Decision Support Systems xxx (xxxx) xxx

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Time-varying effects of search engine advertising on sales–An empirical investigation in E-commerce

Yanwu Yang [a], Kang Zhao [b,*], Daniel Dajun Zeng [c], Bernard Jim Jansen [d]

[a] School of Management, Huazhong University of Science and Technology, Wuhan 430074, China
[b] Department of Business Analytics, University of Iowa, Iowa City, IA 52242, United States of America
[c] Institute of Automation, Chinese Academy of Science, Beijing 100190, China
[d] Qatar Computing Research Institute, HBKU, Doha, Qatar

A R T I C L E I N F O

Keywords:
Online advertising
Electronic commerce
Advertising analytics
Business intelligence

A B S T R A C T

As a mainstream advertising channel, Search Engine Advertising (SEA) has a huge business impact and attracts a plethora of attention from both academia and industry. One important goal of SEA is to increase sales. Nevertheless, while previous research has studied multiple factors that are potentially related to the outcome of SEA campaigns, effects of these factors on actual sales generated by SEA remain understudied. It is also unclear whether and how such effects change over time in dynamic SEA campaigns that last for an extended period of time. As the first empirical investigation of the dynamic advertisement-sales relationship in SEA, this study builds an advertising response model within a time-varying coefficient (TVC) modeling framework, and estimates the model using a unique dataset from a large e-commerce retailer in the United States. Results reveal the effects of the advertising expenditure, consumer behaviors and advertisement characteristics on realized sales, and demonstrate that such effects on sales do change over time in non-linear ways. More importantly, we find that carryover has a stronger effect in generating sales than immediate or direct response does, and advertisers need to carefully decide how much to bid for higher ad positions. These findings have direct implications for business decision-making to launch more effective SEA campaigns and for SEA platforms to improve their pricing mechanism.

1. Introduction

During the past decade, search engine advertising (SEA) has become one of the most prominent outlets for online advertising campaigns. Through SEA, advertisers pay search engines to display their advertisements related to search queries along with organic results on search engine result pages (SERPs). The economic impact of SEA has been well documented. According to the Interactive Advertising Bureau (2018), SEA revenue in the U.S. alone exceeded 22 billions USD during the first half of 2018, accounting for nearly half of the total revenue for online advertising during that period. With the high expenditure on SEA, advertisers are eager to know what drives the outcome of SEA campaigns. Although SEA success can be measured in different ways (e.g., online traffics and brand awareness), sales are typically one of the most important criteria advertisers care about, especially in e-commerce [1]. Indeed, a better understanding of the ad-sales relationship can help advertisers make more effective investment decisions in SEA campaigns and aid SEA platforms in advertising mechanism design.

SEA is a much more dynamic and evolving market [2] than traditional marketing channels (e.g., newspapers and TV). At the core of SEA is real-time position auctions run by search engines to determine which ads to be displayed on a SERP and their rankings. As participants of these auctions, advertisers need to make decisions on expenditures by considering a range of factors related to consumer behaviors (e.g., ad clicks and product purchases), characteristics of advertisement (e.g., keywords and ad positions) as well as competitions from other advertisers. Note that the values of these factors to advertisers could change over time [3,4]. As a result, it has been well-recognized by advertisers that strategies governing SEA campaigns need to be dynamically adjusted in order to achieve more sales [5,6,7,8,9,10].

Although business needs of understanding dynamic ad-sale relationships in SEA are clear, very little research has investigated the drivers of sales volume generated from SEA. Previous studies of SEA have mainly investigated measures related to consumers' clicks on ads

* Corresponding author.
E-mail addresses: yangyanwu.isec@gmail.com (Y. Yang), kang-zhao@uiowa.edu (K. Zhao).

https://doi.org/10.1016/j.dss.2022.113843
Received 10 January 2022; Received in revised form 15 July 2022; Accepted 16 July 2022
Available online 23 July 2022

and conversions. While intuitive and easy to obtain, clicks on ads and conversions are not equal to sales. Even customers who clicked the same ad may purchase the advertised product in different quantities, leading to different amounts of sales. Although earlier studies have attempted to associate the advertising expenditure with sales [11] and quantify the relative effectiveness across different channels [12], these models did not consider other important factors beyond the expenditure, and analyze the ad-sales relationship in a static way.

Therefore, a critical gap exists for researchers to investigate the effects of a comprehensive set of factors on sales from SEA over time. For example, SEA has been commonly recognized as a form of direct response advertising and a short-term investment [13], such that e-commerce firms pay to attract traffics to their own websites and advertised product pages, which may generate online transactions immediately. Many advertisers believe such effects will stop after the expenditure on SEA campaigns stops. However, while researchers have questioned this convectional belief [14,15,16], there lacks a formal study on whether the effectiveness of SEA occurs in a direct (i.e., immediate) or indirect (i.e., time-lagged) manner. Besides the expenditure, other factors related to consumer behaviors (e.g., click-through rate (CTR), conversion rate (CVR), cost-per-click (CPC)), and advertisement characteristics (e.g., keywords and ad positions in SERPs) could also affect sales. Nevertheless, no studies have directly compared their effects with each other or analyzed how their effects change over time.

To address the aforementioned gap and challenges, this paper represents the first effort to empirically explore the dynamic advertising-sales relationship in SEA. Specifically, this research builds an advertising response model within a time-varying coefficient (TVC) modeling framework [17] to capture the dynamic nature of SEA markets. We choose the partial adjustment model [18,19] to examine the carryover effect in SEA. We empirically estimate our response model using a unique panel dataset collected from SEA campaigns by a large U.S. e-commerce retailer.

The contributions of our study and how the paper can support decision-making can be summarized as follows.

First, we are the first study to empirically reveal the dynamic nature of SEA–the effects of various factors on SEA sales do change over time. Our work offers a new perspective compared to previous studies that mainly focus on the overall static effects of different factors. Such dynamics in SEA have important implications for advertisers–they must continuously track and predict the real-time effectiveness of their SEA campaigns, so that they can make better decisions and adjust their advertising strategies. Moreover, the changing pattern of advertising factors over time differs, which may offer valuable insights for advertisers' decision-makings and planning.

Second, one surprising finding by adopting a dynamic perspective is that the ad-sales relationship in SEA demonstrates a strong carryover effect. This contradicts the commonly held view that SEA is a direct-response advertising medium that mainly generates immediate or direct effects (e.g., sales) [20,21,22]. Instead, SEA advertisers may need to be more patient, and their decisions have to be based on longer-term efforts. Meanwhile, without considering the carryover effect, the immediate effect of the advertising expenditure on sales may be either underestimated (for newly launched advertisements) or overestimated (for advertisements that have been delivered for a long time) and could lead to inefficient spending decisions.

Last, this research suggests that advertisers need to pay extra attention to ad positions. While a higher position is associated with more sales, advertisers' costs to bid for better positions may not be covered by the consequent gains in sales, especially when the positive effect of ad positions decays quickly over time. In other words, contrary to the typical SEA strategy of always bidding for the highest ad position on SERPs as previous studies suggested [23,24], we encourage advertisers to base their bidding decisions on expected future sale gains over time.

A better understanding of the dynamic ad-sales relationship in SEA has tremendous values. Because spending more in SEA does not necessarily lead to higher sales [25,8], advertisers need to make their budget allocation taking into account advertising dynamics in SEA to maximize their returns in the ever changing market. Such real-time decision support in SEA is especially important for advertisers from small and medium enterprises, who represent the main revenue sources for search engines but have limited resources to understand such complexity and optimize budgets for their SEA campaigns [26].

The remainder of this paper is organized as follows. Section 2 presents a brief survey of related research. This is followed by descriptions of our data and key variables used in this research in Section 3. In Section 4, we discuss basic principles of the time-varying modeling framework and present a time-varying response model for SEA. Empirical results are listed in Section 5. The last section concludes with managerial implications and theoretical contributions, and future research directions.

2. Related literature and theoretical background

This paper studies the dynamic ad-sales relationship in SEA and is related to literature from three streams of research: (i) factors related to the performance of SEA campaigns, (ii) dynamic processes and advertising decisions in SEA, and (iii) time-varying modeling.

The growth of SEA has motivated studies that investigated factors for the success of SEA campaigns. An advertiser's expenditure is critical for its SEA campaigns to gain more visibility and revenue [27,28]. The advertising expenditure affects sales in two possible ways: (1) Direct via immediate response (a.k.a., short-term advertising elasticity)–the current advertising expenditure affects current sales directly and immediately [29,30]; or (2) Indirect via carryover effect–a considerable time lag exists between the display of an advertisement and sales of the advertised product. In other words, a certain amount of sales generated by an advertisement is not achieved immediately after the expenditure and deployment of the advertisement. Previous research has reported empirical evidence of carryover effect in online advertising channels. For example, Johnson et al. [31] analyzed 432 online display advertising field experiments on the Google Display Network, and found most campaigns have a modest and positive carryover. Previous research has recognized the necessity of considering carryover effect of SEA, as demonstrated in a formal analysis by Archak et al. [32]. However, empirical evidence on carryover effect on actual sale volume generated by SEA is still lacking.

Besides the expenditure, other well-recognized factors for SEA success include ad positions and consumer behaviors over ads. In SEA, ad position is a key factor that advertisers compete for as higher ad positions are expected to generate more traffics and sales [33,34]. Therefore, higher ad positions are usually given to higher bidders if competing ads have the same relevance and quality.

While advertisers care about how many times their ads are displayed to consumers, how consumers interact with their ads is more important. Thus major search engines predominantly adopt the pay-per-click scheme and charge an advertiser only when their ads are clicked by consumers. The PPC scheme has three important measures of consumer behaviors: *Click-through-rate* is the ratio of clicks on an ad over impressions by consumers. *Conversion rate* is the ratio of conversions (e.g., making a purchase) over the total number of ad clicks. In fact, many previous studies have used both measures as proxies to quantify the performance of SEA campaigns [33,35]. The third measure, *cost-per-click*, is directly related to the relationship between the expenditure and outcomes, because it is an advertiser's expenditure on an ad divided by the number of clicks generated from the ad. For an advertiser, the actual CPC serves as a measure of how efficiently the advertising expenditure are generating clicks [36].

Researchers have also investigated relationships beween ad positions and consumer behavior measures [37]. Most studies agreed that CPC and CTR monotonically decrease when ad positions are lower [33,37,34]. However, inconsistent findings exist for CVR. On one hand, some [37,34] agreed that CVR is higher for ads at higher positions and decreases for lower ad positions. On the other hand, Agarwal et al. [33] noticed that CVR could increase with lower ad positions. The reason is that, while ad positions do affect CTR, after a consumer clicks an ad, whether a conversion will occur depends mainly on the website and the product, instead of the ad position in the SERP [34].

Given the strong connection between ad positions and consumer behaviors, researchers have attempted to help advertisers place ads in the right positions. For example, Ghose and Yang [37] revealed that for search engines, bidding prices are more important than prior CTR for the final position of an advertisement. Keywords with more prominent positions are not necessarily the more profitable ones for advertisers. Jeziorski and Moorthy [38] also studied the substitutional relationship between ad positions and advertisers' brand strength, and suggested that advertisers with strong brands do not necessarily need to bid for the highest position. Recently, Zhuang et al. [39] investigated the role of price information in consumers' responses (e.g., clicks and conversions) to product list advertising, and revealed that consumers tend to click on the highest or lowest priced options in the early phases of the purchase funnel, while in later phases, they are more likely to click on moderately priced options.

Due to the highly dynamic nature of SEA, many studies have also modelled dynamic advertising processes and related decisions. One such decision is as bidding for keywords using different strategies at different time [40,2,9,10]. Another dynamic decision in SEA focuses on budget allocation over time [41,8,42].

From the methodological perspective, our study is related to time-varying modeling [17,43]. When dealing with longitudinal data, researchers often want to explicitly capture changes in the association between covariates and the outcome over time in a flexible manner. Thus Tan et al. [43] introduced a time-varying coefficient (TVC) model–a special case of varying-coefficient model [44]. It has been used to explore the changing roles of regulatory regimes, marketing mailers, transaction characteristics and demographic factors on international trades and marketing outcomes [45,46,47].

Specifically, the TVC model has three characteristics that fit this study. *First*, it is capable of estimating time-varying effects of covariates on the dependent variable. Thus TVC models are a generalized form of traditional linear regression models by incorporating time as the third dimension and representing coefficients of covariates with smoothly time-varying functions. *Second*, compared to multi-level (or hierarchical) modeling (MLM) frameworks that can also capture temporal associations between time-varying covariates and the outcome, a TVC model is more flexible and could effectively reveal any arbitrary "data-driven" shapes of covariates' time-varying effects on the outcome, as long as coefficient functions are smooth (i.e., with no sudden jumps or break points).[1] In addition, in the TVC model framework, researchers can also specify a certain functional form when they have sufficient prior knowledge and evidence, while allowing others to change freely. By contrast, an MLM has to assume a specific form of coefficient functions (e.g., linear, quadratic, or cubic) for trajectory shapes. Admittedly, estimating a TVC model needs more data than a parametric model does [43]. *Third*, a TVC model can handle the co-existence of multiple covariates in the same model, including time-varying ones along with time-invariant ones.

Overall, our research is distinct from the extant SEA research in the following ways. *First*, we propose a time-varying response model for SEA and estimate its parameters using a large-scale dataset from a major e-commerce retailer. Compared to advertising models in the literature, our model incorporates a quality-adjustment structure and is more appropriate for the dynamic context of SEA. Therefore, our approach also reveals several key findings that have not been found by or even contradict previous studies.

Second, we are the first study to systematically investigate roles of a comprehensive set of factors, including the advertising expenditure, carryover effect, consumer behaviors (e.g., CTR and CVR), and advertisement-specific characteristics (e.g., ad position) and keyword-specific characteristics (e.g., the length of keywords, and appearances of retailers, brands and holidays), in generating sales from SEA. More importantly, we study these factors' dynamic roles from a longitudinal perspective, so that we can reveal how their effects on sales change over time during an extended period of SEA campaigns.

Third, this is one of the few studies focusing directly on sales generated from SEA campaigns and studies the ad-sales relationship. Compared to marketing outcome measures based on consumers' clicks, the focus on sales can more accurately and directly quantify advertisers' financial gains from SEA campaigns and better inform their decision-makings and planning in such campaigns.

3. Data and variables

This research uses a large-scale panel dataset collected from SEA campaigns by a large U.S. retailer, which offers a wide range of consumer electronics such as home appliances, air purifiers, etc. The retailer owns a large nationwide retail chain with brick-and-mortar stores and an electronic commerce website. The company has continuously conducted SEA campaigns over several years, and recorded data about SEA advertisements and online sales generated by these ads. The dataset we use is about SEA campaigns by this retailer during a 33-month period, spanning 4 calendar years from September 2005 to June 2008. The dataset is valid because the search advertising schema remains (almost) unchanged in the past decade. Moreover, the dataset contains almost 7 million time-stamped records from nearly 40,000 key search phrases and almost 55,000 advertisements, which is quite rich to support our empirical study.

Each record in the dataset is about one advertisement on a given day. Specifically, a record includes keywords that triggered the ad, the number of impressions, the number of clicks, the average CPC, the number of conversions (i.e., purchase or orders), the total number of items ordered, and generated sales. Note that the search query of a keyword may lead to an impression (i.e., display) of a related ad, but not necessarily a click; a click may not lead to a conversion (i.e., an order), and an order may include one or more items. We believe this dataset is appropriate for investigating the time-varying ad-sales relationship in SEA, because sales from SEA are available and the dataset covers a long time period that is sufficient to estimate a time-varying model [43]. There are few empirical studies of SEA using a dataset that has such a large scale, covers such a long time span, or contains such a rich range of advertising and keyword attributes.

Because this paper focuses on the ad-sales relationship, we directly use the number of products (in units) sold online (*Sales*) from each advertisement during a day as the dependent variable. There are several key independent variables whose effects on sales are of interests. Table 1 provides a list of all variables, along with their summary statistics, in this research. Table 2 illustrates the pairwise correlation among these variables.

The first one is the expenditure (*AdExpenditure*) spent on an SEA advertisement on a given day. We also include three independent variables for consumer's click behaviors–*CTR*, *CVR*, and *CPC*–and one independent variable for advertisement characteristics–ad position. At the first glance, it may seem that the transitive relationship from impressions, through clicks to conversions is simply linear in SEA. In other words, the number of clicks on an ad is the product of the number of impressions and CTR. Similarly, the number of conversions is the product of the number of clicks and CVR. However, such linear relationships do not necessarily hold, because CTR and CVR are not constants. Also, the relationship between the advertising expenditure and the number of clicks is essentially nonlinear because cost-per-click (*CPC*) also changes over time. Thus, we investigate the dynamic influence of

[1] A function is smooth if its first-order derivative function is continuous.

Table 1
A Summary of variables.

Variables	Description	Mean	Std. Dev.
Sales(*Sales*)	The total amount of sales (in units)	16.271	3521.487
Lagged Sales($Sales_{t-1}$)	Sales from the previous time step (in units)	16.271	3521.487
Advertising Expenditure (*AdExpenditure*)	Total spending (in dollars) on an ad during a day	149.842	2160.421
Ranking Position (*AdPosition*)	The ranking position of an ad on the SERP	7.894	10.922
Cost-Per-Click (*CPC*)	The cost-per-click of an ad	18.245	53.637
Click-Through-Rate (*CTR*)	The click-through-rate of an ad	0.040	0.131
Conversion Rate (*CVR*)	The conversion-rate of an ad	0.003	0.045
Length of Keywords (*KLength*)	The number of words in a keyword for an ad	2.622	0.803
Brand (*Brand*)	Binary variable–if associated keywords contain one or more specific brand names	0.175	0.380
Retailer (*Retailer*)	Binary variable–if associated keywords contain one or more specific retailer names	0.028	0.167
Holiday (*Holiday*)	Binary variable–if associated keywords contain one or more specific holiday names	0.003	0.054

Table 2
Pairwise correlation coefficients among variables.

	Sales	*Saleslag1*	*Expenditure*	*Position*	*CTR*	*CPC*	*CVR*	*KLength*	*Retailer*	*Brand*	*Holiday*
Sales	1										
Saleslag1	0.954***	1									
Expenditure	0.135***	0.134***	1								
Position	−0.003***	−0.003***	−0.035***	1							
CTR	0.015***	0.015***	0.026***	−0.125***	1						
CPC	−0.000	−0.000	0.186***	−0.156***	0.221***	1					
CVR	0.007***	0.007***	0.023***	−0.033***	0.084***	0.080***	1				
KLength	−0.004***	−0.004***	−0.032***	−0.131***	0.048***	0.004***	−0.003***	1			
Retailer	0.027***	0.027***	0.037***	−0.104***	0.136***	0.019***	0.057***	0.159***	1		
Brand	−0.002***	−0.002***	−0.005***	−0.063***	0.007***	−0.020***	0.006***	0.183***	−0.078***	1	
Holiday	−0.000	−0.000	−0.001*	0.009***	0.009***	−0.000	−0.000	0.042***	−0.009***	−0.025***	1

Computed correlation used pearson-method with listwise-deletion.

each factor on sales over time.

An important concept in SEA is the quality of an ad, which has a significant influence on the ad's performance [48]. For search engines, an ad ranking mechanism that considers advertising quality facilitates better matching between advertisements and queries, and consequently improves revenue [49,50]. For advertisers, a higher-quality ad means the advertiser can pay less for each click, so the same advertising budget can lead to more clicks and potentially higher sales. However, while search engines have widely adopted quality scores for ads, such scores are only available within search engines themselves. In addition, the exact formulas to calculate quality scores vary from one search engine to another, and remain trade secrets.

Although advertisers have no access to their ads' quality scores and how such scores are calculated, Google Adwords reveals that quality score mainly considers three factors[2]: *(1)* the expected CTR, which is also an independent variable in our model; *(2)* advertising relevance, which indicates how closely an ad matches the intention of a consumer's search query or keyword(s); and *(3)* the landing page experience. Note that the CVR does not affect an ad's quality score.[3]

In addition, we also use four control variables related to search keywords: (1) *Brand*–whether the keywords contain any specific brand (e.g., "**Apple** computer"); (2) *Retailer*–whether the keywords contain any specific retailer (e.g., "**BestBuy** smartphone"); (3) *KLength* (Length of keywords)–how many words are there in the search keywords (e.g., "gift" vs "flower gift baskets") [37]. Usually, it is more effective for an advertiser to choose brand-specific, retailer-specific, and longer (i.e., more specific) keywords [40,51,52];(4) *Holiday*–whether the keywords contain any specific holiday. This is because advertisers often promote their products during holidays by using holiday keywords to raise consumers' desire to purchase (e.g., "**Christmas** gift").

Landing page experience refers to how relevant, transparent and easy-to-navigate the web page is for consumers who click an ad. In e-commerce, a landing page generally corresponds to a product page with details of the advertised product. A better landing page experience leads to a higher quality score.[4] In our study, all ads are for the same retailer, which means all the landing pages offer similar experience. Thus we treat landing page quality as a constant for all ads and skip it from our analysis.

4. Model development

Our model, defined in Eq. (1), has one dependent variable (y_{ij}, the sales from subject i at the j-th observation) for advertising performance (i.e., sales from ads) along with a set of independent and control variables (x_{ijk}):

$$y_{ij} = \beta_0(t_{ij}) + \sum_{k=1}^{K} \beta_k(t_{ij}) \cdot x_{ijk} + \varepsilon_{ij}, \quad i = 1, \dots, N;\ j = 1, \dots, M_i;\ k = 1, \dots, K, \tag{1}$$

In Eq. 1, N represents the total number of subjects (i.e., an advertisement), M_i is the total number of measurements (of features) for subject i, and K is the number of explanatory variables; t_{ij} is the measurement time of the $j-th$ observation for the $i-th$ subject.[5] $\beta_0(t_{ij})$ and $\beta_k(t_{ij})$ are the coefficient functions to be estimated: the intercept $\beta_0(t_{ij})$ represents the mean of y when $x_k = 0$ at time t_{ij}; The slope, $\beta_k(t_{ij})$, represents the strength and direction of the influence of x_k on y at time t_{ij}. Note that $\beta_0(t_{ij})$ and $\beta_k(t_{ij})$ are continuous coefficient functions of time t, such that their values change over time. Random errors ε_{ij} in the above equation are assumed to be normally and independently distributed. Although time-varying parameters are treated as non-parametric functions, the class of TVC models is parametric for a specified time t. Thus the TVC model can be considered as conditionally parametric,

[2] https://support.google.com/adwords/answer/6167118

[3] https://www.wordstream.com/blog/ws/2012/06/04/quality-score-landing-pages-faq

[4] https://support.google.com/adwords/answer/2404196

[5] Note that data can be unbalanced with different assessment time within and across individual subjects.

representing a semi-parametric approach [47].

4.1. A time-varying SEA response model

With the nonlinear and temporal ad-sales relationship in SEA [33,37,34], we present an advertising response model for SEA in the TVC modeling framework (shown in Fig. 1). It does not assume any function forms for temporal trajectories of covariate coefficients. We also incorporate a quality-adjusted structure [53,54] to account for latent effects of advertising quality on ad performance.

4.1.1. The basic model

The basic model adopts the advertising response model by Arnold et al. [55]. The model is an advertising spending function adjusted by a quality index based on the hedonic price theory (HPT) [56]. Because the outcome of a given amount of advertising spending depends on quality of the advertising copy, HPT can be naturally adopted to model the parsimonious process from advertising spending to market outcomes (e.g., sales). As discussed earlier in this paper, the quality of an ad explicitly affects the relationship between the advertising expenditure and sales, and needs to be controlled in our model. Therefore, following the Arnold model, we present a time-varying quality-adjusted response model for SEA below:

$$Sales_{ij} = e^{\alpha_0(t_{ij})}\cdot(\psi_{ij})^{\beta(t_{ij})}\cdot D_{ij}\cdot e^{\varepsilon_{ij}}, \tag{2}$$

where $Sales_{ij}$ represents the number of products sold from advertisement i at time t_{ij}; ψ_{ij} is the advertising expenditure adjusted by the quality of advertisement i measured at time t_{ij}. Note that advertising quality is latent and we will discuss how to estimate the quality-adjusted advertising spending function in the next subsection. $\mathbf{D}_{ij}$ represents other covariates for sales. Details of them are in Subsection 4.1.3. In addition, ε_{ij} is the normally distributed error term at time t_{ij}; $\alpha_0(t_{ij})$, the intercept coefficient, and $\beta(t_{ij})$ will be estimated.

4.1.2. The HPT-based advertising spending function

Following the quality-adjusted market price in classical HPT [56], we specify the quality-adjusted advertising expenditure function ψ_{ij} from Eq. (2) as below:

$$\psi_t = \left(B_t\prod_{k=1}^{K'} q_{kt}^{\tau_k}\right), \tag{3}$$

where B_t denotes the advertising spending (measured in dollar) at time t; $q_{kt}(k = 1, \ldots, K')$ is the value of advertisement attribute k that determines an ad's quality score; $\prod_{k=1}^{K'} q_{kt}^{\tau_k}$ is thus the multi-dimensional quality index that is equivalent to the quality adjustment factor in [54](p.85), which adjusts the impact of the actual spending (B_t) based on the advertising quality.

Specifically, in our case, as discussed in Section 3, five attributes of an ad could affect its quality score: CTR, the length of keywords, and appearances of retailers, brands and holidays in keywords. Thus Eq. (3) can be rewritten as a time-varying quality-adjusted advertising spending function defined in Eq. (4):

$$\psi_{ij} = AdExpenditure_{ij}\cdot\theta_{ij} = AdExpenditure_{ij}\cdot\left[CTR_{ij}^{\tau_1(t_{ij})}\cdot\kappa_{ij}\right], \quad \text{with } \kappa_{ij} = e^{(KLength_i)^{\tau_2(t_{ij})}+(Retailer_i)^{\tau_3(t_{ij})}+(Brand_i)^{\tau_4(t_{ij})}+(Holiday_i)^{\tau_5(t_{ij})}}, \tag{4}$$

where $AdExpenditure_{ij}$ denotes the actual advertising expenditure observed at time t_{ij}. θ_{ij} is the quality of advertisement i at time t_{ij}. It is a latent variable that is approximated by the product of time-dependent CTR_{ij} (the CTR for advertisement i measured at time t_{ij}) and time-invariant κ_{ij}, which represents the joint effects of four characteristics of keywords associated with advertisement i: $KLength_i$, $Retailer_i$, $Brand_i$ and $Holiday_i$. Five coefficient functions $\tau_1(t_{ij})$, $\tau_2(t_{ij})$, $\tau_3(t_{ij})$, $\tau_4(t_{ij})$, and $\tau_5(t_{ij})$ will be estimated.

4.1.3. Ad position, CVR and CPC

In addition to the expenditure, CTR, and keyword characteristics, three more independent variables–ad position, CVR, and CPC– are included in D_{ij}, which is defined in Eq. (5), where $AdPosition_{ij}$ is the position of advertisement i on SERPs, CPC_{ij} and CVR_{ij} are the cost-per-click and conversion rate of advertisement i, respectively, measured at time t_{ij}. $\lambda_1(t_{ij})$, $\lambda_2(t_{ij})$ and $\lambda_3(t_{ij})$ are the parameters to be estimated.

$$D_{ij} = \prod_{m=1}^{M} X_{m,ij}^{\lambda_m}\cdot e^{\sigma_{ij}} = e^{(AdPosition_{ij})^{\lambda_1(t_{ij})}}\cdot(CPC_{ij})^{\lambda_2(t_{ij})}\cdot(CVR_{ij})^{\lambda_3(t_{ij})}, \tag{5}$$

4.1.4. The SEA response model

Substituting Eqs. (4) and (5) into Eq. (2), we get Eq. (6). After taking natural logarithm transformations on numeric variables in Eq. (6), we obtain Eq. (7).

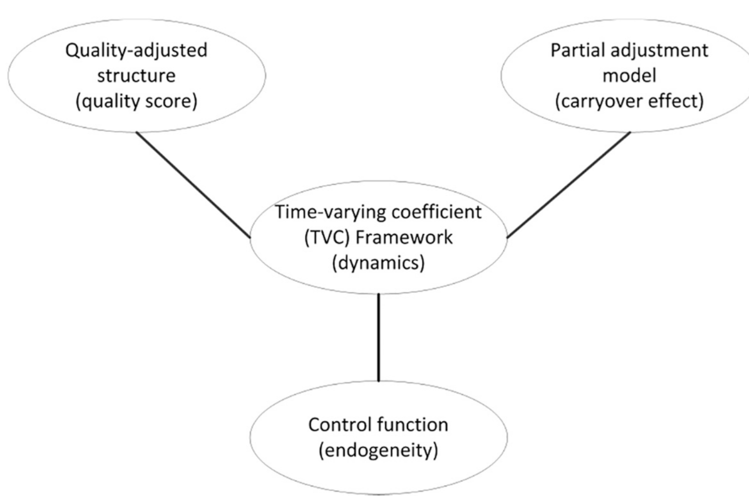

Fig. 1. The conceptual structure of the time-varying SEA response model.

$$Sales_{ij} = e^{\alpha_0(t_{ij})} \cdot \left(AdExpenditure_{ij} \cdot \left[(CTR_{ij})^{\tau_1(t_{ij})} \cdot e^{(KLength_i)^{\tau_2(t_{ij})}} \cdot e^{(Retailer_i)^{\tau_3(t_{ij})}} \cdot e^{(Brand_i)^{\tau_4(t_{ij})}} \cdot e^{(AdPosition_{ij})\lambda_1(t_{ij})} \cdot e^{(Holiday_i)^{\tau_5(t_{ij})}} \right] \right)^{\beta(t_{ij})} \cdot (CPC_{ij})^{\lambda_2(t_{ij})} \cdot (CVR_{ij})^{\lambda_3(t_{ij})} \cdot e^{\varepsilon_{ij}} \tag{6}$$

The control function approach is essentially a Two Stage Least Squares (2SLS) estimator. In the first stage, the correction term is estimated by regressing the advertising expenditure ($AdExpenditure_{ij}$) on a set of exogenous variables. In SEA, advertisers might plan their budget on an ad according to three factors [65]–search demand ($Demand_{ij}$), CTR, and CPC. Search demand is related to search users' behaviors (reflected by the number of queries) as well as the match between queries and advertisements. Basically, search demand can be considered as environmental factors in sponsored search advertising. Thus, search demand is an exogenous variable to the advertiser's budgeting process.

$$lnSales_{ij} = \alpha_0(t_{ij}) + \beta(t_{ij})\left(lnAdExpenditure_{ij} + \tau_1(t_{ij})lnCTR_{ij} + \tau_2(t_{ij})KLength_i + \tau_3(t_{ij})Retailer_i + \tau_4(t_{ij})Brand_i + \tau_5(t_{ij})Holiday_i\right) + \lambda_1(t_{ij})AdPosition_{ij} + \lambda_2(t_{ij})lnCPC_{ij} + \lambda_3(t_{ij})lnCVR_{ij} + \varepsilon_{ij}, \tag{7}$$

For an advertiser, when search demand for a keyword associated with her advertisements is higher, her advertisements will have more opportunity to be displayed to search users and then clicked by them. Thus, she will be charged a larger amount of expenditure by the search engine. In turn, when more search users (i.e., potential consumers) reached the advertised product page or the advertiser's website, in general, more sales will occur. CPC is a proxy for the bidding price under the underlying mechanism implemented by search engines. In other words, CPC is a product of advertisers' bidding decisions and the SEA mechanism. Generally, advertisers make bidding decisions based their maximum willing to pay for a click [9,10]. Moreover, bidding decisions were made before the auction process run by search engines and the payment made by advertisers. In sponsored search advertising, as prior research (e.g., [37]) reported, higher CPC usually results in higher CTR. In this sense, CPC and CTR are most likely exogenous to the budgeting process. For an advertiser, when the cost for each click (i.e., CPC) increases, she may generally pay more to search engines; in a similar vein, the probability that search users will click on her advertisements (i.e., CTR) may result in a higher expenditure. Therefore, we can conclude that, in the control function for budget endogeneity correction (i.e., Eq. 9), coefficients of search demand, CPC and CTR are not zero, i.e. $z_1 \neq 0$, $z_2 \neq 0$, and $z_3 \neq 0$. This meets the requirement that at least one exogenous variable that is omitted from Eq. (8). That is, Eq. (8) is partially correlated with the variable of advertising expenditure, thus the rank condition for identification holds. Thus, we specify the budgeting process in Eq. (9).

4.1.5. Adding carryover effects

To account for the dynamic carryover effect of past advertising outcomes on current outcomes [57,58], we add time-lagged independent variable (i.e., $Sales_{ij-1}$) to Eq. (7), as in dynamic linear models. Specifically, we choose the partial adjustment model [59], which describes a dynamic response process where a variable adjusts over time to a series of desired values [19]. In other words, only some fraction of the desired adjustment is accomplished within a time period. The partial adjustment model has been widely adopted to describe the dynamic response process of sales to advertising and capture carryover effect of current advertising on future sales [60,18,54]. Different from the carryover parameter in the Koyck model, the carryover parameter in the partial adjustment framework characterizes the complete dynamic nature of the advertising response [19]. Then Eq. (7) is transformed into Eq. (8):

$$\begin{aligned} lnSales_{ij} &= \eta(t_{ij})\alpha_0(t_{ij}) + (1-\eta(t_{ij}))lnSales_{ij-1} \\ &+\eta(t_{ij})\beta(t_{ij})\begin{pmatrix} lnAdExpenditure_{ij} + \tau_1(t_{ij})lnCTR_{ij} \\ +\tau_2(t_{ij})KLength_i + \tau_3(t_{ij})Retailer_{ij} + \tau_4(t_{ij})Brand_i \\ +\tau_5(t_{ij})Holiday_i \end{pmatrix} + \eta(t_{ij})\lambda_1(t_{ij})AdPosition_{ij} \\ &+\eta(t_{ij})\lambda_2(t_{ij})lnCPC_{ij} + \eta(t_{ij})\lambda_3(t_{ij})lnCVR_{ij} + \varepsilon_{ij}, \end{aligned} \tag{8}$$

where $\eta(t_{ij})$ is the partial adjustment coefficient, and $(1 - \eta(t_{ij}))$ denotes the carryover effect at time t_{ij}. As $\eta(t_{ij}) \to 1$, the effect of advertising on sales is mainly instantaneous and the carryover effect hardly exists; conversely, as $\eta(t_{ij}) \to 0$, sales become increasingly persistent.

4.1.6. Accounting for the endogeneity of budgeting policies

In general, advertisers need to allocate their expenditures over advertisements strategically to achieve marketing objectives (e.g., maximizing revenues from SEA campaigns) [8,61]. Such budgeting policies could lead to the endogeneity problem [62]: the estimated effect of advertising budget on sales might be biased by the correlation between advertising budget and one or more unobserved latent factors in the error term of Eq. (8). To account for such endogeneity, we use the control function approach, which has been widely used to eliminate the endogeneity bias with marketing mix variables in marketing research [63,64].

$$\begin{aligned} lnAdExpenditure_{ij} &= \varphi_{ij}^B \cdot z_{ij}^B + \mu_{ij}^B \\ &= z_1(t_{ij})lnDemand_{ij} + z_2(t_{ij})lnCPC_{ij} \\ &+z_3(t_{ij})lnCTR_{ij} + \mu_{ij}^B, \end{aligned} \tag{9}$$

where z_{ij}^B indicates the vector of exogenous variables (i.e., $Demand_{ij}$, CTR_{ij}, and CPC_{ij}) for the advertising expenditure, φ_{ij}^B is the unknown parameter vector, and the random error μ_{ij}^B is assumed to be independently and normally distributed.

In the second stage, we include estimated residual $\widehat{\mu}_{ij}^B$ as an addi-

Table 3
First stage results of the control function for budget endogeneity correction.

Coefficients	*Dependent Variables*		
	Ad Expenditure		
	Estimates	*Conf. Int.*	*Std. Errors*
(Intercept)	−0.414***	−0.414 - - 0.413	0.000
Demand	0.249***	0.248–0.249	0.000
CPC	1.070***	1.070–1.071	0.000
CTR	0.636***	0.631–0.641	0.003
Observations	5,649,220		
R^2/adj. R^2	0.950 / 0.950		
AIC	7,976,845.454		

*** $p < .001$.

tional variable in Eq. (8). Note that we also remove CPC_{ij} from the second stage model because CPC is related to sales indirectly via its relationship with the advertising expenditure. This can be confirmed by the zero and insignificant correlation between CPC and sales, as compared to the positive and significant correlation between CPC and the advertising expenditure (see Table 2). Note that the advertising response model in Eq. (8) is a nonlinear regression with regard to parameters $\eta(t_{ij})$, $\alpha_0(t_{ij})$, $\beta(t_{ij})$, $\tau_1(t_{ij})$, $\tau_2(t_{ij})$, $\tau_3(t_{ij})$,$\tau_4(t_{ij})$, $\tau_5(t_{ij})$, $\lambda_1(t_{ij})$ and $\lambda_3(t_{ij})$. It is more convenient to consider it as a linear form for model estimation, and then use estimation results to identify these original parameters. The linear regression form of the final advertising response model is specified in Eq. (10).

$$\begin{aligned} lnSALES_{ij} &= \alpha_0^*(t_{ij}) + \gamma^*(t_{ij})lnSales_{ij-1} \\ &+\beta^*(t_{ij})lnAdExpenditure_{ij} + \tau_1^*(t_{ij})lnCTR_{ij} \\ &+\tau_2^*(t_{ij})KLength_i + \tau_3^*(t_{ij})Retailer_i + \tau_4^*(t_{ij})Brand_i \\ &+\tau_5^*(t_{ij})Holiday_i + \lambda_1^*(t_{ij})AdPosition_{ij} \\ &+\lambda_3^*(t_{ij})lnCVR_{ij} + \alpha_1^*\mu_{ij}^B + \varepsilon_{ij}^*, \end{aligned} \tag{10}$$

where $\alpha_0^*(t_{ij}) = \eta(t_{ij})\alpha_0(t_{ij})$, $\gamma^*(t_{ij}) = 1 - \eta(t_{ij})$, $\beta^*(t_{ij}) = \eta(t_{ij})\beta(t_{ij})$, $\tau_1^*(t_{ij}) = \eta(t_{ij})\beta(t_{ij})\tau_1(t_{ij})$, $\tau_2^*(t_{ij}) = \eta(t_{ij})\beta(t_{ij})\tau_2(t_{ij})$, $\tau_3^*(t_{ij}) = \eta(t_{ij})\beta(t_{ij})\tau_3(t_{ij})$, $\tau_4^*(t_{ij}) = \eta(t_{ij})\beta(t_{ij})\tau_4(t_{ij})$, $\tau_5^*(t_{ij}) = \eta(t_{ij})\beta(t_{ij})\tau_5(t_{ij})$, $\lambda_1^*(t_{ij}) = \eta(t_{ij})\lambda_1(t_{ij})$, $\lambda_3^*(t_{ij}) = \eta(t_{ij})\lambda_3(t_{ij})$. The budget correction term (μ_{ij}^B) can be viewed as an additional explanatory variable in Eq. (10).

5. Results

Following previous studies [47,43], we leverage the penalized spline (P-spline) smoothing approach introduced by Eilers and Marx [66] to estimate unknown coefficient functions in Eq. (10). See Appendix **A.1** for details on the P-spline approach. This section first presents results of endogeneity correction of advertising budget policies. Then we compare the fit of various model specifications, including the time-invariant model (baseline) and three variants of our proposed time-varying SEA response model. Finally, we report results of our advertising model and discuss potential implications. All covariates are standardized in our models.

5.1. Budget endogeneity correction

Table 3 presents the first stage results of our control function approach (i.e., Eq. 9), which corrects the potential endogeneity from strategical budget allocation policies. Results confirm advertisers' strategical budgeting decisions in SEA campaigns. Specifically, three exogenous variables–search demand, CPC and CTR–are all positive and statistically significant predictors for the advertising expenditure.

Theoretically, in the PPC scheme, the influence of search demand, CPC and CTR on the advertising expenditure should be similar and close to 1.0, because the expenditure can be computed as the product of these three factors. However, in our results, CPC appears to be the most influential factor for the advertising expenditure, followed by CTR and search demand. In other words, advertisers tend to emphasize on CPC and pay the least attention to search demand. This phenomenon is in line with the principles of information obtainability and least effort in information seeking behaviors [67]. On one hand, the principle of information obtainability states that, information that is more accessible to people is the more likely to be used by people, and vice versa. Similarly, according to the principle of least effort, when solving problems, a person tends to minimize her effort (over time). In the case of SEA, CPC has the highest obtainability for advertisers among the three factors, which is the most intuitive for them to understand and improve. This is because, for keywords with higher CPC (and bid prices), advertisers have to invest more in order to get sufficient opportunities to be displayed on SERPs and then clicked by search users.

By contrast, although advertisers often realize CTR's importance and have strong motivations to improve it, it takes much more time and effort to achieve a higher CTR and predict its temporal changes [36]. Also, precise information about search demand is challenging for advertisers to obtain during their campaigns. Even though some search engines or third-party companies (e.g., WordTracker) provide potential information about search demand in a certain market, it is generally difficult for ordinary advertisers to predict the future search demand on daily basis and adjust advertising policies accordingly in a real-time way.

5.2. Model fit comparisons

Instead of a specific knot selection process, P-Spline-based approaches only need a large enough knot number (see Appendix for details), yet there is no agreement on the optimal number of knots (K). Wand [68] suggested the lower number between 35 and $T/4$, where T denotes the number of distinctive measurement times. Ruppert [69] recommended that K around 10 is enough to estimate monotonic functions and K around 20 is needed for complicated functions. Our dataset is unbalanced with different assessment time points within and across individual ads (i.e., $1 \leq T \leq 958$). In order to estimate parameters of our SEA response model (Eqs. 9 and 10), we start with the B-Spline-based approach to fine-tune the analysis by incrementally increasing or decreasing the number of knots, and eventually use $K = 30$ in the P-Spline-based approach to estimate our model. We choose P-Splines over B-Splines because P-Splines can produce smoother estimates of the coefficient functions.

Next, we evaluate our time-varying model in terms of model fit by comparing it with several alternative specifications. The first alternative is a time-invariant model (MODEL-Time-Invariant), which treats coefficients of covariates in our full model (i.e., Eqs. 9 and 10) as time-invariant constants. We also compare three variants of our time-

Table 4
Model fit statistics of the proposed model and alternative specifications.

Model specifications	Trend specification	−2 Res Log Likelihood	AIC	BIC
MODEL-Time-Invariant	NA	−4,226,454	−4,226,431	4,226,268
MODEL-Time-Varying-linear	linear spline	−4,290,931	−4,290,335	−4,286,298
MODEL-Time-Varying-quadratic	quadratic spline	−4,297,012	−4,296,394	−4,296,394
MODEL-Time-Varying-cubic	cubic spline	**−4,300,328**	**−4,299,688**	**−4,295,353**

Table 5
Estimated parameters for the time-invariant model with budget endogeneity correction.

Coefficients	*Dependent Variables*		
	Sales		
	Estimates	*Conf. Int.*	*Std. Errors*
(Intercept)	0.015***	0.014–0.015	0.000
lagged Sales	0.695***	0.694–0.696	0.000
Ad Expenditure	0.001***	0.001–0.001	0.000
Ad Position	0.000***	0.000–0.000	0.000
CTR	0.035***	0.033–0.036	0.001
CVR	2.392***	2.387–2.396	0.002
Keyword Length	−0.007***	−0.007 to −0.007	0.000
Brand	0.002***	0.001–0.002	0.000
Retailer	0.053***	0.053–0.054	0.000
Holiday	−0.001	−0.003 to 0.002	0.001
B-Residuals	0.069***	0.068–0.069	0.000
Observations	5,649,220		
R^2/adj. R^2	0.693 / 0.693		

*** *p < .001.*

varying model specified with linear (MODEL-Time-Varying-linear), quadratic (MODEL-Time-Varying-quadratic) and cubic (MODEL-Time-Varying-cubic) spline functions, respectively. Table 4 illustrates model fit statistics for these various specifications, including twice the negative of the residual log likelihood (−2 Res Log Likelihood), the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).

As Table 4 shows, our time-varying advertising model specified with cubic splines provides the best fit, followed by MODEL-Time-Varying-quadratic and MODEL-Time-Varying-linear, while MODEL-Time-Invariant has the worst fit. In other words, including temporal dynamics helps time-varying models significantly improve their model fit compared to the time-invariant model. Theoretically, time-varying models break the study period into more fine-grained time intervals (rather than treat it as a single interval) and can reveal much more information about relationships between the explanatory variables and the dependent variable [43]. Also, the best fit by cubic spline function, compared to the linear and quadratic functions, highlights the dynamic complexity of SEA markets [70].

To examine the effect of each covariate on sales, we also analyze estimated parameters from our time-invariant model with budget endogeneity correction (i.e., treating coefficients of Eqs. 9 and 10 as constants) because such coefficients are easier to be understood and interpreted than coefficient functions. Results in Table 5 reveal two interesting findings: First, the budget correction term (μ_{ij}^{B}) has a statistically significant effect ($\alpha_1{}^* = 0.069$, $p < 0.001$), which justifies the addition of budget control function (i.e., Eq. 9) to the model. Second, compared to other covariates, the budget correction term explains a substantial part of the variance in the dependent variable (*Sales*). This suggests that, there are indeed some unobserved factors associated with advertisers' budgeting decisions. The positive effect of the budget correction term also implies that, without the budget correction process, parameter estimates of the advertising expenditure will be biased upwards, because the original model (in Eq. 8) omits unobserved factors that correlate with the advertising expenditure.

5.2.1. Carryover effect

From Table 5, we can see that the variable for lagged sales ($Sales_{ij-1}$) has a statistically significant and positive effect on sales ($\gamma^* = 0.695$, $p <$ 0.001). Also, the coefficient of lagged sales fluctuates between 0.648 and 0.774 over time (Fig. 2). Similar to traditional advertising, lagged sales remain a significant predictor for current sales in the context of SEA. This is in line with what has been reported in the literature [71,72,73]. The carryover effect is also higher in SEA than those reported in traditional advertising channels (e.g., via newspapers, radio, TV, and billboards) [55,29,74]. Although Dinner et al. [65] argued that the

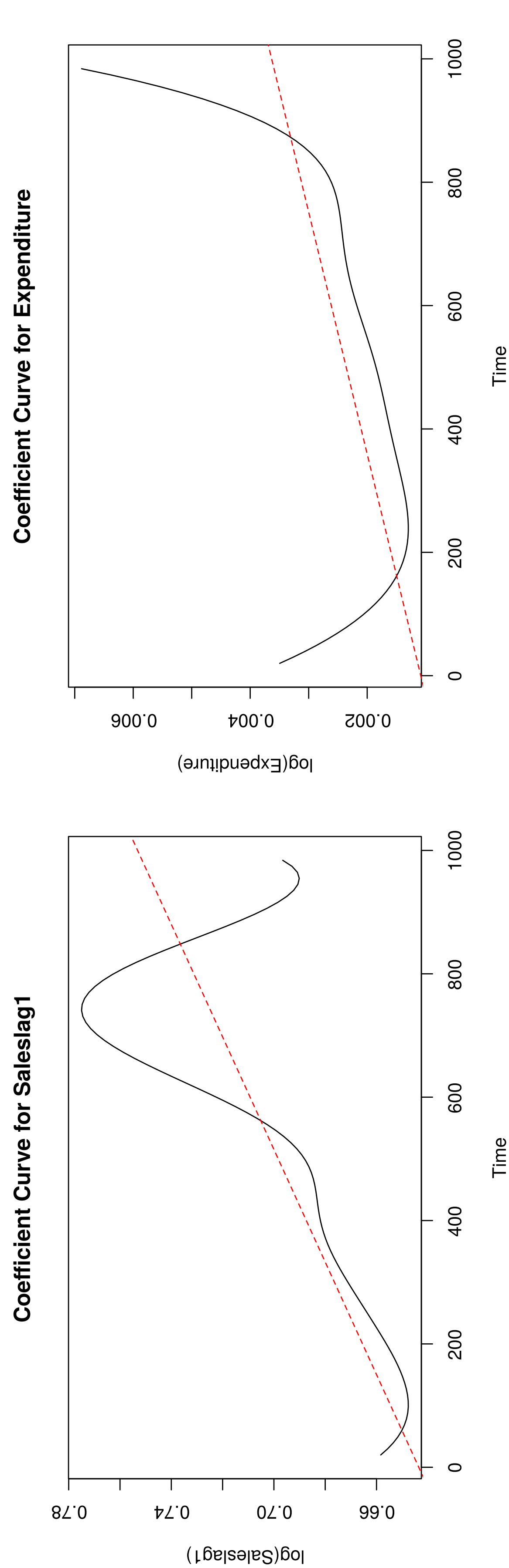

Fig. 2. Coefficient functions of lagged sales (left) and advertising expenditure (right). Straight lines represent linear curve fitting of the time-varying coefficient and illustrates the overall trend of the coefficient over time.

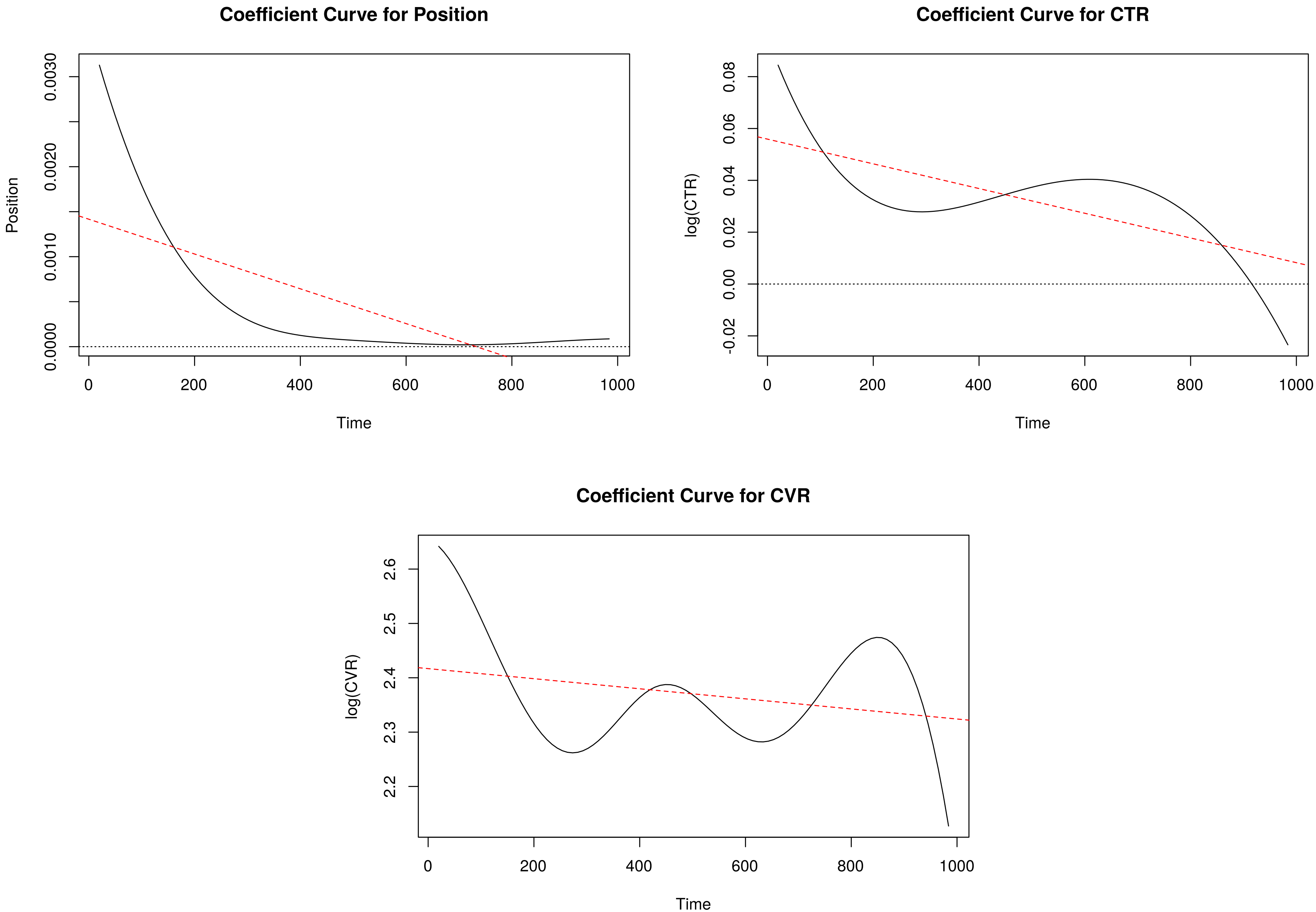

Fig. 3. Estimated coefficient functions of ad-specific factors (Straight lines represent linear curve fitting).

carryover effect is almost zero in SEA, their study was based on one apparel retailer, whose main revenue (nearly 85%) is generated through offline channels. However, in the e-commerce dataset we use, all sales driven by SEA are fulfilled online, which makes the estimation of carryover effect less biased. From a temporal perspective, the carryover effect is persistent and strong in SEA over time. In general, the carryover effect has an upward trend in SEA, as shown in Fig. 2. In addition, the trend is not monotonic in SEA, highlighting the complex dynamics in carryover effect.

5.2.2. The ad-sales relationship

The coefficient of independent variable *AdExpenditure* represents the short-term advertising elasticity. According to Table 5, the advertising expenditure has a statistically significant and positive effect on sales ($\beta^* = 0.001$, $p < 0.001$). The coefficient of the advertising expenditure fluctuates between 0.001 and 0.007 over time (Fig. 2). It drops first before increases and fits a linear line with a positive slope.

Overall, despite the statistical significance, the magnitude of the current advertising expenditure's effect on sales is small. This contradicts the commonly-held views in traditional advertising that the advertising expenditure is the major driving force to generate direct-response sales [29,30]. One possible reason for the difference is that millions of competing advertisers in SEA lead to more intense competitions [75,76] and thus lower advertising elasticity [65,30].

5.2.3. Ad positions, CTR and CVR

In this section, we investigate effects of advertisement characteristics (i.e., ad position) and consumer behavior measures (i.e., CTR and CVR) on sales (see Fig. 3).

Ad position is a statistically significant predictor of sales, but has a low coefficient ($\lambda_1^* = 0.0001397$ and $p < 0.001$). In other words, if an ad's position goes up by one unit, the contemporary sales is expected to increase by only 0.014%. According to Fig. 3, the effect of ad positions drops very fast over time. We also estimate a model with an additional quadratic term of ad position on ads listed on the first search engine results page (SERP). However, the new model still has close to zero coefficients for both ad position and its quadratic term. We will discuss the implications of this finding later in this paper.

Click-through rate (CTR) is a statistically significant and positive predictor of sales ($\tau_2^* = 0.035$, $p < 0.001$). Such a low coefficient is not surprising, because click-through is primarily a consequence of the brand building in online advertising. Moreover, Fig. 3 shows that the influence of CTR on sales declines over the promotion period. Its effect on sales even becomes negative at the final stage–as the campaigns move on, that is, ads with a lower CTR can produce more sales than those with a higher CTR. This is probably because at the later stage of a SEA campaign, certain search users get to know the advertiser and their ads better, and consider the advertiser's SEA ads as a quality source for products they desire. Consequently, when they do click an ad from the advertiser, they tend to purchase a higher amount of products [77].

Conversion rate (CVR) is a statistically significant predictor of sales, with a stronger effect on sales ($\lambda_3^* = 2.392$, $p < 0.001$) than CTR. Meanwhile, CVR's influence on sales fluctuates over time.

5.2.4. Control variables

To represent latent advertising quality, we add control variables (Fig. 4) to the model. **Keywords length** negatively influences sales ($\tau_2^* = -0.007$, $p < 0.001$)–a longer and more specific keyword leads to fewer sales than shorter and more general keywords. Even though consumers who search for more general keywords usually have lower CVR, a more general keyword can trigger much more clicks, leading to more sales. As for what is in a keyword, containing a **brand** and containing a **retailer** in an ad are positive predictors of sales while containing a **holiday** is not. Also, all control variables' coefficient functions trend downward over time.

5.3. Accounting for unobservable factors at the individual level

To account for unobservable factors at the individual level, we ran a separate panel model with fixed effects. Even though the model does not provide time-varying coefficients, it would serve as a good robustness check for results from our TVC framework. Note that, in our model, i refers to the index of subjects (i.e., advertisements) and j refers to the index of observations. In other words, for different advertisements, the same j may correspond to different time of observations. For example, one ad was observed during the period from Jan. 1 to Jan. 10, while another ad was observed during the period of from Jan. 3 to Jan. 15. For example, $j = 1$ would correspond to Jan. 1 for the former ad and Jan. 3 for the latter ad. Hence, adding fixed effects on j makes little sense. Therefore, we estimate fixed-effects and random-effects models on advertisements (i.e., i) and conducted the Hausman test to choose between the two models. The results reveal that the Hausman test statistic is significant with $p < 0.001$, suggesting that the fixed-effects model is an appropriate choice for our research. Results of the panel model with fixed effects on advertisements (Table 6) are consistent with the time-invariant model and the TVC model.

5.4. Accounting for heteroscedasticity and nonindependence among observations

To account for heteroscedasticity and nonindependence among observations, we cluster standard errors by advertisements for the panel model with fixed effects. Our dataset is organized by each advertisement on a daily basis and our model (Eq. 10) is estimated at the individual advertisement level. While it makes little sense to cluster errors at the level of j, in order to further check the results of our model, we cluster the standard errors at the level i to account for heteroscedasticity and nonindependence among observations across days for each advertisement. Table 7 presents the estimated coefficients of the fixed effects model and robust standard errors clustered by advertisement. The results show that the estimated coefficients remain consistent, while robust standard errors clustered by advertisement become larger, compared to those of the fixed effects model.

6. Conclusions

6.1. Theoretical implications

Our research has theoretical implications to the research of SEA.

First, this study contributes to the SEA literature by taking into account the temporal variations in the effectiveness of various types of SEA factors and addressing an important gap in the literature. SEA is one of the most dynamic advertising environments with interactive behaviors between users and advertisers, auction processes and mechanisms. This research contributes to our understanding of dynamics in SEA. Although previous research (e.g., [33,37,34]) reported non-linear effects of ad advertising variables, our research is the first to reveal the time-varying pattern of these effects. Moreover, our model implicitly encapsulates the concept of advertising quality score as a latent variable by adopting a quality-adjustment structure. This allows us to explore influence trajectories of advertising expenditure and various related factors on the expected market outcome (i.e., sales) over time. In addition, our results reveal non-linear pattern in the temporal effects of various key factors in SEA on sales. This finding can inform future studies of temporal dynamics in SEA.

Second, this study adds to the broader line of research on advertising response models by capturing the dynamic ad-sales relationship in the SEA context. Conceptually, different advertising forms may be described by different ad-response models distinct by the underlying mechanisms and inherent advertising variables. Although prior research has adapted the Vidale-Wolfe model [28] to the SEA context for the purpose of supporting optimal budget allocation, their model fails to incorporate

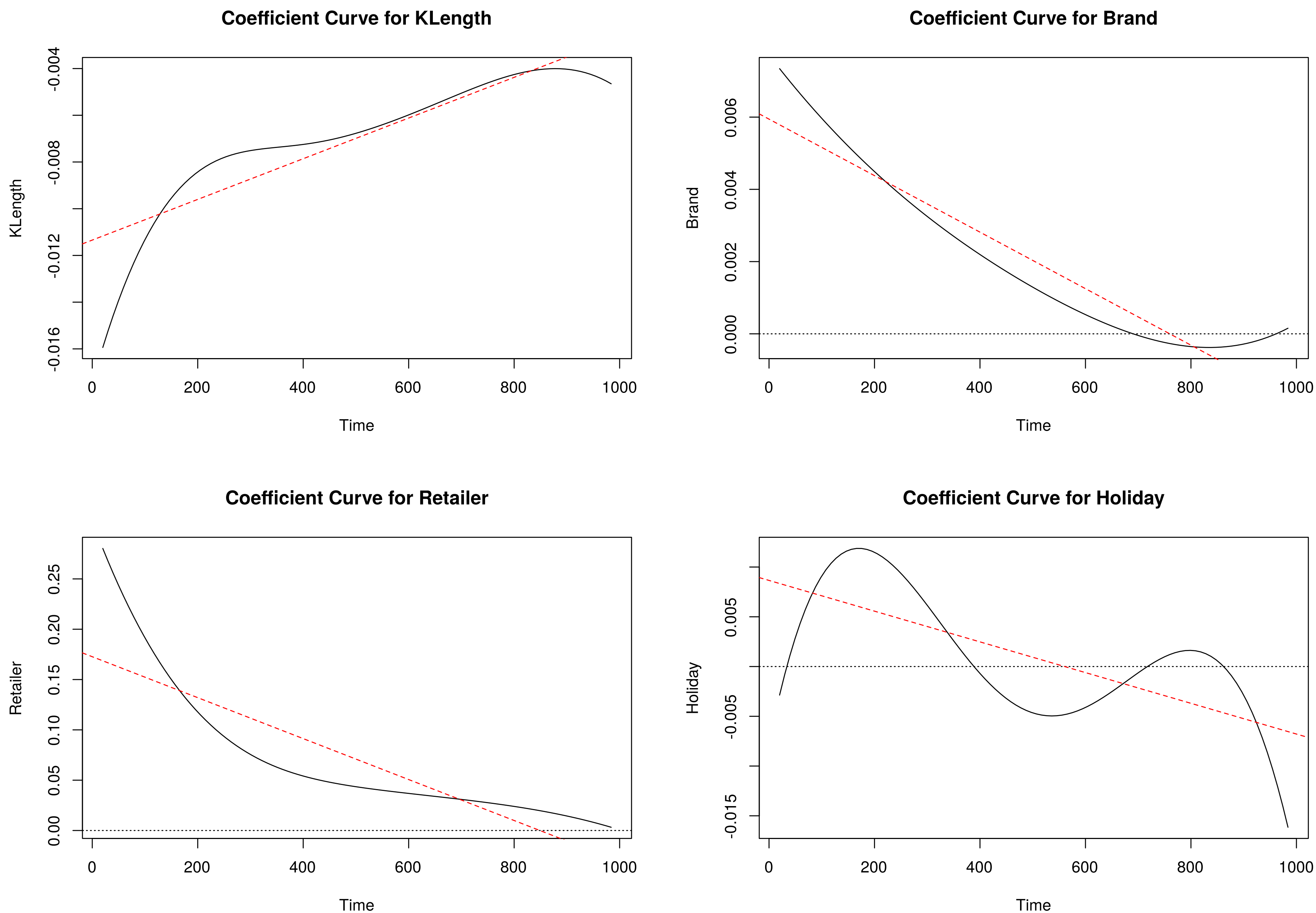

Fig. 4. Estimated coefficient functions of control variables (Straight lines represent linear curve fitting).

Table 6
Results of the panel model with fixed effects on advertisements.

Coefficients	*Dependent Variables*		
	Sales		
	Estimates	*Conf. Int.*	*Std. Errors*
lagged Sales	0.396***	0.396–0.397	0.000
Ad Expenditure	0.002***	0.002–0.002	0.000
Ranking Position	−0.000***	−0.000 to −0.000	0.000
CTR	0.024***	0.022–0.025	0.001
CVR	2.404***	2.400–2.408	0.002
B-Residuals	0.064***	0.063–0.064	0.000
Observations	5,649,220		
R^2/adj. R^2	0.39053 / 0.38534		

*** $p < .001$.

Table 7
Results of the panel model with fixed effects and robust standard errors clustered by advertisements.

Coefficients	*Dependent Variables*		
	Sales		
	Estimates	*Conf. Int.*	*Std. Errors*
lagged Sales	0.396***	0.351–0.442	0.023
Ad Expenditure	0.002***	0.001–0.002	0.000
Ranking Position	−0.000***	−0.000 to −0.000	0.000
CTR	0.024***	0.016–0.032	0.004
CVR	2.404***	2.345–2.463	0.030
B-Residuals	0.064***	0.059–0.068	0.002
Observations	5,649,220		
R^2/adj. R^2	0.39053 / 0.38534		

*** $p < .001$.

more advertising variables except for a quality score index, due to the limitation of differential-equation modeling structure. Our response model not only incorporates rich advertising features of SEA, but also is capable of handling the time-varying effects of various factors.

Moreover, our research empirically compares short-term advertising elasticity and carryover effect on SEA sales based on a large-scale dataset from a major U.S. e-commerce retailer. Results show that the carryover effect is stronger than the short-term elasticity and suggest that SEA is not a direct-response advertising medium. Instead, advertisers need to be patient and make a longer-term advertising investment before getting returns in sales. This also calls for more research on long-term strategies for SEA including budget allocation, bid pricing and keywords selections, because ordinary advertisers have little knowledge and time to operate such sophisticated and dynamical campaigns in the long run.

Third, our research also finds important patterns on how advertisers make budget decisions in SEA. We find that SEA advertisers mainly consider CPC, instead of CTR and search demand, when making such decisions. More explorations on advertisers' behaviors in different advertising schemes can help search engines improve their market design.

Last, our research helps to better understand the inefficiency in the current SEA scheme. Prior studies [33,37] have found that higher positions on SERPs are not necessarily the more profitable ones for advertisers. Our research finds one potential reason for this. Advertising performance evaluation based on CTR is inevitability biased, because SEA campaigns experience a significant, positive and increasing carryover effect. In addition, the effects of CVR and CTR on sales make it possible to design a hybrid advertising scheme that combines pay-per-click and pay-per-action.

6.2. Practical implications

This research provides several practical implications for SEA advertisers. *First*, our findings could serve as the basis for designing a decision-support tool that helps advertisers better understand the ad-sales relationships, especially the influence of various factors on sales, and how their influences change over time. More importantly, the tool based on our model entitles advertisers to predict advertising performance and allocate their advertising resources accordingly in a real-time fashion for their SEA campaigns. For instance, our time-varying response model can be used to generate close-loop budget strategies over time via developing an optimal control model of budget planning (e.g., [8]). However, budget optimization is beyond the scope of this research.

Second, for advertisers, focusing only on the direct and immediate effect of the expenditure on sales would underestimate the performance of their SEA campaigns in terms of sales, because SEA features a significant carryover effect that is more influential than the immediate effect. In other words, the effect of the advertising expenditure on sales would be overestimated without considering lagged sales (Table 2). In practice, given the temporal dynamics in the carryover effect in SEA, advertisers could consider increasing/decreasing their advertising budget when the coefficient of the carryover effect is on the rise/decline, in order to get bigger "bang of the buck".

Third, our research also reveals the effects of advertisement characteristics and consumer behavior measures on SEA sales. For instance, advertisers may want to carefully evaluate the return on investment of bidding for higher ad positions on SERPs. Our finding offers empirical evidence that always bidding for higher ad positions may lead to limited increase in sales and possibly hurt return on investment. Different from search engines' traditional stance, our results could offer alternative advertising strategies.

Among measures of consumers behaviors, besides CTR, SEA advertisers should also pay extra attention to CVR, because it is directly associated with sales. Moreover, it is more important for an advertiser to improve its CVR during the initial stage of its SEA campaigns, because the influence of CVR gradually declines over time.

Last but not the least, advertisers can adjust keyword selection strategies over time–they can focus on shorter and more general keywords in the initial stage of a campaign, and then increase the portion of longer and more specific keywords over time. It also helps to improve sales if an advertiser can include more retailer-specific and brand-specific keywords during the initial stage of a campaign.

6.3. Limitations and future research

We acknowledge several limitations of our research. First, similar to most, if not all, studies using advertising response models, we investigate the ad-sales relationship in SEA at the campaign level, rather than consumer behaviors at the individual level. The former is about how to allocate resources on advertising campaigns, while the latter focuses on how an advertiser should bid for a keyword in an auction against rivals. Our focus on the former means that our model cannot discern the heterogeneity inherent in behaviors of individual advertisers and their competitors.

Second, our study is limited by the dataset we used. For example, the dependent variable in our study is sales measured by the units of products sold from SEA campaigns. Sales is certainly important for advertisers and is often considered more valuable than clicks [1]. However, return on investment is often more important in business [78], because it combines the advertising expenditure and the profit from transactions, and can be a more straightforward way to measure an advertisers' financial gains. Besides monetary outcomes, some advertisers may also value the positive image of their brand gained from SEA campaigns. At the same time, even though our dataset is large in scale and covers an extended period of time, whether our conclusions apply to SEA in other contexts needs further investigations.

This research can be extended in several ways. One direction is to systematically understand the temporal pattern of each factor on sales, so that dynamic strategies for optimal resource allocation can be designed. Based on such a time-varying advertising response model, the discrete-time optimal control technology can be utilized to get a close-loop solution for resources allocation for SEA campaigns. Then formal field experiments can be designed and executed to improve model fits and support advertising resource allocation in real time. Second, it will be interesting to adopt the panel vector autoregression (PVAR) model to explore dynamic inter-dependency between a set of advertising factors in the SEA context, which could significantly enhance our understanding of the underlying mechanism currently adopted by major search engines. In particular, the PVAR analysis can also be supplemented with generalized forecast error variance decomposition to shed light on the relative power of each advertising variable and impulse response function to visualize the dynamical relationship between each pair of variables. Third, we plan to extend our model to include the continual bidding process and advertisers' behaviors at the individual level. In addition, time-varying interactions between advertisers and consumers should be an interesting topic to explore in the field of SEA as well.

CRediT authorship contribution statement

Yanwu Yang: Conceptualization, Methodology, Investigation, Writing – original draft, Project administration, Resources. **Kang Zhao:** Conceptualization, Investigation, Writing – review & editing, Project administration. **Daniel Dajun Zeng:** Conceptualization, Writing – review & editing. **Bernard Jim Jansen:** Conceptualization, Writing – review & editing, Resources.

Data availability

The authors do not have permission to share data.

Acknowledgements

We are thankful to the associate editor and anonymous reviewers who provided valuable suggestions that led to a considerable improvement in the organization and presentation of this manuscript. This work is partially supported by the (NSFC National Natural Science Foundation of China) grants (72171093, 71672067).

Appendix A

A.1. Estimation of the Time-varying search advertising response model

In the following we provide the estimation of the time-varying search advertising response model described in Section 4.1. P-splines have several very attractive merits. First, they do not impose any assumption on the changing pattern of a given explanatory variable with respect to time t (e.g., linear, quadratic, or cubic), which makes the estimated model immune to the misspecification problem [43]. Second, compared to smoothing approaches (e.g., regression splines, B-splines), P-splines have no boundary effects, can conserve moments of data and have polynomial fits as limits, and their computation are relatively inexpensive [66]. Accordingly, P-splines have been widely used in marketing literature on semi-parametric models (e.g., Stremersch and Lemmens [47], Saboo et al. [46]).

The general idea behind splines-based smoothers is that any smoothly varying (coefficient) function (e.g., $f(t)$) defined on a certain interval can be approximated by a linear combination of lower order polynomial base functions. Specifically, the interval is partitioned into $K+1$ smaller intervals, which are determined by K dividing points (i.e., knots), $\tau_1, \tau_2, \ldots, \tau_K$; then we can approximate $f(t)$ within each small interval $[\tau_r, \tau_{r+1})$, $0 \le r \le K$ with lower order polynomial functions. In the case of time-varying coefficient functions [43], the q-order truncated power basis can be specified as

$$t^0, t^1, t^2, \ldots, t^q, (t-\tau_1)_+^q, \ldots, (t-\tau_K)_+^q \\ with \ \ (t-\tau)_+^q = \begin{cases} 0 \ \text{ if } \ t \le \tau \\ (t-\tau)^q \ \text{ if } \ t > \tau \end{cases}, \tag{A1}$$

where the first $q+1$ functions are the 0,1,2, …, q order power functions of t, and the other K functions are truncated q order power functions determined by the k knots, respectively.

In practice, researchers need to specify the number of knots K. However, the number of knots is less crucial in P-splines based estimation approaches, because they can optimally estimate the coefficients using the linear mixed-effects model. Thus, theoretically we can only choose a large enough K (e.g., 10, which also depends on the number of distinctive measurement times) for P-splines [43].

As an example, the coefficient function $\alpha_0(t_{ij})$ can be approximately represented as

$$\alpha_0(t_{ij}) = a_0 + a_1 t_{ij} + a_2 t_{ij}^2 + \sum_{k=1}^{K} a_{q+k}(t_{ij} - \tau_k)_+^q. \tag{A2}$$

By substituting a set of coefficient functions of t into the original model to be estimated, we can get a linear regression model with these base functions (such as 1, t_{ij}, t_{ij}^2, …, $(t_{ij}-\tau_k)_+{}^q$) as covariates and $a_0, a_1, a_2, \ldots, a_{q+k}$ as coefficients, which can be easily estimated with ordinary least square (OLS). P-splines combine B-splines with different penalties on estimated coefficients, i.e., using "a simple difference penalty on the coefficients themselves of adjacent B-splines" [66], in order to address the overfitting problem. The approach suggested by Ruppert [69] and Wand [68] shrinks the coefficients of coefficient functions (e.g., a_{q+k}, $k = 1, 2, \ldots, K$ in Eq. 10) towards zero, by minimizing the sum of SSE (sum of squared errors) and the penalty term (defined as the summation of a series of products of coefficients and corresponding tuning parameters), i.e., $SSE + \lambda_1 \sum_{k=1}^{K} a_{q+k}(t_{ij}-\tau_k)_+{}^q + \ldots$. The resulting optimal tuning terms (e.g., λ_1) balance the tradeoff between the goodness of fit and the smooth of the estimated functions. Thus, the penalty term could prevent these coefficients from being too large in absolute value. Wand [68] developed an approach that treats these coefficients as random variable with normal distribution, and expands the model to be estimated into a linear mixed-effect model, which can be estimated with the restricted maximum likelihood (REML) to the optimal balance. For more details on the P-splines estimation of non- and semi-parametric models, see [66,69,68].

References

[1] H. Sun, M. Fan, Y. Tan, An empirical analysis of seller advertising strategies in an online marketplace, Inf. Syst. Res. 31 (1) (2020) 37–56.
[2] S. Yao, C.F. Mela, A dynamic model of sponsored search advertising, Mark. Sci. 30 (3) (2011) 447–468.
[3] Comscore, Why Google's Surprising Paid Click Data Are Less Surprising, Accessed September 3, 2021,, https://www.comscore.com/lat/Prensa-y-Eventos/Blog/Why-Google-s-surprising-paid-click-data-are-less-surprising, 2008.
[4] Google Adwords, About Quality Score, Accessed September 3, 2021, https://support.google.com/google-ads/answer/7050591?hl=en, 2019.
[5] Adaplo, Google Shopping Campaigns Optimisation, Accessed September 3, 2021,, https://adaplo.com/google-shopping-optimisation, 2019.
[6] C.N. Da Silva, Best Practices for Calculating and Mastering Your PPC Budget, Accessed September 3, 2021, https://www.acquisio.com/blog/agency/best-practices-calculating-mastering-ppc-budget/, 2018.
[7] T. George, How to use amazon advertising's dynamic bidding feature, Accessed September 3, 2021, https://searchengineland.com/how-to-use-amazon-advertisings-dynamic-bidding-feature-320505, 2019.
[8] Y. Yang, D. Zeng, Y. Yang, J. Zhang, Optimal budget allocation across search advertising markets, INFORMS J. Comput. 27 (2) (2015) 285–300.
[9] S. Ye, G. Aydin, S. Hu, Sponsored search marketing: dynamic pricing and advertising for an online retailer, Manag. Sci. 61 (6) (2015) 1255–1274.
[10] X. Zhang, J. Feng, Cyclical bid adjustments in search-engine advertising, Manag. Sci. 57 (9) (2011) 1703–1719.
[11] T. Blake, C. Nosko, S. Tadelis, Consumer heterogeneity and paid search effectiveness: A large-scale field experiment, Econometrica 83 (1) (2015) 155–174.
[12] J. Liu, S. Hill, Frontiers: moment marketing: measuring dynamics in Cross-Channel ad effectiveness, Mark. Sci. 40 (1) (2021) 13–22.
[13] N. Pabich, 5 Long Term Benefits of PPC Advertising, Accessed September 3, 2021, https://www.digitalthirdcoast.com/blog/5-long-term-benefits-of-ppc-advertising, 2011.
[14] J. Baadsgaard, What Can CTR Tell Me About My Campaigns?, Accessed September 3, 2021,, https://www.disruptiveadvertising.com/adwords/what-is-ctr-click-through-rate/, 2017.
[15] J. Challis, Short Term & Long Term PPC Advertising Goals, Accessed September 3, 2021,, https://www.koozai.com/blog/pay-per-click-ppc/short-term-long-term-pay-per-click-advertising-goals/, 2014.
[16] A. Membrillo, 13 PPC Trends to Increase Your CTR and Sales, Accessed September 3, 2021,, https://www.cardinaldigitalmarketing.com/blog/13-ppc-trends-to-increase-your-ctr-and-sales/, 2018.
[17] P.A. Naik, Marketing dynamics: A primer on estimation and control, Foundations and Trends in Marketing 9 (3) (2015) 175–266.
[18] C. Köhler, M.K. Mantrala, S. Albers, V.K. Kanuri, A Meta-analysis of marketing communication carryover effects, J. Mark. Res. 54 (6) (2016) 990–1008.
[19] W.R. Vanhonacker, Carryover effects and temporal aggregation in a partial adjustment model framework, Mark. Sci. 2 (3) (1983) 297–317.
[20] E. Bayer, S. Srinivasan, E.J. Riedl, B. Skiera, The impact of online display advertising and paid search advertising relative to offline advertising on firm performance and firm value, Int. J. Res. Mark. 37 (4) (2020) 789–804.
[21] N. Brooks, H. Magun, Navigational behaviour and sponsored search advertising, Int. J. Electron. Bus. 6 (2) (2008) 132–148.
[22] S. Ceri, A. Bozzon, M. Brambilla, E.D. Valle, P. Fraternali, S. Quarteroni, Advertising in search, in: Web Information Retrieval, Springer, Berlin, Heidelberg, 2013, pp. 121–133.
[23] J. Feldman, S. Muthukrishnan, Algorithmic methods for sponsored search advertising, in: Performance Modeling and Engineering, Springer, Boston, MA, 2008, pp. 91–122.
[24] L. Xu, J. Chen, A. Whinston, Effects of the presence of organic listing in search advertising, Inf. Syst. Res. 23 (4) (2012) 1284–1302.
[25] M. Fischer, S. Albers, N. Wagner, M. Frie, Dynamic marketing budget allocation across countries, products, and marketing activities, Mark. Sci. 30 (4) (2011) 568–585.
[26] C. Anderson, Google's Long Tail, Accessed May 10, 2021, http://www.longtail.com/the_long_tail/2005/02/googles_long_ta.html, 2005.
[27] R.Y. Du, Y. Hu, S. Damangir, Leveraging trends in online searches for product features in market response modeling, J. Mark. 79 (1) (2015) 29–43.
[28] Y. Yang, B. Feng, D. Zeng, Learning parameters for a generalized Vidale-Wolfe response model with flexible ad elasticity and word-of-mouth, IEEE Intell. Syst. 36 (5) (2021) 69–79.
[29] G. Assmus, J.U. Farley, D.R. Lehmann, How advertising affects sales: meta-analysis of econometric results, J. Mark. Res. (1984) 65–74.
[30] R. Sethuraman, G.J. Tellis, R.A. Briesch, How well does advertising work? Generalizations from Meta-analysis of brand advertising Elasticities, J. Mark. Res. 48 (3) (2011) 457–471.
[31] G. Johnson, R.A. Lewis, E. Nubbemeyer, The Online Display ad Effectiveness Funnel & Carryover: Lessons from 432 Field Experiments, 2017 (Available at SSRN 2701578).
[32] N. Archak, V. Mirrokni, S. Muthukrishnan, Budget optimization for online campaigns with positive carryover effects, in: International Workshop on Internet and Network Economics, Springer, Berlin, Heidelberg, 2012, pp. 86–99.
[33] A. Agarwal, K. Hosanagar, M.D. Smith, Location, location, location: an analysis of profitability of position in online advertising markets, J. Mark. Res. 48 (6) (2011) 1057–1073.
[34] B.J. Jansen, Z. Liu, Z. Simon, The effect of ad rank on the performance of keyword advertising campaigns, J. Assoc. Inf. Sci. Technol. 64 (10) (2013) 2115–2132.
[35] S. Yang, A. Ghose, Analyzing the relationship between organic and sponsored search advertising: positive, negative, or zero interdependence? Mark. Sci. 29 (4) (2010) 602–623.
[36] Y. Yang, P. Zhai, Click-through rate prediction in online advertising: A literature review, Inf. Process. Manag. 59 (2) (2022), 102853.
[37] A. Ghose, S. Yang, An empirical analysis of search engine advertising: sponsored search in electronic markets, Manag. Sci. 55 (10) (2009) 1605–1622.
[38] P. Jeziorski, S. Moorthy, Advertiser prominence effects in search advertising, Manag. Sci. 64 (3) (2017) 1365–1383.
[39] M. Zhuang, E. Fang, J. Lee, X. Li, The effects of price rank on clicks and conversions in product list advertising on online retail platforms, Inf. Syst. Res. 32 (4) (2021) 1412–1430.
[40] V. Abhishek, K. Hosanagar, Optimal bidding in multi-item multislot sponsored search auctions, Oper. Res. 61 (4) (2013) 855–873.
[41] Y. Yang, J. Zhang, R. Qin, J. Li, F.Y. Wang, W. Qi, A budget optimization framework for search advertisements across markets, IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and Humans 42 (5) (2012) 1141–1151.
[42] J. Zhang, Y. Yang, X. Li, R. Qin, D. Zeng, Dynamic dual adjustment of daily budgets and bids in sponsored search auctions, Decis. Support. Syst. 57 (2014) 105–114.
[43] X. Tan, M.P. Shiyko, R. Li, Y. Li, L. Dierker, A time-varying effect model for intensive longitudinal data, Psychol. Methods 17 (1) (2012) 61–77.
[44] T. Hastie, R. Tibshirani, Varying-coefficient models, Journal of the Royal Statistical Society. Series B (Methodological) (1993) 757–796.
[45] E.C. Osinga, P.S. Leeflang, J.E. Wieringa, Early marketing matters: A time-varying parameter approach to persistence modeling, J. Mark. Res. 47 (1) (2010) 173–185.
[46] A.R. Saboo, V. Kumar, I. Park, Using big data to model time-varying effects for marketing resource (Re) allocation, MIS Q. 40 (4) (2016) 911–939.
[47] S. Stremersch, A. Lemmens, Sales growth of new pharmaceuticals across the globe: the role of regulatory regimes, Mark. Sci. 28 (4) (2009) 690–708.
[48] Z. Katona, Y. Zhu, Quality Score That Makes You Invest, Available at SSRN. https://ssrn.com/abstract=2954707, 2018.
[49] J. Chen, J. Stallaert, An economic analysis of online advertising using behavioral targeting, MIS Q. 38 (2) (2014) 429–A7.
[50] J. Feng, H.K. Bhargava, D.M. Pennock, Implementing sponsored search in web search engines: computational evaluation of alternative mechanisms, INFORMS J. Comput. 19 (1) (2007) 137–148.
[51] O.J. Rutz, R.E. Bucklin, From generic to branded: A model of spillover in paid search advertising, J. Mark. Res. 48 (1) (2011) 87–102.
[52] S. Yang, J. Pancras, Y. Song, Broad or exact? Search ad matching decisions with keyword specificity and position, Decis. Support. Syst. 143 (2021), 113491.
[53] J.D. Little, BRANDAID: A marketing-mix model, part 1: structure, Oper. Res. 23 (4) (1975) 628–655.
[54] L.J. Parsons, R.L. Schultz, Marketing Models and Econometric Research, North-Holland Publishing Company, New York, 1976.
[55] S.J. Arnold, T.H. Oum, B. Pazderka, D.W. Snetsinger, Advertising quality in sales response models, J. Mark. Res. (1987) 106–113.
[56] M. Ohta, Production technologies of the US boiler and Turbogenerator industries and hedonic Price indexes for their products: a cost-function approach, J. Polit. Econ. 83 (1) (1975) 1–26.
[57] D.G. Clarke, Econometric measurement of the duration of advertising effect on sales, J. Mark. Res. (1976) 345–357.
[58] D.S. Tull, The carry-over effect of advertising, J. Mark. (1965) 46–53.
[59] R.J. Caballero, E.M. Engel, Beyond the partial-adjustment model, Am. Econ. Rev. (1992) 360–364.
[60] D.G. Clarke, Sales-advertising cross-elasticities and advertising competition, J. Mark. Res. (1973) 250–261.
[61] Y. Yang, B. Feng, J. Salminen, B.J. Jansen, Optimal advertising for a generalized Vidale–Wolfe response model, Electronic Commerce Research (2021), https://doi.org/10.1007/s10660-021-09468-x forthcoming.
[62] P.E. Rossi, Even the rich can make themselves poor: A critical examination of IV methods in marketing applications, Mark. Sci. 33 (5) (2014) 655–672.
[63] Y.J. Luan, K. Sudhir, Forecasting marketing-mix responsiveness for new products, J. Mark. Res. 47 (3) (2010) 444–457.
[64] A. Petrin, K. Train, A control function approach to Endogeneity in consumer choice models, J. Mark. Res. 47 (1) (2010) 3–13.
[65] I.M. Dinner, H. Heerde Van, J., & Neslin, S. A., Driving online and offline sales: the cross-channel effects of traditional online display and paid search advertising, J. Mark. Res. 51 (5) (2014) 527–545.
[66] P.H. Eilers, B.D. Marx, Flexible smoothing with B-splines and penalties, Stat. Sci. (1996) 89–102.
[67] B.J. Jansen, S.Y. Rieh, The seventeen theoretical constructs of information searching and information retrieval, J. Assoc. Inf. Sci. Technol. 61 (8) (2010) 1517–1534.
[68] M.P. Wand, Smoothing and mixed models, Comput. Stat. 18 (2) (2003) 223–249.
[69] D. Ruppert, Selecting the number of knots for penalized splines, J. Comput. Graph. Stat. 11 (4) (2002) 735–757.
[70] Y. Yang, X. Li, D. Zeng, B.J. Jansen, Aggregate effects of advertising decisions: A complex systems look at search engine advertising via an experimental study, Internet Res. 28 (4) (2018) 1079–1102.
[71] P.A. Naik, K. Raman, Understanding the impact of synergy in multimedia communications, J. Mark. Res. 40 (4) (2003) 375–388.

[72] A. Babić Rosario, F. Sotgiu, K. De Valck, T.H. Bijmolt, The effect of electronic word of mouth on sales: A Meta-analytic review of platform, product, and metric factors, J. Mark. Res. 53 (3) (2016) 297–318.
[73] D.L. Weiss, C.B. Weinberg, P.M. Windal, The effects of serial correlation and data aggregation on advertising measurement, J. Mark. Res. (1983) 268–279.
[74] P.J. Danaher, A. Bonfrer, S. Dhar, The effect of competitive advertising interference on sales for packaged goods, J. Mark. Res. 45 (2) (2008) 211–225.
[75] J. Golden, J.J. Horton, The effects of search advertising on competitors: an experiment before a merger, Manag. Sci. 67 (1) (2021) 342–362.
[76] Y. Yang, Y.C. Yang, D. Liu, D.D. Zeng, Dynamic Budget Allocation in Competitive Search Advertising, Available at SSRN. https://ssrn.com/abstract=2912054, 2016.
[77] H. Garcia-Molina, G. Koutrika, A. Parameswaran, Information seeking: convergence of search, recommendations, and advertising, Commun. ACM 54 (11) (2011) 121–130.
[78] M. Lash, K. Zhao, Early predictions of movie success: the who, what, and when of profitability, J. Manag. Inf. Syst. 33 (3) (2016) 874–903.

Yanwu Yang is a full professor and the head of the Internet Sciences and Economic Computing research center at the School of Management, Huazhong University of Science and Technology, China. He received his Ph.D. degree in computer science from the graduate school of École Nationale Supérieure des Arts et Métiers, France. His research interests include e-commerce, recommender systems, web personalization, and computational advertising.

Kang Zhao is an Associate Professor of Business Analytics at Tippie College of Business, The University of Iowa. His current research focuses on data science and social computing, especially the mining, predictive modeling, and simulation of social, business, and collaboration networks. His research has been featured in public media from more than 25 countries. He also served as the Chair for INFORMS Artificial Intelligence Section 2014–2016. He earned his Ph.D. from Penn State University.

Danile Dajun Zeng received the M.S. and Ph.D. degrees in industrial administration from Carnegie Mellon University. He is a professor at the Institute of Automation, Chinese Academy of Sciences. He was a Gentile Family Professor of MIS at the Department of Management Information Systems, the University of Arizona. His research interests include social computing, recommender systems, intelligence and security informatics, infectious disease informatics, applied operations research, and game theory. He has published one monograph as well as more than 300 peer-reviewed articles and coedited 22 books and proceedings. His-work has been published in major information systems journals such as MISQ, INFORMS JOC, DSS, JASIST, IEEE TKDE, among others. Now he serves as the editor-in-chief of ACM Transactions on Management Information Systems. He is a fellow of the IEEE and the AAAS.

Bernard Jim Jansen is a Principal Scientist in the Social Computing group of the Qatar Computing Research Institute, and a Professor with the College of Science and Engineering, Hamad Bin Khalifa University, and an Adjunct Professor with the College of Information Sciences and Technology at The Pennsylvania State University. He has authored or co-authored 300 or so research publications, with articles appearing in a multi-disciplinary range of journals and conferences. He is editor-in-chief of the journal Information Processing & Management.