Uncovering Causal Effects of Online Short Videos on Consumer Behaviors

Ziqi Tan[1], Shengyu Zhang[1], Nuanxin Hong[1], Kun Kuang[1†], Yifan Yu[2†], Jin Yu[3], Zhou Zhao[1], Hongxia Yang[3], Shiyuan Pan[4], Jingren Zhou[3], Fei Wu[1]
[1] College of Computer Science and Technology, Zhejiang University
[2] School of Economics, Zhejiang University, [3] Alibaba Group
[4] Center for Research on Private Economy, Zhejiang University
{tanziqi,sy_zhang,nuanxin,kunkuang,yifanyu,zhaozhou,shiyuanpan,wufei}@zju.edu.cn
{kola.yu,yang.yhx,jingren.zhou}@alibaba-inc.com

ABSTRACT

In recent years, online short videos have become more popular, especially as an online advertising intermediary. To better understand their effects as advertisements, it is essential to analyze the causal relations of online short videos on consumer behaviors. Our study is based on fine-grained consumer behavior data from a world-leading e-commerce platform, *i.e.*, Taobao.com[1]. We first decompose the total causal effects into informative effects and persuasive effects following a common practice in the economic literature. Moreover, we extract the subjectivity scores of short videos through a dictionary-based subjectivity analysis model and evaluate the correlation between the subjectivity scores and each causal effect. The findings of this paper are as follows: First, both causal effects (*i.e.*, informative and persuasive effects) are significant. Second, these effects have a strong correlation with the short videos' subjectivity scores. Third, the signs of these correlations vary with the prices of the products. Our results not only shed light on the research of how short videos exert influence on online consumers, but also give sellers advice on better video design and recommendation.

CCS CONCEPTS

• **Applied computing** → **Online shopping**; **Marketing**; *Economics*.

KEYWORDS

Short video, Advertising effects, Video subjectivity, Doubly robust

ACM Reference Format:
Ziqi Tan, Shengyu Zhang, Nuanxin Hong, Kun Kuang, Yifan Yu, Zhou Zhao, Jin Yu, Hongxia Yang, Shiyuan Pan, Jingren Zhou, Fei Wu. 2022. Uncovering Causal Effects of Online Short Videos on Consumer Behaviors. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining (WSDM '22), February 21–25, 2022, Tempe, AZ, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3488560.3498513

[1]Dataset available at https://github.com/tanziqi1234/AdvertisingEffects

WSDM '22, February 21–25, 2022, Tempe, AZ, USA.

ACM ISBN 978-1-4503-9132-0/22/02…$15.00
https://doi.org/10.1145/3488560.3498513

1 INTRODUCTION

In China, the online short videos industry is growing rapidly for the past decade. By June 2020, the number of online video viewers has reached 818 million [12]. With the market expansion of smartphones, people tend to access the internet by their mobile devices and watch short videos during their fragmented leisure time. Short videos are much more competitive to grasp users' attention on mobile devices than regular length videos [32]. Because of its popularity, the short videos industry has become an exceedingly promising area for many investors and large Internet corporations, (*e.g.*, Alibaba and ByteDance). For example, many sellers on Taobao.com, the largest B2C platform in China, produce short videos to promote their products. They expect the viewers to be more interested in the products demonstrated in the short video and may actually increase the sales. In many people's minds, short videos are, by their nature, advertisements. By 2019, the market size of short videos in China has reached 130.24 billion yuan, of which advertisements contribute more than half (79.95 billion yuan [19]). How well the money is spent and how to use the money more effectively remains unclear. However, there are only a handful of researches focusing on to what extent the short videos influence the consumers and on the analysis of causal effects of online short videos on consumer behaviors. These research questions are also essential for sellers to produce more attractive short videos and for online platforms to design more effective recommendation algorithm.

To study the advertising effects, Chamberlin [8] proposed a framework to distinguish the informative effects and persuasive effects. This framework is still the leading paradigm in economics and marketing research today. Conceptually, advertising affects demand because (i) it conveys information to consumers about the existence of sellers and the prices and qualities of products in the marketplace, and (ii) it alters consumers' wants or tastes, which could create a spurious product and brand preference [5]. Following his paper, a great number of researches in the field of economics and marketing continue to dedicate to further research on the same topic [14, 20, 31, 33]. Recently, Ackerberg [1, 2] found that advertising's primary effects are informing consumers about product information in the yogurt market, and Ching and Ishihara

[†]Corresponding Authors.
[*]Work was performed when Z. Tan and S. Zhang were interns at Alibaba Group.

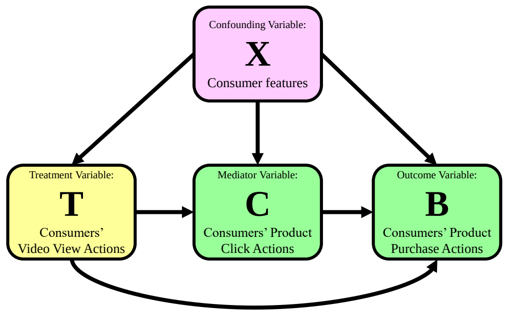


Figure 1: The intuitive illustration of our proposed causal framework w.r.t. a scenario of consumers' behaviors, *i.e.*, video view actions T as a treatment, product click actions C and product purchase actions B as outcomes, and consumer features X as a confounder. $T \to C$ stands for short videos' informative effects. $T \to B|C = 1$ stands for short videos' persuasive effects. $T \to B$ stands for the total effects.

[10] found that persuasive effect only plays a minor role in the pharmaceutical industry. Consistent with these literature, advertising can be interpreted as taking informative effects if they affect awareness and consideration (not distinguished in many research [3, 13, 40]), or taking persuasive effects if they affect choice conditional on awareness. However, current research is still limited in that, (i) there is hardly any study that focuses on the advertising effects of online short videos which might have different results; (ii) only a few papers have been dedicated to the detailed analysis of the potential source of these effects. Our paper is trying to fill these research gaps.

Our paper is based on a large-scale fine-grained user behavior dataset from Taobao.com, which contains users' view, click and purchase history as well as the consumer, video, and product characteristics. The watch behaviors of short videos affect the click rates of the products as well as the purchase decisions. In the whole process, confounders (*e.g.*, consumer and product characteristics) exist and affect the consumer's decisions as illustrated in Figure 1. We measure the causal effects of watching short videos on product click rates ($T \to C$) as the informative effects. Intuitively, consumers with more information tend to take the product into consideration with a higher probability. After forming his consideration set, the probability that a certain product is purchased is directly related to the utility it generates to the consumer. Thus, we take the causal effects of watching short videos on product purchase rates conditional on consumers who have clicked the products ($T \to B|C = 1$) as persuasive effects. Besides, we also estimate the total effects by estimating the effects of watching short videos on the product purchase rates unconditionally. To estimate these effects precisely, we use a well-established causal inference tool, *i.e.*, doubly robust estimator [6].

Furthermore, we want to trace the potential source of informative effects and persuasive effects. One of our conjectures is that they might arise from subjective information (*e.g.*emotional expressions conveying personal feelings) and objective information (*e.g.*descriptions of the product's properties or parameters) respectively. Though there are also many other video analysis tasks as video captioning [42, 43], object detection [15, 54], video titling [49, 50], video pre-training [48], they seem to have nothing to do with videos' advertising effects. To test such conjecture, we extract the short videos' subjectivity scores through speech recognition and dictionary-based textual subjectivity extraction. Finally, we analyze the relationship between short videos' subjectivity scores and the advertising effects.

According to our experiments, we find that (i) all the advertising effects are significant and (ii) the relationship between short videos' subjectivity scores and the advertising effects varies by product prices. Based on our findings, we summarize some suggestions for (i) sellers to produce short videos to promote their products and (ii) online platforms to improve their recommendation systems which are of great interest in both academic and industrial community [51, 53]. Details could be found in Section 5.

To accommodate our research, we have also collected billions of fine-grained consumer behavior records and we will make our data publicly available for further research after desensitization. To the best of our knowledge, we are the initial to estimate the advertising effects of online short videos and to analyze the relationship between these advertising effects and automatically extract short videos' subjectivity score and find some implications. In summary, the main contributions for this work are as follows:

- We propose a procedure to estimate online short videos' informative effects, persuasive effects, and total effects and correlate these effects with automatically extracted short videos' content, *i.e.*, subjectivity.
- We find online short videos' informative, persuasive and total effects all significant and summarize the properties of short videos with different levels of advertising effects.
- We find these advertising effects have a strong correlation with the short videos' subjectivity scores while the correlation signs vary with the prices of the products.
- We provide some practical advice for sellers to produce better online short videos and for online platforms to design better short video recommendation algorithms.

2 RELATED WORKS

Our paper is an interdisciplinary study involving both computer science and economics and is related to several streams of research. We conclude them into (i) online short video's informative effect and persuasive effect framework; (ii) a well-designed causal inference tool doubly robust and (iii) video subjectivity analysis methods.

Videos' Advertising Effects. Since Chamberlin [8] first proposed a framework of informative effects and persuasive effects of advertising in his classical book, many researchers have made empirical researches to further study these two effects of advertising in various markets [1, 10, 28, 30, 37, 41]. Here informative effects refer to informing consumers about the products while persuasive effects refer to altering consumers' preferences. Generally, consumers take three stages, awareness, consideration and, choice, before they reach final purchase decisions [44]. Informative effects communicate both information that operate at the awareness and consideration stage [3, 13, 40] while persuasive effects refer to the influence that operates at the choice stage [16, 29].

Although researchers have made great efforts on various sorts of advertisements, the advertising effects of online short videos as a new media of advertisement are still rarely discovered. Some researches targeting dynamic visual media focused on television advertisements and conclude that informative effects are significant [28, 41]. Our paper differs in the following three aspects: (i) We focus on online short videos which are distinct from television advertisements in their more close-to-life content, heterogeneous user groups, and costs of action after receiving the advertisements. (ii) We focus on more fundamental video property (*i.e.*, subjectivity) which could be extracted automatically from existing algorithms instead of laborsome human annotation. (iii) We utilize a different causal inference tool doubly robust to adapt to the nature of our fine-grained data. (iv) We collect millions of fine-grained consumer behavior data from one of the largest e-commerce platform, Taobao.com, which could reveal much more information.

Causal Inference. Causal inference is an effective statistical tool to uncover causal effects of treatments from observation data[24]. Existing tools can be classified into two prominent frameworks, structural causal models (SCM) [34] and potential outcome [18, 36]. Despite the difference in their framework, the objectives are identical, aiming at removing spurious correlation (*e.g.*, confounding bias) from observation data. Graphs are powerful tools in many tasks [47], and causal graphs are often used to describe the assumptions of data generation processes.

To eliminate bias in the treatment group and control group to simulate the environment of randomized experiments, propensity score [35] computes the probability that an individual takes treatment [11]. Propensity score matching (PSM), inverse of propensity weighting (IPW), and doubly robust (DR) [4, 6, 22, 23] are all applicable estimators to utilize propensity score to estimate causal effects. Especially, doubly robust combines maximum likelihood regression for outcome and IPW approach so that it has more stable performance and is more robust to missing data. Recently, doubly robust has been used to estimate online advertisements [9, 26]. In this paper, we also select doubly robust as the causal inference tool due to its robustness nature and accordance with our data.

Videos' Subjectivity Analysis. To extract videos' subjectivity, sentiment classification is a valid approach in that both strong positive and strong negative could be deemed as subjective. In the field of video sentiment classification [7, 38, 45], researchers usually leverage extracted image and audio features to make sentiment classification. For example, Hu et al. [17] fused the DNN and LSTM extracted facial and audio features and used SVM for classification, and Zhao et al. [52] used a CNN-based approach to identify videos' emotions in an end-to-end manner. However, these machine learning-based methods rely on appropriate datasets and most of which are in English. The only Chinese video sentiment classification datasets are CH-SIMS proposed by Yu et al. [46] and CHEAVD proposed by Li et al. [27].

Another videos' subjectivity extraction approach is text subjectivity classification. However, there are much fewer studies focusing on the subjective and objective classification of texts. Kwon et al. [25] have done some work on the identification and classification of subjective claims. Rather than classifying documents or sentences into binary polarities as in much previous work, they identify the main claim or assertion of the writer and classify it into the predefined classes of opinion (attitude) over the topic.

Since there are great gaps between CH-SIMS, CHEAVD, and our dataset, models trained on these datasets show poor performance. To ensure the robustness of the subjectivity extraction, we use a dictionary-based text subjectivity extraction method for recognizing speeches, which shows excellent performance.

3 ESTIMATING ADVERTISING EFFECTS OF SHORT VIDEOS

3.1 Overview

Given the data of N_v short videos with associated products and N_u consumers, there are N_e^v consumers exposed to short video v. We estimate the advertising effects for each short video separately. For notation conciseness, we drop the index v of short video v in our following discussion as we compute the advertising effects independently in the same way for all short videos. Considering the great endogenous problem between consumer groups exposed and unexposed, we only consider the N_e exposed consumers.

After data pre-processing (in section 5), for consumer u, we have the following data: (i) consumer's information as a feature vector $\mathbf{X}_u$ (*e.g.*, consumer's click status for the product before the short video's exposure); (ii) consumer's short video exposure status E_u in the homepage, with $E_u = 1$ if exposed and $E_u = 0$ if unexposed; (iii) consumer's behaviors as demonstrated in Figure 1, including (iii.a) consumer's video watching status T_u, with $T_u = 1$ only if the consumer is exposed to the short video ($E_u = 1$, a consumer can watch a short video only if he is exposed to it) and watches it; (iii.b) consumers' product click status C_u, with $C_u = 1$ only if the consumer is exposed to the short video ($E_u = 1$, we don't consider consumers who are not exposed) and clicks the product within five days after exposure; (iii.c) consumer's product purchase status B_u, with $B_u = 1$ only if consumer u clicks the product ($C_u = 1$, a consumer must click a product before they purchase it) and purchases the product within five days after exposure.

The rest of this section is organized as follows: Section 3.2 introduces the measurements of the advertising effects and Section 3.3 uses doubly robust to estimate these causal effects.

3.2 Advertising Effects

To study the advertising effects of short videos, this paper borrows ideas from the well-known book by Chamberlin [8], in which he decomposed the total effects of advertising on consumer behaviors into informative effects and persuasive effects. He argued that advertising affects demand because (i) it conveys information to consumers, about the existence of sellers and the prices and qualities of products in the marketplace, and (ii) it alters consumers' wants or tastes, which could create a spurious product and brand preference [5]. Thus, informative effects work better for inexperienced consumers. According to [16], there are three stages in the consumer purchase decision, awareness, consideration, and choice. Informative effects mainly affect the awareness and consideration stage and increase the probability certain product enters the consumers' consideration sets [3, 13, 40]. While persuasive effects take place in the choice stage and alter consumers' tastes.

In our scenario, a short video often contains a product's information about price, attributes, usage, experience, and so on. As long as a consumer watches the video, this information will raise his awareness of the existence of such a product. Thus, it will increase its chance to end up in the consumer's consideration sets which we characterized by the action of clicking the link to the product purchase page. Since informative effects are measured by how short videos attract consumers' awareness and consideration, a natural way to measure informative effects is to estimate the ATE (average treatment effects on population) of watching a short video on the product's click rate (*i.e.*, $T \longrightarrow C$ in Figure 1):

$$\mathbf{ATE}_{info} = \frac{1}{N_c}\sum_{u=1}^{N_e} T_u C_u - \frac{1}{N_{nc}}\sum_{u=1}^{N_e}(1-T_u)C_u, \tag{1}$$

where $N_c = \sum_{u=1}^{N_e} T_u$ is the number of the consumers who clicked the exposed short video, $N_{nc} = \sum_{u=1}^{N_e}(1-T_u)$ is the number of consumers who didn't click the exposed short video. It can be intuitively interpreted that consumers with more product awareness tend to put them into their consideration sets.

The persuasive effects can create spurious love for the advertised product and, thus, increase its chance of being purchased given the consumers actually consider it. In our settings, the viewer of a short video may gain extra utility from purchasing the product, which leads to an increase in the purchase rate. Therefore, we can quantify the persuasive effects by the ATE of watching a short video on the product purchase rate given that the consumers have clicked the link to the product purchase page ($T \longrightarrow B|C=1$ in Figure 1):

$$\mathbf{ATE}_{pers} = \frac{1}{N_{cb}}\sum_{u=1}^{N_e} T_u C_u B_u - \frac{1}{N_{ncb}}\sum_{u=1}^{N_e}(1-T_u)C_u B_u, \tag{2}$$

where $N_{cb} = \sum_{u=1}^{N_e} T_u C_u$ is the number of the consumers who clicked the exposed short video and then click the product, $N_{ncb} = \sum_{u=1}^{N_e}(1-T_u)C_u$ is the number of consumers who didn't click the exposed short video but clicked the product.

Besides informative and persuasive effects, it is natural to wonder how short videos affect the final sales volume of the associated products. To this end, we also estimate the total effect of watching short videos, deemed as the composite effects of informative effects and persuasive effects. Sellers and online platforms care more about the question whether produce or recommend short videos could actually make a profit for the sellers. We use the ATE of watching a short video on the product purchase rate increase in all consumers who are exposed to the short video as a metric ($T \longrightarrow B$ in Figure 1) :

$$\mathbf{ATE}_{total} = \frac{1}{N_c}\sum_{u=1}^{N_e} T_u B_u - \frac{1}{n_{nc}}\sum_{u=1}^{N_e}(1-T_u)B_u, \tag{3}$$

3.3 Doubly Robust Estimator

However, these three ATEs above reveal simply statistical effects instead of causal effects. Due to the existence of confounding variables, whether a consumer clicks a piece of a short video also depends on the consumer himself, that we need to control for those potential confounding variables. Confounding variables refer to variables that affect both treatment and effects and overlooking those confounders will lead to spurious correlations due to the backdoor path [34]. In our task, the feature of consumers act as confounding variables, as shown in the casual graph in Figure 1. In this paper, we make use of the doubly robust estimator [6] as an efficient and robust causal inference tool to estimate the causal effects. Doubly robust, as its name implies, the estimation is unbiased as long as the result of either approach is correct in that it combines inverse propensity weighting (IPW) approach for outcome and maximum likelihood regression.

To do inverse propensity weighting, we need to compute the propensity score first. Given a set of confounding variables (*i.e.*, consumer feature, $\mathbf{X}_u$ for consumer u) that affect both treatment (*i.e.*, watch status T_u) and effect Y_u (*i.e.*, click status C_u or purchase status B_u), we use linear regression to estimate the probability an individual receives treatment as propensity score $e(\mathbf{X}_u)$:

$$e(\mathbf{X}_u) = P(T_u = 1|\mathbf{X}_u, E_u = 1). \tag{4}$$

Since the propensity score represents treatment probability, we can use $e(X_u)$ as the weight of a sample to do IPW for (1)-(3). Take (1) as an example and we can obtain:

$$\mathbf{ATE}_{info}^{IPW} = \frac{1}{N_e}\sum_u \frac{T_u C_u}{e(\mathbf{X}_u)} - \frac{1}{N_e}\sum_u \frac{(1-T_u)C_u}{1-e(\mathbf{X}_u)}. \tag{5}$$

Though ATE_{info}^{IPW} is an unbiased estimator of ATE_{info}, the result highly depends on the accuracy of $e(X_u)$. So doubly robust also combines maximum likelihood regression to make the causal effect estimation more robust. We use confounding variables $\mathbf{X}_u$ to regress effect Y_u (*i.e.*, click status C_u or purchase status B_u) for both treated and not treated:

$$m_0(\mathbf{X}_u) = \mathbf{E}(Y_u|T_u = 0, \mathbf{X}_u, E_u = 1). \tag{6}$$

$$m_1(\mathbf{X}_u) = \mathbf{E}(Y_u|T_u = 1, \mathbf{X}_u, E_u = 1). \tag{7}$$

Finally, we combine the aforementioned two parts of estimation and we can get double robustness for the causal effects of informative:

$$\begin{aligned}\mathbf{ATE}_{info}^{DR} = &\frac{1}{N_e}\sum_u\Big(\frac{T_u C_u}{e(\mathbf{X}_u)} - \frac{T_u - e(\mathbf{X}_u)}{e(\mathbf{X}_u)}m_1(\mathbf{X}_u)\Big) - \\ &\frac{1}{N_e}\sum_u\Big(\frac{(1-T_u)C_u}{1-e(\mathbf{X}_u)} + \frac{T_u - e(\mathbf{X}_u)}{1-e(\mathbf{X}_u)}m_0(\mathbf{X}_u)\Big)).\end{aligned} \tag{8}$$

Similarly, we can get the doubly robust causal effects of persuasive and total. However, due to the space limitation, we don't write them here.

4 SUBJECTIVITY ANALYSIS

4.1 Overview

After obtaining the advertising effects, a direct question is: Where do they come from? We want to trace the potential source of these effects. One of our conjectures is that the decomposed two effects might arise from subjective information and objective information respectively in that, (i) subjective information often attempts to subjectively induce consumers which might have some certain relationships with persuasive effects and (ii) subjective information also attempts to tell consumers about the information which could

Algorithm 1 (Short video's causal pattern discovery)

Input: Consumers' short video watch status T_u, product click status C_u, product purchase status B_u and confounders X_u for each short video.
Ouput: Short videos' informative effects $\mathbf{ATE}^{DR}_{info}$, persuasive effects $\mathbf{ATE}^{DR}_{pers}$, total effects $\mathbf{ATE}^{DR}_{total}$, correlation between subjectivity scores and three advertising effects $Corr_{info}, Corr_{pers}, Corr_{total}$
Step 1: find a propensity score $e(X_u)$ for each consumer u: do maximum likelihood regression to estimate effects
Step 3: compute ATE^{DR} for each effect
Step 4: extract subjectivity score S_{text} for each short video according to subjective words dictionary
Step 5: compute Pearson, Spearman and Kendall correlation coefficients $Corr_{info}, Corr_{pers}, Corr_{total}$ between the subjectivity scores and three advertising effects

not be seen directly and differs from person to person, which might have something to do with the informative effects. However, the influence direction and strength still remain undiscovered. In this section, we attempt to uncover the patterns behind short videos' subjectivity scores and the advertising effects through short videos' subjectivity extraction and correlation analysis.

4.2 Short Videos' Subjectivity Extraction

Since the speech content is the most important part of the short videos and there is no appropriate training dataset to determine video subjectivity, we choose to extract short videos' subjectivity scores from the speech texts by a subjective word dictionary-based approach. The extraction procedure can be divided into two steps: (i) speech recognition, (ii) textual subjectivity extraction.

Speech recognition. First, we extract the audio of the short videos. Then we use an existing Chinese speech recognition model to obtain the speech texts of the speakers.

Textual subjectivity analysis. The dictionary-based subjectivity extraction approach includes two phases. First, we use subjective and objective word dictionary where different categories of words (*e.g.*, emotional words, negative words, and interjection) have different weights (in the range of [0, 1]). Higher weights represent stronger subjectivity. Then we tokenize each sentence in the speeches. Assuming that the subjectivity scores satisfy the linear superposition principle, for each sentence in a speech, we compute the subjectivity score S_{sent} as the weighted average of the words' subjectivity scores S_{word_i} (**i**th word in a sentence) score according to the dictionary:

$$S_{sent} = \frac{1}{L_{sent}} \sum_{i=1}^{L_{sent}} S_{word_i}, \tag{9}$$

where L_{sent} refers to the length of the sentence. Then, for each speech text, we compute the speech subjectivity score S_{text} as the weighted average of all the sentences' subjectivity scores:

$$S_{text} = \frac{1}{L_{text}} \sum_{i=1}^{L_{text}} S^{i}_{sent}, \tag{10}$$

where L_{text} refers to the length of the speech text. All of the subjectivity scores take values in a range between [0, 1], and the closer the score is to 1 the more subjective the text is, and vice versa.

4.3 Correlation Analysis

With the subjectivity score, *i.e.*, S_{text} and three advertising effects, *i.e.*, $\mathbf{ATE}^{DR}_{info}$, $\mathbf{ATE}^{DR}_{pers}$ and $\mathbf{ATE}^{DR}_{total}$ for each short video, we are able to analyze the correlation between them. To ensure the robustness and capture both linear and non-linear correlations, we compute not only Pearson correlation coefficient but also Spearman rank correlation coefficient [39] and Kendall rank correlation coefficient [21]. However, relationship discovery and mining are intrinsically challenging, in experiments, we find a non-negligible confounder product price in this problem. For simplicity, we stratify the short video data into several price ranges according to the price distribution so that the correlation significance starts to emerge.

Our overall causal effects analysis and subjectivity relationship mining algorithm are summarized in Algorithm 1.

5 EXPERIMENTS

5.1 Data Preparation

Since there is little research on the advertising effects of online short videos (to the best of our knowledge), no public dataset is available. To exactly evaluate both informative effects and persuasive effects, we collected a large-scale fine-grained consumer behavior dataset from a world-leading e-commerce scenario, *i.e.*Taobao.com. With detailed information, we are in a privileged position to study these effects with few assumptions. Details of our dataset can be found in Appendix.

5.2 Results of Videos' Advertising Effects

We now discuss the causal effects of watching short videos on consumers' behaviors, including both clicking and purchasing the associated products. Specifically, we are interested in whether the following effects exist: (i) informative effects, *i.e.*, $T \to C$; (ii) persuasive effects, *i.e.*, $T \to B|C = 1$; (iii) total effects, *i.e.*, $T \to B$. We first quantitatively discuss the significance of the above advertising effects (with numerical assessments in Table 1). Following [16], we use * to indicate significance at the 0.10 level (p<0.10), ** the 0.05 level (p<0.05), and *** the 0.01 level (p<0.01). We then give illustrations of the significant results from the perspective of sellers and online platforms.

We also provide evidence for these illustrations by analyzing the main content of short videos. We conduct a manual analysis and find several types of short videos with salient features (Table 2):

- KOLs' videos. To better promote their products, many sellers invite Internet celebrities (KOLs, *i.e.*, Key Opinion Leaders) to introduce their products. In our experiments, we include Weiya and Jiaqi Li since they are the top two Internet celebrities on Taobao.com.
- Tutorial videos. To better arouse consumers' attention, sellers make tutorials of some skills, providing HowTo instructions (*e.g.*, sticking eyelashes). In these videos, sellers make advertisements by using well-selected products to make tutorials.

Table 1: Average advertising effects of watching short videos on different consumer behaviors.

Model	Informative	Persuasive	Total
Average Effects	0.1648***	0.0641***	0.0052***

- Story videos. To get higher video click rates, many sellers also shoot small interesting stories and highlight their products that appear in the videos as advertisements.

Eventually, according to the significance of the advertising effects of different types of short videos, we provide some practical advice for online platforms and sellers based on our findings.

5.2.1 Informative Effects ($T \to C$). We observe significant average informative effects (0.1648 with p<0.01) of watching a short video on the short video's associated product click rate increase. These short videos generally enlarge consumers' product awareness which confirms the role of online short video in informing consumers about the existence and availability of the associated products. Intuitively, we attribute this phenomenon to the fact that short videos have a stronger visual impact than traditional information media (*e.g.*, text and images). Short videos, consisting of multi-modal contents, including audio, images, motions, and titles, can better serve as an informative tool and enlarge consumers' awareness set.

We can find evidence for the above analysis from Table 2. Videos from KOLs bring a more significant click-through rate boost for the recommended products by providing product-oriented informative descriptions. However, tutorial and story videos generally provide less comprehensive product-oriented information and thus bringing significantly less product click rate increase than other kinds of videos, though with high video click rates.

Advice: Therefore, for the online platforms, it is suggested to recommend more product-oriented informative videos for consumers that are less likely to be aware of the mentioned products, which should result in higher product click rates after consumers watch them. For sellers, it is suggested that they could consider adding more informative content describing the new product and less product-irrelevant content, such as skill tutorials or stories, when creating short videos about new products that people barely know about. For large sellers, inviting KOLs is definitely a practical way to make people consider more about their products.

5.2.2 Persuasive Effects ($T \to B|C = 1$). Among consumers who had clicked on the product, watching short videos has a significant mean persuasive effect (0.0641 with p<0.01) on the increase of the associated product purchase rate. It indicates that short videos can function as persuasive tools that alter consumers' original preferences by generating utility for them. In addition to providing comprehensive information, short videos can also show novel features of the tools or novel makeup skills that exceed consumers' expectations. This can make consumers get more delighted and psychologically increase their favor of the products.

However, results in Table 2 demonstrate that tutorial videos are significantly more persuasive than other short videos. Neither KOLs' nor storytelling short videos show a significant effect. It is also natural since tutorials may give consumers new skills that will give consumers a favorable feeling and alter their preferences. The insignificance also signals that it is difficult to alter consumers' preferences by simple strategies in that consumers differ in their preferences and few strategies can influence all of them.

Table 2: The advertising effects of different kinds of short videos.

	KOLs	Tutorial	Story
Informative	0.0720***	-0.1066***	-0.1440***
Persuasive	0.0042	0.0449**	0.0285
Total	0.0063***	0.0025	-0.0045

Advice: Despite the significance of persuasive effects, there reveal only a few patterns. This might imply the difficulty to design a certain preference-altering strategy. The only suggestion we can offer might be producing and recommending tutorial short videos for the usage of the products.

5.2.3 Total Effects ($T \to B$). We take a further step on analyzing the total effects of watching short videos on the associated product purchase rate increase. We also observe a significant average causal effect (0.0052 with p<0.01). This result verifies that short videos as a form of advertisement can improve the purchase rates of the associated products which might be a piece of good news for sellers.

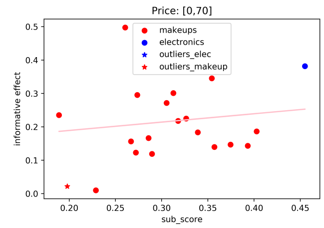
Advice: Results in Table 2 demonstrate that KOLs have the ability to improve product sales volume. As a result, we can suggest the online platforms give the KOLs' videos more exposure, and the sellers invite KOLs to promote their products.

5.2.4 Summary. Overall, we find that: (i) online short videos have significant advertising effects and can improve the click rates and purchase rates of products; (ii) short videos with more product information have more informative effects and would bring more clicks; (iii) persuasive effect reveals few patterns that only tutorial videos are significant; (iv) KOLs' videos have both more informative effects and total effects. Therefore, we suggest sellers and online platforms produce and recommend (i) short videos with more product-oriented information, (ii) tutorial videos of sellers' products, (iii) KOLs' short videos.

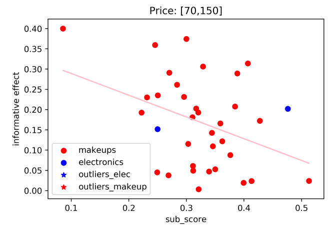
5.3 Correlation between Advertising Effects and Short Videos' Subjectivity

Furthermore, we aim to obtain a more fine-grained understanding of how short videos affect consumers' behaviors and identify some key factors. We examine whether the subjectivity scores of short videos are correlated with the above three effects.

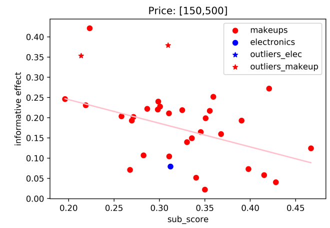
Figure 2m-2o and the first row in Table 3 lists the correlation tests between short videos' subjectivity scores and three advertising effects. In a nutshell, we can hardly identify any correlation between them on a global scale. However, due to the existence of potential confounders, we still cannot draw a conclusion here. After heavy investigation, we finally find an important confounder, *i.e.*, product price, which might be the common cause of the above-examined factors. To deconfound the influence of product price, we build different price ranges and examine the correlation respectively.



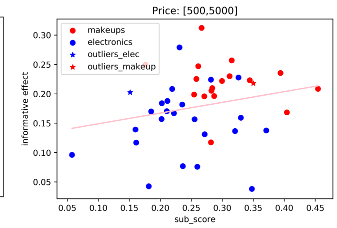
(a) Informative, Price 0-70



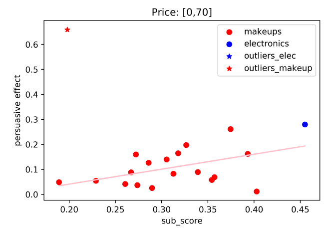
(b) Informative, Price 0-70



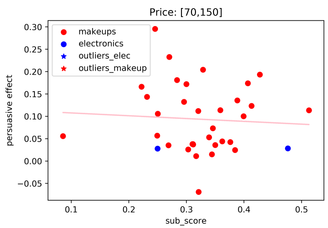
(c) Informative, Price 0-70



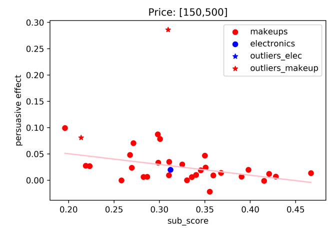
(d) Informative, Price 0-70



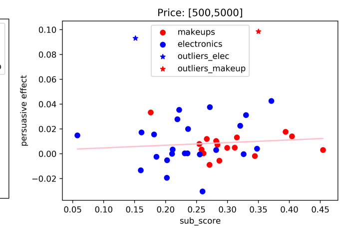
(e) Informative, Price 0-70



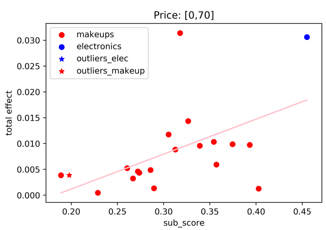
(f) Informative, Price 0-70



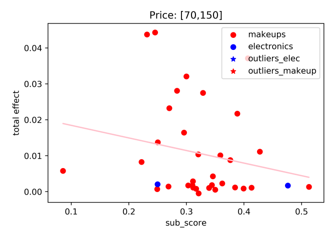
(g) Informative, Price 0-70



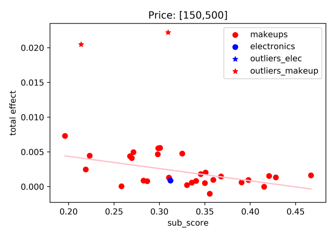
(h) Informative, Price 0-70



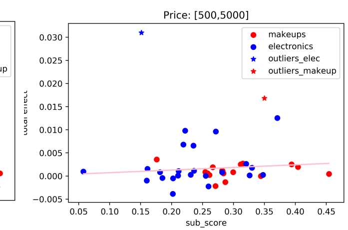
(i) Informative, Price 0-70



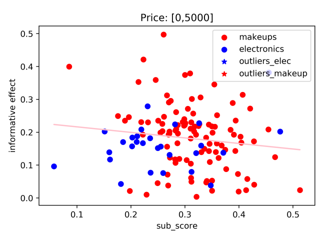
(j) Informative, Price 0-70



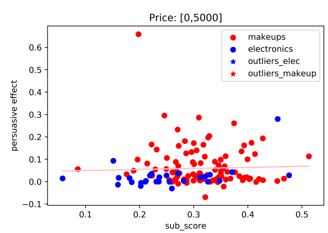
(k) Informative, Price 0-70



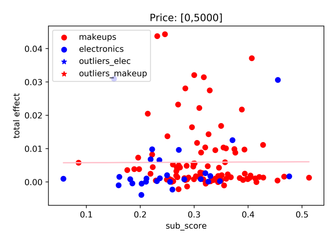
(l) Informative, Price 0-70



(m) Informative, Price 0-70



(n) Informative, Price 0-70



(o) Informative, Price 0-70

Figure 2: The correlation tests between short videos' subjectivity scores and effects. Figure 2m-2o show the global correlation. Figure 2a-2l individually discuss the correlation at different price ranges.

Considering the distribution of prices in the experiment, we choose to make a discussion in price ranges including 0-70 (in Chinese Yuan), 70-150, 150-500, and 500-5000 respectively, and examine the correlations separately. Each subfigure in Figure 2 plots sampled short videos as dots, of which the X axis denotes their subjectivity scores and Y axis denotes their corresponding advertising effects. The color of each dot indicates the category of product associated with the corresponding short video. The star-shaped dots indicate outliers which we dropped in correlation computing. For robustness, we test three kinds of correlation coefficients, *i.e.*, Pearson, Spearman, and Kendall, as listed in Table 3.

5.3.1 Price 0-70. At this price range, we observe a significant positive correlation between the subjective level of short videos and their advertising effects except for informative effects, *i.e.*, persuasive effects 0.4966 with $p<0.05$ and total effects 0.5047 with $p<0.05$. Consumers' purchase behaviors are more likely to be affected by subjective information (*e.g.*, personalized judgments or emotional induction) in online short videos when the promoted products are cheap enough. The significance of the positive correlation between persuasive effects and subjectivity indicates that these videos might create a spurious preference for the products. Facing cheap products, people generally have more tolerance and less psychological burden and are more likely to make impulse purchases. Informative effects' insignificance may also be due to the low psychological burden that consumers don't care about whether the information comes from the subjective part or objective part of a short video.

Table 3: The correlation tests between short videos' subjectivity scores and advertising effects concerning different types of correlation coefficients. We further individually discuss the correlation at different price ranges.

	Informative			Persuasive			Total		
	Pearson	Spearman	Kendall	Pearson	Spearman	Kendall	Pearson	Spearman	Kendall
All	-0.1381	-0.1364	-0.0931	0.0409	0.1046	0.0733	0.0055	0.0432	0.0299
Price 0-70	0.1443	0.0807	0.0643	0.4966**	0.3892*	0.2866*	0.5047**	0.5386**	0.3801**
Price 70-150	-0.3818**	-0.3193*	-0.2101*	-0.0649	-0.0938	-0.0555	-0.2097	-0.2591	-0.1630
Price 150-500	-0.4368**	-0.3843**	-0.2774**	-0.4740***	-0.4246**	-0.2903***	-0.5406***	-0.4089**	-0.2559**
Price 500-5000	0.2240	0.2283	0.1821*	0.1052	0.1385	0.1128	0.1309	0.1696	0.1154

Advice: We suggest that online platforms could consider recommending subjective short videos associated with cheap products to consumers. For sellers, they may consider emotionally promoting their cheap products by giving a high evaluation.

5.3.2 Price 70-150 and 150-500. When concerning products with prices at these ranges, consumers become less affected by subjective content in short videos. Especially in price range 150-500, we can observe significant negative correlations between the subjective level of short videos and all three kinds of advertising effects, *i.e.*, informative effects -0.4368 with $p<0.05$, persuasive effects -0.4740 with $p<0.01$, and total effects -0.5406 with $p<0.01$. It indicates that subjective content does not have the effect of increasing sale volume, changing consumers' consideration set, or altering consumers' preferences anymore. When it comes to relatively upscale products, with limited incomes, consumers generally become more cautious and less tolerant when they eventually waste their money, *i.e.*, a non-ignorable sunk cost. Consumers are more likely to ignore information that is intended to affect them subjectively. Objective content act an effective role in reducing consumers' uncertainty when they are not sure whether the associated products fit their needs. By comparing the correlation coefficients in the price range 70-150 and 150-500, we find that consumers get much more cautious and more undesired to watch subjective content in the range 150-500. It can be reflected by both stronger intensity and significance of correlations in the range 150-500 ($-0.4368^{**}<-0.3818^{**}$, $-0.4740^{***}<-0.0649$, $-0.5406^{***}<-0.2097$). Higher prices lead to more difficult choices for consumers.

Advice: We suggest that short videos associated with relatively upscale products that appear to have more objective descriptions could be recommended to consumers by platforms. Sellers may consider preventing subjective information (*e.g.*, emotional expressions like "you cannot miss it this time") and faithfully presenting the advantages and disadvantages of such products.

5.3.3 Price 500-5000. As for upscale products not routinely needed, we can hardly observe any significant correlation between subjectivity scores and short videos' advertising effects. There might be several explanations. For such products, people might collect comprehensive information from multiple sources and are less affected by a particular source that is known to have the functionality of product promotion. This phenomenon might also be due to the product categories' shift, as most products are electronics in this price range while makeups in other price ranges. However, because of data limitations, we don't have approach to verify this assumption.

5.3.4 Summary. Overall, we find that: (i) the relationships between short videos' subjectivity and advertising effects vary with prices; (ii) subjectivity scores show positive correlations with advertising effects when products are cheap; (iii) when prices rise, the correlations gradually turn into negative; (iv) when prices keep rising to a very high range, the correlations disappear. Therefore, we suggest sellers and online platforms produce and recommend (i) subjective short videos for cheap products, (ii) objectivity short videos for relatively upscale products.

6 DISCUSSION AND CONCLUSION

In this paper, we estimate the advertising effects, *i.e.*, informative effects, persuasive effects, and total effects, of online short videos as an online advertising intermediary through a well-established causal inference tool doubly robust. To track the source of these effects, we extract the subjectivity scores of the short videos by speech recognition and dictionary-based subjectivity analysis and mine the correlations between these effects and short videos' subjectivity. To accommodate our research, we collect a fine-grained consumer behavior dataset from a world-leading e-commerce platform, *i.e.*, Taobao.com. In experiments we first find all three kinds of advertising effects significant. Then we observe significant correlations between short videos' subjectivity and advertising effects. However, the signs and intensity vary with product prices. According to our findings, we provide some advice for both sellers and online platforms. To the best of our knowledge, this is the first work that mines both the strength and sources of advertising effects of online short videos. To extend this work, future directions include (i) verifying our findings across more short videos in more categories and price ranges, (ii) investigating the effect of other properties of short videos, and (iii) estimating the causal effect of some certain video content.

7 ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China (No. 61625107, No. 62006207), Key Research and Development Projects of the Ministry of Science and Technology (No. 2020YFC0832500), National Key Research and Development Program of China (No. 2018AAA0101900), the Fundamental Research Funds for the Central Universities and Zhejiang Province Natural Science Foundation (No. LQ21F020020).

REFERENCES

[1] Daniel A Ackerberg. 2001. Empirically distinguishing informative and prestige effects of advertising. *RAND Journal of Economics* (2001), 316–333.

[2] Daniel A Ackerberg. 2003. Advertising, learning, and consumer choice in experience good markets: an empirical examination. *International Economic Review* 44, 3 (2003), 1007–1040.

[3] Greg M Allenby and James L Ginter. 1995. Using extremes to design products and segment markets. *Journal of Marketing Research* 32, 4 (1995), 392–403.

[4] Peter C Austin. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research* 46, 3 (2011), 399–424.

[5] Kyle Bagwell. 2007. The economic analysis of advertising. *Handbook of industrial organization* 3 (2007), 1701–1844.

[6] Heejung Bang and James M Robins. 2005. Doubly robust estimation in missing data and causal inference models. *Biometrics* 61, 4 (2005), 962–973.

[7] Erik Cambria, Yang Li, Frank Z Xing, Soujanya Poria, and Kenneth Kwok. 2020. SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 105–114.

[8] Edward Hastings Chamberlin. 1933. *The Theory of Monopolistic Competition: A Re-orientation of the Theory of Value.* Cambridge, Massachusetts: Harvard University Press.

[9] David Chan, Rong Ge, Ori Gershony, Tim Hesterberg, and Diane Lambert. 2010. Evaluating online ad campaigns in a pipeline: causal models at scale. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. 7–16.

[10] Andrew T Ching and Masakazu Ishihara. 2012. Measuring the informative and persuasive roles of detailing on prescribing decisions. *Management Science* 58, 7 (2012), 1374–1387.

[11] Zhixuan Chu, Stephen L Rathbun, and Sheng Li. 2020. Matching in Selective and Balanced Representation Space for Treatment Effects Estimation. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 205–214.

[12] China Internet Network Information Center (CINIC). 2020. The 46th China statistical report on Internet development.

[13] C Robert Clark, Ulrich Doraszelski, and Michaela Draganska. 2009. The effect of advertising on brand awareness and perceived quality: An empirical investigation using panel data. *Qme* 7, 2 (2009), 207–236.

[14] William S Comanor and Thomas A Wilson. 1974. *Advertising and market power.* Number 144. Harvard University Press.

[15] Hanming Deng, Yang Hua, Tao Song, Zongpu Zhang, Zhengui Xue, Ruhui Ma, Neil Robertson, and Haibing Guan. 2019. Object guided external memory network for video object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 6678–6687.

[16] Elisabeth Honka, Ali Hortaçsu, and Maria Ana Vitorino. 2017. Advertising, consumer awareness, and choice: Evidence from the US banking industry. *The RAND Journal of Economics* 48, 3 (2017), 611–646.

[17] Min Hu, Haowen Wang, Xiaohua Wang, Juan Yang, and Ronggui Wang. 2019. Video facial emotion recognition based on local enhanced motion history image and CNN-CTSLSTM networks. *Journal of Visual Communication and Image Representation* 59 (2019), 176–185.

[18] Guido W Imbens and Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences.* Cambridge University Press.

[19] iResearch. 2020. China's online advertising market annual monitoring report 2020.

[20] Nicholas Kaldor. 1950. The economic aspects of advertising. *The review of economic studies* 18, 1 (1950), 1–27.

[21] Maurice George Kendall. 1948. Rank correlation methods. (1948).

[22] Kun Kuang, Meng Jiang, Peng Cui, Hengliang Luo, and Shiqiang Yang. 2017. Effective promotional strategies selection in social media: A data-driven approach. *IEEE transactions on big data* 4, 4 (2017), 487–501.

[23] Kun Kuang, Meng Jiang, Peng Cui, and Shiqiang Yang. 2016. Steering social media promotions with effective strategies. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, 985–990.

[24] Kun Kuang, Lian Li, Zhi Geng, Lei Xu, Kun Zhang, Beishui Liao, Huaxin Huang, Peng Ding, Wang Miao, and Zhichao Jiang. 2020. Causal inference. *Engineering* 6, 3 (2020), 253–263.

[25] Namhee Kwon, Liang Zhou, Eduard Hovy, and Stuart W Shulman. 2007. Identifying and classifying subjective claims. In *DG. O.* 76–81.

[26] Diane Lambert and Daryl Pregibon. 2007. More bang for their bucks: Assessing new features for online advertisers. In *Proceedings of the 1st international workshop on Data mining and audience intelligence for advertising*. 7–15.

[27] Ya Li, Jianhua Tao, Linlin Chao, Wei Bao, and Yazhu Liu. 2017. CHEAVD: a Chinese natural emotional audio–visual database. *Journal of Ambient Intelligence and Humanized Computing* 8, 6 (2017), 913–924.

[28] Jura Liaukonyte, Thales Teixeira, and Kenneth C Wilbur. 2015. Television advertising and online shopping. *Marketing Science* 34, 3 (2015), 311–330.

[29] Sridhar Narayanan and Puneet Manchanda. 2009. Heterogeneous learning and the targeting of marketing communication for new products. *Marketing science* 28, 3 (2009), 424–441.

[30] Sridhar Narayanan, Puneet Manchanda, and Pradeep K Chintagunta. 2005. Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research* 42, 3 (2005), 278–290.

[31] Phillip Nelson. 1974. Advertising as information. *Journal of political economy* 82, 4 (1974), 729–754.

[32] Kenton O'Hara, April Slayden Mitchell, and Alex Vorbau. 2007. Consuming video on mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 857–866.

[33] S Andrew Ozga. 1960. Imperfect markets through lack of knowledge. *The quarterly journal of Economics* (1960), 29–52.

[34] Judea Pearl. 2009. *Causality.* Cambridge university press.

[35] Paul R Rosenbaum and Donald B Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 1 (1983), 41–55.

[36] Donald B Rubin. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology* 66, 5 (1974), 688.

[37] Navdeep S Sahni, S Christian Wheeler, and Pradeep Chintagunta. 2018. Personalization in email marketing: The role of noninformative advertising content. *Marketing Science* 37, 2 (2018), 236–258.

[38] Chen Shiyun, Lin Xin, Xiao Yanghua, and He Liang. 2019. Sentiment commonsense induced sequential neural networks for sentiment classification. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 1021–1030.

[39] Charles Spearman. 1961. The proof and measurement of association between two things. (1961).

[40] Nobuhiko Terui, Masataka Ban, and Greg M Allenby. 2011. The effect of media advertising on brand consideration and choice. *Marketing science* 30, 1 (2011), 74–91.

[41] Yi-Lin Tsai and Elisabeth Honka. 2020. Informational and Non-Informational Advertising Content. *Available at SSRN 3706665* (2020).

[42] Bairui Wang, Lin Ma, Wei Zhang, and Wei Liu. 2018. Reconstruction network for video captioning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7622–7631.

[43] Xin Wang, Wenhu Chen, Jiawei Wu, Yuan-Fang Wang, and William Yang Wang. 2018. Video captioning via hierarchical reinforcement learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4213–4222.

[44] Russell S Winer, Dhar Ravi, Fabrizio Mosca, et al. 2013. *Marketing management.* Apogeo.

[45] Fangzhao Wu, Chuhan Wu, and Junxin Liu. 2018. Imbalanced sentiment classification with multi-task learning. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 1631–1634.

[46] Wenmeng Yu, Hua Xu, Fanyang Meng, Yilin Zhu, Yixiao Ma, Jiele Wu, Jiyun Zou, and Kaicheng Yang. 2020. Ch-sims: A chinese multimodal sentiment analysis dataset with fine-grained annotation of modality. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 3718–3727.

[47] Ningyu Zhang, Shumin Deng, Zhanlin Sun, Guanying Wang, Xi Chen, Wei Zhang, and Huajun Chen. 2019. Long-tail Relation Extraction via Knowledge Graph Embeddings and Graph Convolution Networks. In *NAACL*, Jill Burstein, Christy Doran, and Thamar Solorio (Eds.). 3016–3025.

[48] Shengyu Zhang, Tan Jiang, Tan Wang, Kun Kuang, Zhou Zhao, Jianke Zhu, Jin Yu, Hongxia Yang, and Fei Wu. 2020. DeVLBert: Learning Deconfounded Visio-Linguistic Representations.. In *MM '20: The 28th ACM International Conference on Multimedia, Virtual Event / Seattle, WA, USA, October 12-16, 2020.*

[49] Shengyu Zhang, Ziqi Tan, Jin Yu, Zhou Zhao, Kun Kuang, Jie Liu, Jingren Zhou, Hongxia Yang, and Fei Wu. 2020. Poet: Product-oriented Video Captioner for E-commerce. In *Proceedings of the 28th ACM International Conference on Multimedia*. 1292–1301.

[50] Shengyu Zhang, Ziqi Tan, Zhou Zhao, Jin Yu, Kun Kuang, Tan Jiang, Jingren Zhou, Hongxia Yang, and Fei Wu. 2020. Comprehensive information integration modeling framework for video titling. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2744–2754.

[51] Shengyu Zhang, Dong Yao, Zhou Zhao, Tat-Seng Chua, and Fei Wu. 2021. CauseRec: Counterfactual User Sequence Synthesis for Sequential Recommendation.. In *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021.*

[52] Sicheng Zhao, Yunsheng Ma, Yang Gu, Jufeng Yang, Tengfei Xing, Pengfei Xu, Runbo Hu, Hua Chai, and Kurt Keutzer. 2020. An End-to-End visual-audio attention network for emotion recognition in user-generated videos. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 303–311.

[53] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1059–1068.

[54] Xizhou Zhu, Yujie Wang, Jifeng Dai, Lu Yuan, and Yichen Wei. 2017. Flow-guided feature aggregation for video object detection. In *Proceedings of the IEEE International Conference on Computer Vision*. 408–417.

A APPENDIX

A.1 Details of Data Preparation

A.1.1 Data Collection. Our data is composed of several segments, including (i) short videos and their information (associated products and page view number, etc.), (ii) consumers' information (gender, age class, and purchase power, etc.), (iii) products' information (L1 category, leaf category[2], and price, etc.), (iv) our data source platform's homepage exposure records (short video index, consumer index, and watch status, etc.) and (v) consumers' click (consumer index, product index and click time, etc.) and purchase records (consumer index, product index and purchase time, etc.). We take several steps to select our data. First, we randomly select 517 short videos in categories of electronics and makeups with cumulative page views over 100,000, click numbers over 300 from October 4th, 2020 to November 5th, 2020, and relatively good texts extracted by speech recognition. Then we could get a set of 397 associated products and 308 sellers. We collect the information of these products, sellers, and short videos. For each short video, we randomly select homepage exposures with at most 1,000 views as well as at most 1,000 un-views to make sure that there are both view behaviors and un-view behaviors which concern about 1.2 million consumers in total. We collect these consumers' information. Next, we collect the click records and purchase records from September 1st, 2020 to November 10th, 2020 of products with the same leaf categories as the target 397 products and there are about 0.3 billion click records, 6 million purchase records in total. We also collect homepage exposure records from September 20th, 2020 to November 5th, 2020 of the target 397 products and 517 short videos. Considering the space and time efficiency, we randomly select at most 10,000 un-clicked exposures and 10,000 clicked exposures for each short video and there are 13 million exposure records in total.

A.1.2 Data Preprocessing. To preprocess the large-scale consumer behavior records, we first disentangle them by consumers and get the exposure records, click records and purchase records for each consumer. Then we union the three kinds of records and rearrange them by the timestamp and get the timeline of the consumers' behaviors. Next, for each short video, we scan over the timeline to check several necessary properties.

- Whether a consumer was exposed to the short video. We only consider consumers who are all exposed to a certain short video to remove endogenous bias.
- Whether a consumer watched the short video. We use consumers who didn't watch the exposed short video as a control group.
- Whether a consumer clicked the associated product in five days after the short video exposure.
- Whether a consumer purchased the product in five days after the short video exposure.
- Whether a consumer had clicked the products of the same leaf category as the associated product of the short video. This is a property used for propensity score calculation since consumers who have clicked those same-leaf-category products know more about the associated product.
- Whether a consumer had been recommended the same products as the associated products of the short videos in the past. This is also a property used for propensity score calculation since consumers who have been recommended the same product before might also know more about the product and have a stronger interest in the product.

Finally, together with the consumers' information, sellers' information, and products' information, we obtain the data necessary for section 3 and section 4.

A.2 Experimental Settings

For propensity score calculation, we need to specify the confounders by our prior knowledge. In our problem, the information of the consumer might be the confounders we have to balance. We use the click status of the products of the same leaf categories as the associated products in a consumer's last 30 clicks, the expose status of the associated products in the past, the purchase power, gender, whether the leaf categories and the L1 categories of the associated products is in the consumer's preferred leaf categories and L1 categories, the operating system, age class as the confounders that make the two consumer groups, *i.e.*, consumers who watch the exposed short videos and consumers who don't watch the exposed short videos, differs.

Because of online consumer behaviors' sparsity, even if we have at least 300 views for each short video, there is still a great probability that few consumers finally click or even purchase the associated products. To remove the randomness brought by the sparsity, we only consider short videos with consumer click and purchase numbers both over two. After this filtering, there are finally 163 valid short videos. We analyze these short videos' advertising effects. Especially, in subjectivity correlation analysis, considering tutorial videos (defined in section 5.3) could be both subjective and objective, we choose to drop them together with several short videos with still relatively low text quality and two short videos with associated products' price more than 5,000. Finally, 131 short videos remain for correlation analysis.

In the process of short videos' subjectivity extraction, we utilize the **ffmpeg**[3] toolkit to extract acoustic features of our short videos and resample the audio data to mono 16 kHz. Then we call Baidu's speech recognition api[4] to complete the speech recognition process. In textual subjectivity extraction, we refer to the Chinese subjective word dictionary and subjectivity extraction code on https://github.com/liuhuanyong/ZhuguanDetection. For speech text preprocessing, we tokenize the sentences with Jieba Chinese Tokenizer[5] which is suitable for text analysis.

[2]The categories of products on Taobao.com have a tree structure with L1 (level 1) category as the highest in the category tree. Besides, there are also L2, L3, L4, L5 categories as different levels of categories in the tree. However, not all of the products have all these five levels of categories and the lowest level of category is named as leaf category.

[3]https://ffmpeg.org/

[4]https://ai.baidu.com/tech/speech/aasr

[5]https://github.com/fxsjy/jieba