impact is smaller than a lower threshold of maintaining the current self-efficacy level, the student can move to a lower self-efficacy state. An ordered Logit model allows us to model such transitions which are not limited to adjacent states. The state transition matrix is defined as $P = \{p_{it}(s, s')\}$, where $p_{it}(s, s')$ denotes the probability of transition of student *i* at time *t* from state s to state s':

$$p_{it}(s, 1) = \frac{e^{\mu_{s\to 1} - \beta_s W_{it}}}{1 + e^{\mu_{s\to 1} - \beta_s W_{it}}};$$

$$\vdots$$

$$p_{it}(s, s-1) = \frac{e^{\mu_{s\to s-1} - \beta_s W_{it}}}{1 + e^{\mu_{s\to s-1} - \beta_s W_{it}}} - \frac{e^{\mu_{s\to s-2} - \beta_s W_{it}}}{1 + e^{\mu_{s\to s-2} - \beta_s W_{it}}};$$

$$p_{it}(s, s) = \frac{e^{\mu_{s\to s+1} - \beta_s W_{it}}}{1 + e^{\mu_{s\to s+1} - \beta_s W_{it}}} - \frac{e^{\mu_{s\to s-1} - \beta_s W_{it}}}{1 + e^{\mu_{s\to s-1} - \beta_s W_{it}}};$$

$$p_{it}(s, s+1) = \frac{e^{\mu_{s\to s+2} - \beta_s W_{it}}}{1 + e^{\mu_{s\to s+2} - \beta_s W_{it}}} - \frac{e^{\mu_{s\to s+1} - \beta_s W_{it}}}{1 + e^{\mu_{s\to s+1} - \beta_s W_{it}}};$$

$$\vdots$$

$$p_{it}(s, n) = 1 - \frac{e^{\mu_{s\to n} - \beta_s W_{it}}}{1 + e^{\mu_{s\to n} - \beta_s W_{it}}};$$

Where $\mu_{s \to s'}$ denotes the threshold for the transition from state s to state s'. For a given state s, the thresholds follow an order of $\mu_{s\to 1} \leq \mu_{s\to 2} \leq \cdots \leq \mu_{s\to s-1} \leq \mu_{s\to s+1} \leq \cdots \leq \mu_{s\to n-1} \leq \mu_{s\to n}$. The vector W_{it} contains the covariates that affect transitions in students' self-efficacy state, such as the students' self-observation of their own behavior patterns and the amount of social feedback activities from other users on the platform. The vector β_s is a set of state-dependent coefficients to be estimated. $W_{it}\beta_s$ is the propensity value of student i at time t to move to a higher state. A transition to a higher state occurs if the $W_{it}\beta_s$ value is higher than a certain threshold required for the transition, whereas a transition to a lower state happens if the $W_{it}\beta_s$ value falls below a threshold value for the transition to the lower states. Any transitions between two nonadjacent self-efficacy states can happen in the ordered Logit model. For example, a student in the lowest self-efficacy state can stay in the state or move to any one of n – 1 higher states.

4.3.3. State-Dependent Outcome

In the framework, the outcome Y is a count variable that corresponds to individual students' engagement in online educational social platforms, which is the number of social activities made by students in each period. Following Singh et al. (2011) and Yan and Tan (2014), I adopt a negative binomial (NB) model to represent the state-dependent outcome generation process:

$$\Pi(Y_{it} \mid S_{it}) = \frac{\Gamma(Y_{it} + \theta_s)}{\Gamma(Y_{it}+1)\Gamma(\theta_s)} \left(\frac{e^{\delta_s X_{it}}}{\theta_s + e^{\delta_s X_{it}}}\right)^{Y_{it}} \left(\frac{\theta_s}{\theta_s + e^{\delta_s X_{it}}}\right)^{\theta_s},$$

where Y_{it} is the number of social activities made by student *i* at time *t,* S_{it} is the self-efficacy state of student i at time t, and θ_s is the state-dependent parameter to account for the possible over dispersion in where Y_{it}. The vector X_{it} contains the variables that affect the outcome for student *i* at time *t*. The vector δ_s is a set of state-dependent coefficients to be estimated.

4.3.4. Estimation

Maximum likelihood estimation is used to estimate the HMM. The likelihood of an observed sequence of outcomes $L(Y(i) = \{Y_{i1}, Y_{i2}, \cdots, Y_{iT}\})$ for student *i* is the sum of likelihood values of all possible paths of hidden states over time periods:

$$L\big(Y(i)\big) = \sum_{s_1=1}^{n}\sum_{s_2=1}^{n}\cdots\sum_{s_T=1}^{n} P(S_{i1} = s_1)\prod_{t=2}^{T} P(S_{it} = s_t \mid S_{it-1} = s_{t-1})\prod_{t=1}^{T} P(Y_{it} \mid S_{it} = s_t)$$

where S_{it} is the self-efficacy state of student *i* at time *t*. The overall likelihood is the product of all individual students' likelihood values $L(Y) = \prod_{i=1}^{N} L\big(Y(i)\big)$.

4.4. Data

4.4.1. Research Context

In this study, I investigate a leading educational social platform in South Korea. The platform is intended for providing separate online communities dedicated to schools and their classes. Once a school joins the platform, all the associated people can join the platform. Therefore, there are three types of users on the platform: (1) teachers, (2) students, and (3) parents. Teachers can create virtual classrooms for their classes, then members (i.e., students and their parents) join the classroom. Users can interact with others by posting, writing comments, and liking postings and comments in a classroom. For example, teachers can make announcements, post homework, collect feedback from students, and share photos and videos. Students can submit their homework and communicate with classmates. Parents can also participate in the platform. The focal platform provides a unique setting that allows me to observe social interactions among teachers, students, and parents.

Additionally, this study also investigates how the recent COVID-19 pandemic changed the dynamics of students' online social behavior. In February 2020, the COVID-19 pandemic started in Korea. As the number of COVID-19 patients rapidly increased in late February 2020, Korean governments regulated schools to delay their semester starting from March to April. Moreover, the semester started online instead of the usual face-to-face manner. Face-to-face schooling started in late May. The face-to-face school experience became different during the pandemic. Students and teachers were required to wear masks and required to have more physically-distanced interaction compared to the past. Moreover, if the number of new COVID-19 patients in a school's area exceeded a certain threshold, the school went back to the online schooling method for a while.

Because of the sudden outbreak of COVID-19, This change happened to every school in Korea and created an external shock that made all schools in Korea have fewer opportunities to have face-to-face physical interactions among members. Therefore, this natural experiment setting allows me to observe how the dynamics of students' online social behavior change when there are reduced opportunities for face-to-face interactions among school members.

4.4.2. Variable construction

The data are collected over two years before and after the Covid-19 pandemic: (1) pre-pandemic data on 527 students are collected weekly for 25 weeks from February 2019 to July 2019, and (2) post-pandemic data on 889 students are collected weekly for 25 weeks from February 2020 to July 2020. Both data collection time periods represent the same semester in the academic calendar of Korea. The data cover student, teacher, and parent behavior data in four elementary schools in two major cities in Korea: Seoul and Daegu.

4.4.2.1. Variables impacting state transition

The variables in W_{it} are used to calculate the state transition probabilities among multiple self-efficacy states. Previous literature on self-efficacy has shown that individuals' self-efficacy can be dynamically altered by social influences and their personal observation of their past behavior patterns (Bandura, 1997). First, to account for the social influence, I construct three variables, *Teacher cumulative feedback*, *Peer cumulative feedback*, and *Parent cumulative feedback*. These variables show the cumulative amount of social feedback provided by teachers, peers, and parents on the platform. Social feedback activities include all reactive social activities made on the platform, such as commenting and liking. The cumulative feedback amount from the community can form students' expectations of the level of social feedback that they can receive after they exert social interaction on the platform, which in turn, alter students' self-efficacy of building good relationships with others by exerting social interactions on the platform. I measure the impact of self-observation by including students' personal past behavior patterns, *Cumulative social activity*, which is the cumulative number of social activities made by each student. I calibrate the above cumulative variables by applying exponential smoothing since a more recent activity tends to have a stronger impact on individual behavior compared to older activities (Bandura 1991, Zhou et al. 2021). For example, the activity that happened Δt ago, its effect is discounted by $e^{-r\Delta t}$, where r is the time discounting factor, which needs to be optimized based on the model fit.

4.4.2.2. Variables impacting state-dependent outcome

The variables in X_{it} are used to calculate the state-dependent outcome probabilities. To measure the short-term activity shocks, the amount of proactive social activities (i.e., posting on the platform) at time t is collected for the teacher, peer, and parent groups. *Teacher Post*, *Peer Post*, and *Parent Post* are the variables that represent the amount of recent proactive social activities made by each user group. Additionally, the binary variables *High Grade* and *School* are constructed to include students' personal characteristics in the model. The variable *High Grade* takes one if a student is in a higher grade (i.e., 4^{th}. 5^{th}, and 6^{th} grades), or it takes zero (i.e., 1^{st}, 2^{nd}, and 3^{rd} grades). There are three binary *School* variables to represent the four schools in the dataset. Each *School* variable takes one if a student is attending the school that is associated with the *School* variable, or it takes zero. Finally, as an environmental factor, I collect the size of a student's virtual classroom (i.e., the number of members in the classroom). The variable *Class size* is the number of users in a student's classroom.

4.4.2.3. Outcome variable and treatment variable

As the outcome variable of the model, the number of observable social activities made by a student is used to measure the amount of the student's social engagement during a certain time period. A student's observable social activities include posting, commenting, and liking activities on the focal platform. Also, a binary variable *Pandemic,* which takes one for the year 2020 data and takes zero for the year 2019 data, was constructed to examine the impact of the external shock (i.e., the sudden decrease of face-to-face interactions in schools) that the pandemic brought on the dynamics of students' social behavior on the focal platform. The variable *Pandemic* was interacted with the social and personal influence variables to explore the difference between the patterns of student online social engagement in the two years. Table 4.1 shows variable descriptions, and Table 4.2 shows their descriptive statistics.

Table 4.1: Variable Description

Variables	Description
Outcome variable (Y_{it})	
Social activity$_{it}$	Number of social activities made by student *i* at time *t*
Variables impacting state transition (W_{it})	
Teacher cumulative feedback$_{it}$	Cumulative number of social feedback made by teachers in the student *i*'s classroom
Peer cumulative feedback$_{it}$	Cumulative number of social feedback made by classmates in the student *i*'s classroom at time *t*
Parent cumulative feedback$_{it}$	Cumulative number of social feedback made by parents in the student *i*'s classroom at time *t*
Cumulative social activity$_{it}$	Cumulative number of social activities made by student *i*
Variables impacting state-dependent outcome (X_{it})	
Sign-in$_{it}$	Number of sign-ins made student *i* at time *t*
Class size$_{i}$	Number of users in student *i*'s classroom
Teacher Post$_{it}$	Number of proactive social activities made by teachers in the student *i*'s classroom at time *t*
Peer Post$_{it}$	Number of proactive social activities made by classmates in the student *i*'s classroom at time *t*
Parent Post$_{it}$	Number of proactive social activities made by parents in the student *i*'s classroom at time *t*
High Grade$_{i}$	Grade 1-3 = 0; 4-6 = 1
School_X$_{i}$	School dummy variables (X = A, B, or C)
Treatment Variable	
Pandemic	Year before COVID-19 (2019) = 0; after COVID-19 (2020) = 1

Table 4.2: Descriptive Statistics

Variables	Mean	S.D.	Min.	Max.
Outcome variable (Y_{it})				
Social activity$_{it}$	1.896	8.057	0	357
State transition (W_{it})				
Teacher cumulative feedback$_{it}$	117.399	448.266	0	3704.628
Peer cumulative feedback$_{it}$	37.035	84.920	0	746.634
Parent cumulative feedback$_{it}$	23.817	50.238	0	519.180
Cumulative social activity$_{it}$	2.095	8.632	0	393.225
State-dependent outcome (X_{it})				
Teacher Post$_{it}$	10.435	41.867	0	355
Peer Post$_{it}$	2.843	8.888	0	106
Parent Post$_{it}$	2.032	7.720	0	106
Sign-in$_{it}$	1.915	5.358	0	190
Class size$_{i}$	43.152	9.920	19	70
Categorical variables				
High Grade$_{i}$	0: 244	1:1416		
School_X$_{i}$	A: 610	B:389	C:238	D:179
Treatment Variable				
Pandemic	0: 527	1:889		

4.5. Estimation Results

I use maximum likelihood estimation to estimate the HMM. To improve computational efficiency, I randomly sampled 400 students from the dataset. A robustness check of the sample selection was performed by estimating the HMM on different samples. I start the estimation process with a latent class model to estimate the initial probability for the latent self-efficacy states. Then, I estimate the model parameters of HMM by maximizing the likelihood function described in Section 3.4. To avoid the label switching problem in HMM estimation (Jasra et al. 2005), additional constraints on the constant terms for the state-dependent outcome is added: $\delta_{0s_L} < \delta_{0s_H}$, where δ_{0s} is the constant term for a state s, and s_L and s_H denotes a lower self-efficacy state and a higher self-efficacy state, respectably. The constraints imply that a higher self-efficacy state produces more social activities than a lower self-efficacy state when other variables remain the same. Prior to the estimation of the main model, I optimize the time discounting factor by conducting a one-dimensional search. I evaluate the likelihood function with different values with a fixed interval and obtain 0.3 as the optimal value of the time discounting factor. Appendix G provides the detail of the process.

Then, with the optimized time discounting factor, I run models with different number of latent states to select the optimal number of states. Bayesian Information Criterion (BIC) is used as the selection criteria to choose the number of latent states n. Table 4.3 summarizes the model selection criteria values for estimated models. The result shows that the two-state HMM outperforms other models.

Table 4.3: Selection of the Number of States

	Log-likelihood	Number of parameters	AIC	BIC
1-state	-11280.366	15	22590.732	22650.604
2-state	**-9955.232**	**48**	**20006.464**	**20198.054**
3-state	-10455.873	75	21061.746	21361.106

Table 4.4 presents the estimated parameters for the two-state HMM with two self-efficacy states, which are low and high. For each parameter estimate, the corresponding standard error is in parentheses.

Table 4.4: Estimates for 2-state HMM

Parameter	State 1 (L)		State 2 (H)	
θ Over dispersion	0.190	(0.638)	0.775	(1.110)
State-dependent outcome (X_{it})				
Teacher Post$_{it}$	10.831	(11.893)	24.171***	(7.570)
Peer Post$_{it}$	103.168***	(3.102)	50.830***	(1.069)
Parent Post$_{it}$	18.008***	(1.280)	24.697***	(2.607)
Teacher Post$_{it}$ × Pandemic$_t$	-3.486	(11.217)	-22.130***	(7.616)
Peer Post$_{it}$ × Pandemic$_t$	-4.611***	(1.714)	-42.103***	(3.285)
Parent Post$_{it}$ × Pandemic$_t$	4.210***	(1.131)	23.370***	(3.347)
Pandemic$_t$	-0.814	(3.987)	1.211	(1.462)
High Grade$_i$	18.408***	(1.089)	44.452***	(4.396)
School_A$_i$	3.653**	(1.543)	-4.055	(2.597)
School_B$_i$	0.111	(1.140)	36.957***	(2.700)
School_C$_i$	13.310***	(1.048)	47.463***	(1.593)
Class size$_i$	-19.394***	(3.517)	-28.866***	(7.315)
Sign-in$_{it}$	192.464***	(3.776)	54.529***	(10.016)
Constant	-24.771***	(3.768)	14.077***	(5.252)
State transition (W_{it-1})				
Teacher cumulative feedback$_{it-1}$	4.084*	(2.217)	8.044***	(2.240)
Peer cumulative feedback$_{it-1}$	-8.032***	(2.883)	39.676***	(2.457)
Parent cumulative feedback$_{it-1}$	9.607**	(5.028)	33.569***	(1.134)
Cumulative social activity$_{it-1}$	10.869***	(1.573)	37.760***	(1.849)
Teacher cumulative feedback$_{it-1}$ × Pandemic$_{t-1}$	-4.432*	(2.453)	-8.214***	(2.326)
Peer cumulative feedback$_{it-1}$ × Pandemic$_{t-1}$	10.732***	(3.265)	-5.499**	(2.495)
Parent cumulative feedback$_{it-1}$ × Pandemic$_{t-1}$	-7.973***	(2.549)	-24.151***	(1.286)
Cumulative social activity$_{it-1}$ × Pandemic$_{t-1}$	10.572***	(1.469)	25.791***	(1.401)
Threshold				
	2.920***	(0.956)	0.411	(2.075)

Notes. Standard errors are reported in parentheses. The following rescaling is performed: *Teacher Post*, *Peer Post*, *Parent Post*, *Sign-in*, *Teacher cumulative feedback*, *Peer cumulative feedback*, *Parent cumulative feedback*, and *Cumulative social activity* are scaled down by a factor of 1000; *High Grade*, *School_A*, *School_B*, and *School_C* are scaled down by a factor of 100; *Class size* is log transformed and scaled down by a factor of 10.

4.5.1. State-dependent outcome

The estimated parameters for the variables affecting state-dependent outcome describe which factors influence students' social activities in online learning communities based on a given self-efficacy state. The result shows that recent proactive social activities from all of the three user groups (teacher, peer, and parents) have a positive influence on students' social activities on the focal platform. For example, the estimated coefficient for the variable *Peer Post* is positive and statistically significant for a student in any state (103.168, $p < 0.01$ for state L; 50.830, $p < 0.01$ for state H). This result suggests that peers' recent proactive social activity on the platform has a significant positive impact on students' social engagement regardless of their state. Similarly, the significant positive estimates for the other two social influence variables, *Teacher Post* (24.171, $p < 0.01$ for state H) and *Parent Post* (18.008, $p < 0.01$ for state L; 24.697, $p < 0.01$ for state H), reveals that students' social activity count increases when there are more proactive social activities from teachers and parents on the platform. However, the effect of teacher activity on state-dependent outcomes of students in state L is not statistically significant.

The estimates for the interaction terms between three social influence variables and the treatment variable, *Pandemic*, show the change in magnitude of each group's social influence on students' social engagement. While the parental influence on the student's social engagement increases during the pandemic setting, the social influences from teachers and peers decreased during the pandemic. The result suggests that, during the pandemic, parents' influence on a student's social engagement is significantly increased regardless of the student's efficacy state (4.210, $p < 0.01$ for state L; 23.370, $p < 0.01$ for state H). In contrast, the teacher and peer influences on student online social engagement are significantly decreased for students in state H

(-22.130, $p < 0.01$ for teachers; and -42.103, $p < 0.01$ for peers), and the peer influence is significantly decreased for students in state L (-4.611, $p < 0.01$ for peers). This result suggests that, during the pandemic, social influence from parents in deriving more social engagement from students got relatively stronger compared to the social influences from teachers and peers. One possible explanation for this result is the weakened social bond among students and teachers during the pandemic. During the pandemic period, students had fewer opportunities to have face-to-face interactions with teachers and peers. Indeed, from interviews with teacher users in the focal platform, I found that they felt they were less socially connected with their students during the pandemic compared to the pre-pandemic time:

"Communication with students was very limited compared to the same academic weeks before the pandemic. It was hard to learn about each student's characteristics before schools resumed face-to-face schooling."

In contrast, students may spend more time with their parents during the pandemic than during the pre-pandemic periods since they take online classes from home instead of going to schools. This environmental change suggests that students' perceived social distances with peers and teachers may increase, whereas the perceived social distance with parents may decrease during the pandemic, which, in turn, strengthens the parental influence on students' online social engagement.

4.5.2. State transition

The estimated parameters for the variables affecting state transitions describe the self-efficacy state dynamics. Table 4.5 represents the intrinsic state transition probabilities between the states L and H. For example, the intrinsic probability of transit from state L to H is 0.0512. Following Yan and Tan (2014), the intrinsic probabilities are calculated with the estimated threshold values for state transition. All the other variables that affect state transition were set to zero. Therefore, the intrinsic transition probabilities represent the transition probabilities without other influences. The intrinsic transition probabilities indicate that students in state L have a very high tendency to stay in state L (0.9488), while students in state H have a low tendency to stay in state H (0.3987). This result suggests that students have a strong tendency to be in State L without other influences.

Table 4.5: Intrinsic state transition matrix

	L	H
L	0.9488	0.0512
H	0.6013	0.3987

4.5.2.1. Social feedback and state transition

The result shows that the effect of social feedback on a student's social behavior changes based on the student's self-efficacy state. For students in state H, all three user group's social feedback has a significant positive effect on state transition (8.044, $p < 0.01$ for teachers; 39.676, $p < 0.01$ for peers; 33.569, $p < 0.01$ for parents). This result suggests that a student in state H has a higher propensity to stay in H when the social feedback from teachers, peers, and parents increases. For students in state L, teachers' and parents' feedback has a significant positive effect on state transition (4.084, $p < 0.1$ for teachers; 9.607, $p < 0.05$ for parents), but peer feedback has a significant negative effect on state transition (-8.032, $p < 0.01$).

Peer dynamics can explain the negative peer effect for students in state L. According to Sunwolf and Leets (2004), peer group boundaries are socially constructed through interactions, and the boundaries are penetrable and continually redefined. They found that peer group rejection can take various forms ranges from ignoring to insulting. When a student tries to penetrate the boundary of an existing peer group, there is a potential risk of experiencing rejections from others. Peer group rejection creates substantial social stress for students. Previous literature on peer rejection has shown that students feel more isolated, have fewer opportunities to develop their social skills, and have a higher tendency to have antisocial behavior after experiencing repetitive rejections from other students (Dodge et al. 1983, Hartup 1995, Hartup 1998). In the context of this study, students may feel that the existing peer group boundaries on the platform are stronger as other students are actively supporting each other on the platform. As the perceived strength of the boundaries gets stronger, students with lower self-efficacy may feel less confident about penetrating the pre-existing peer group boundaries. As a result, they feel less confident in building good relationships with others on the platform by exerting social interactions. However, the significant positive effect of teacher feedback on students' self-efficacy suggests that teachers and parents can mediate the lack of confidence that students in state L may suffer from. This finding aligns with the prior literature on the social role of adults, such as teachers and parents (e.g., Hamm et al. 2011, Farmer et al. 2011).

The estimated parameters on the interaction terms suggested that the pandemic reduces the baseline impact of social feedback on state transition. For state H, the estimated coefficients on the interaction terms are negative and statistically significant (-8.214, $p < 0.01$ for teachers; -5.499, $p < 0.05$ for peers; -24.151, $p < 0.01$ for parents), which are opposite to the positive baseline effects. Similarly, for state L, the estimated coefficients on the interaction terms for teacher, peer, and parents feedback show significant opposite effects (-4.432, $p < 0.1$ for teachers; 10.732, $p < 0.01$ for peers; -7.973, $p < 0.01$ for parentss) to the baseline effects, which are positive for teachers and parents, and negative for peers.

The estimated results show that the social influences from other users on state transition are reduced during the pandemic compared to the pre-pandemic period. Table 4.6 shows the changed transition probabilities after having the average level of social feedback. To calculate the difference in transition probabilities of a state, I set the social feedback variables, which are *Teacher cumulative feedback*, *Peer cumulative feedback*, and *Parent cumulative feedback,* and their corresponding interaction terms to their average values in the dataset. The corresponding intrinsic probabilities are in the parenthesis. The result shows that the magnitude of the impact of others' social feedback on students' self-efficacy state transition probabilities is smaller during the pandemic.

Table 4.6: Changes in state transition matrix – Social feedback

Pre-pandemic			Pandemic		
	L	H		L	H
L	0.9248	0.0752	*L*	0.9439	0.0561
	(0.9488)	(0.0512)		(0.9488)	(0.0512)
H	0.0572	0.9428	*H*	0.2575	0.7425
	(0.6013)	(0.3981)		(0.6013)	(0.3981)

4.5.2.2. Self-observation and state transition

The positive estimates on the variable *Cumulative social activity* (10.869, $p < 0.01$ for state L; 37.760, $p < 0.05$ for state H) suggest that the cumulative number of social activities of a student have a positive effect on state transition regardless of the student's self-efficacy state. This result implies that students' self-observation on their past engagement in social interactions with others can improve students' self-efficacy in exerting new social interactions and building good relationships with others. Such a positive effect of self-observation on one's self-efficacy is in line with the previous literature on self-efficacy (Bandura, 1997; Zhou et al. 2021)

The results also indicate that the influence of self-observation on student social behavior increases during the pandemic. The estimated coefficients on the interaction terms between *Cumulative social activity* and *Pandemic* are positive and statistically significant (10.572, $p < 0.01$ for state L; 25.791, $p < 0.01$ for state H), which indicates that the positive effect of self-observation on student social participation becomes stronger during the pandemic. Table 4.7 demonstrates the changes in transition probabilities with the presence of self-observation. Similar to the construction of Table 4.6, I calculate the transition probabilities by setting the variable *Cumulative social activity* to its average value in the dataset. The result shows that the magnitude of the positive impact of self-observation on state transition probabilities becomes more substantial during the pandemic compared to the pre-pandemic period.

Table 4.7: Changes in state transition matrix – Self-observation

Pre-pandemic				Pandemic		
	L	H			L	H
L	0.9477	0.0523		*L*	0.9466	0.0534
	(0.9488)	(0.0512)			(0.9488)	(0.0512)
H	0.5822	0.4178		*H*	0.5690	0.4310
	(0.6013)	(0.3981)			(0.6013)	(0.3981)

From the estimated results, I observe the weaker social influence and the stronger personal influence on state transition during pandemic regardless of students' self-efficacy states. This tendency can potentially be explained by the change in students' perceived group membership during the pandemic. Prior literature on group membership has shown that group membership affects individuals' behavior and their preferences over outcomes, and saliency of a group affects individuals' perception of the surrounding environment. However, a weak group membership only has a marginal impact on individual behavior (Charness et al. 2007). The pandemic reduced the opportunities for face-to-face interactions between a student and her school members, and this may impair students' perceived group membership. As a result, the relative importance of personal influence on students' self-efficacy increases during the pandemic.

4.6. Conclusions

In this study, I proposed a Hidden Markov Model (HMM) to investigate the dynamics of students' online social engagement and to what extent the external shock from the recent COVID-19 pandemic brought in the dynamics. The result suggests that recent proactive social activities from all user groups (i.e., teachers, peers, and parents) have positive influences on students' social engagement. The analysis on state transition suggests that the social influences from peers, teachers, and parents on students' self-efficacy are positive for students with higher self-efficacy. For students with lower self-efficacy, the social influence from peers is negative, which can be potentially explained by peer dynamics. However, the social influence from teachers and parents remain positive for the students with lower self-efficacy. This result suggests that the positive impact from teachers and parents can mediate the negative impact from peers and highlights the importance of the social role of adults in online learning environments. The result also implies that students' self-observation on their past engagement in social interactions can improve students' self-efficacy.

The findings also provide an insight into students' online social behavior in the post-pandemic era. During the pandemic, the relative effect of parents' social proactive activities on students' social engagement increases compared to that of teachers and peers. Also, the relative effect of students' self-observation on students' self-efficacy increases during the pandemic, whereas that of other users' social feedback decreases during the pandemic. The reduced face-to-face schooling can explain such change during the pandemic.

This study can be extended in several ways. First, the research setting is an online learning platform designed for schools, which means the focal platform is an online dimension of a hybrid learning system that consists of both online and face-to-face learning. Therefore, it is not clear if the findings can be generalized to other independent online learning environments without face-to-face learning. Future research can explore the dynamics of students' social engagement in independent online learning environments and compare the result with the current findings. Second, the sentiment of social activities on the platform was not considered in this study. Prior student engagement literature suggests that the positive and negative experiences with others have different effects on student engagement (e.g., Juvonen et al. 2014). Therefore including the sentiment of social activities would enrich the analysis and give insightful implications. Lastly, future research can extend this study by including academic achievement measures. Including academic achievement measures would allow researchers to examine the differing effects of different groups' (i.e., teachers, peers, and parents) social activities on students' social engagement and academic achievement.

Chapter 5. CONCLUDING REMARKS

The rapid development in digital technologies has enabled us to do various everyday activities in online environments. Ranging from shopping to learning, our collective reliance on digital technologies has continuously increased in recent years. Especially, the COVID-19 pandemic has intensified the penetration of the internet in our daily activities. Considering the unique characteristics of online settings, understanding online-specific user behavior is crucial to successfully managing online business platforms. In this dissertation, I explored user behavior in two domains: online shopping and online learning, and how environmental and technological changes affect online user behavior.

In the first essay, I investigate the provision of helpful product video reviews. Video reviews have become a major Word-of-Mouth (WOM) channel thanks to the remarkable improvement on the internet and the increased popularity of user-created video platforms. According to Haller (2019), more than 50 percent of consumers today utilize online videos before making any purchase decisions. As the provision of online review videos grows and consumers increasingly rely on them for their purchase decisions, understanding factors that contribute to the perceived helpfulness of review videos becomes critical for review video management. This study examined how various visual, speech, and content characteristics are associated with perceived helpfulness. This study contribute to the online review literature by investigating and suggesting the guideline of the provision of helpful review videos, which has been underexplored. To the best of my knowledge, this study is one of the earliest studies investigating the review video helpfulness.

My second essay studied the dynamics of students' online social engagement and how the recent COVID-19 pandemic affected it. Online learning has shown continuous growth in recent years, and such a trend was already present before the COVID-19 pandemic. In 2015, the global online learning market was expected to reach $350 Billion by 2025 from its $107 Billion value in 2015. The projection was updated to $458 billion after witnessing the unexpectedly rapid growth during the pandemic (Koksal 2020). Students experienced a sudden shift away from the traditional face-to-face schooling during the pandemic. Accordingly, the demand for online learning rapidly increased since using online education tools became necessary for schools to deal with the pandemic-initiated school closures. This study explored elementary school students' social behavior in the context of an online learning platform. The result suggested that the relative importance of parental influence on students' online social engagement and students' personal influence on their self-efficacy increase during the pandemic. The reduced face-to-face schooling can explain such change during the pandemic. This study contributes to the online learning literature by conducting an extensive empirical analysis of students' online behavior dynamics. Further, The findings provide a set of insights on students' online behavior in the post-pandemic era.

This dissertation intends to provide both theoretical and practical insights into user behavior in different online platforms. Further, it addresses how the individuals' behavior patterns change in response to various technological and environmental changes. The findings suggest that online user behavior is distinctive from their corresponding offline behavior and subject to change as the relevant technologies and environmental settings change in the future. Therefore, this dissertation encourages future research to explore the change in online user behavior when systematical changes occur.

BIBLIOGRAPHY

Ansong, D., Okumu, M., Bowen, G. L., Walker, A. M., and Eisensmith, S. R. 2017. "The role of parent, classmate, and teacher support in student engagement: Evidence from Ghana," *International Journal of Educational Development* (54), pp. 51-58.

Appelman, A., and Sundar, S. S. 2016. "Measuring message credibility: Construction and validation of an exclusive scale," *Journalism & Mass Communication Quarterly* (93:1), pp.59-79.

Apple, W., Streeter, L. A., and Krauss, R. M. 1979. "Effects of pitch and speech rate on personal attributions," *Journal of personality and social psychology* (*37*:5), pp.715.

Bandura, A. 1991. "Social cognitive theory of self-regulation," *Organizational behavior and human decision processes* (50:2), pp. 248-287.

Bandura, A. 1977. "Self-efficacy: toward a unifying theory of behavioral change," *Psychological review* (84:2), pp. 191.

Bandura, A. 1989. "Human agency in social cognitive theory," *American psychologist* (44:9), pp. 1175.

Bandura, A. 2001. "Social cognitive theory: An agentic perspective," *Annual review of psychology* (52:1), pp. 1-26.

Bandura, A., Freeman, W. H., and Lightsey, R. 1999. "Self-efficacy: The exercise of control."

Berger, J. A., Sorensen, A. T., and Rasmussen, S. J. 2010. "Positive Effects of Negative Publicity: When Negative Reviews Increase Sales," *Marketing Science* (29:5), pp. 815-827.

Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. "Latent dirichlet allocation," *Journal of Machine Learning research* (3), pp. 993-1022.

Borji, A., and Itti, L. 2012. "State-of-the-art in visual attention modeling," *IEEE transactions on pattern analysis and machine intelligence* (35:1), pp.185-207.

Broadbent, D. E. 1958. "Perception and communication," *London: Pergamon press*

Brindley, J. E., Blaschke, L. M., and Walti, C. 2009. "Creating effective collaborative learning groups in an online environment," *International Review of Research in Open and Distributed Learning* (10:3).

Burgoon, J. K., Birk, T., and Pfau, M. 1990. "Nonverbal behaviors, persuasion, and credibility," *Human communication research* (17:1), pp.140-169.

Carabajal, K., Lapointe, D., and Gunawardena, C. N. 2003. "Group development in online learning communities," *Handbook of distance education*, pp. 217-234.

Cacioppo. J. T. and Petty, R. E. 1982. "The need for cognition," *Journal of Personality and Social Psychology* (42), pp. 116-13

Chang, J. J., Lin, W. S., and Chen, H. R. 2019. "How attention level and cognitive style affect learning in a MOOC environment? Based on the perspective of brainwave analysis," *Computers in Human Behavior* (100), pp. 209-217.

Charness, G., Rigotti, L., and Rustichini, A. 2007. "Individual behavior and group membership," *American Economic Review* (97:4), pp. 1340-1352.

Chebat, J. C., Hedhli, K. E., Gélinas-Chebat, C., and Boivin, R. 2007. "Voice and persuasion in a banking telemarketing context," *Perceptual and motor skills* (104:2), pp. 419-437.

Chevalier, J. A., and Mayzlin, D. 2006. "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research* (43:3), pp.345-354.

Chen, C. M., and Wu, C. H. 2015. "Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance," *Computers & Education* (80), pp. 108-121.

Chen, Y. and Xie, J. 2008. "Online consumer review: Word-of-mouth as a new element of marketing communication mix," *Management Science* (54:3), pp.477-491.

Cho, M.-H., Shen, D., and Laffey, J. 2010. “Relationships between self-regulation and social experiences in asynchronous online learning environments,” *Journal of Interactive Learning Research* (21), pp. 297–316.

Connors, L., Mudambi, S. M., and Schuff, D. 2011. “Is It the Review or the Reviewer? A Multi-Method Approach to Determine the Antecedents of Online Review Helpfulness,” in *Proceedings of the 44th Hawaii International Conference on System Sciences*, Los Alamitos, CA: IEEE Computer Society.

Cox, B., and Cox, B. 2008. “Developing interpersonal and group dynamics through asynchronous threaded discussions: The use of discussion board in collaborative learning,” *Education* (128:4).

Daft, R.L. and Lengel, R.H. 1986. “Organizational information requirements, media richness and structural design,” *Management Science* (32:5), pp.554-571.

Dennis, A.R. and Kinney, S.T. 1998. “Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality,” *Information Systems Research* (9:3), pp.256-274.

Dos Santos, C., and Gatti, M. 2014. “Deep convolutional neural networks for sentiment analysis of short texts,” *In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers,* pp. 69-78.

Dodge, K. A., Schlundt, D. C., Schocken, I., and Delugach, J. D. 1983. “Social competence and children's sociometric status: The role of peer group entry strategies,” *Merrill-Palmer Quarterly*, pp. 309-336.

Farmer, T. W., Lines, M. M., and Hamm, J. V. 2011. “Revealing the invisible hand: The role of teachers in children's peer experiences,” *Journal of Applied Developmental Psychology* (32:5), pp. 247-256.

Fasel, B., and Luettin J. 2003. “Automatic Facial Expression Analysis: A Survey,” *Pattern Recognition* (36:3), pp. 259–275.

Ford, G. T., Smith, D. B., and Swasy, J. L. 1990. “Consumer skepticism of advertising claims: Testing hypotheses from economics of information,” *Journal of consumer research* (16:4), pp. 433-441.

Forman, C., Ghose, A., and Wiesenfeld, B. 2008. “Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," *Information systems research* (19:3), pp. 291-313.

Fulk, J., Schmitz, J., and Steinfield, C. 1990. "A social influence model of technology use," *Organizations and Communication Technology*, pp. 117-140.

Gefen, D., and Detmar W. S. 2004. "Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services," *Omega* (32:6), pp. 407-424.

Ghose, A. and Ipeirotis, P.G. 2011. "Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp. 1498-1512.

Goes, P. B., Lin, M., and Au Yeung, C. M. 2014. "Popularity effect" in user-generated content: Evidence from online product reviews," *Information Systems Research* (25:2), pp. 222-238.

Gu, B., Park, J. and Konana, P. 2012. "Research note—the impact of external word-of-mouth sources on retailer sales of high-involvement products," *Information Systems Research* (23:1), pp.182-196.

Gu, B. and Ye, Q. 2014. "First step in social media: Measuring the influence of online management responses on customer satisfaction," *Production and Operations Management* (23:4), pp. 570-582.

Gunning, R. 1969. "The Fog Index After Twenty Years," *Journal of Business Communication* (6:2), pp. 3–13.

Hall, J. A. 1980. "Voice tone and persuasion," *Journal of Personality and Social Psychology* (38:6), pp. 924.

Haller, M. 2019 "Search to video and back again: Explore how customers move through the purchase journey," *Think with Google* (https://www.thinkwithgoogle.com/advertising-channels/search-video-purchase-journey/).

Hamm, J. V., Farmer, T. W., Dadisman, K., Gravelle, M., and Murray, A. R. 2011. "Teachers' attunement to students' peer group affiliations as a source of improved student experiences of the school social–affective context following the middle school transition," *Journal of Applied Developmental Psychology* (32:5), pp. 267-277.

Harrisim, L., Hiltz, S. R., Teles, L., and Turoff, M. 1995. *Learning networks: A field guide to teaching and learning online*. Boston: MIT Press.

Hartup, W. W. 1995. "The three faces of friendship," *Journal of Social and personal Relationships* (12:4), pp. 569-574.

Hartup, W. W. 1998, "10 Cooperation, close relationships, and cognitive development," *The company they keep: Friendships in childhood and adolescence,* pp. 213.

Hass, R. G. 1981. "Effects of source characteristics on cognitive responses in persuasion," *Cognitive responses in persuasion*, pp.141-172.

Haugtvedt, C. P., and Petty, R. E. 1992. "Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes," *Journal of Personality and Social Psychology* (63:2), pp. 308–319.

Ho, Y. C., Wu, J., and Tan, Y. 2017. "Disconfirmation effect on online rating behavior: A structural model," *Information Systems Research* (28:3), pp. 626-642.

Homer, P. M., and Kahle, L. R. 1990. "Source expertise, time of source identification, and involvement in persuasion: An elaborative processing perspective," *Journal of advertising* (19:1), pp. 30-39.

Hong, W., Thong, J.Y. and Tam, K.Y. 2004. "Does animation attract online users' attention? The effects of flash on information search performance and perceptions," *Information Systems Research* (15:1), pp.60-86.

Huang, A. H., Lehavy, R., Zang, A. Y., and Zheng, R. 2018. "Analyst information discovery and interpretation roles: A topic modeling approach," *Management Science* (64:6), pp.2833-2855.

Inoue, K., and Takeda, Y. 2012. "The role of attention in the contextual enhancement of visual memory for natural scenes," *Visual Cognition* (20:1), pp. 94-107.

Itti, I., Koch, C., and Niebur, E. 1998. "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," *IEEE Transactions on PAMI* (20:11), pp. 1254-1259.

Jasra, A., Holmes, C. C., and Stephens, D. A. 2005. "Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling," *Statistical Science*, pp. 50-67.

Jarboe, G. 2015. “64 Percent of Consumers Use YouTube to Review Electronics Before Purchase,” *Tubular Insights* (https://tubularinsights.com/64-percent-consumers-youtube-to-review-electronics-purchase/)

Jiang, Z. and Benbasat, I., 2007. “The effects of presentation formats and task complexity on online consumers' product understanding,” *MIS Quarterly*, pp.475-500.

Jones, Q., Ravid, G., and Rafaeli, S. 2004. “Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration,” *Information Systems Research* (15:2), pp. 194-210.

Joseph, W. B. 1982. "The credibility of physically attractive communicators: A review," *Journal of advertising* (11:3), pp 15-24.

Juvonen, J., Espinoza, G., and Knifsend, C. 2012. “The role of peer relationships in student academic and extracurricular engagement,” in *Handbook of research on student engagement* (pp. 387-401). Springer, Boston, MA.

Ke, F. 2010. “Examining online teaching, cognitive, and social presence for adult students,” *Computers & Education* (55:2), pp. 808-820.

Khan, M. L. 2017. “Social media engagement: What motivates user participation and consumption on YouTube?” *Computers in Human Behavior* (*66)*, pp. 236-247.

Kiefer, S. M., Alley, K. M., and Ellerbrock, C. R. 2015. “Teacher and peer support for young adolescents’ motivation, engagement, and school belonging,” *Rmle Online* (38:8), pp. 1-18.

Kim, S. and Choi, S.M. 2012. “Credibility cues in online shopping: an examination of corporate credibility, retailer reputation, and product review credibility,” *Int. J. Internet Marketing and Advertising* (3), pp.217–236.

King, R. B. 2016. “Gender differences in motivation, engagement and achievement are related to students' perceptions of peer—but not of parent or teacher—attitudes toward school,” *Learning and Individual Differences* (52), pp. 60-71.

Kizilcec, R. F., Papadopoulos, K., and Sritanyaratana, L. 2014. “Showing face in video instruction: effects on information retention, visual attention, and affect,” *Proceedings of the SIGCHI* *conference on human factors in computing systems*.

Kohli, R., Devaraj, S. and Mahmood, M.A. 2004. "Understanding determinants of online consumer satisfaction: A decision process perspective," *Journal of Management Information Systems* (21:1), pp.115-136.

Koksal, I. 2020. "The Rise Of Online Learning", *Forbes*, (https://www.forbes.com/sites/ilkerkoksal/2020/05/02/the-rise-of-online-learning/)

Kuan, K. K. Y., Hui, K-L., Prasarnphanich, P., and Lai, H-Y. 2015. "What Makes a Review Voted? An Empirical Investigation of Review Voting in Online Review Systems," *Journal of the Association for Information Systems* (16:1), pp. 48-71.

Lancaster, T. 2000. "The incidental parameter problem since 1948," *Journal of Econometrics* (95:2), pp. 391–413.

Laustsen, L., Petersen, M. B., and Klofstad, C. A. 2015. "Vote choice, ideology, and social dominance orientation influence preferences for lower pitched voices in political candidates," *Evolutionary Psychology* (13:3), pp. 1-13.

Li, M., Huang, L., Tan, C.H. and Wei, K.K. 2013. "Helpfulness of online product reviews as seen by consumers: source and content features," *International Journal of Electronic Commerce* (17:4), pp. 101-136.

Liu, Y. 2006. "Word of mouth for movies: Its dynamics and impact on box office revenue," *Journal of Marketing* (70:3), pp. 74-89.

Lu, B., Fan, W., and Zhou, M. 2016. "Social presence, trust, and social commerce purchase intention: An empirical research," *Computers in Human Behavior* (56), pp. 225-237.

Lyons, A., Reysen, S., and Pierce, L. 2012. "Video lecture format, student technological efficacy, and social presence in online courses," *Computers in Human Behavior* (28:1), pp. 181–186.

Martin, D. P., and Rimm-Kaufman, S. E. 2015. "Do student self-efficacy and teacher-student interaction quality contribute to emotional and social engagement in fifth grade math?" *Journal of school psychology* (53:5), pp. 359-373.

Mayer, R.E., 1997. "Multimedia learning: Are we asking the right questions?," *Educational Psychologist* (32:1), pp. 1-19.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., and Burgette, L. F. 2013. “A tutorial on propensity score estimation for multiple treatments using generalized boosted models,” *Statistics in medicine* (32:19), pp. 3388–3414.

McDonald, J. F., and Moffitt, R. A. 1980. “The uses of Tobit analysis,” *The review of economics and statistics*, pp. 318-321.

McGlohon, M., Glance, N., and Reiter, Z. 2010. “Star quality: Aggregating reviews to rank products and merchants,” *Fourth International AAAI Conference on Weblogs and Social Media*.

McInnerney, J. M., and Roberts, T. S. 2004. “Online learning: Social interaction and the creation of a sense of community,” *Journal of Educational Technology & Society* (7:3), pp. 73-81.

Miller, A. 1987. “Cognitive Styles: an integrated model,” *Educational Psychology* (7:4), pp. 251-268.

Moreno, R. and Mayer, R.E. 1999. “Cognitive principles of multimedia learning: The role of modality and contiguity,” *Journal of Educational Psychology* (91:2), pp. 358.

Mudambi, S. M., and Schuff, D. 2010. “What makes a helpful review? A study of customer reviews on Amazon.com,” *MIS quarterly* (34:1), pp. 185-200.

Muilenburg, L. Y., and Berge, Z. L. 2005. “Student barriers to online learning: A factor analytic study,” *Distance education* (26:1), pp. 29-48.

Nelson, P. 1970. “Information and Consumer Behavior,” *Journal of Political Economy* (78:20), pp. 311-329.

Neyman, J., and Scott, E. 1948. “Consistent Estimates Based on Partially Consistent Observations,” *Econometrica* (16:1), pp. 1–32.

Palloff, R. M., and Pratt, K. 1999. “Building Learning Communities in Cyberspace: Effective strategies for theonline classroom,” *San Francisco: Jossey-Bass*.

Parsons, J., and Taylor, L. 2011. “Improving student engagement,” *Current issues in education* (14:1).

Pavlou, P. A., and Dimoka, A. 2006. “The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation," *Information Systems Research* (17:4), pp. 392-414.

Peterson, R. A., Cannito, M. P., and Brown, S.P. 1995. "An exploratory investigation of voice characteristics and selling effectiveness," *Journal of Personal Selling & Sales Management* (15:1), pp.1-15.

Pianta, R. C., Hamre, B. K., and Allen, J. P. 2012. "Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions," In *Handbook of research on student engagement* (pp. 365-386). Springer, Boston, MA.

Qazi, A., Syed, K.B.S., Raj, R.G., Cambria, E., Tahir, M. and Alghazzawi, D. 2016. "A concept-level approach to the analysis of online review helpfulness," *Computers in Human Behavior* (58), pp. 75-81.

Rabiner L. R. 1989. "A tutorial on hidden Markov-models and selected applications in speech recognition," *Proceedings of the IEEE* (77:2), pp. 257–286.

Rajashekar, U., Van Der Linde, I., Bovik, A. C., and Cormack, L. K. 2008. "GAFFE: A gaze-attentive fixation finding engine," *IEEE transactions on image processing* (17:4), pp. 564-573.

Rice, R. E. 1993. "Media appropriateness: Using social presence theory to compare traditional and new organizational media," *Human communication research* (19:4), pp. 451-484.

Richardson, J. C., and Swan, K. 2003. "Examining social presence in online courses in relation to students' perceived learning and satisfaction," *Journal of Asynchronous Learning Networks* (7), pp. 68–88.

Rimm-Kaufman, S. E., Baroody, A. E., Larsen, R. A., Curby, T. W., and Abry, T. 2015. "To what extent do teacher–student interaction quality and student gender contribute to fifth graders' engagement in mathematics learning?" *Journal of Educational Psychology* (107:1), pp. 170.

Rogers, E. M., and Agarwala-Rogers, R. 1975. "Organizational communication," *Communication Behaviour*, pp. 218-239.

Röder, M., Both, A., and Hinneburg, A. 2015. "Exploring the space of topic coherence measures," *Proceedings of the eighth ACM international conference on Web search and data mining*, pp.

399-408.

Rosenbaum, P. R., and Rubin, D. B. 1983. "The central role of the propensity score in observational studies for causal effects," *Biometrika* (70:1), pp. 41-55.

Rosenberg, A., and Hirschberg, J. 2009. "Charisma perception from text and speech," *Speech Communication* (51:7), pp. 640-655.

Sawyer, A. G., Laran, J., and Xu, J. 2008. "The readability of marketing journals: are award-winning articles better written?" *Journal of Marketing*, (72:1), pp. 108-117.

Shao, G. 2009. "Understanding the appeal of user-generated media: a uses and gratification perspective," *Internet research* (19:1), pp. 7-25.

Short, J., E. Williams, B. Christie. B., 1976. "The social psychology of telecommunications," *Wiley, New York*.

Singh, P. V., Tan, Y., and Youn, N. 2011. "A hidden Markov model of developer learning dynamics in open source software projects," *Information Systems Research* (22:4), pp. 790-807.

Sit, J. W., Chung, J. W., Chow, M. C., and Wong, T. K. 2005. "Experiences of online learning: students' perspective," *Nurse education today* (25:2), pp. 140-147.

Skowronski, J. J., and Carlston, D. E. 1987. "Social judgment and social memory—the role of cue diagnosticity in negativity, positivity, and extremity biases," *Journal of Personality and Social Psychology* (52:4), pp. 689-699.

Simpson, C. W., and Prusak, L. 1995. "Troubles with information overload-moving from quantity to quality in information provision," *International Journal of Information Management* (15:6), pp. 413-425.

Smedt, T. D., and Daelemans, W. 2012. "Pattern for python," *Journal of Machine Learning Research*, (13:2), pp. 2063-2067.

Spool, J. M. 2009. "The Magic Behind Amazon's 2.7 Billion Dollar Question," User Interface Engineering (http://www.uie. com/articles/magicbehindamazon/).

Sun, A., and Chen, X. 2016. "Online education and its effective practice: A research review," *Journal of Information Technology Education* (15)

Sun, H. and Xu, L. 2018. "Online reviews and collaborative service provision: A signal-jamming model," *Production and Operations Management* (27:11), pp.1960-1977.

Sunwolf and Leets, L. 2004. "Being left out: Rejecting outsiders and communicating group boundaries in childhood and adolescent peer groups," *Journal of Applied Communication Research* (32:3), pp. 195-223.

Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., and Feinberg, D. R. 2012. "Voice pitch influences voting behavior," *Evolution and Human Behavior* (33:3), pp. 210-216.

Tormala, Z. L., and Petty, R. E. 2002. "What doesn't kill me makes me stronger: The effects of resisting persuasion on attitude certainty," *Journal of Personality and Social Psychology* (83:6), pp. 1298-1313.

Trowler, V. 2010. "Student engagement literature review," *The higher education academy* (11:1), pp. 1-15.

Tu, C., and Corry, M. 2002. "E-Learning communities," *The Quarterly Review of Distance Education* (3:2), pp. 207-218

Wang, T. Y., Kawaguchi, I., Kuzuoka, H., and Otsuki, M. 2018. "Effect of Manipulated Amplitude and Frequency of Human Voice on Dominance and Persuasiveness in Audio Conferences," *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), pp. 1-18.

Wang, Y., Goes, P., Wei, Z. and Zeng, D. 2019. "Production of online word-of-mouth: Peer effects and the moderation of user characteristics," *Production and Operations Management*, (28:7), pp. 1621-1640.

Xu, P., Chen, L. and Santhanam, R., 2015. "Will video be the next generation of e-commerce product reviews? Presentation format and the role of product type," *Decision Support Systems*, 73, pp. 85-96.

Yan, L., and Tan, Y. 2014. "Feeling blue? Go online: An empirical study of social support among patients," *Information Systems Research* (25:4), pp. 690-709.

Yang, J., Kinshuk, Yu, H., Chen, S. J., and Huang, R. 2014. "Strategies for smooth and effective cross-cultural online collaborative learning," *Journal of Educational Technology & Society* (17:3), pp. 208-221.

Yin, D., Bond, S., and Zhang, H. 2014. "Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews," *MIS quarterly* (38:2), pp. 539-560.

Zhao, W., Chellappa R., Phillips, J., and Rosenfeld, A. 2003. "Face Recognition: A Literature Survey," *ACM Computing Surveys (CSUR)*, (35:4), pp. 399-458.

Zhou, T., Yan, L., Wang, Y., and Tan, Y. 2021. "Turn Your Online Weight Management from Zero to Hero: A Multidimensional, Continuous-Time Evaluation," *Management Science*. Advance online publication.

Appendix A: Data Collection Process

I collect data on electronic product review videos uploaded on YouTube. I focus on electronic product review videos because it is one of the most popular product types for review videos. I identify and collect data on product review videos uploaded on YouTube by following steps. First, I select 40 keywords to identify and search for electronic product review videos that are published between February 2016 and January 2019. I set up 40 keywords that cover electronic products from seven different manufacturers: Apple, Google, Huawei, LG, Oppo, Samsung, and Xiaomi. The seven manufacturers contribute more than 90% of mobile vendor market share in North America.[18] For example, the set of search keywords for the manufacturer Apple includes iPhone, iPad, MacBook, and Apple Watch. The set of search keywords for the manufacturer Samsung includes Samsung Galaxy, Samsung Notebook, and Samsung Smart TV. Based on the 40 keywords, I combine each keyword with the word "review" to search for electronic product review videos that are published between February 2016 and January 2019. Examples of search queries include "MacBook review," which combines a keyword "MacBook" with the word "review". The resulting video reviews contain the search query in their titles, but the words in a search query can be discontinuous in titles. For instance, a product review video titled "Review of the MacBook Pro with Touch Bar" published on November 14, 2016, which is identified from the search query "MacBook review," reviews an electronic product, Apple MacBook Pro (2016).

To ensure the quality of auto-generated transcription, I focus on review videos created by English-speaking reviewers because YouTube's speech recognition has the best performance when detecting English. After keyword-based searches, I exclude videos with less than 1000 views

[18] Statcounter, Mobile Vendor Market Share in North America - June 2020, https://gs.statcounter.com/vendor-market-share/mobile/north-america

to ensure that each video had enough viewers to determine the helpfulness of the review videos. This results in 17,675 review videos. Then, I analyze the audio and transcription of the collected 17,675 review videos. As I focus on review videos that primarily use speech to deliver their core information, I exclude videos that convey their contents by primarily using written language or visual aids without using many verbal descriptions. To do so, I exclude 2,119 videos with an average duration-to-word larger than 1 second per word. The threshold is determined based on the observation that the average conversation rate for English speakers in the U.S. is 150 words per minute. Finally, exclude videos with less than 10 total votes to ensure that like and dislike votes can represent the perceived represent the perceived review helpfulness of review videos. As a result, the dataset includes 13,840 review videos.

Appendix B: Visual Variability construction

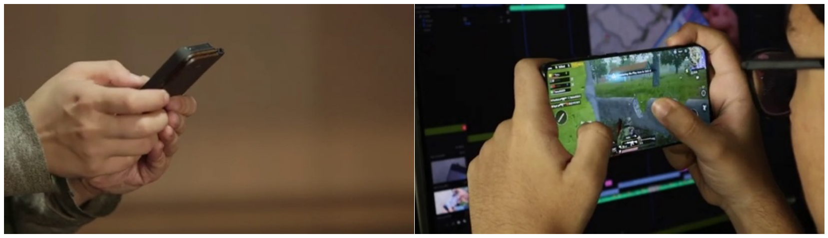
Following Li et al. (2019), I construct the variable *Visual Variation*, which measures the dynamic changes in review videos, by calculating the pixel-level distance between adjacent video frames. When a review video includes more visual changes, such as moving objects or background changes, the review video would have a higher visual variation. Appendix Figure B.1 shows how I process each frame to construct the variable *Visual Variation*. First, I convert RGB-coded color frames into grayscale frames.[19] Then, I normalize each greyscale frame to account for potential light exposure differences. For a frame with n pixels of grayscale value of $x_1, \ldots, x_n$, the normalized grayscale value for the pixel i is $x_i = (x_i - x_{min})/(x_{max} - x_{min})$, where $x_{max} = \max\{x_1, \ldots, x_n\}$, and $x_{min} = \max\{x_1, \ldots, x_n\}$. The normalization transforms the pixel value from the interval [0, 255] to [0, 1]. Afterward, I calculate the pixel-level difference between the nearby frames based on the Manhattan distance, $d(x_i', y_i') = |x_i' - y_i'|$. The difference between two nearby video frames is the average difference of every pair of pixels in the two frames. Appendix Figure B.2 illustrates examples with different value of frame difference, where Example 1 does not show significant changes between the two video frames whereas Example 2 shows a greater variation between frames in terms of both background and main object. The frame difference for Example1 (19406.61) is much smaller than the frame difference for Example 2 (537660.40). Lastly, the variable *Visual Variation* is the average frame distance between all consecutive frames included in a video:

$$Visual\ Variation = \frac{1}{N-1}\sum_{i=1}^{N-1} d(F_i, F_{i+1})$$

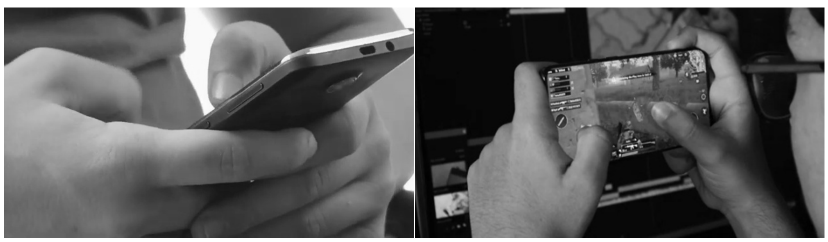
where $d(F_{i,} F_{i+1})$ is the frame distance between the *i*-th frame and the (*i*+1)-th frame in a review video, and N is the number of video frames included in the review video. Two consecutive video frames (i.e., F_i and F_{i+1}) are 10 seconds apart from each other.

[19] The conversion, which is done by taking the average value of the RGB values, reduces the computational overhead (Li et al. 2019),

Appendix Figure B.1: Visual Variability construction



Original RGB-coded color frames



Converted greyscale frames

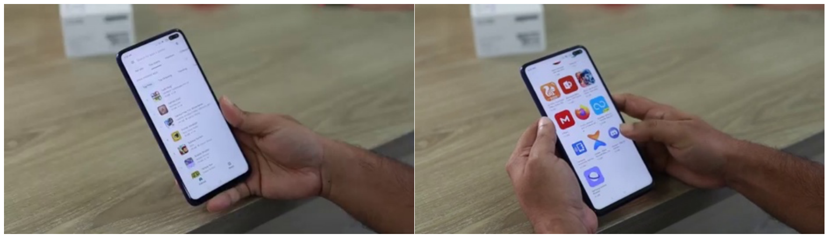
$$x_i = (x_i - x_{min})/(x_{max} - x_{min}),$$

where $x_{max} = \max\{x_1, \dots, x_n\}$, and $x_{min} = \max\{x_1, \dots, x_n\}$.

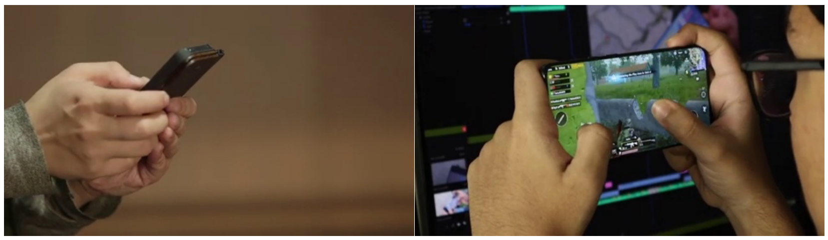
Intensity normalization

$$\boldsymbol{Visual\,Variation} = \log\left(\frac{1}{N-1}\sum_{i=1}^{N-1} d(F_i, F_{i+1})\right)$$

Appendix Figure B.2: Examples of frame distance



Example 1: distance between two above frames = 19406.61

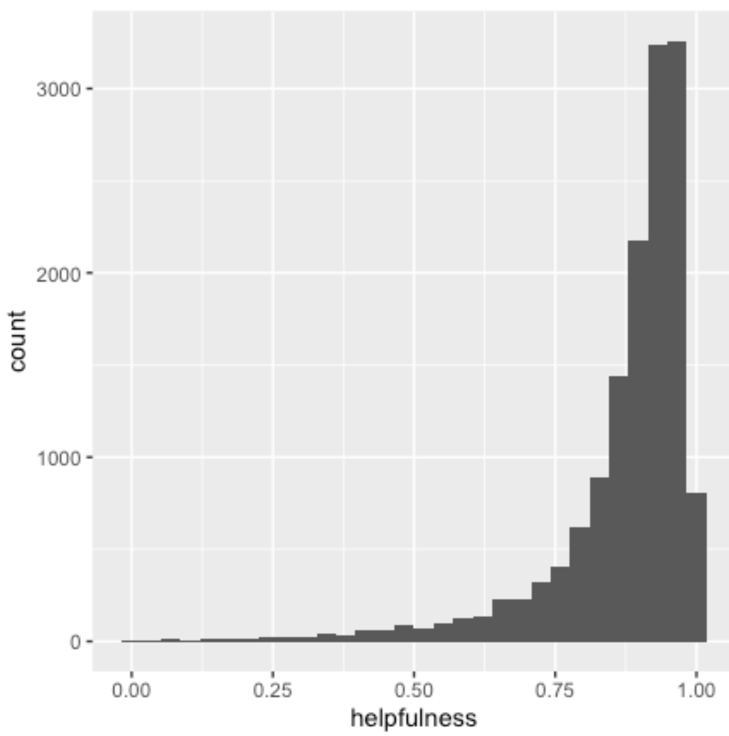


Example 2: distance between two above frames = 537660.40

Appendix C: Distribution of Helpfulness

The variable *Helpfulness* is defined as *Number of Likes* divided by *Total Votes*, where total votes refer to the total number of like and dislike votes. The distribution of the variable *Helpfulness* is a skewed *J* shape distribution between 0 and 1, and such type of distribution is commonly observed in various rating platforms (Ho et al. 2017, McGlohon et al. 2010).

Appendix Figure C.1: Distribution of Helpfulness



Notes: This figure plots the distribution of the variable *Helpfulness*. The x-axis shows the *Helpfulness* value, and the y-axis represents the number of review videos with corresponding value of *Helpfulness*.

Appendix D: Topic Modeling with Latent Dirichlet Allocation (LDA) Algorithm

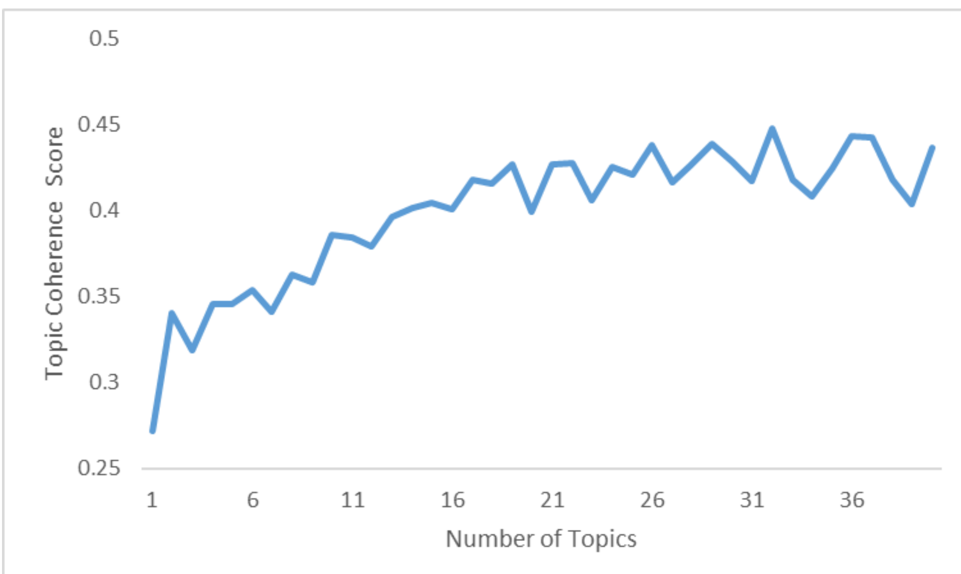
I apply Latent Dirichlet Allocation (LDA), which was established by Blei et al. (2003), on the review speech transcript data to identify the underlying topics in the reviews in the dataset. LDA is a widely used algorithm for topic modeling that develops a generative probabilistic model to discover the abstract topics that occur in the collection of text documents. LDA is an algorithm that resembles the writing process of a human (Huang et al. 2018). LDA assumes two distributions; (1) topic distribution in each document, and (2) word distribution in each topic. Then, LDA assumes the two steps involved in the generation of each word in a document. First, a topic is randomly drawn from the topic distribution of the document. Then, a word is randomly drawn from the word distribution of the topic. The document is generated by repeating these two steps. Based on this assumption, LDA estimates the topic distribution for each document and the word distribution of each topic by applying the Expectation-maximization (EM) algorithm. In short, LDA estimates the topic distribution of each document by discovering the probabilistic relations among topics and words.

I clean the review speech transcript data to feed them into the LDA algorithm. I parse the transcript data into the word-level tokens, lemmatize each word token, and remove stop words from the tokens. I use *Natural Language Toolkit (NLTK),* a suite of Python libraries and programs for natural language processing for English, to lemmatize each word token and remove stop words. These steps convert each transcript to a list of tokens based on the meaning of words. Then, I create a dictionary of the words in the entire transcript data and convert each transcript to bag-of-words corpus.

With a fixed number of topics (k), I run LDA by using *Gensim*, a library for unsupervised topic modeling. As a result, I get the topic probabilities of each review speech. To determine the optimal *k*, I estimate several LDA models with different values of *k* and pick the optimal value based on the topic coherence score. Topic coherence score measures the degree of semantic similarity among highly influencing words of each topic. The advantage of using coherence scores to optimize k is that it usually yields human interpretable topics. I use *Cv* coherence score, which was developed by Röder and Hinneburg (2015). Röder and Hinneburg (2015) examined five existing coherence scores and two new coherence scores, including *Cv*, and find that *Cv* showed the best performance among those seven coherence scores in terms of the correlations with human ratings.

I pick the optimal k that gives the highest *Cv* before the graph shows a flattened trend. After examining the topic coherence scores of resulted LDA models with different values of *k* (See Appendix Figure D.1), I pick *k = 19*. The LDA model with k = 19 (*Cv* = 0.427) increased *Cv* by 57% compared to the baseline LDA model with *k* = 1 (*Cv* = 0.272).

Appendix Figure D.1: Topic Coherence Scores



As a result of topic modeling, I get the percentage contribution of 19 topics in review videos. The result shows that the majority of reviews does not have one dominant topic that has a contribution value larger than 0.5. The median value of each review's highest topic contribution value is 0.4584. If I select one dominant topic in each review and assumes the topic represents the content of the review, I would not be able to capture large portion of the entire content in the majority of reviews. Thus, I use the contribution values directly, instead of determining one dominant topic in each review. Appendix Table D.1 shows the ten most related words in each topic.

Appendix Table D.1: 19 Topics and Their Top 10 Relevant Keywords

Topics	Top 10 Relevant Keywords
1	watch, heart, smartwatch, fitness, wrist, strap, workout, face, health, tracker
2	sound, speaker, music, little, Bluetooth, headset, audio, really, volume, earbuds
3	tablet, stylus, drawing, writing, note, brush, procreate, pressure, sketch, pencil
4	camera, photo, image, light, video, shot, focus, picture, capture, stabilization
5	laptop, macbook, drive, windows, computer, editing, trackpad, intel, speed, machine
6	monitor, input, remote, hertz, panel, stand, content, movie, ultra, television
7	setting, option, change, application, screen, swipe, select, notification, press, download
8	case, protect, cutout, cover, button, leather, bottom, raise, rubber, material
9	device, phone, camera, battery, game, performance, gaming, display, megapixel, pretty
10	phone, galaxy, samsung, feature, screen, battery, display, fingerprint, camera, still
11	huawei, smartphone, price, sensor, display, design, honor, support, megapixel, monochrome
12	apple, keyboard, pencil, macbook, touch, iphone, generation, accessory, ipad, charge
13	pixel, google, excel, assistant, android, leica, nexus, iphone, supercharge, phone
14	machine, wash, cycle, washer, clothes, freesync, water, microwave, laundry, clean
15	right, ahead, alright, phone, going, charge, unbox, video, charger, battery
16	really, think, little, actually, thing, would, something, going, maybe, people
17	temperature, degree, fridge, celsius, freezer, printer, refrigerator, compressor, centigrade, cooling
18	protector, screen, glass, temper, installation, adhesive, cloth, bubble, clean, microfiber
19	mobile, music, mother, number, friend, coffee, attack, another, could, child

Appendix E: Propensity Score Weighting Validation

To examine the validity of the propensity score weighting method, I evaluate overall balance across the multiple sample groups with different review product types. The distribution of baseline observable reviewer characteristic and video publication time covariates should be independent of review product types conditional on the true propensity score. Therefore, it is important to evaluate the covariate balance to assess whether the propensity score model is adequately specified (Austin 2011).

Following McCaffrey et al. (2014), I examine the population standardized bias (PSB) and population Kolmogorov-Smirnov statistic (PKS) to compare the means and the distributions of the covariates among review product types. For product type j ($j \in$ {wearable device, smartphone, tablet, laptop, other portable device, larger device}) and covariate k (k = 1, … K), population standardized bias is defined:

$$PSB_{jk} = \left|\bar{X}_{jk} - \bar{X}_k\right| / \hat{\sigma}_k$$

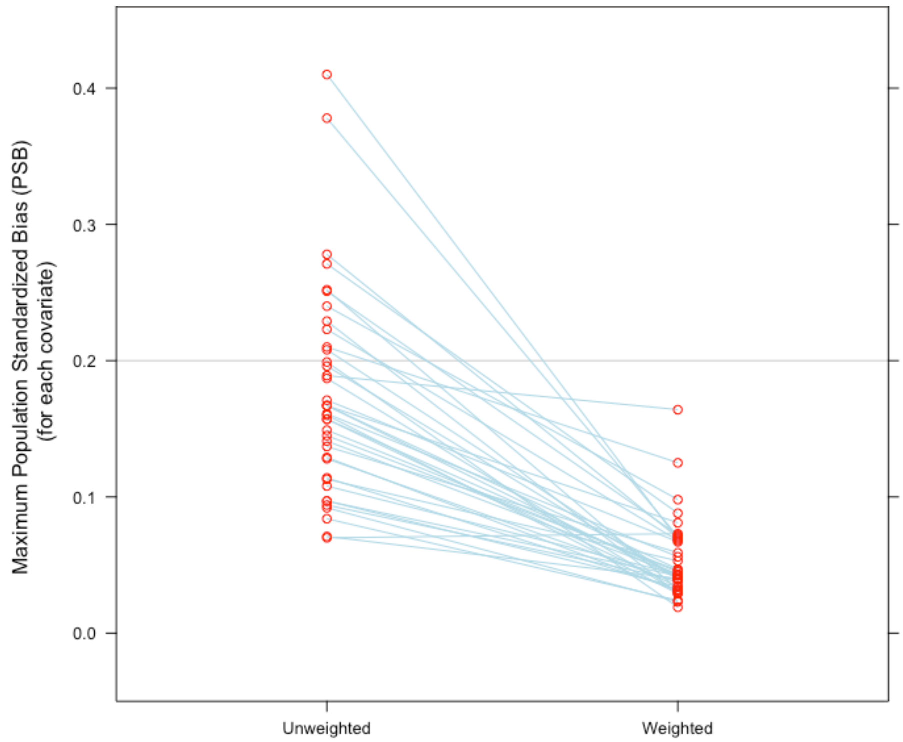
where PSB_{jk} is the population standardized bias (PSB) of covariate k for the product type j reviews, $\bar{X}_{kj}$is the propensity score weighted mean of the covariate j for the product type j reviews, and $\bar{X}_k$ and $\hat{\sigma}_k$ denote the unweighted mean and standard deviation of the covariate k for the pooled reviews across all product types. Similarly, population Kolmogorov-Smirnov (PKS) statics is given by:

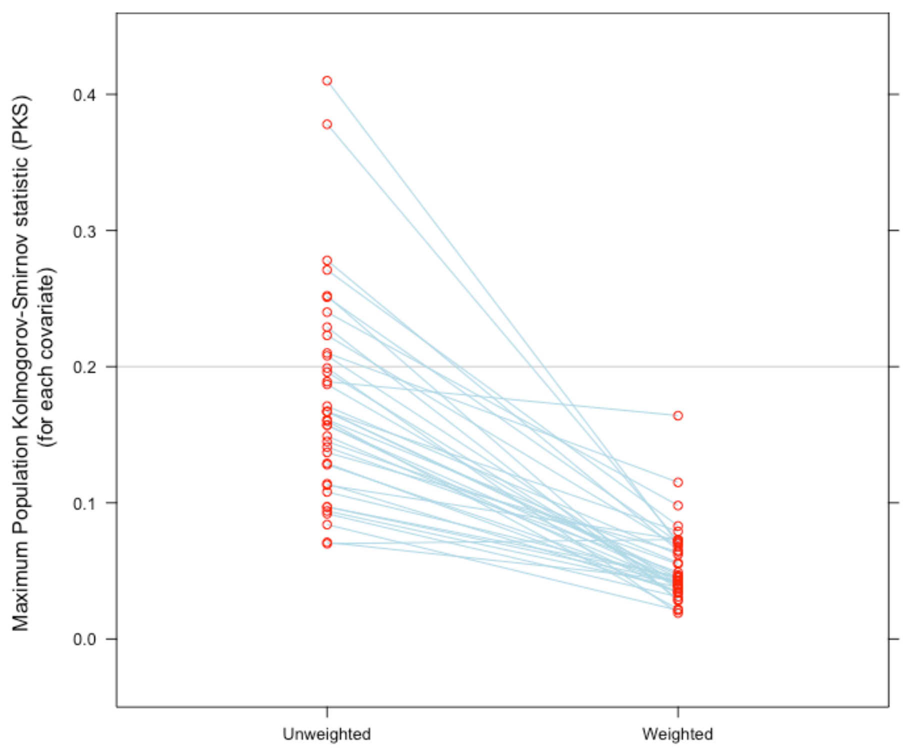
$$PKS_{jk} = sup_x\left|EDF_{jk}(x) - EDF_k(x)\right|$$

where PKS_{jk} is the population Kolmogorov-Smirnov statistic (PKS) of covariate k in product type j reviews, $EDF_k(x) = \sum_{i=1}^{n} I(X_{ik} \leq x)\,/n$ is the unweighted empirical distribution function for the pooled reviews across all product types, $EDF_{jk}(x)$ is the propensity score weighted empirical distribution function for the product type j reviews. PSB_{jk} and PKS_{jk} allow us to evaluate how similar the each product type's weighted sample is to the pooled review sample in terms of covariate means and entire distributions. For both PSB_{jk} and PSB_{jk}, values less than 0.20 are considered to be sufficiently small and an indication of successful reweighted sample (McCaffrey et al. 2014).

Appendix Figure E.1 show the overall summary of balance across review product types, the maximum of the PSB_{jk} and PKS_{jk} among review product types for each covariate k, and compare the balance metrics before and after the propensity score weighting method. I find that the propensity score weighting method improves the covariate balance across different review product types. Also, for all covariates, both balance metrics are smaller than the threshold. Therefore, the balancing hypothesis holds and the propensity score weighting method balances the observed covariates among different review product types.

Appendix Figure E.1: Summary of Balance Metrics





Appendix F: Robustness Checks using Alternative Samples

During the data collection process, I exclude review videos with less than 1000 total views or 10 total like/dislike votes. I conduct a robustness check of the sample filtering criteria by differing the cut values of *Total Views* and *Total Votes* of videos. Appendix Table F.1 shows the estimation results of the the reviewer fixed effect model with different cut values of *Total Views* and *Total Votes*. The coefficients on the explanatory variables are mostly statistically significant and show the consistent results. Therefore, the findings are robust to employing different cut-off values of *Total Views* and *Total Votes*.

Appendix Table F.1: Robustness Check – Estimation Result Employing Different Samples

	$Total\ Views \geq 2{,}000$	$Total\ Votes \geq 20$
Visual Variability	1.330***	1.278***
	(0.350)	(0.324)
Brightness	0.006	0.007*
	(0.004)	(0.004)
Face Presence	0.432*	0.388*
	(0.241)	(0.229)
Speech Rate	0.824***	0.641***
	(0.205)	(0.193)
Voice Pitch	0.031***	0.037***
	(0.008)	(0.007)
Pitch Variability	0.004	0.011
	(0.008)	(0.008)
Positivity	10.729***	10.203***
	(1.796)	(1.702)
Subjectivity	-2.879	-3.416*
	(2.023)	(1.920)
Control Variables	Yes	Yes
N	12,594	13,469

Notes: *p<0.1; **p<0.05; ***p<0.01. This table reports results based on different sample selection criteria. The model additionally controls for word counts, total votes, product category fixed effect, content-level topic control, and monthly time fixed effect. Standard errors are reported in parentheses under coefficients.

Appendix G: Time Discounting Factor in Exponential Smoothing

I use the exponential smoothing method to calibrate the cumulative variables. Exponential smoothing allows a more recent activity to have a stronger impact on individual behavior compared to older activities. For example, the activity that happened Δt ago, its effect is discounted by $e^{-r\Delta t}$, where *r* is the time discounting factor. I conduct a one-dimensional search to find the optimal value of r. In detail, I evaluate the log-likelihood of the model at different values of r ranging from 0 to 1. Appendix table G.1 represents the evaluation result. The result shows the optimal value of *r* is 0.3 as it shows the highest value of log-likelihood. Therefore, I select 0.3 as the optimal time discounting factor.

Appendix Table G.1: Evaluation of different discount factor

r	Log-likelihood
0.1	-10406.435
0.2	**-9955.232**
0.3	-9974.165
0.4	-9986.731
0.5	-10023.172
0.6	-10029.394
0.7	-10069.183
0.8	-10124.063
0.9	-10181.012